Centrum voor Wiskunde en Informatica

REPORT *RAPPORT*

The algebra of modal logic

P. Blackburn, M. de Rijke, Y. Venema

Computer Science/Department of Software Technology

Report CS-R9463 November 1994

CWI is the National Research Institute for Mathematics and Computer Science. CWI is part of the Stichting Mathematisch Centrum (SMC), the Dutch foundation for promotion of mathematics and computer science and their applications.
SMC is sponsored by the Netherlands Organization for Scientific Research (NWO). CWI is a member of ERCIM, the European Research Consortium for Informatics and Mathematics.

P.O. Box 94079, 1090 GB Amsterdam (NL)
Kruislaan 413, 1098 SJ Amsterdam (NL)
Telephone +31 20 592 9333
Telefax +31 20 592 4199

The Algebra of Modal Logic

Patrick Blackburn[1], Maarten de Rijke[2] and Yde Venema[3]

[1] *Computerlinguistik, Universität des Saarlandes*
Postfach 1150, D-66041 Saarbrücken, Germany
patrick@coli.uni-sb.de

[2] *CWI, P.O. Box 94079, 1090 GB Amsterdam, The Netherlands*
mdr@cwi.nl

[3] *Dept. of Mathematics and Computer Science, Free University*
De Boelelaan 1081a, 1081 HV Amsterdam, The Netherlands
yde@cs.vu.nl

Abstract

Our main aim is to review the frame semantics and axiomatics of modal logic from the perspective of the duality between (Kripke) frames and boolean algebras with operators as defined by Jónsson and Tarski.

To this end, we introduce modal languages and their interpretation in models and frames in Part II. We define and discuss the notion of a modal formula characterizing a class of frames or models, and give the Sahlqvist algorithm which yields, given a suitable modal formula as input, the corresponding first-order condition on the class of frames characterized by the formula. We define the concept of a normal modal logic and explain the canonical frame method for proving completeness of a logic with respect to classes of frames.

In Part III we develop the algebraic perspective on modal logic. We introduce boolean algebras with operators and show how they arise naturally in both the semantic and the axiomatic approach towards algebraizing modal logic. We discuss in detail how the category of boolean algebras with operators and homomorphisms links up with the category of frames with so-called bounded morphisms. Finally, we apply this duality to give easy proofs for some important and well-known results from modal logic.

AMS Subject Classification (1991): 03B45, 03G25, 03G15, 06E25.
CR Subject Classification (1991): F.4.0, F.4.1, I.2.3, I.2.4.
Keywords & Phrases: Modal logic, algebraic logic, boolean algebras with operators, duality theory, deduction, knowledge representation formalisms.
Note: This report is loosely based on parts of Blackburn, de Rijke and Venema [4]; the material was used as lecture notes for a course on Modal Logic and Boolean Algebras with Operators given by the third author during the summer school on Algebraic Logic and the Method of Applying It (Budapest, July 11–17 1994), which was held as part of the TEMPUS summer school series 'Algebraic and Categorial Methods in Computer Science.'

Part I: Introduction

It has been a long time since modal logic involved only the study of extensions of classical propositional logic with one operator '$\Box$' referring to the *necessity* of truth of a formula. Nowadays, modal logicians are studying a variety of modal languages and interpretations for them, each system designed to formalize some aspect of linguistic, computational, mathematical or philosophical reasoning, and it has become increasingly hard to indicate what these

Report CS-R9463
ISSN 0169-118X
CWI
P.O. Box 94079, 1090 GB Amsterdam, The Netherlands

formalisms have in common. Let us instead give a few examples of applications of modal logic in computer science, deferring the general definition to Section 1.

Epistemic logic, which originates in philosophical logic, is used in a formal analysis of knowledge, involving either human agents or for instance processors in a distributed system. Given a set A of such agents, the basic language has an operator K_a for each agent a; the intended reading of $K_a\phi$ is 'agent a knows ϕ', or 'agent a has access to the information that ϕ'. One aim of epistemic logicians is to try and find the 'right' *logic* of such formalizations, i.e., to develop deductive systems in which laws about knowledge arise as theorems of the deductive apparatus. In order to do so, it turns out to be very convenient to have a proper *semantics* for such logics. Most approaches in the literature involve some kind of Kripke semantics, in which two knowledge states s and t stand in a relation R_a if and only if t is, from the perspective of agent a, a possible 'epistemic alternative' to s. The natural truth definition of the knowledge operators is the basis of the semantics of modal logic:

$$s \Vdash K_a\phi \text{ if } t \Vdash \phi \text{ for all } t \text{ such that } R_a st.$$

A second incarnation of modal logic in computer science is formed by *dynamic logic*, a formalism developed to reason about processes. Here the structures are given in the form of labeled transition systems: a labeled transition system is nothing but a set of states, related by transitions that are labeled with atomic actions. The idea is that for any program α (either an atomic action or a more complex program involving for instance `while`-statements), a binary relation R_α is defined, where $R_\alpha st$ holds if it is possible in state s to perform α and thus reach t. The language of dynamic logic (which is just one alternative to talk about labeled transition systems) then has operators $\langle\alpha\rangle$ for every program α, and the truth definition of these operators is

$$s \Vdash \langle\alpha\rangle\phi \text{ iff there is a } t \text{ with } R_\alpha st \text{ and } t \Vdash \phi.$$

The *logics* developed for this semantics find applications in the theory of specification and verification of programs. Another branch of modal logic used to reason about the behavior of processes is *temporal logic*, which also finds applications in the formal semantics of natural languages and in artificial intelligence.

As a last example we mention the *calculus of binary relations*, which is applied (in various disguises) in almost any field of theoretical computer science. Here it is perhaps less easy to see the modal connection; the basic idea is that we can read the statement $(x, y) \in R; S$ (the composition of the relations R and S) as follows: 'there are pairs (x, z) and (z, y) such that $(x, z) \in R$ and $(y, z) \in S$'. Thus we can see the composition operator as a binary operator, with a ternary accessibility relation C defined on pairs by

$$C((u, v), (w, x), (y, z)) \text{ iff } u = w, v = z \text{ and } x = y.$$

Many of the questions arising from such applications have been studied from a more general, theoretical and abstract perspective. In the literature on modal logic, the following areas can be identified (among others):

Definability. Given a class K of frames, can we *define* K by a (set of) modal formula(s)? Or, conversely, given a modal formula, in which frames will it be valid?

Axiomatics. Given a class K of frames, can we find a nice axiomatization of the set of formulas that are valid in K?

Complexity. Is it possible to decide whether a given formula is valid in a given class of frames, or a theorem of a given modal logic? And if so, how effective an algorithm can we give?

Proof theory. Which kind of deductive system is suited best to find out whether a given formula is a theorem of a logic, and what are the properties of such a system?

In each of these areas, many years of research have lead to a great number of results, some of which have been obtained by applying results from other areas of logic (and mathematics). It is the aim of these notes to show how in two fields of modal logic, viz. definability theory and axiomatics, techniques and results can be used from *universal algebra.* In this respect, modal logic is a good witness of the successful strategy of *algebraic logic.* We summarize the program of algebraic logic here as to translate logical problems into algebraic questions, then to use the machinery of universal algebra to solve the problem algebraically, and finally, to translate the solution back to logic. In the particular case of modal logic, a bridge between logic and algebra is formed by the duality theory between boolean algebras with operators and modal frames[1]. In its full version, this duality theory is an extension of the duality between boolean algebras and Stone spaces. In these notes, we can only treat the basic issues, omitting any reference to topology.

Let us give a representative example of this strategy (treated in more detail in Section 7). An important problem in the definability theory of modal logic is the question whether there is, for a given class K of frames, a set Δ of modal formulas which defines this class in the sense that for any frame $\mathfrak{F}$, Δ is valid on $\mathfrak{F}$ iff $\mathfrak{F}$ belongs to K. The analogous question in universal algebra is answered by Birkhoff's Theorem stating that a class of algebras is equationally definable iff the class is a variety, i.e., closed under taking subalgebras, homomorphic images and direct products. Now, applying Birkhoff's result through the channel of duality theory, (and with some additional modal reasoning), Goldblatt and Thomason [13] give an analogous structural characterization of the frame classes that are modally definable (Theorem 7.2).

Finally, we should mention that results are also being transferred in the opposite direction, i.e., from modal logic to the theory of boolean algebras with operators. A second important result in modal logic, due to Fine, states that a modal logic which is complete with respect to a first-order definable class, has the additional desirable property of being canonical. This result can be applied, together with a theorem by Sahlqvist, to show that many varieties of boolean algebras with operators are closed under taking canonical embedding algebras (cf. Theorem 7.8).

We hope to persuade the reader that it is worthwhile to study modal frames and boolean algebras with operators from *both* the logical *and* the algebraic perspective. These notes contain the basic tools for such a study.

[1]This way of introducing algebraic logic is historically unjustified. Curiously, in the case of modal logic, the bridge was constructed, in the paper "Boolean Algebras with Operators" by Jónsson and Tarski, about a decade before Kripke and others developed and studied the possible world semantics for modal logic.

Outline of the paper

The paper is divided into four parts: after this introductory Part I, the second part forms a concise introduction to modal logic, highlighting two of its main themes: definability theory and axiomatics. In Part III we develop the algebraic perspective on modal logic; here we concentrate on the basic duality theory between frames and boolean algebras with operators. (For a more detailed overview of Parts II and III we refer to the introducing paragraphs of both parts.) We conclude in Part IV by briefly commenting on topics we had to omit for lack of space, and by pointing out some more advanced literature.

Part II: Modal logic

As was pointed out in the Introduction, modal formalisms are applied in a great number of disciplines. The two important uses of modal logic are (i) as a tool for analyzing non truth-functional sentential operators, and (ii) as a description language for relational structures. Examples of the former use include epistemic logic, the calculus of binary relations, but also provability logic where modal operators are used to study constructions such as 'it is provable that ...' (see [24]). Examples of modal logic as a description language can be found in computational linguistics where modal languages are used to single out trees corresponding to grammatical strings (see [3]), and in computer science where dynamic and temporal languages are used to pin down the desired execution structures of programs (see [19]).

Often the main motivation for using (propositional) modal logic in either of the above two uses is their flexibility, their naturalness and the fact that in many cases they have better computational properties than richer formalisms such as first-order logic.

Here's the plan for Part II; the contents of Part III are described just before Section 5 on algebraizing modal logic. In Section 1 we introduce modal languages and the semantic structures used to interpret them. In Section 2 (on definability) we analyze the expressive power of modal languages, and Section 3 describes an algorithm for automatically obtaining the properties defined by modal formulas. Section 4 gives an introduction to completeness theory by using canonical models.

1. Preliminaries

In this section we review basic definitions about modal logic.

1.1 Modal languages

In these notes we will mainly be working in a very simple modal language with just a single modal operator $\Diamond$ ('diamond') and its dual $\Box$ ('box'). For the record we will also define more general modal languages, but for didactical purposes statements of results (and proofs) will most often be given for the simple modal language only.

Definition 1.1 The *standard modal language* is defined using a set of proposition letters Φ whose elements are denoted p, q, ..., and a unary sentential operator $\Diamond$ ('diamond'). Formulas of the standard modal language are given by the rule

$$\phi ::= p \mid \bot \mid \neg\phi \mid \phi \wedge \phi \mid \Diamond\phi,$$

where p ranges over elements of Φ. This definition means that a formula is either a proposition letter, the propositional constant falsum, a negated formula, a conjunction of formulas, or a formula prefixed by a the diamond.

Just as the existential and universal quantifier are duals to each other (in the sense that $\forall x\,\alpha \leftrightarrow \neg\exists x\,\neg\alpha$), we have a dual operator $\Box$ ('box') for our diamond. It is defined by $\Box\phi := \neg\Diamond\neg\phi$. Further, the usual classical abbreviations for disjunction and implication apply: $\phi \vee \psi := \neg(\neg\phi \wedge \neg\psi)$, $\phi \rightarrow \psi := \neg\phi \vee \psi$.

To give the general notion of a modal language we need the following.

Definition 1.2 A *modal similarity type* is a pair $\tau = (O, \rho)$ where O is a non-empty set, and ρ is a function $O \rightarrow \mathbb{N}$. The elements of O are called *modal operators*; we use Δ ('*triangle*'), Δ_0, Δ_1, ... to denote elements of O. The function ρ assigns to each operator $\Delta \in O$ a finite *arity*, indicating the number of arguments Δ can be applied to.

We usually refer to unary triangles as *diamonds*, and denote them with $\Diamond$ or $\langle a\rangle$, for a in some index set. We often assume that the arity of modal operators is known, and make no distinction between τ and O.

Definition 1.3 A *modal language* $ML(\tau, \Phi)$ is built up from a modal similarity type $\tau = (O, \rho)$ and a set of proposition letters Φ. The set $Form(\tau, \Phi)$ of *modal formulas* over τ and Φ is given by the rule

$$\phi ::= p \mid \bot \mid \neg\phi \mid \phi_1 \wedge \phi_2 \mid \Delta(\phi_1, \ldots, \phi_n),$$

where p ranges over elements of Φ, and Δ is an n-ary modal operator in τ.

1.2 Models and frames; truth and validity

We will first define models and frames for the standard modal language.

Definition 1.4 A *frame* for the standard modal language is a pair $\mathfrak{F} = (W, R)$ such that

1. W is a nonempty set.
2. R is a binary relation on W.

A *model* for the standard modal language is a pair $\mathfrak{M} = (\mathfrak{F}, V)$, where $\mathfrak{F}$ is a frame, and V is a function assigning to each proposition letter p a subset $V(p)$ of W. Informally, we think of $V(p)$ as the set of points where p is true. V is called a *valuation*.

The notion of a standard modal formula ϕ being *true at a state w in a model* $\mathfrak{M} = (W, R, V)$, notation $\mathfrak{M}, w \Vdash \phi$, is defined inductively:

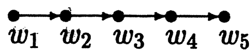
$$\begin{array}{rll}
\mathfrak{M}, w \Vdash p & \text{iff} & w \in V(p) \\
\mathfrak{M}, w \Vdash \bot & \text{iff} & w \neq w \\
\mathfrak{M}, w \Vdash \neg\phi & \text{iff} & \text{not } \mathfrak{M}, w \Vdash \phi \\
\mathfrak{M}, w \Vdash \phi \wedge \psi & \text{iff} & \mathfrak{M}, w \Vdash \phi \text{ and } \mathfrak{M}, w \Vdash \psi \\
\mathfrak{M}, w \Vdash \Diamond\phi & \text{iff} & \text{for some } v \in W \text{ with } Rwv \text{ we have } \mathfrak{M}, v \Vdash \phi.
\end{array}$$

When $\mathfrak{M}$ is clear from the context, we write $w \Vdash \phi$ for $\mathfrak{M}, w \Vdash \phi$. A formula ϕ is *true in a model* $\mathfrak{M}$ (notation: $\mathfrak{M} \Vdash \phi$) if it is true at all points in $\mathfrak{M}$.

It is often convenient to extend the valuation V from proposition letters to arbitrary formulas so that $V(\phi)$ is the set of states at which ϕ is true: $V(\phi) = \{w \mid \mathfrak{M}, w \Vdash \phi\}$.

We write $\mathfrak{M}, w \equiv \mathfrak{N}, v$ to denote that w and v verify the same formulas.

Example 1.5 Consider the structure $\mathfrak{F} = (\{w_1, w_2, w_3, w_4, w_5\}, R)$, where Rw_iw_j iff $j = i+1$ $(1 \leq i \leq 4)$.



If we choose a valuation V on $\mathfrak{F}$ with $V(p) = \{w_2, w_3\}$, then the model $\mathfrak{M} = (\mathfrak{F}, V)$ has $\mathfrak{M}, w_1 \nVdash p$, but $\mathfrak{M}, w_1 \Vdash \Diamond\Box p$, and so $\mathfrak{M}, w_1 \nVdash \Diamond\Box p \to p$.

Whereas a diamond $\Diamond$ corresponds to making a single R-step in a model, stacking diamonds corresponds to making a sequence of R-steps. We write $\Diamond^n\phi$ for ϕ preceded by n occurrences of $\Diamond$; correspondingly, R^n is defined inductively by R^0xy iff $x = y$, and $R^{n+1}xy$ iff $\exists z\,(Rxz \wedge R^nzy)$. Then, for any model $\mathfrak{M}$ and state w in $\mathfrak{M}$ we have $\mathfrak{M}, w \Vdash \Diamond^n\phi$ iff there exists v such that R^nwv and $\mathfrak{M}, v \Vdash \phi$.

We now define frames, models and truth for arbitrary modal languages.

Definition 1.6 Let τ be a modal similarity type. A *τ-frame* is a frame $\mathfrak{F}$ consisting of the following ingredients 1. and 2.:

1. a non-empty set W,
2. for each $n \geq 0$ and each n-ary modal operator Δ in the similarity type τ an $(n+1)$-ary relation R_Δ.

If τ contains finitely many modal operators $\Delta_1, \ldots, \Delta_n$, we write $\mathfrak{F} = (W, R_{\Delta_1}, \ldots, R_{\Delta_n})$; otherwise we write $\mathfrak{F} = (W, R_\Delta)_{\Delta \in \tau}$.

A *τ-model* is a pair $\mathfrak{M} = (\mathfrak{F}, V)$ where $\mathfrak{F}$ is a τ-frame, and V is a valuation.

The notion of a formula ϕ being *true at a state* w in a model $\mathfrak{M} = (W, \{R_\Delta \mid \Delta \in \tau\}, V)$ (notation: $\mathfrak{M}, w \Vdash \phi$) is defined inductively. The clauses for the atomic and boolean cases are the same as for the standard modal language (Definition 1.4); for the modal case, we define

$$\mathfrak{M}, w \Vdash \Delta(\phi_1, \ldots, \phi_n) \text{ iff for some } v_1, \ldots, v_n \in W \text{ with } R_\Delta wv_1 \ldots v_n \text{ we have } \mathfrak{M}, v_i \Vdash \phi_i \ (1 \leq i \leq n).$$

As before, we leave out $\mathfrak{M}$ if it is provided by the context. Also, the notion of truth in a model can be defined as before, as can the use of V applied to arbitrary formulas.

Example 1.7 Let τ be a similarity type with three unary operators $\langle a \rangle$, $\langle b \rangle$ and $\langle c \rangle$. Then, a τ-frame has three binary relations R_a, R_b and R_c, that is, it is a labeled transition system. To give an example, let W, R_a, R_b and R_c be as in Figure 1(a), and consider the formula $\langle a \rangle p \to \langle b \rangle p$. Informally, this formula is true at a state, if it has an R_a-successor satisfying p only if it has an R_b-successor satisfying p. Let V be a valuation with $V(p) = \{w_2\}$. Then the model $\mathfrak{M} = (W, R_a, R_b, R_c, V)$ has $\mathfrak{M}, w_1 \nVdash \langle a \rangle p \to \langle b \rangle p$ as $\mathfrak{M}, w_1 \Vdash \langle a \rangle p$, but $\mathfrak{M}, w_1 \nVdash \langle b \rangle p$.

In *arrow logic* one thinks of the objects of models as arrows or transitions rather than states or points. An important binary modal operator in arrow logic is composition $\circ$; intuitively, $\phi \circ \psi$ is true at an arrow if the arrow can be decomposed into two arrows satisfying ϕ and ψ, respectively. Formally, one introduces a ternary relation C and defines $x \Vdash \phi \circ \psi$ to hold if for some y and z, $Cxyz$ and $y \Vdash \phi$ and $z \Vdash \psi$. Familiar properties of composition can then be obtained by imposing additional constraints on C. For example, associativity of $\circ$ does

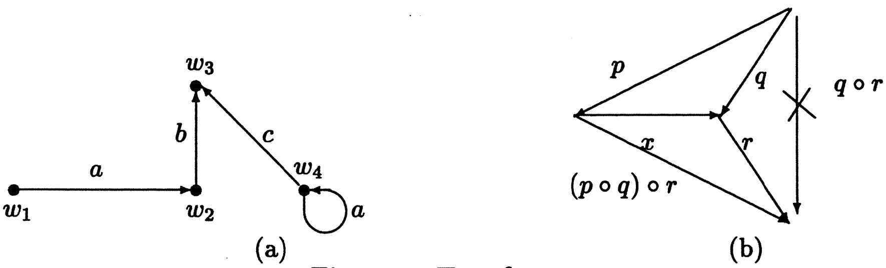


Figure 1: Two frames.

not hold automatically; in Figure 1(b) we have $x \Vdash (p \circ q) \circ r$, but $x \not\Vdash p \circ (q \circ r)$: x can be decomposed into an arrow labeled $p \circ q$ and an arrow labeled r, but it can't be decomposed into arrows labeled p and $q \circ r$.

Quite often we have a special reading in mind for our modal operators, or for the models on which we interpret them.

Example 1.8 In the standard modal language the diamond $\Diamond$ is sometimes interpreted as 'it is *possibly* the case that ... '; $\Box\phi$ the stands for 'necessarily ϕ.' Given a state w, let us call states v for which Rwv holds, states that are *possible* for w; then a formula $\Diamond\phi$ is true at w whenever ϕ is true at some state that is possible from the point of view of w. A typical example of a complex statement here is 'whatever is necessary, is possible': $\Box\phi \to \Diamond\phi$. The terms (*possible*) *world* and *possible world model* often found in the literature, derive from this particular language with this particular reading of the modal operator.

As we saw in the introduction, *epistemic logic* is a branch of modal logic used for reasoning about the knowledge an agent has; instead of $\Box\phi$ or $[a]\phi$ one writes $K_a\phi$ for 'the agent a *knows that* ϕ.' The intuitive reading for $w \Vdash K_a\phi$ is: the agent a knows ϕ in a situation w iff ϕ is true in all situations v that are compatible with a's knowledge (that is, if $v \Vdash \phi$ for all v such that R_awv). A formula whose truth seems a minimal requirement to be able to talk about knowledge (as opposed to, say, belief or rumor) is $K_a\phi \to \phi$: if a knows that ϕ, then ϕ must be true.

Assume that the set of operators $O = \{\langle F\rangle, \langle P\rangle\}$, and that R_P is the converse of R_F, that is $\forall wv\,(R_Fwv \leftrightarrow R_Pvw)$. If we interpret R_Fwv as 'v is later in time than w,' then $\langle F\rangle\phi$ is true at a point in time whenever ϕ is true at some *f*uture point, and $\langle P\rangle\phi$ is true at a point whenever ϕ is true at some *p*ast point. The operators $\langle F\rangle$ and $\langle P\rangle$ are usually written as F and P; they form the core of a special branch of modal logic called *tense* or *temporal logic*. The duals of F and P are written as G and H, respectively. Complex tense logical statements describe interesting properties of time; $P\phi \to GP\phi$, for instance, says 'what has happened will always have happened.'

Another example concerns the earlier arrow logic. In addition to a binary operation denoting composition, arrow logic has a unary operator $\otimes$ to talk about the *converse* of arrows, and a constant δ to talk about *identity* arrows. A reasonable axiom in this language is $\otimes(p \circ q) \leftrightarrow (\otimes q \circ \otimes q)$: the converse of a composition is the composition of the converses of the component arrows in reverse order.

In addition to interpreting modal formulas on models, we can also interpret modal formulas on frames, namely by quantifying over all valuations.

Definition 1.9 A formula ϕ is *valid at a state w in a frame* $\mathfrak{F}$ (notation: $\mathfrak{F}, w \models \phi$) if ϕ is true at w in every model $(\mathfrak{F}, V)$ based on $\mathfrak{F}$; ϕ is *valid in a frame* $\mathfrak{F}$ (notation: $\mathfrak{F} \models \phi$) if it is valid at every state in $\mathfrak{F}$.

Example 1.10 The formula $\Diamond(p \vee q) \to (\Diamond p \vee \Diamond q)$ is valid on all frames. To see this, take any frame $\mathfrak{F}$ and state w in $\mathfrak{F}$, and let V be a valuation on $\mathfrak{F}$. We have to show that if $(\mathfrak{F}, V), w \Vdash \Diamond(p \vee q)$, then $(\mathfrak{F}, V), w \Vdash \Diamond p \vee \Diamond q$. So assume that $(\mathfrak{F}, V), w \Vdash \Diamond(p \vee q)$. Then there is a state v such that Rwv and $(\mathfrak{F}, V), v \Vdash p \vee q$. If $v \Vdash p \vee q$ then either $v \Vdash p$ or $v \Vdash q$. Hence, either $w \Vdash \Diamond p$ or $w \Vdash \Diamond q$ — but in both cases it follows that $w \Vdash \Diamond p \vee \Diamond q$.

The formula $\Diamond q \to \Diamond\Diamond q$ is not valid on all frames. To see this we need to come up with a frame $\mathfrak{F}$, a state w in $\mathfrak{F}$, and a valuation on $\mathfrak{F}$ that falsifies the formula at w. Take a two-point frame $\mathfrak{F}$ whose universe is $\{0, 1\}$, and whose relation is $\{(0, 1)\}$. Define a valuation by putting $V(p) = \{1\}$. Then $(\mathfrak{F}, V), 0 \Vdash \Diamond p$, but obviously $(\mathfrak{F}, V), 0 \nVdash \Diamond\Diamond p$.

Here is a frame on which the above formula $\Diamond p \to \Diamond\Diamond p$ is valid. As the universe of the frame take the set of all rational numbers, $\mathbb{Q}$, and let R denote the usual $<$-ordering on $\mathbb{Q}$. To show that $\Diamond p \to \Diamond\Diamond p$ is valid on this frame, take any state w in it, and any valuation V such that $(\mathbb{Q}, R, V), w \Vdash \Diamond p$; we have to show that $w \Vdash \Diamond\Diamond p$. But this is easy: as $w \Vdash \Diamond p$, there exists v with Rwv and $v \Vdash p$. Because we are working on the rationals, there must be a z with Rwz and Rzv. So, $z \Vdash \Diamond p$, but then $w \Vdash \Diamond\Diamond p$.

As we have seen in the previous example, when constraints are imposed on frames more formulas may become valid: on arbitrary frames the formula $\Diamond q \to \Diamond\Diamond q$ may be falsified, but if we restrict ourselves to dense orderings such as $(\mathbb{Q}, <)$ we are no longer able to falsify it. This is a general point: $\Diamond q \to \Diamond\Diamond q$ is valid on a frame iff (the ordering of) the frame is dense. We will return to issues such as these in Sections 2 and 3 below.

2. Definability and its limits

In this section we study the expressive power of modal languages as description languages for relational structures. Corresponding to the two ways of interpreting modal formulas that we discussed in Section 1, we will carry out this study at two levels: at the level of models, and at the level of frames. We will study the expressive power of modal languages by considering the classes of models and frames that modal languages can single out.

Definition 2.1 Let C be a class of structures (either models or frames), and Γ a set of modal formulas. We say that Γ *defines* or *characterizes* a class K of structures *within* C if for all structures $\mathfrak{S}$ in C we have that $\mathfrak{S}$ is in K iff Γ is true/valid on $\mathfrak{S}$. If C is the class of all structures we will drop the clause 'within C' and simply say that Γ *defines* K. A class K is *definable* (within C) if there is a set of modal formulas that defines it (within C).

2.1 Models

In this subsection we consider definability of properties of *models*; that is, the classes C and K in Definition 2.1 are now taken to be classes of models. Two tools are extremely important in studying definability issues: the *standard translation* and *bisimulations*. The standard translation takes modal formulas to first-order formulas; the following definition specifies the relevant first-order language.

Definition 2.2 Given a modal similarity type τ, and a set of proposition letters Φ, we write $\mathcal{L}^1_\tau(\Phi)$ for the *first-order correspondence language*. This language has identity, unary predicates $P_0, P_1, \ldots$ corresponding to the proposition letters $p_0, p_1, \ldots$ in Φ, and $(n+1)$-ary relation symbols R_Δ for every n-ary modal operator Δ in τ.

It is important to observe that we can view our modal τ-models $\mathfrak{M} = (W, R_\Delta, V)_{\Delta\in\tau}$ as models for the first-order correspondence language $\mathcal{L}^1_\tau(\Phi)$. To do so we need to say how the relation symbols R_Δ should be interpreted — but that is obvious —, and we need to say how to interpret the unary predicate symbols, and here we use the valuation V: we will say that a state w is in the extension of the predicate P_i iff $w \in V(p_i)$.

Definition 2.3 Fix an individual variable x. We let $[y/z]\alpha$ denote the result of substituting the individual variable y for all free occurrences of z in α. The *standard translation* ST takes modal formulas to first-order formulas as follows:

$$\begin{aligned}
ST(p_i) &= P_i x\\
ST(\bot) &= (x \neq x)\\
ST(\neg\phi) &= \neg ST(\phi)\\
ST(\phi\wedge\psi) &= ST(\phi)\wedge ST(\psi)\\
ST(\Delta(\phi_1,\ldots,\phi_n)) &= \exists y_1\ldots\exists y_n\,(R_\Delta x y_1\ldots y_n \wedge\\
&\qquad [y_1/x]ST(\phi_1)\wedge\ldots\wedge[y_n/x]ST(\phi_n)),
\end{aligned}$$

where $y_1, \ldots, y_n$ are fresh variables.

As an example we compute the standard translation of $\Diamond(p\wedge\Box q)$:

$$\begin{aligned}
ST(\Diamond(p\wedge\Box q)) &= \exists y\,(Rxy\wedge[y/x]ST(p\wedge\Box q))\\
&= \exists y\,(Rxy\wedge[y/x]ST(p)\wedge[y/x]ST(\Box q))\\
&= \exists y\,(Rxy\wedge Py\wedge[y/x]ST(\forall z\,(Rxz\to[z/x]ST(q))))\\
&= \exists y\,(Rxy\wedge Py\wedge\forall z\,(Ryz\to Qz)).
\end{aligned}$$

Proposition 2.4 *Let ϕ be a modal formula. Then, for any model $\mathfrak{M}$ and state w we have $\mathfrak{M}, w \Vdash \phi$ iff $\mathfrak{M} \models ST(\phi)[w]$.*

As a corollary we find that on models every modal formula is equivalent to a first-order formula. Does the converse hold as well? If not, what fragment of first-order logic does modal logic correspond to?

To answer these questions we will use bisimulations. Before introducing them we look at a simple motivating example of two models that are different, but that verify the same modal formulas. We restrict ourselves to the standard modal language with $\Diamond$, $\Box$ with a single proposition letter p. Consider the two models depicted in Figure 2. That is, all points in both models verify the proposition letter p. We will show that w verifies exactly the same modal formulas as each of the v_is; we will do this by induction. The atomic and boolean cases are trivial. As to the modal case, assume $w \Vdash \Diamond\phi$. This means that we can move along an arrow to a state where ϕ holds — this can only be w itself. Hence $w \Vdash \phi$, and therefore, by induction hypothesis, each v_i has $v_i \Vdash \phi$. But this means that from any state in the model on

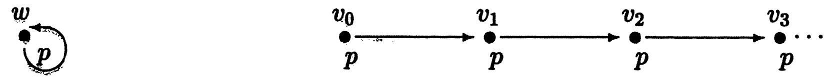


Figure 2: Equivalent models

the right-hand side we can make a move to a state with ϕ. So $v_i \Vdash \Diamond\phi$. Conversely, assume that every v_i has $v_i \Vdash \Diamond\phi$. Then, from every v_i a move can be made to a state where ϕ holds, but this implies that all v_i verify ϕ. Hence, by induction hypothesis, $w \Vdash \phi$, and as w can 'see' itself, $w \Vdash \Diamond\phi$, as required.

The important thing about the above example is the intuition underlying the proof: every 'modal step' in the one model in Figure 2 must be matched with a move in the other; the same intuition underlies bisimulations.

Definition 2.5 We will first define *bisimulations for the standard modal language*. Let $\mathfrak{M} = (W, R, V)$ and $\mathfrak{M}' = (W', R', V')$ be two models. A non-empty relation $Z \subseteq W \times W'$ is called a *bisimulation between* $\mathfrak{M}$ and $\mathfrak{M}'$ if the following holds:

1. If uZu' then u and u' verify the same proposition letters.
2. If uZu' and Ruv, then there exists v' in $\mathfrak{M}'$ such that vZv' and $R'u'v'$.
3. If uZu' and $R'u'v'$, then there exists v in $\mathfrak{M}$ such that vZv' and Ruv.

When Z is a bisimulation linking two states u and u' we say that u and u' are *bisimilar*, and write uZu' or $Z : \mathfrak{M}, u \leftrightarrows \mathfrak{M}', u'$. As an example, observe that the two models in Figure 2 are bisimilar via the relation $Z = \{(w, v_i) \mid i \in \mathbb{N}\}$.

We now define bisimulations for arbitrary modal languages. Fix a similarity type τ, and two τ-models $\mathfrak{M} = (W, R_\Delta, V)_{\Delta\in\tau}$ and $\mathfrak{M}' = (W', R'_\Delta, V')_{\Delta\in\tau}$. A non-empty relation $Z \subseteq W \times W'$ is a *τ-bisimulation* if it satisfies condition 1. as before, and conditions 2'. and 3'. below.

2'. For any Δ in τ, if uZu' and $R_\Delta uv_1 \ldots v_n$, then there exist $v'_1, \ldots, v'_n$ in $\mathfrak{M}'$ such that $R'_\Delta u'v'_1 \ldots v'_n$ and $v_iZv'_i$ (for $1 \leq i \leq n$).

3'. For any Δ in τ, if uZu' and $R'_\Delta u'v'_1 \ldots v'_n$, then there exist $v_1, \ldots, v_n$ in $\mathfrak{M}$ such that $R_\Delta uv_1 \ldots v_n$ and $v_iZv'_i$ (for $1 \leq i \leq n$).

Proposition 2.6 *Fix a modal similarity type τ. Let $\mathfrak{M}$, $\mathfrak{M}'$ be two τ-models, and let w, w' be states in $\mathfrak{M}$, $\mathfrak{M}'$ respectively. If there is a bisimulation $Z : \mathfrak{M}, u \leftrightarrows \mathfrak{M}', u'$, then, for all τ-formulas ϕ we have $\mathfrak{M}, u \Vdash \phi$ iff $\mathfrak{M}', u' \Vdash \phi'$.*

Proof. The proof is by induction on ϕ. $\dashv$

By Proposition 2.6 we have a tool to test for modal undefinability: a first-order condition is modally definable only if it is invariant for bisimulations in the following sense.

Definition 2.7 We say that a first-order formula $\alpha(x)$ in $\mathcal{L}^1_\tau(\Phi)$ is *invariant for bisimulations* if for all τ-models $\mathfrak{M}$, $\mathfrak{M}'$ and states w and w' in $\mathfrak{M}$ and $\mathfrak{M}'$, respectively, and all bisimulations Z we have that

$$Z : \mathfrak{M}, w \leftrightarroweq \mathfrak{M}', w' \text{ implies that } \mathfrak{M} \models \alpha[w] \text{ iff } \mathfrak{M}' \models \alpha[w'].$$

Example 2.8 Using Figure 2 we see that the first-order condition Rxx is not modally definable: it is not invariant for bisimulations because it holds for w but it does not hold for v_0, for example. The same argument also shows that $\forall y\, Rxy$ is not modally definable.

For a slightly more complicated example, consider a similarity type with a binary modal operator Δ based on a ternary relation S_Δ. We claim that the condition

$$\forall yzz'\,(Sxyz \land Sxyz' \to z = z')$$

is not modally definable. To see this, consider the two models in Figure 3.

Figure 3: Bisimilar models

The obvious bisimulation between the models in Figure 3 proves the claim.

We can use bisimulations to detect *un*definability, but can we also use it to find out whether a condition is definable? To find the answer, we return to Proposition 2.6; recall that according to this result bisimilarity implies modal equivalence. The following example shows that modal equivalence does *not* imply bisimilarity. Consider the two models for the standard modal language (over an empty set of proposition letters) depicted in Figure 4.

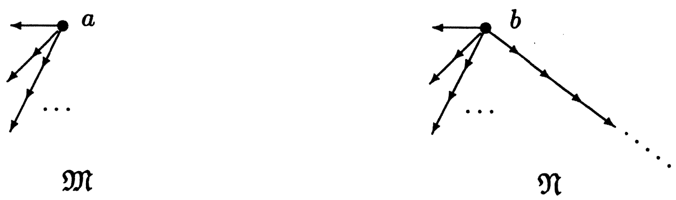


Figure 4: Equivalent but not bisimilar.

The claim is that a and b in Figure 4 are equivalent but not bisimilar. To see that they are not bisimilar, consider a point $b' \neq b$ on the infinite branch in $\mathfrak{N}$. Whatever point in $\mathfrak{M}$ we try to link b' to via a candidate bisimulation, it has to be different from a (why?), and hence it will only have finitely many successors. But b' has an infinite chain of successors, and in the long run this will violate the bisimulation condition. To see that, nevertheless, a and b are equivalent, one can show by induction on formulas that for every state b' on the infinite branch of $\mathfrak{N}$, and every modal formula ϕ with $b' \Vdash \phi$, there is a (finite) branch in $\mathfrak{M}$ and a state a' on that branch with $a' \Vdash \phi$; from this it follows that a and b are modally equivalent.

The above discussion motivates the following definition.

Definition 2.9 Let τ be a modal similarity type, and K a class of τ-models. K is a *Hennessy-Milner* class if for every two models $\mathfrak{M}$, $\mathfrak{M}' \in \mathsf{K}$ and any two states u and u' in $\mathfrak{M}$ and $\mathfrak{M}'$, respectively, $\mathfrak{M}, u \equiv \mathfrak{M}', u'$ implies $\mathfrak{M}, u \leftrightarroweq \mathfrak{M}', u'$.

Example 2.10 By our earlier examples, the class of all models is not a Hennessy-Milner class. Here are examples of Hennessy-Milner classes:

- Finite models, and more generally, image-finite models (these are models such that for every state w and every relation R_Δ ($\Delta \in \tau$) the set $\{(v_1, \ldots, v_n) \mid R_\Delta w v_1 \ldots v_n\}$ is finite).
- Canonical models (cf. Section 4 below for a definition).
- Saturated models (cf. Chang and Keisler [8, Chapter 6] for a definition).

We will show that the class of finite models is a Hennessy-Milner class.

Proposition 2.11 *Fix a modal similarity type τ. The class of finite τ-models is a Hennessy-Milner class.*

Proof. We prove the result for the standard modal language. Assume $\mathfrak{M}, a \equiv \mathfrak{N}, b$, where $\mathfrak{M}$ and $\mathfrak{N}$ are finite models. We have to show that $\mathfrak{M}, a \leftrightarroweq \mathfrak{N}, b$. The natural candidate bisimulation is

$$xZy \quad \text{iff} \quad \text{for all modal formulas } \phi\text{: } \mathfrak{M}, x \Vdash \phi \text{ iff } \mathfrak{N}, y \Vdash \phi.$$

Let us show that Z is a bisimulation. Z is non-empty, and it trivially fulfills the condition on proposition letters. Next, assume that xZy and Rxx' hold; we have to find a y' with $x'Zy'$ and $R'yy'$ (in $\mathfrak{N}$). Assume that no such y' exists; we will derive a contradiction. Note that $X = \{z \mid R'yz \text{ in } \mathfrak{N}\}$ is non-empty (otherwise $x \Vdash \Diamond\top$, but $y \nVdash \Diamond\top$, contradicting xZy). As $\mathfrak{N}$ is finite, so is X, say $X = \{z_1, \ldots, z_n\}$. Hence, for every z_i there is a formula ϕ_i such that $x' \Vdash \phi_i$, but $z_i \nVdash \phi_i$. Let $\Phi := \bigwedge_i \phi_i$. Then $x \Vdash \Diamond\Phi$, but $y \nVdash \Diamond\Phi$ — contradicting xZy.

The final bisimulation condition is proved entirely analogously. ⊣

Theorem 2.12 *Fix a similarity type τ. A first-order formula $\alpha(x)$ in $\mathcal{L}^1_\tau(\Phi)$ is invariant for τ-bisimulations iff it is equivalent to (the translation of) a modal formula.*

Proof. The direction from right to left is Proposition 2.6. Proving the converse requires more work; we will sketch a proof for the standard modal language. Assume $\alpha(x)$ is invariant for bisimulations; to exclude trivial cases we will also assume that α is consistent. Consider the set of modal consequences of α:

$$\text{MOD-CON}(\alpha) = \{ST(\phi) \mid \alpha \models ST(\phi), \quad \phi \text{ is a modal formula}\}.$$

By compactness it suffices to show that MOD-CON(α) $\models \alpha$. For then there exists a finite $\Gamma \subseteq$ MOD-CON(α) such that $\Gamma \models \alpha$ (and conversely) and $\bigwedge \Gamma$ is a modal formula.

Assume $\mathfrak{M} \models$ MOD-CON(α)[w]. We have to show that $\mathfrak{M} \models \alpha[w]$. Our first observation is that by a simple compactness argument the set $X := \{\alpha\} \cup \{ST(\psi) \mid \mathfrak{M}, w \Vdash \psi\}$ is consistent. Let $\mathfrak{N}$ be a model with $\mathfrak{N} \models X[v]$, for some v. Note that $\mathfrak{M}, w \equiv \mathfrak{N}, v$.

If $\mathfrak{M}$ and $\mathfrak{N}$ both lived in a Hennessy-Milner class, $\mathfrak{M}, w \equiv \mathfrak{N}, v$ would imply $\mathfrak{M}, w \leftrightarroweq \mathfrak{N}, v$, and from this we would be able to infer $\mathfrak{M} \models \alpha[w]$, which would complete the proof. We can get away with slightly less: it's enough to make a detour through a Hennessy-Milner class, as follows. By general model-theoretic considerations from first-order logic, both $\mathfrak{M}$ and $\mathfrak{N}$ have saturated elementary extensions $\mathfrak{M}^*$ and $\mathfrak{N}^*$; it follows that $\mathfrak{M}^*, w \equiv \mathfrak{N}^*, v$ and that $\mathfrak{N}^* \models \alpha[w]$. The class of saturated models is a Hennessy-Milner class, hence $\mathfrak{M}^*, w \leftrightarrows \mathfrak{N}, v$. By invariance under bisimulations we get $\mathfrak{M}^* \models \alpha[w]$. And as $\mathfrak{M}^*$ is an elementary extension of $\mathfrak{M}$ we infer that $\mathfrak{M} \models \alpha[w]$ — and we are done. $\dashv$

Corollary 2.13 *Fix a modal similarity type τ. Let K be a class of τ-models that is defined by a set of first-order formulas. Then K is modally definable iff it is closed under bisimulations.*

Corollary 2.14 *Fix a modal similarity type τ. A class of τ-models is modally definable iff it is closed under bisimulations and ultraproducts, and its complement is closed under ultrapowers.*

2.2 Frames

We now turn to a brief study of the expressive power of modal languages on the level of frames. As in the case of models we will approach the issue by looking at definable classes of structures, that is: definable classes of frames. Informally, we shall say that a formula ϕ defines a *property* of frames whenever ϕ defines the class of frames satisfying that property. As many such properties are expressed in first-order logic, the following is convenient.

Definition 2.15 For a modal similarity type τ, we denote by $\mathcal{L}^1_\tau$ the first-order *frame language* of τ. This language has identity and a ($n+1$-ary) relation symbol R_Δ for each (n-ary) modal operator Δ in τ.

If E is a frame property (for instance, reflexivity of the relation $R_\Diamond$) which can both be expressed by a first/second-order formula α and defined by a modal formula ϕ, then we say that α and ϕ are each others *correspondents*.

There are two approaches to modal definability: most often, one is interested in a particular class of frames and wants to find out whether the modal language can distinguish the 'good frames' inside the class from the 'bad ones' outside the class. Conversely, sometimes the syntax of the modal language comes first, for instance when a set of laws is given in a modal language. here one wants to develop a natural semantics for these laws.

Example 2.16 We first consider the example of *epistemic logic*, cf. Example 1.8. For the moment, it suffices to confine ourselves to a system with a single agent a. Many axioms have been proposed as laws governing the cognitive behavior of agents, including

($A1$) $K_a p \to p$
($A2$) $K_a p \to K_a K_a p$
($A3$) $\neg K_a p \to K_a \neg K_a p$.

($A1$) says that one can only know true things; ($A2$) and ($A3$) are the so-called introspection axioms: by ($A2$), if a knows something, then he knows that he knows it; and by the *negative* introspection law ($A3$) one also knows that one does not know things.

Let us see which frame conditions these axioms define. Our first claim is that for any frame $\mathfrak{F} = (W, R_a)$, the axiom ($A1$) corresponds to *reflexivity* of the relation R_a:

$$\mathfrak{F} \models (A1) \text{ iff } \mathfrak{F} \models \forall x\, R_a xx. \tag{2.1}$$

The proof of the right to left direction of (2.1) is easy: let $\mathfrak{F}$ be a reflexive frame, and take an arbitrary valuation V on $\mathfrak{F}$, and an arbitrary state w in $\mathfrak{F}$ such that $(\mathfrak{F}, V), w \Vdash K_a p$. That is, p holds at all states v that are compatible with a's knowledge in w. However, w itself meets this condition, as R_a is reflexive. So, $w \Vdash p$.

For the other direction we use contraposition: suppose that R_a is *not* reflexive, i.e., there is a world w such that w is not compatible with a's knowledge in w. To falsify $(A1)$ in $\mathfrak{F}$, it suffices to find a valuation V and a state x such that $K_a p$ holds at x, but p does not. It is obvious that for x we can take our irreflexive state w. The valuation V has to satisfy two conditions: (1) $w \notin V(p)$ and (2) $\{x \in W \mid Rwx\} \subseteq V(p)$. Consider the *maximal* valuation V satisfying condition (1), i.e., take

$$V(p) = W \setminus \{w\}.$$

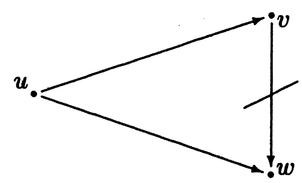
Clearly, $(\mathfrak{F}, V), w \not\Vdash p$. Let v be any R_a-successor of w. As $R_a ww$ does not hold, v must be distinct from w, so $v \Vdash p$. We find that $w \Vdash K_a p$. Hence $w \not\Vdash K_a p \to p$. This proves (2.1).

Likewise, one can prove that for any frame $\mathfrak{F} = (W, R_a)$:

$$\mathfrak{F} \models (A2) \text{ iff } R_a \text{ is transitive, and,} \tag{2.2}$$

$$\mathfrak{F} \models (A3) \text{ iff } R_a \text{ is euclidean,} \tag{2.3}$$

where a relation is *euclidean* if it satisfies $\forall xyz\, (Rxy \land Rxz \to Ryz)$. We leave the proofs of (2.2) and the easy (right to left) direction of (2.3) to the reader. For the left to right direction of (2.3), we again argue by contraposition. Assume that $\mathfrak{F}$ is a non-euclidean frame; then there exist u, v and w such that $R_a uv$, $R_a uw$, but not $R_a vw$:



We will try to falsify $(A3)$ in u; to this end we have to find a valuation V such that $(\mathfrak{F}, V), u \Vdash \neg K_a p \land \neg K_a \neg K_a p$. That is, we have to make p *false* at an R_a-successor x of u, and *true* at all R_a-successors of some R_a-successor y of u. Some reflection shows that appropriate candidates for x and y are w and v, respectively. The constraints on V are twofold: (1) $w \notin V(p)$ and (2) $\{z \mid R_a vz\} \subseteq V(p)$. Take a *minimal* V satisfying condition (2), i.e. define

$$V(p) = \{z \in W \mid R_a vz\}.$$

Clearly $v \Vdash K_a p$, so $u \Vdash \neg K_a \neg K_a p$. On the other hand, we have $w \not\Vdash p$, since w is *not* in the set $\{z \in W \mid R_a vz\}$. So $u \Vdash \neg K_a p$. In other words, we have found a valuation V and a state u such that $(A3)$ does not hold in u. Therefore, $(A3)$ is not valid in $\mathfrak{F}$. This proves (2.3).

The above example may be a bit deceptive, as it might suggest that on frames all modal formulas correspond to first-order conditions. By the following corollary to Proposition 2.6, however, this is the exception rather than the rule.

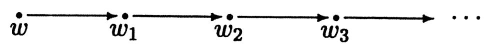
Proposition 2.17 *Let τ be a modal similarity type, and ϕ a τ-formula. Then for any τ-frame $\mathfrak{F}$ and any τ-formula ϕ:*

$$\mathfrak{F} \models \phi \text{ iff } \mathfrak{F} \models \forall P_1 \ldots \forall P_n \forall x\, ST(\phi).$$

Here the second-order quantifications $\forall P_1, \ldots, \forall P_n$ take place over the monadic predicates P_i such that the propositional variable p_i occurs in ϕ.

Example 2.18 Consider the formula $\Box(\Box p \to p) \to \Box p$, which we will call L for brevity. This formula is important in *provability logic*, a branch of modal logic where $\Box\phi$ is read as 'it is *provable* (in some formal system) that ϕ'. The formula L is named after Löb, who proved L as a theorem of the provability logic of Peano arithmetic. We will show that L characterizes the frames (W, R) where R is transitive and its converse is well-founded. (A relation R is well-founded if there are no infinite sequences $\ldots Rw_2Rw_1Rw_0$.)

First assume that $\mathfrak{F} = (W, R)$ is a frame with a transitive and conversely well-founded relation R, and suppose that L is not valid in $\mathfrak{F}$ (in order to arrive at a contradiction). This means that there are a valuation V and a point w such that $(\mathfrak{F}, V), w \not\Vdash \Box(\Box p \to p) \to \Box p$. In other words, $w \Vdash \Box(\Box p \to p)$, but $w \not\Vdash \Box p$. Then w must have a successor w_1 with $w_1 \not\Vdash p$; as $\Box p \to p$ holds at all successors of w, we find that $w_1 \not\Vdash \Box p$. This implies that w_1 must have a successor w_2 where p is false. By transitivity of R, w_2 is a successor of w. Now we repeat the same argument to show that w_2 must have a successor w_3 where p is false, etc. Following this procedure, we can find an infinite path $wRw_1Rw_2Rw_3R\ldots$, contradicting the converse well-foundedness of R.



For the other direction, we use contraposition. That is, we assume that either R is not transitive or its converse is not well-founded; in both cases we will then try to find a valuation V and a point w such that $(\mathfrak{F}, V), w \not\Vdash$ L. We leave the case where R is not transitive to the reader, and only consider the second case: assume that R is not conversely well-founded. In other words, there is an infinite sequence $w_0Rw_1Rw_2R\ldots$. Define a valuation V as follows:

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

We leave it to the reader to verify that with this valuation, $\Box(\Box p \to p)$ is true *everywhere* in the model; the claim then follows by the fact that $(\mathfrak{F}, V), w_0 \not\Vdash \Box p$.

Finally, to show that the class of frames defined by L is not first-order definable, an easy compactness argument suffices: let $\sigma_n(x_0, \ldots, x_n)$ be the formula expressing that there is a path of length n through $x_0, \ldots, x_n$: $\sigma_n(x_0, \ldots, x_n) := Rx_0x_1 \wedge Rx_1x_2 \wedge \ldots \wedge Rx_{n-1}x_n$. Every finite subset of

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

is satisfiable in a finite linear order and hence, in our class. However, it is clear that $\Sigma(x_0)$ itself is not satisfiable in a conversely well-founded frame.

We will now take a more systematic look at the phenomena illustrated in the above examples, and try to capture the expressive power of modal languages on frames in terms of preservation and closure properties, as we did with bisimulations on the level of models in the previous subsection. The invariance of modal truth under bisimulations makes that modal languages are blind for some kinds of distinctions between frames.

Definition 2.19 We will first say what a disjoint union of standard modal frames is, and then give the general concept. Given a collection of standard modal frames $\mathfrak{F}_i = (W_i, R_i)$, for $i \in I$, their *disjoint union* is the frame $\biguplus \mathfrak{F}_i = (W, R)$, where W is the disjoint union of the sets W_i, and R is the disjoint union of the relations R_i.

Generally, if τ is a modal similarity type, and $\mathfrak{F}_i = (W_i, R_{\Delta i})_{\Delta \in \tau}$ $(i \in I)$ a collection of τ-frames, then the *disjoint union* $\biguplus \mathfrak{F}_i = (W, R_\Delta)_{\Delta \in \tau}$ has the disjoint union of all the sets W_i as its domain W, and for each relation R_Δ is simply the disjoint union of all the relations $R_{\Delta i}$ (where Δ is kept constant, and only the index i is quantified over).

Definition 2.20 Again, we first define the notion of a generated subframe for the standard modal language. Given two frames $\mathfrak{F} = (W, R)$ and $\mathfrak{F}' = (W', R')$ we say that $\mathfrak{F}'$ is a *generated subframe* of $\mathfrak{F}$ (notation: $\mathfrak{F}' \rightarrowtail \mathfrak{F}$) if

1. $\mathfrak{F}'$ is a subframe of $\mathfrak{F}$, that is: $W' \subseteq W$ and $R' = R \cap (W' \times W')$.
2. If Rxy and $x \in W'$ then $y \in W'$.

Let X be a subset of the universe of a frame $\mathfrak{F}$; by $\mathfrak{F}_X$ we denote the *subframe generated by* X, that is: the smallest generated subframe of $\mathfrak{F}$ that contains X. If X is a singleton a, we write $\mathfrak{F}_a$ for the *subframe generated by* a; if a frame $\mathfrak{F}$ is generated by a singleton subset of its universe, we call $\mathfrak{F}$ *point-generated*.

For the general case, let $\mathfrak{F}'$, $\mathfrak{F}$ be two τ-frames. $\mathfrak{F}'$ is a *generated subframe* of $\mathfrak{F}$ if $\mathfrak{F}'$ is a subframe of $\mathfrak{F}$ as before (that is: $W' \subseteq W$ and for each $\Delta \in \tau$, $R'_\Delta = R_\Delta \cap (W \times \cdots \times W)$), and if for each $\Delta \in \tau$, $R_\Delta x y_1 \ldots y_n$ and $x \in W'$ implies $y_1, \ldots, y_n \in W'$.

Definition 2.21 We first define bounded morphisms for the standard language. Let $\mathfrak{F}'$, $\mathfrak{F}$ be two frames; a function $f : W' \to W$ is a *bounded morphism* if it satisfies

(*zig*) f is a homomorphism, that is: $R'xy$ implies $Rf(x)f(y)$.

(*zag*) If $Rf(x)y$ then there exists z in $\mathfrak{F}'$ such that $R'xz$ and $f(z) = y$.

We write '$\mathfrak{F} \twoheadrightarrow \mathfrak{F}'$' if $\mathfrak{F}'$ is a bounded morphic image of $\mathfrak{F}$.

Generally, if $\mathfrak{F}$, $\mathfrak{F}'$ are τ-frames, then $f : W' \to W$ is a *bounded morphism* if (zig) f is a homomorphism as before (that is: for each $\Delta \in \tau$, $R'_\Delta x y_1 \ldots y_n$ implies $R_\Delta f(x) f(y_1) \ldots f(y_n)$), and (zag) if for each $\Delta \in \tau$, we have $R_\Delta f(x) y_1 \ldots y_n$ only if there exist $z_1, \ldots, z_n$ such that $R'_\Delta x z_1 \ldots z_n$ and $f(z_1) = y_1, \ldots, f(z_n) = y_n$.

Validity of modal formulas is preserved under the above operations:

Proposition 2.22 *Let τ be a modal similarity type, and ϕ a τ-formula.*

1. *Let $\{\mathfrak{F}_i \mid i \in I\}$ be a family of frames. Then $\biguplus \mathfrak{F}_i \models \phi$ if $\mathfrak{F}_i \models \phi$ for every i in I.*
2. *Assume that $\mathfrak{F}' \rightarrowtail \mathfrak{F}$. Then $\mathfrak{F}' \models \phi$ if $\mathfrak{F} \models \phi$.*

3. Assume that $\mathfrak{F} \twoheadrightarrow \mathfrak{F}'$. *Then* $\mathfrak{F}' \models \phi$ *if* $\mathfrak{F} \models \phi$.

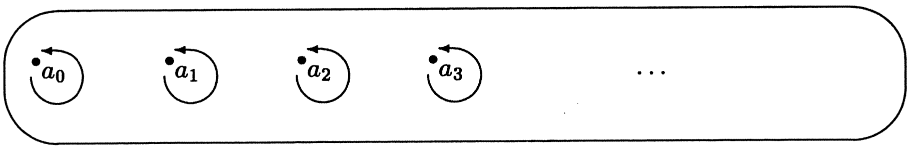
Proof. We only prove (3), the preservation result for taking bounded morphic images, and leave the other cases to the reader. So, assume that f is a surjective bounded morphism from $\mathfrak{F}$ onto $\mathfrak{F}'$, and that $\mathfrak{F} \models \phi$. We have to show that $\mathfrak{F}' \models \phi$. To arrive at a contradiction, suppose that ϕ is *not* valid in $\mathfrak{F}'$. Then there must be a valuation V' and a state w' such that $(\mathfrak{F}', V'), w' \not\Vdash \phi$. Define the following valuation V on $\mathfrak{F}$:

$$V(p_i) = \{x \in W \mid f(x) \in V'(p_i)\}.$$

This definition is tailored to turn f into a bisimulation between the models $(\mathfrak{F}, V)$ and $(\mathfrak{F}', V')$ — the reader is asked to verify the details. As f is surjective there is a w such that $f(w) = w'$. It is an easy exercise to show that $(\mathfrak{F}, V), w \not\Vdash \phi$. That is, we have falsified ϕ in $\mathfrak{F}$. ⊣

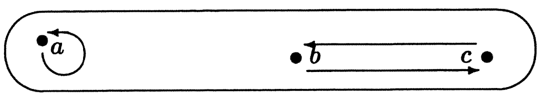
We may view these frame constructions as *testing material* for the definability of a frame property. If the property is not preserved under one (or more) of these frame constructions, then it cannot be modally definable. Let us consider some examples of this testing.

Example 2.23 We first show that the class of finite frames is not modally definable. For, suppose that there were a set of formulas Δ (in the basic modal similarity type) characterizing the finite frames. Then Δ would be valid in every one-point frame $\mathfrak{F}_i = (\{a_i\}, \{(a_i, a_i)\})$ $(i < \omega)$. By Proposition 2.22(1) this implies that Δ would also be valid in the disjoint union $\biguplus_i \mathfrak{F}_i$:



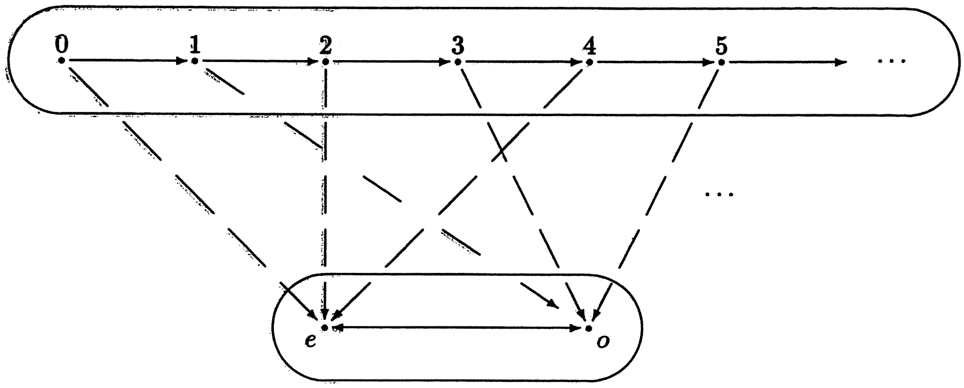
But clearly, this cannot be the case, as $\biguplus_i \mathfrak{F}_i$ is infinite.

Next we consider the class of frames having a reflexive point ($\exists x\, Rxx$); this class does not have a modal characterization either (we are considering the basic modal similarity type). For, suppose Δ characterizes this class. Consider the following frame $\mathfrak{F}$:



As a is a reflexive point, we find $\mathfrak{F} \models \Delta$. Consider the generated subframe $\mathfrak{F}_b$ of $\mathfrak{F}$. Clearly, Δ cannot be valid in $\mathfrak{F}_b$, since b, c are irreflexive. This contradicts Proposition 2.22(2).

Our last example involves the use of bounded morphisms. Consider the following two frames: $\mathfrak{F} = (\omega, S)$, the natural numbers with the successor relation (Smn iff $m = n + 1$), $\mathfrak{G} = (\{e, o\}, \{(e, o), (o, e)\})$, viz.



We leave it to the reader to verify that the map f sending even numbers to e and odd numbers to o, is a surjective bounded morphism. One can prove that no property E is modally definable if $\mathfrak{F}$ has E and $\mathfrak{G}$ lacks it. For instance, there is no set of formulas characterizing the asymmetric frames ($\forall xy\,(Rxy \rightarrow \neg Ryx)$).

A natural question at this point is, whether it is *sufficient* to test a property on its preservation under these structural operations, in order to decide whether the property is modally definable. The answer depends on the 'global' class of frames of which we want to determine the definable subclasses.

Theorem 2.24 *Let τ be the basic modal similarity type. A class K is definable within the class of transitive finite τ-frames if and only if K is closed under taking bounded morphic images, generated subframes and (finite) disjoint unions.*

Proof. See Blackburn, de Rijke and Venema [4, Chapter 3]. $\dashv$

In general however, a frame class has to satisfy more closure conditions in order to be modally definable. In particular we will need a very important new frame construction, namely that of the *ultrafilter extension* of a frame; its introduction and its use in obtaining further definability results are postponed until Section 6 on basic duality.

3. Definability: automatic correspondence

Even though modal formulas express second-order properties of frames, there are cases in which a reduction to first-order properties is possible. Some examples to this effect were given in Example 2.16. Is there any system to when a modal formula expresses a first-order condition on frames? And if it exists, how can we find the corresponding first-order condition? In this section we will define an important class of formulas for which the corresponding first-order correspondent can be computed *effectively*. We will build up the definition of this class in stages — a fairly general version will appear towards the end of this section.

Example 3.1 Consider the language of ordinary temporal logic. The formula $PH\bot$ expresses that before any point in the flow of time there is a beginning. For, recall that by Proposition 2.17, for any temporal frame $\mathfrak{F}$

$$\mathfrak{F} \models PH\bot \text{ iff } \mathfrak{F} \models \forall P_1 \ldots \forall P_n \forall x\, ST(PH\bot),$$

where $p_1, \ldots, p_n$ are the propositional variables occurring in $PH\bot$. But $PH\bot$ contains *no* propositional variables, so the second-order quantification is vacuous. The claim follows from the fact that the standard translation of $PH\bot$ is the formula $\exists y\,(Ryx \wedge \neg\exists z\, Rzy)$.

Definition 3.2 Let τ be a similarity type. A *closed modal formula* is any modal formula in which no proposition letters occur. That is: closed formulas are built up using only $\top$, $\bot$, $\neg$, $\wedge$ and modal operators in τ (plus their duals).

Proposition 3.3 *Let τ be a similarity type, and ϕ a closed τ-formula. Then ϕ expresses a first-order condition c_ϕ which is effectively obtainable from ϕ.*

Proof. By Proposition 2.17 we have for any frame $\mathfrak{F}$,

$$\mathfrak{F} \models \phi \text{ iff } \mathfrak{F} \models \forall P_1 \ldots \forall P_n \forall x\, ST(\phi), \tag{3.4}$$

where $P_1, \ldots, P_n$ are unary predicates corresponding to the proposition letters in ϕ. But if ϕ is closed, $ST(\phi)$ does not contain any unary predicates, and (3.4) reduces to

$$\mathfrak{F} \models \phi \text{ iff } \mathfrak{F} \models \forall x\, ST(\phi). \quad \dashv$$

The classes of positive and negative formulas provide two less trivial examples of why a modal formula may reduce to a first-order condition on frames.

Definition 3.4 An occurrence of a proposition letter p is a *positive* occurrence if it is in the scope of an even number of negation signs; it is a *negative* occurrence if it is in the scope of an odd number of negation signs. A modal formula ϕ is said to be *positive in p* (*negative in p*) if all occurrences of p in ϕ are positive (negative). A formula is called *positive* (*negative*) if it is positive (negative) in all proposition letters occurring in it.

An occurrence of a unary predicate in a first-order formula is *positive* (*negative*) if it is in the scope of an even (odd) number of negation signs.

Lemma 3.5 *Let τ be a similarity type, and let ϕ be a τ-formula.*

1. *Then ϕ is positive in p iff $ST(\phi)$ is positive in the corresponding unary predicate P.*
2. *If ϕ is positive (negative) in p, then $\neg\phi$ is negative (positive) in p.*

Positive (negative) formulas enjoy special properties captured by the following definition.

Definition 3.6 Fix a similarity type τ, and let p be a proposition letter. A modal formula ϕ is *upward monotone in p* if its truth is preserved under extending the valuation of p, or more precisely, if for every model $(W, R_\Delta, V)_{\Delta\in\tau}$, state $w \in W$ and valuation V' such that for all $V(p) \subseteq V'(p)$ and for all $q \neq p$, $V(q) = V'(q)$:

$$\text{if } (W, R_\Delta, V)_{\Delta\in\tau}, w \Vdash \phi, \text{ then } (W, R_\Delta, V')_{\Delta\in\tau}, w \Vdash \phi.$$

In words, a formula ϕ is upward monotone in p if increasing $V(p)$ (and not affecting the interpretation of any other propositional variable) always has the effect of increasing $V(\phi)$.

Likewise, a formula ϕ is *downward monotone in p* if its truth is preserved under shrinking the valuation of p.

The notions of a first-order formula being *upward* and *downward monotone in* a unary predicate P are defined analogously.

Lemma 3.7 *Let τ be a similarity type, and let ϕ be a τ-formula.*

1. *If ϕ is positive in p, then it is upward monotone in p.*

2. *If ϕ is negative in p, then it is downward monotone in p.*

Example 3.8 The following simple formula in standard modal logic corresponds to a first-order condition: $\Diamond\Box p$. For, suppose that $\mathfrak{F} \models \Diamond\Box p$. Consider a state w of $\mathfrak{F}$. Regardless of the valuation at hand, the formula $\Diamond\Box p$ holds at w. Now consider a *minimal* valuation on $\mathfrak{F}$, i.e., a V_m with $V_m(p) = \emptyset$. Then $w \Vdash \Diamond\Box p$ implies the existence of a successor v of w such that $\Box p$ holds at v. However, there are no p-states, so v must be 'blind' (i.e., without successors). So $\mathfrak{F} \models \forall x \exists y\,(Rxy \wedge \neg\exists z\, Ryz))$. In other words, we showed that

$$\mathfrak{F}, a \models \Diamond\Box p \text{ implies } \mathfrak{F} \models \exists y\,(Ray \wedge \neg\exists z\, Ryz)).$$

For the converse direction, assume that every world of a frame $\mathfrak{F}$ has a blind successor. It follows immediately that $(\mathfrak{F}, V_m), w \Vdash \Diamond\Box p$ where V_m is the minimal valuation. We claim that the formula $\Diamond\Box p$ is valid on the frame. To see this, consider an arbitrary valuation V and a point w of $\mathfrak{F}$. Note that $V_m(p) \subseteq V(p)$, since V_m was *minimal*. Now $(\mathfrak{F}, V), w \Vdash \Diamond\Box p$ follows from $(\mathfrak{F}, V_m), w \Vdash \Diamond\Box p$ and Lemma 3.7.

Theorem 3.9 *Fix a similarity type τ, and let ϕ be a τ-formula. If every proposition letter occurring in ϕ occurs only positively or only negatively in ϕ, then ϕ corresponds to a first-order condition c_ϕ on frames. Moreover, c_ϕ can be effectively obtained from ϕ.*

Proof. Consider the universally quantified second-order equivalent of ϕ:

$$\forall P_1 \dots \forall P_n \forall x\, ST(\phi), \tag{3.5}$$

where $P_1, \dots, P_n$ correspond to the proposition letters occurring in ϕ. Our aim is to show that (3.5) is equivalent to a first-order condition by performing appropriate instantiations for the universally quantified variables $P_1, \dots P_n$.

As ϕ is positive or negative in each of its proposition letters, $ST(\phi)$ is positive or negative in each of its unary predicates $P_1, \dots, P_n$. We will instantiate unary predicates that occur only positively with as small a set as possible (viz. the empty set), and we will use as big a set as possible (viz. the whole domain) to instantiate unary predicates that occur only negatively in $ST(\phi)$. Formally, for every P occurring in $ST(\phi)$ define

$$\sigma(P) \equiv \begin{cases} \lambda u.\, u \neq u, & \text{if } ST(\phi) \text{ is positive in } P \\ \lambda u.\, u = u, & \text{if } ST(\phi) \text{ is negative in } P. \end{cases}$$

Now consider the following instance of (3.5) in which every unary predicate P has been replaced by $\sigma(P)$:

$$[\sigma(P_1)/P_1, \dots, \sigma(P_n)/P_n] \forall x\, ST(\phi). \tag{3.6}$$

We will show that (3.6) is equivalent to (3.5). Observe that (3.5) trivially implies (3.6) as the latter is merely an instantiation of the former. For the converse we assume that

$$\mathfrak{M} \models [\sigma(P_1)/P_1, \dots, \sigma(P_n)/P_n] \forall x\, ST(\phi), \tag{3.7}$$

and we have to show that

$$\mathfrak{M} \models \forall P_1 \ldots \forall P_n \forall x\, ST(\phi).$$

By the choice of $\sigma(P)$, for predicates P that occur only positively in $ST(\phi)$ we have that $\mathfrak{M} \models \forall y\, (\sigma(P)(y) \rightarrow P(y))$, and for predicates P that occur only negatively in $ST(\phi)$, we have $\mathfrak{M} \models \forall y\, (P(y) \rightarrow \sigma(P)(y))$. As $ST(\phi)$ is positive or negative in all unary predicates P occurring in it, (3.7) together with Lemma 3.7 implies that for any choice of $P_1, \ldots, P_n$,

$$(\mathfrak{M}, P_1, \ldots, P_n) \models \forall x\, ST(\phi), \quad \text{or} \quad \mathfrak{M} \models \forall P_1 \ldots \forall P_n \forall x\, ST(\phi). \quad \dashv$$

The important point about the proof of Theorem 3.9 is the general idea underlying it: we showed that the formula ϕ in Theorem 3.9 corresponds to a first-order condition on frames by finding a suitable instantiation for its second-order translation. We will now extend the class of formulas covered by the Theorem 3.9 to a class of Sahlqvist formulas to which this method can also be applied — although the instantiations needed will be more complex than the ones used in Theorem 3.9.

Roughly, Sahlqvist formulas are built up from implications of the form

$$\phi \rightarrow \psi,$$

where ψ is positive and ϕ is of a restricted form (to be specified below) from which the required instantiations can be read off. We will first define a limited Sahlqvist fragment of the standard modal language; generalizations and extensions will be discussed later.

Definition 3.10 Consider the standard modal language. A *very simple Sahlqvist antecedent* over this language is a formula built up from $\top$, $\bot$ and proposition letters, using only $\wedge$ and $\Diamond$. A *very simple Sahlqvist formula* is an implication $\phi \rightarrow \psi$ in which ψ is positive and ϕ is a very simple Sahlqvist antecedent.

Example 3.11 Consider the following 'mirror image' of the formula expressing transitivity: $\Diamond p \rightarrow \Diamond\Diamond p$. It expresses *denseness* of the underlying relation R: $\forall xy\, (Rxy \rightarrow \exists z\, (Rxz \wedge Rzy)$.

First, assume that $\mathfrak{F} \models \Diamond p \rightarrow \Diamond\Diamond p$. Suppose that a point a in the frame has a successor b. To show that a and b satisfy the denseness condition, consider the following *minimal* valuation V_m guaranteeing that $(\mathfrak{F}, V_m), a \Vdash \Diamond p$: define

$$V_m(p) = \{b\}.$$

By the assumption, $a \Vdash \Diamond\Diamond p$; so, a must have a successor c such that $c \Vdash \Diamond p$. As b is the *only* point where p holds, this implies that Rcb. The crucial observation is that

$$(\mathfrak{F}, V_m), a \Vdash \Diamond\Diamond p \text{ iff } \mathfrak{F} \models \exists z\, (Raz \wedge Rzb). \tag{3.8}$$

Note that the choice of V_m depends on a *and* on b.

Conversely, let $\mathfrak{F}$ be a dense frame, and assume that under some valuation V, $\Diamond p$ holds at some a in $\mathfrak{F}$. Then a has a successor b such that $b \Vdash p$. Let V_m be the minimal valuation as defined above. As $\mathfrak{F}$ is dense, $(\mathfrak{F}, V_m), a \Vdash \Diamond\Diamond p$ by (3.8). However, $b \Vdash p$ implies $V_m(p) \subseteq V(p)$, so $(\mathfrak{F}, V), a \Vdash \Diamond\Diamond p$ follows from Lemma 3.7 and the fact that $\Diamond\Diamond p$ is positive in p.

The following theorem is the central one in understanding what Sahlqvist correspondence is all about. The reader is advised to follow the proof and understand the algorithm given in it, with a glance at the examples following the theorem.

Theorem 3.12 *Let $\chi = \phi \to \psi$ be a very simple Sahlqvist formula in the standard modal language. Then χ corresponds to a first-order condition c_χ on frames. Moreover, c_χ is effectively obtainable from χ.*

Proof. Consider the universally quantified second-order transcription of χ:

$$\forall P_1 \ldots \forall P_n \forall x\, (ST(\phi) \to ST(\psi)). \tag{3.9}$$

We can make sure that no two quantifications bind the same variable. In a number of steps we will rewrite (3.9) to a formula from which we read off instantiations that yield a first-order equivalent of (3.9).

Step 1. Pull out diamonds.
Use equivalences of the form

$$\forall \cdots ((\ldots \land \exists x_i\, \alpha(x_i) \land \ldots) \to \beta) \quad \leftrightarrow \quad \forall \cdots \forall x_i\, ((\ldots \land \alpha(x_i) \land \ldots) \to \beta)$$

to move all existential quantifiers occurring in the antecedent $ST(\phi)$ of (3.9) out in front. Observe that this is unproblematic as the existential quantifiers only have to cross disjunctions. This process results in a formula of the form

$$\forall P_1 \ldots \forall P_n \forall x \forall x_1 \ldots \forall x_m\, (\text{REL} \land \text{AT} \to ST(\psi)), \tag{3.10}$$

where REL is a conjunction of atomic first-order statements of the form Rx_ix_j corresponding to occurrences of diamonds, and AT is a conjunction of (translations of) atomic formulas.

Step 2. Read off instances.
We can assume that every unary predicate P that occurs in the consequent of the matrix of (3.10), also occurs in the antecedent of the matrix of (3.10): otherwise (3.10) is positive in P and we can substitute $\lambda u.\, u \neq u$ for P to obtain a formula without occurrences of P.

Let P_i be a unary predicate occurring in (3.10), and let $P_ix_{i_1}, \ldots, P_ix_{i_k}$ be all the occurrences of the predicate P_i in the antecedent of (3.10). Define

$$\sigma(P_i) \equiv \lambda u.\, (u = x_{i_1} \lor \ldots \lor u = x_{i_k}).$$

The intuitive idea is that $\sigma(P_i)$ is the *minimal* instance making the antecedent REL $\land$ AT true. It is *essential* to observe that for any model $\mathfrak{M}$:

$$\mathfrak{M} \models \text{AT}[ww_1 \ldots w_n] \text{ implies } \mathfrak{M} \models \forall y\, (\sigma(P_i)(y) \to P_iy)[ww_1 \ldots w_n] \tag{3.11}$$

Step 3. Instantiating.
We now use the formulas of the form $\sigma(P_i)$ found in Step 2 as instantiations; we substitute $\sigma(P_i)$ for each occurrence of P_i in the first-order matrix of (3.10). This results in a formula of the form

$$[\sigma(P_1)/P_1, \ldots, \sigma(P_n)/P_n] \forall x \forall x_1 \ldots \forall x_m\, (\text{REL} \land \text{AT} \to \text{POS}).$$

By the choice of our $\sigma(P)$'s, the formula $[\sigma(P_1)/P_1, \ldots, \sigma(P_n)/P_n]\mathrm{AT}$ will be trivially true. So we end up with an equivalent formula of the form

$$\forall x \forall x_1 \ldots \forall x_m (\mathrm{REL} \rightarrow [\sigma(P_1)/P_1, \ldots, \sigma(P_n)/P_n]\mathrm{POS}). \tag{3.12}$$

As we assumed that every unary predicate occurring in the consequent of (3.10) also occurs in its antecedent, (3.12) must be a first-order formula involving only $=$ and the relation symbol R. So, to complete the proof of the theorem it suffices to show that (3.12) is equivalent to (3.10). The implication from (3.10) to (3.12) is simply an instantiation. To prove the other implication, assume that (3.12) and the antecedent of (3.10) are true:

$$\mathfrak{M} \models \forall x \forall x_1 \ldots \forall x_m (\mathrm{REL} \rightarrow [\sigma(P_1)/P_1, \ldots, \sigma(P_n)/P_n]\mathrm{POS}), \text{ and}$$

$$\mathfrak{M} \models \mathrm{REL} \wedge \mathrm{AT}[ww_1 \ldots w_m].$$

We need to show that $\mathfrak{M} \models \mathrm{POS}[ww_1 \ldots w_m]$. First of all, the above assumptions imply

$$\mathfrak{M} \models [\sigma(P_1)/P_1, \ldots, \sigma(P_n)/P_n]\mathrm{POS}[ww_1 \ldots w_m].$$

As POS is positive, it is upwards monotone in all unary predicates occurring in it, so it suffices to show that $\mathfrak{M} \models \forall y (\sigma(P_i) \rightarrow P_i)[ww_1 \ldots w_m]$. But, by the essential observation in Step 2, this is precisely what the assumption $\mathfrak{M} \models \mathrm{AT}[ww_1 \ldots w_m]$ amounts to. ⊣

Example 3.13 First consider the formula $p \rightarrow \Diamond p$. Its second-order equivalent is the formula

$$\forall P \forall x (\underbrace{Px}_{\mathrm{AT}} \rightarrow \exists z (Rxz \wedge Pz)).$$

There are no diamonds to be pulled out here, so we can read off the minimal instance $\sigma(P) \equiv \lambda u.\, u = x$ immediately. Instantiation gives

$$\forall x (x = x \rightarrow \exists z (Rxz \wedge z = x)),$$

which reduces to the formula $\forall x\, Rxx$.

Our second example is the density formula $\Diamond p \rightarrow \Diamond\Diamond p$, having a second-order equivalent

$$\forall P \forall x (\exists x_1 (Rxx_1 \wedge Px_1) \rightarrow \exists z_0 (Rxz_0 \wedge \exists z_1 (Rz_0z_1 \wedge Pz_1))).$$

Here we can pull out the diamond $\exists x_1$:

$$\forall P \forall x \forall x_1 (\underbrace{Rxx_1}_{\mathrm{REL}} \wedge \underbrace{Px_1}_{\mathrm{AT}} \rightarrow \exists z_0 (Rxz_0 \wedge \exists z_1 (Rz_0z_1 \wedge Pz_1))).$$

Instantiating with $\sigma(P) \equiv \lambda u.\, u = x_1$ gives

$$\forall x \forall x_1 (Rxx_1 \wedge x_1 = x_1 \rightarrow \exists z_0 (Rxz_0 \wedge \exists z_1 (Rz_0z_1 \wedge z_1 = x_1))),$$

which can be simplified to $\forall x \forall x_1 (Rxx_1 \rightarrow \exists z_0 (Rxz_0 \wedge Rz_0x_1))$.

Our last example is the formula $(p \wedge \Diamond\Diamond p) \rightarrow \Diamond p$. Its second-order equivalent is

$$\forall P \forall x (Px \wedge \exists x_1 (Rxx_1 \wedge \exists x_2 (Rx_1x_2 \wedge Px_2)) \rightarrow \exists z_0 (Rxz_0 \wedge Pz_0)).$$

Pulling out the diamonds $\exists x_1$ and $\exists x_2$ results in

$$\forall P \forall x \forall x_1 \forall x_2 \,(\underbrace{Rxx_1 \wedge Rx_1x_2}_{\text{REL}} \wedge \underbrace{Px \wedge Px_2}_{\text{AT}} \rightarrow \exists z_0\,(Rxz_0 \wedge Pz_0))).$$

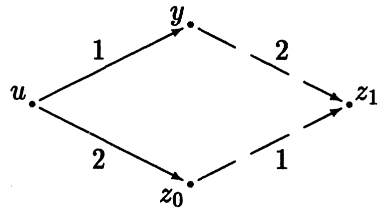
The minimal instantiation here is $\sigma(P) \equiv \lambda u.\,(u = x \vee u = x_2)$. After instantiating we find

$$\forall x \forall x_1 \forall x_2\,(Rxx_1 \wedge Rx_1x_2 \wedge (x = x \vee x = x_2) \wedge (x_2 = x \vee x_2 = x_2) \rightarrow \\ \exists z_0\,(Rxz_0 \wedge (z_0 = x \vee z_0 = x_2)))$$

This formula simplifies to $\forall x \forall x_1 \forall x_2\,(Rxx_1 \wedge Rx_1x_2 \rightarrow (Rxx \vee Rxx_2))$.

In the last part of this section we will show that the strategy of finding minimal instantiations works for more complex Sahlqvist antecedents as well.

Example 3.14 Consider the formula $\Diamond_1\Box_2 p \rightarrow \Box_2\Diamond_1 p$; we will show that this formula corresponds to some kind of *confluence* property of R_1 and R_2: $\forall xyz_0\,(R_1xy \wedge R_2xz_0 \rightarrow \exists z_1(R_2yz_1 \wedge R_1z_0z_1))$. The name 'confluence' is explained by the following picture:



First, let $\mathfrak{F} = (W, R_1, R_2)$ be a frame such that $\mathfrak{F} \models \Diamond_1\Box_2 p \rightarrow \Box_2\Diamond_1 p$, and let a and b be states in $\mathfrak{F}$ such that R_1ab. For a valuation to verify $\Diamond_1\Box_2 p$ at a it suffices that p holds at all R_2-successors of b. So a *minimal* such valuation can be defined as

$$V_m(p) = \{u \in W \mid R_2bu\}.$$

It follows that $(\mathfrak{F}, V_m), a \Vdash \Box_2\Diamond_1 p$. The crucial observation is that by the choice of V_m:

$$(\mathfrak{F}, V_m), a \Vdash \Box_2\Diamond_1 p \text{ iff } \forall z_0\,(R_2az_0 \rightarrow \exists z_1\,(R_2bz_1 \wedge R_1z_0z_1)). \tag{3.13}$$

It follows that $\mathfrak{F} \models \forall xyz_0\,(R_1xy \wedge R_2xz_0 \rightarrow \exists z_1\,(R_2yz_1 \wedge R_1z_0z_1))$.

Conversely, assume that $\mathfrak{F}$ has the confluence property. In order to show that $\mathfrak{F} \models \Diamond_1\Box_2 p \rightarrow \Box_2\Diamond_1 p$, let V be a valuation on $\mathfrak{F}$ and a an arbitrary state of $\mathfrak{F}$ such that $(\mathfrak{F}, V), a \Vdash \Diamond_1\Box_2 p$. We have to prove that $a \Vdash \Box_2\Diamond_1 p$. It immediately follows by the truth definition of $\Diamond_1$ that a has an R_1-successor b satisfying R_1ab and $b \Vdash \Box_2 p$. Now we use the minimal valuation V_m again; first note that by the definition of V_m, we have $V_m(p) \subseteq V(p)$. Therefore, Lemma 3.7 ensures that it suffices to show that $\Box_2\Diamond_1 p$ holds at a *under the valuation* V_m. But this is immediate by the assumption that $\mathfrak{F}$ is confluent and (3.13).

Definition 3.15 Let τ be a modal similarity type. A *boxed atom* is a formula of the form $\Box_{i_1} \cdots \Box_{i_k} p$ $(k \geq 0)$, where $\Box_{i_1}, \ldots \Box_{i_k} p$ are (not necessarily distinct) boxes of the language. In the case where $k = 0$, the boxed atom $\Box_{i_1} \cdots \Box_{i_k} p$ is just the atom p.

It is convenient to treat sequences of boxes as single boxes. We will therefore denote the formula $\Box_{i_1} \cdots \Box_{i_k} p$ by $\Box_\beta p$, where β is the sequence $i_1 \ldots i_k$ of indices. Analogously, we will pretend to have a corresponding dyadic predicate R_β in the frame language $\mathcal{L}^1_\tau$. Thus the expression $R_\beta xy$ abbreviates

$$\exists y_1 (R_{i_1} x y_1 \wedge \exists y_2 (R_{i_2} y_1 y_2 \wedge \ldots \wedge \exists y_{k-1} (R_{i_{k-1}} y_{k-2} y_{k-1} \wedge R_{i_k} y_{k-1} y) \ldots)).$$

Definition 3.16 Let τ be a modal similarity type. A *simple Sahlqvist antecedent* over this similarity type is a formula built up from $\top$, $\bot$ and boxed atoms, using only $\wedge$ and existential modal operators ($\Diamond$ and Δ). A *simple Sahlqvist formula* is an implication $\phi \to \psi$ in which ψ is positive and ϕ is a simple Sahlqvist antecedent.

Example 3.17 Typical examples of simple Sahlqvist formulas are the following: $\Diamond p \to \Diamond\Diamond p$, $\Box p \to \Box\Box p$, $\Box_1\Box_2 p \to \Box_3 p$, $\Diamond_1\Box_2 p \to \Box_2\Diamond_1 p$ and $(\Box_1\Box_2 p)\Delta(\Diamond_3 p \wedge \Box_2\Box_1 q) \to \Diamond_3(q\Delta p)$.

Typically forbidden in a Sahlqvist antecedent are:

- 'boxes over disjunctions', as in $H(r \vee Fq) \to G(Pr \wedge Pq)$
- 'boxes over diamonds', as in $\Box\Diamond p \to \Diamond\Box p$
- 'dual-triangled atoms', as in $p\nabla p \to p$.

Theorem 3.18 *Let τ be a modal similarity type, and let $\chi = \phi \to \psi$ be a simple Sahlqvist formula over τ. Then χ corresponds to a first-order condition c_χ on frames. Moreover, c_χ is effectively obtainable from χ.*

Proof. The proof of this theorem is an adaptation of the proof of Theorem 3.12. We consider the universally quantified second-order transcription of χ, and identify appropriate instantiations that turn it into an equivalent first-order statement. After we've pulled out diamonds (Step 1) we end up with a formula of the form

$$\forall P_1 \ldots \forall P_n \forall x \forall x_1 \ldots \forall x_m (\text{REL} \wedge \text{BOX-AT} \to ST(\psi)), \tag{3.14}$$

with REL as before, and BOX-AT a conjunction of (translations of) boxed atoms.

Let P be a unary predicate occurring in (3.14), and let $\pi_1(x_{i_1}), \ldots, \pi_k(x_{i_k})$ be all the (translations of the) boxed atoms in the antecedent of (3.10) in which P occurs. Observe that every π_j is of the form $\forall y (R_{\beta_j} x_{i_j} y \to Py)$, where β_j is a sequence of diamond indices (cf. Definition 3.15). Define

$$\sigma(P) \equiv \lambda u. (R_{\beta_1} x_{i_1} u \vee \ldots \vee R_{\beta_k} x_{i_k} u).$$

Again, the intuitive idea is that $\sigma(P_1), \ldots, \sigma(P_n)$ form the *minimal* instances making the antecedent REL $\wedge$ BOX-AT true. The remainder of the proof is the same as the proof of Theorem 3.12 (of course, all occurrences of 'AT' should be replaced with 'BOX-AT'). $\dashv$

As in the case of very simple Sahlqvist formulas, the algorithm is best understood by inspecting some examples; due to space limitations we have to leave this to the reader.

To conclude the section we briefly describe the full Sahlqvist fragment. First, a Sahlqvist antecedent is a formula which is built from constants, boxed atoms and negative formulas, using only $\wedge$, $\vee$ and existential modal operators. Then, a formula is a *Sahlqvist formula* if it is built from implications $\phi \to \psi$ in which ϕ is a Sahlqvist antecedent and ψ a positive formula, using only conjunctions, disjunctions between formulas that don't share proposition letters, and boxes. The result is that all Sahlqvist formulas express first-order conditions on frames, and that these conditions can be effectively obtained via the substitution method of Theorems 3.12 and 3.18.

4. Completeness through canonicity

In this section we discuss *normal modal logics*. Such logics can be defined both syntactically and semantically, thus giving rise to the questions which dominate this section: given a syntactically specified logic Λ and a semantically specified logic Λ_{S}, have we got *soundness* (that is: $\Lambda \subseteq \Lambda_{\mathsf{S}}$) and *completeness* (that is: $\Lambda_{\mathsf{S}} \subseteq \Lambda$)? Soundness results tend to be routine, and the main goal of this section is to develop a general tool for establishing completeness, namely the use of *canonical models*.

4.1 Preliminaries

Throughout this section we assume we are working with languages with a countably infinite collection of proposition symbols. We say that a modal formula is a *tautology* if it has the form of a propositional tautology when all its modal subformulas are viewed as atomic symbols.

Definition 4.1 A *logic* Λ is a set of modal formulas that contains all tautologies and is closed under modus ponens (that is, if $\phi \in \Lambda$ and $\phi \to \psi \in \Lambda$ then $\psi \in \Lambda$). The formulas in a logic are its *theorems*. If ϕ is a theorem of Λ we write $\vdash_\Lambda \phi$, and if not, $\not\vdash_\Lambda \phi$.

To give some examples, the collection of all formulas is a logic, the *inconsistent logic*. Also, if $\{\Lambda_i \mid i \in I\}$ is a collection of logics then $\bigcap_{i \in I} \Lambda_i$ is a logic too. And if we define Λ_{S} to be $\{\phi \mid \mathfrak{S} \models \phi$, for all structures $\mathfrak{S} \in \mathsf{S}\}$, where S is a class of frames (or models), then Λ_{S} is a logic.

It follows that for any collection of formulas Γ there is a smallest logic containing Γ, namely $\bigcap\{\Lambda \mid \Lambda$ is a logic $\Gamma \subseteq \Lambda\}$. This intersection cannot be empty, for Γ contains the inconsistent logic. The smallest logic containing Γ is called the *logic generated by* Γ. For example, the logic generated by $\emptyset$ contains all instances of propositional validities in the modal language, and nothing else.

Definition 4.2 Let $\psi_1, \ldots, \psi_n, \phi$ be modal formulas. We say that ϕ is *deducible in propositional calculus from assumptions* $\psi_1 \ldots \psi_n$ if $(\psi_1 \to (\psi_2 \to (\psi_n \to \phi))\cdots)$ is a tautology.

It is easily verified that all logics are closed under deduction in propositional calculus, in the sense that if ϕ is deducible in propositional calculus from $\psi_1, \ldots, \psi_n$, and $\vdash_\Lambda \psi_1, \ldots, \vdash_\Lambda \psi_n$, then $\vdash_\Lambda \phi$.

Definition 4.3 If $\Gamma \cup \{\phi\}$ is a set of formulas then we say ϕ *is deducible in* Λ *from* Γ (or more simply: ϕ *is* Λ*-deducible from* Γ) if there are finitely many formulas $\psi_1, \ldots, \psi_n \in \Gamma$ such that

$$\vdash_\Lambda (\psi_1 \to (\psi_2 \to (\psi_n \to \phi))\cdots).$$

If this is the case we write $\Gamma \vdash_\Lambda \phi$, and if not, $\Gamma \not\vdash_\Lambda \phi$. A set of formulas is Λ-consistent if $\Gamma \not\vdash_\Lambda \bot$, and inconsistent otherwise. A formula ϕ is consistent if $\{\phi\}$ is, and inconsistent otherwise.

We leave it to the reader to check that a set of sentences Γ is Λ-inconsistent iff there is a formula ϕ such that $\Gamma \vdash_\Lambda \phi \wedge \neg\phi$, iff for all formulas ψ, $\Gamma \vdash_\Lambda \psi$. Also, a set of sentences Γ is Λ-consistent iff every finite subset of Λ is.

The definitions and results we have encountered so far apply to modal languages of any similarity type. Now we introduce *normal modal logics*. To keep matters simple, we restrict our initial discussion to languages of the basic modal similarity type, deferring the full definition till the end of the section.

Definition 4.4 A modal logic Λ is *normal* if it contains the axiom

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

and is closed under *universal generalization*, that is, $\vdash_\Lambda \phi$ implies $\vdash_\Lambda \Box\phi$, and under *uniform substitution*. The latter is defined as follows. If $\vdash_\Lambda \phi$, and $q_1, \ldots, q_n$ are atomic symbols, and $\sigma_1, \ldots, \sigma_n$ are arbitrary formulas, then $\vdash_\Lambda \phi'$, where ϕ' is the formula obtained by simultaneously replacing all occurrences of q_i in ϕ by σ_i $(1 \leq i \leq n)$.

The K schema is sometimes called the *distribution schema*. Intuitively, it enables us to transform a formula in which a modality is the main connective (that is, $\Box(p \to q)$) into a formula $\Box p \to \Box q$ in which a propositional connective (namely, $\to$) is the main connective. This makes it possible to apply further purely propositional reasoning. For example, we can apply modus ponens 'under the scope of a box'. Suppose we are given $\Box(p \to q)$ and $\Box p$. We know our normal modal logic contains $\Box(p \to q) \to (\Box p \to \Box q)$. One application of (ordinary) modus ponens yields $\Box p \to \Box q$, thus making a second implication available for an application of (ordinary) modus ponens. This application yields $\Box q$.

Remark 4.5 It is sometimes convenient to consider an equivalent formulation of normal modal logics involving diamond instead of boxes. As a lemma, a logic Λ is normal iff it is closed under the rule of substition and satisfies the following:

1. $\Diamond\bot \to \bot$
2. $\Diamond(p \vee q) \to \Diamond p \vee \Diamond q$
3. $\vdash_\Lambda \phi \to \psi$ implies $\vdash_\Lambda \Diamond\phi \to \Diamond\psi$.

To see, for example that any normal modal logic derives $\Diamond\bot \to \bot$, observe that $\vdash \top$ implies $\vdash \Box\top$, or $\vdash \neg\Diamond\neg\top$ or $\vdash \Diamond\bot \to \bot$. And conversely, if Λ satisfies 1–3 above, then if $\vdash_\Lambda \phi$, then $\vdash_\Lambda \neg\phi \to \bot$, so $\vdash_\Lambda \Diamond\neg\phi \to \Diamond\bot$, so $\vdash_\Lambda \neg\Diamond\neg\phi \to \bot$, so $\vdash_\Lambda \Box\phi$, which means that Λ is closed under generalization.

Example 4.6 The inconsistent logic is a normal logic. Also, if $\{\Lambda_i \mid i \in I\}$ is a collection of normal modal logics then so is $\bigcap_{i \in I} \Lambda_i$. Further, if S is a class of frames, then Λ_{S} is a normal modal logic. However, if S is a class of models, then Λ_{S} need not be closed under uniform substitution; consider a model $\mathfrak{M}$ in which p is true at all nodes but q is not. Then $\vdash_{\Lambda_{\mathfrak{M}}} p$ but $\not\vdash_{\Lambda_{\mathfrak{M}}} q$; but q can be obtained from p by uniform substitution.

These examples demonstrate that there are both semantic and syntactic perspectives on normal modal logics. The third example tells us that the formulas validated by any class of frames is a normal modal logic.

The syntactic perspective arises via the first two examples. They jointly guarantee that for any set of formulas Γ, it makes sense to talk of the smallest normal modal logic containing Γ; this logic is just

$$\bigcap\{\Lambda \mid \Lambda \text{ is a normal modal logic and } \Gamma \subseteq \Lambda\}.$$

We call this the *normal modal logic generated by* Γ. For example, the smallest normal modal logic is the one generated by $\emptyset$; it is called **K**, in honor of Saul Kripke.

The generative perspective is essentially syntactic. We can regard Γ as a set of *axioms*, and modus ponens, generalization and uniform substitution as *rules of inference*. The logic generated by Γ is simply the set of all formulas deducible from the axioms using the rules of inference.

Example 4.7 The following lists some of the better known axioms in standard modal logic, together with their traditional names:

(T) $\Box p \to p$
(4) $\Box p \to \Box\Box p$
(B) $p \to \Box\Diamond p$
(5) $\Diamond p \to \Box\Diamond p$
(D) $\Box p \to \Diamond p$
(.3) $\Diamond p \wedge \Diamond q \to ((\Diamond(p \wedge \Diamond p) \vee (\Diamond(p \wedge q) \vee (\Diamond(q \wedge \Diamond p))$
(L) $\Box(\Box p \to p) \to \Box p$
(M) $\Box\Diamond p \to \Diamond\Box p$.

There is a convention for talking about the logics generated by such axioms: if $S_1, \ldots, S_n$ are axioms then $\mathbf{K}\mathbf{S}_1, \ldots, \mathbf{S}_n$ is the normal logic generated by $S_1, \ldots, S_n$. But irregularities abound. Many historical names are firmly entrenched, thus modal logicians tend to talk of the logics **T**, **S4**, **B**, and **S5** instead of **KT**, **KT4**, **KB** and **KT45** respectively. Moreover, many schemas have multiple names. For example, the schema we call L (for Löb) is also known as G (for Gödel) and W (for well-founded).

As with the basic normal modal logic **K**, for the above extensions it is sometimes convenient to consider diamond versions of the axioms instead of the more traditional box versions.

We turn now to the fundamental concepts linking the syntactic and semantic perspectives: *soundness* and *completeness*.

Definition 4.8 Let S be a class of frames (or models). A normal modal logic Λ is *sound* with respect to S if $\Lambda \subseteq \Lambda_{\mathsf{S}}$. If Λ is sound with respect to S we say that S *is a class of frames* (or models) *for* Λ.

We will usually be interested in proving soundness for syntactically specified logics Λ, that is, for logics Λ generated by a collection of formulas Γ.

Example 4.9 Here's a list of soundness results; each of the logics on the left is sound with respect to the class of frames (models) satisfying the condition on the right.

K	no condition
T	reflexivity
KB	symmetry
K4	transitivity
S4	transitivity and reflexivity
K5	euclidicity
S5	transitivity, reflexivity and symmetry
KD	right-unboundedness
K4.3	transitivity and right-linearity
KLöb	finite trees
K4McKinsey	atomicity.

Here, a frame is called *right-unbounded* if it satisfies $\forall w \exists w' \, Rww'$. And a frame is *right-linear* if it satisfies $\forall w \forall w' \forall w'' \, (Rww' \wedge Rww'' \rightarrow (Rw'w'' \vee w' = w'' \vee Rw''w'))$. A frame is *atomic* if it satisfies $\forall w \exists w' \, (Rww' \wedge Rw'w')$.

The above claims (with the exception of the last two) are easily demonstrated. Indeed, soundness proofs are often routine; we rarely bother to explicitly state or prove the soundness theorem. The sister concept, *completeness*, leads to much harder problems.

Definition 4.10 To define completeness, we first define semantic consequence '$\Gamma \models_{\mathsf{S}} \phi$'. If S is a class of models, then '$\Gamma \models_{\mathsf{S}} \phi$' means that for all $\mathfrak{M} \in \mathsf{S}$ and all w, if $\mathfrak{M}, w \Vdash \Gamma$, then $\mathfrak{M}, w \Vdash \phi$. And if S is a class of frames, then it means: for all $\mathfrak{F} \in \mathsf{S}$, for all valuations V on $\mathfrak{F}$, and for all states w in $\mathfrak{F}$, if $(\mathfrak{F}, V), w \Vdash \Gamma$, then $(\mathfrak{F}, V), w \Vdash \phi$.

Let S be a class of frames (or models). A logic Λ is *strongly complete* with respect to S if for all sets of formulas $\Gamma \cup \{\phi\}$, $\Gamma \models_{\mathsf{S}} \phi$ implies $\Gamma \vdash_{\Lambda} \phi$. A logic Λ is *weakly complete* with respect to S iff for any formula ϕ, if $\models_{\mathsf{S}} \phi$ then $\vdash_{\Lambda} \phi$.

Λ is strongly (weakly) complete with respect to a single structure $\mathfrak{S}$ iff Λ is strongly (weakly) complete with respect to $\{\mathfrak{S}\}$.

Note that weak completeness is the special case of strong completeness in which $\Gamma = \emptyset$, thus strong completeness with respect to some class of structures implies weak completeness with respect to that same class. The definition of weak completeness can be reformulated to parallel the definition of soundness: Λ is weakly complete with respect to S iff $\Lambda_{\mathsf{S}} \subseteq \Lambda$.

Example 4.11 The following completeness results will be proved in the next subsection; each of the logics **K**, **T**, **KB**, **K4**, **S4**, and **S5** is strongly complete with respect to the class of frames satisfying the properties mentioned in Example 4.9.

These completeness results (together with their soundness counterparts in Example 4.9) give simple semantic characterizations of normal modal logics such as **T**, **K4**, **S4** and **S5**; hitherto we only had syntactic definitions of these systems.

Below we will make use of the following characterization.

Proposition 4.12 *A logic Λ is strongly complete with respect to a class of structures S iff every Λ-consistent set of sentences is satisfiable on some $\mathfrak{S} \in \mathsf{S}$. Λ is weakly complete with respect to a class of structures S iff every Λ-consistent sentence is satisfiable on some $\mathfrak{S} \in \mathsf{S}$.*

Proof. The result for weak completeness follows from the characterization of strong completeness, so we examine only the latter. To prove the right to left implication we argue by contraposition. Suppose Λ is not strongly complete with respect to S. Thus there is a set of formulas $\Gamma \cup \{\phi\}$ such that $\Gamma \models_{\mathsf{S}} \phi$ but $\Gamma \not\vdash_{\Lambda} \phi$. Then $\Gamma \cup \{\neg\phi\}$ is Λ-consistent but not satisfiable on any structure in S. The left to right implication is left to the reader. ⊣

To conclude this subsection, we generalize the definition of normal modal logics.

Definition 4.13 Let τ be a similarity type. A *modal logic* in this language is a set of formulas containing all tautologies that is closed under modus ponens. A modal logic Λ is *normal* if it contains all instances of the following axiom, for all operators ∇ in the language:

$$(\mathrm{K}_\tau) \quad \nabla(p_1, \ldots, q \to r, \ldots, p_{\rho(\nabla)}) \to (\nabla(p_1, \ldots, q, \ldots, p_{\rho(\nabla)}) \to \nabla(pa_1, \ldots, r, \ldots, p_{\rho(\nabla)})),$$

and is closed under *uniform substitution* and *generalization* for all operators:

$$\vdash_\Lambda \sigma_1, \ldots, \vdash_\Lambda \sigma_{\rho(\nabla)} \text{ implies } \vdash_\Lambda \nabla(\sigma_1, \ldots, \sigma_{\rho(\nabla)}).$$

4.2 Canonical models

By Proposition 4.12, to prove completeness results it suffices to build models, and to do so we use maximal consistent sets of formulas as building blocks.

Definition 4.14 A set of formulas Γ is called *maximal Λ-consistent* if Γ is Λ-consistent, and any set of formulas properly containing Γ is Λ-inconsistent. If Γ is maximal Λ-consistent then we say it is a Λ-MCS.

Lemma 4.15 (Lindenbaum's lemma) *If Σ is a Λ-consistent set of sentences then there is a Λ-MCS Σ^+ such that $\Sigma \subseteq \Sigma^+$.*

Proof. Let ϕ_0, ϕ_1, $\phi_2, \ldots$ be an enumeration of the formulas of our language. We will define the set Σ^+ as the union of a chain of consistent sets $\Delta_0, \ldots, \Delta_n, \ldots$ as follows:

$$\begin{aligned} \Delta_0 &= \Sigma \\ \Delta_{n+1} &= \begin{cases} \Delta_n \cup \{\phi_n\}, & \text{if } \Delta_n \vdash_\Lambda \phi_n \\ \Delta_n \cup \{\neg\phi_n\}, & \text{otherwise.} \end{cases} \\ \Sigma^+ &= \textstyle\bigcup_{n \geq 0} \Delta_n. \end{aligned}$$

The proof of the following properties of Σ^+ is left as an exercise: (1) Δ_n is Λ-consistent, for all n; (2) exactly one of ϕ and $\neg\phi$ is in Σ^+, for every formula ϕ; (3) if $\Sigma^+ \vdash_\Lambda \phi$, then $\phi \in \Sigma^+$; and finally (4) Σ^+ is maximal Λ-consistent. ⊣

Proposition 4.16 *If Λ is a logic and Γ is a Λ-MCS then:*

1. $\Lambda \subseteq \Gamma$.
2. *For all formulas ϕ: $\phi \in \Gamma$ or $\neg\phi \in \Gamma$.*

3. *For all formulas ϕ, ψ: $\phi \wedge \psi \in \Gamma$ iff $\phi \in \Gamma$ and $\psi \in \Gamma$.*

Note that if $\mathfrak{M}$ is any model for a logic Λ, and w is any node in $\mathfrak{M}$, then $\{\phi \mid \mathfrak{M}, w \models \phi\}$ is a Λ-MCS. It must be consistent. And as for all formulas ψ either ψ or $\neg\psi$ is true in $\mathfrak{M}$ at w, we cannot add any formula without introducing inconsistency; this proves maximality. Thus MCSs correspond to complete states of affairs. The completeness proof builds on this intuition.

Definition 4.17 The *canonical model* $\mathfrak{M}^\Lambda$ for a normal modal logic Λ is defined as the triple $(W^\Lambda, R^\Lambda, V^\Lambda)$ where:

1. W^Λ is the set of all Λ-MCSs.
2. $R^\Lambda \subseteq W^\Lambda \times W^\Lambda$ is defined by $R^\Lambda wu$ if for all formulas ψ, $\psi \in u$ implies $\Diamond\psi \in w$. (Equivalently: $R^\Lambda wu$ if for all formulas ψ, $\Box\psi \in w$ implies $\psi \in u$.)
3. V^Λ is the valuation defined by $V(p) = \{w \in W^\Lambda \mid p \in w\}$.

The pair $\mathfrak{F}^\Lambda = (W^\Lambda, R^\Lambda)$ is called the *canonical frame* for Λ.

Lemma 4.18 (Truth lemma) *For any normal modal logic Λ and any formula ϕ we have $\mathfrak{M}^\Lambda, w \Vdash \phi$ iff $\phi \in w$.*

Proof. By induction on ϕ. The base case follows from the definition of V^Λ. The boolean cases follow from Proposition 4.16. It remains to deal with the modalities. One direction follows from the definition of R^Λ:

$$\begin{array}{lll} \mathfrak{M}^\Lambda, w \models \Diamond\phi & \text{implies} & \exists v\,(R^\Lambda wv \wedge \mathfrak{M}^\Lambda, v \Vdash \phi) \\ & \text{implies} & \exists v\,(R^\Lambda wv \wedge \phi \in v) \qquad \text{(Inductive Hypothesis)} \\ & \text{implies} & \Diamond\phi \in w \qquad \text{(Definition } R^\Lambda\text{)}. \end{array}$$

The converse is more interesting. Suppose $\Diamond\phi \in w$. We want to show that $\mathfrak{M}, w \Vdash \Diamond\phi$. That is, we want to find a node v such that $R^\Lambda wv$ and $\mathfrak{M}, v \Vdash \phi$. By the inductive hypothesis it suffices to find an MCS v such that $R^\Lambda wv$ and $\phi \in v$. We will construct such an MCS.

Let v^- be $\{\phi\} \cup \{\psi \mid \Box\psi \in w\}$. Then v^- is consistent. For suppose not. Then there are $\psi_1, \ldots, \psi_n \in w$ such that: $\vdash_\Lambda (\psi_1 \to (\psi_2 \to \cdots (\psi_n \to (\phi \to \bot))\cdots)$. Abbreviate $\psi_2 \to \cdots (\psi_n \to (\phi \to \bot)\cdots)$ to θ. Thus we have $\vdash_\Lambda \psi_1 \to \theta$. Now argue as follows. As Λ is normal it is closed under generalization, hence $\vdash_\Lambda \Box(\psi_1 \to \theta)$. Moreover, Λ contains all instances of K, so $\vdash_\Lambda \Box(\psi_1 \to \theta) \to (\Box\psi_1 \to \Box\theta)$. As Λ is closed under modus ponens, $\vdash_\Lambda \Box\psi_1 \to \Box\theta$.

Repeating this argument $n-1$ times we find that

$$(\Box\psi_1 \to (\Box\psi_2 \to \cdots (\Box\psi_n \to \Box(\phi \to \bot))\cdots)) \tag{4.15}$$

is a theorem of Λ. As w is an a Λ-MCS by Proposition 4.16 we conclude that the formula (4.15) is in w. In addition, $\Box\psi_1, \ldots \Box\psi_n \in w$, and so $\Box\neg\phi \in w$. But this is impossible: by assumption w is consistent and contains $\Diamond\phi$. We conclude that v^- is consistent.

Let v be any MCS extending v^-; such extensions exist by Lindenbaum's Lemma. By construction $\phi \in v$. Furthermore, for all formulas ψ, $\Box\psi \in w$ implies $\psi \in v$. So $R^\Lambda wv$. $\dashv$

Theorem 4.19 (Canonical Model Theorem) *Every normal modal logic Λ is strongly complete with respect to its canonical model $\mathfrak{M}^\Lambda$.*

Proof. Suppose Σ is Λ consistent. By Lindenbaum's Lemma there is a Λ-MCS Σ^+ extending Σ. By the previous lemma, $\mathfrak{M}^\Lambda, \Sigma^+ \Vdash \Sigma$. $\dashv$

Although Theorem 4.19 is a universal completeness result for normal modal logics, it has the drawback of being rather abstract. However, for many important logics the canonical model theorem contains all the information needed to give simple proofs of more concrete completeness results such as those mentioned in Example 4.11.

We are particularly interested in frame completeness results: given a normal modal logic Λ and a class F of frames for Λ, show that Λ is strongly complete with respect to F. (For example, we might want to prove that **T** is strongly complete with respect to the class of reflexive frames.) Proofs of such results must establish two things:

1. That there is a model $\mathfrak{M}$ for any Λ-consistent set of sentences Σ, and

2. That the frame underlying $\mathfrak{M}$ belongs to F.

The basic idea below is to use the canonical models to short circuit this process. We are simply going to use the canonical model for Λ to establish step 1. Thus proving completeness reduces to establishing step 2, that is, showing that the canonical frame for Λ belongs to F. The following definition captures this idea.

Definition 4.20 A normal modal logic Λ is *canonical* if its canonical frame is a frame for Λ. That is, Λ is canonical if for all ϕ such that $\vdash_\Lambda \phi$, ϕ is valid on the canonical frame for Λ. A set of formulas Σ is canonical if the logic $\mathbf{K}\boldsymbol{\Sigma}$ is.

Clearly, every canonical logic is strongly complete. We will use this new terminology to prove some of the results listed in Example 4.11.

Theorem 4.21 *Each of the logics* **K**, **K4**, **T**, **KB**, **KD**, **S4**, *and* **S5** *is canonical, and hence strongly complete.*

Proof. For each of the logics mentioned, one can show that the canonical frame for the logic has the properties mentioned in Example 4.11. By way of illustration we will do this for **K** and for **K4**, leaving the other cases to the reader.

We start with **K**. This case is really trivial, as any frame is a frame for **K**, in particular the canonical frame $\mathfrak{F}^{\mathbf{K}}$. As to **K4**, we have to show that the canonical frame $(W^{\mathbf{K4}}, R^{\mathbf{K4}}, V^{\mathbf{K4}})$ for **K4** is transitive. So suppose w, v and u are points in this frame such that $R^{\mathbf{K4}}wv$ and $R^{\mathbf{K4}}vu$. We wish to show that $R^{\mathbf{K4}}wu$. Suppose $\phi \in u$. As $R^{\mathbf{K4}}vu$, $\Diamond\phi \in v$. So as $R^{\mathbf{K4}}wv$, $\Diamond\Diamond\phi \in w$. But as w is a **K4**-MCS it contains $\Diamond\Diamond\phi \to \Diamond\phi$, hence by closure under modus ponens it contains $\Diamond\phi$. Thus $R^{\mathbf{K4}}wu$. $\dashv$

A lot more can be said about canonicity than we have room for here; in the following subsection we briefly discuss the limitations of the canonical model method, and we also mention some alternatives; in Section 6 and 7 of Part III an algebraic perspective on canonicity will be offered.

To conclude this subsection we briefly sketch a generalization of the canonical model theorem to languages of arbitrary similarity types. Let τ be a modal similarity type, and Λ a normal modal logic in the language over τ. The canonical model $\mathfrak{M}^\Lambda = (W^\Lambda, R^\Lambda_\Delta, V^\Lambda)_{\Delta\in\tau}$ for Λ has W^Λ and V^Λ defined as in Definition 4.17, while for an n-ary operator $\Delta \in \tau$ the relation $R^\Lambda_\Delta \subseteq (W^\Lambda)^{n+1}$ is defined by

$$R^\Lambda_\tau \quad \text{iff} \quad \text{for all formulas } \psi_1 \in u_1, \ldots, \psi_n \in u_n \text{ we have } \Delta(\psi_1, \ldots, \psi_n) \in w.$$

Given this definition a Truth Lemma (Lemma 4.18) can be established. The only complication here is the step where, starting from the assumption that $\Delta(\psi_1, \ldots, \psi_n) \in w$, we need to show the existence of $u_1, \ldots, u_n$ with $\psi_1 \in u_1, \ldots, \psi_n \in u_n$. For simplicity assume that $n = 2$, that is: Δ is binary, and $\Delta(\psi_1, \psi_2) \in w$. Let $\phi_0, \phi_1, \ldots$ enumerate all formulas. We construct two sequences of sets of formulas

$$\{\psi_1\} = \Pi_0 \subseteq \Pi_1 \subseteq \cdots \quad \text{and} \quad \{\psi_2\} = \Sigma_0 \subseteq \Sigma_1 \subseteq \cdots$$

such that all Π_i and Σ_i are finite and consistent, Π_{i+1} is either $\Pi_i \cup \{\phi_i\}$ or $\Pi_i \cup \{\neg\phi_i\}$, and similarly for Σ_{i+1}. Moreover, putting $\pi_i := \bigwedge_i \Pi_i$ and $\sigma_i := \bigwedge_i \Sigma_i$, we will have that $\Delta(\pi_i, \sigma_i) \in w$. The key step in the inductive construction is

$$\begin{aligned}
\Delta(\pi_i, \sigma_i) \in w \quad &\Rightarrow \quad \Delta(\pi_i \wedge (\phi_i \vee \neg\phi_i), \sigma_i \wedge (\phi_i \vee \neg\phi_i)) \in w \\
&\Rightarrow \quad \Delta((\pi_i \wedge \phi_i) \vee (\pi_i \wedge \neg\phi_i), (\sigma_i \wedge \phi_i) \vee (\sigma_i \wedge \neg\phi_i)) \in w \\
&\Rightarrow \quad \text{one of the formulas } \Delta(\pi_i \wedge (\neg)\phi_i, \sigma_i \wedge (\neg)\phi_i) \text{ is in } w.
\end{aligned}$$

If, for example, $\Delta(\pi_i \wedge \phi_i, \sigma_i \wedge \neg\phi_i) \in w$, we take $\Pi_{i+1} := \Pi_i \cup \{\phi_i\}$, $\Sigma_{i+1} := \Sigma_i \cup \{\neg\phi_i\}$. Under this definition, all Π_i, Σ_i will have the required properties. Finally, let $u_1 = \bigcup_i \Pi_i$ and $u_2 = \bigcup_i \Sigma_i$. Then u_1, u_2 are Λ-MCSs and $R^\Lambda_\Delta w u_1 u_2$, as required.

Given the Truth lemma for general modal languages thus adapted, completeness follows as in Theorem 4.19.

4.3 Alternatives and limitations

How general is the method of proving (strong) completeness through canonicity? Let a *Sahlqvist logic* be a logic of the form $\mathbf{KS}_1 \ldots \mathbf{S}_n \ldots$, where $S_1, \ldots, S_n, \ldots$ are Sahlqvist formulas. In Section 7 on applications of duality we will show by algebraic means that every Sahlqvist logic is canonical (cf. Theorem 7.8). Combined with the results on correspondence from Section 3, this implies that every Sahlqvist logic is strongly complete with respect to the class of frames defined by the first-order conditions to which the Sahlqvist formulas correspond.

However, unlike the general canonical model theorem (Theorem 4.19), there is no general canonical frame theorem. The logic **KLöb** provides an example. It can be shown that every strongly complete logic is *compact* in the following sense; if K is the class of frames on which the logic is valid, then for any set of formulas $\Sigma \cup \{\phi\}$, if $\Sigma \models_{\mathsf{S}} \phi$, then $\Sigma_0 \models_{\mathsf{S}} \phi$, for some finite $\Sigma_0 \subseteq \Sigma$. But **KLöb** is not compact, as is witnessed by the set

$$\Gamma = \{\Diamond q_1\} \cup \{\Box(q_i \to \Diamond q_{i+1}) \mid i \in \mathbb{N}\}.$$

Now, as **KLöb** is not compact, it is not strongly complete. However, it is weakly complete, but by the above argument we cannot use the canonical model method to prove this. Instead, one can use a *finite* canonical model construction in which the states are maximal consistent subsets of a finite, subformula closed set.

Another alternative method for proving (weak) completeness results is what we call the Segerberg method. This method is often used to prove completeness with respect to frames satisfying certain modally undefinable conditions. It consists in first taking the canonical model, and then applying various constructions to massage it into a model based on a frame that has the required properties. The method has proved particularly useful in temporal logic (cf. Segerberg [23]).

Using the finite canonical model method and the Segerberg method, it is possible to establish many completeness results that escape the ordinary canonical model method. However, these methods have limitations too. Worse still, there are modal logics for which no completeness result with respect to frames can be given at all. Thomason [25] was the first to come up with an example of an incomplete tense logic; then Fine [11] provided an incomplete extension of **S4**. Finally, Blok [5] shows that there is a continuum of distinct incomplete extensions of **KT**.

Part III: An algebraic perspective

In the second Part of our notes we will develop an alternative, *algebraic* semantics for modal logic. The basic idea is to extend the algebraic treatment of classical propositional logic in the framework of *boolean algebras* to the setting of modal logic. The main reason for studying logics from an algebraic perspective is that it allows us to apply some powerful techniques and results from the theory of *universal algebra* to logic. In the case of modal logic, an additional reason is that the algebraic semantics behaves better than the frame semantics in the sense that one can prove a general *completeness theorem* for the algebraic semantics, while we saw in Section 4 that in general, modal logics are *incomplete* with respect to the frame semantics.

To give an overview of Part III: in Section 5 we introduce the algebraic approach towards modal logic. Section 6 describes some basic relations between boolean algebras with operators and relational frames. Finally, in Section 7 we give applications of the duality theory developed in the earlier sections. These applications include relatively easy proofs of the Goldblatt-Thomason Theorem characterizing the first-order definable frame classes that are modally definable, and the result due, in essence, to Fine that the modal theory of a frame class is a canonical logic whenever the class is closed under ultraproducts.

We assume some basic working knowledge on Universal Algebra and Boolean Algebras, not going beyond (proofs of) the following: Birkhoff's Theorem identifying varieties as equational classes, Birkhoff's Completeness Theorem for equational logic and Stone's Representation Theorem. These prerequisites can be found in any text book on Universal Algebra like Burris and Sankappanavar [7].

5. Algebraizing modal logic

As usual in algebraic logic, there are two approaches towards the algebraization of modal logic: a semantic approach based on the connections between frames and boolean algebras with operators, and a more syntactic, axiomatic approach relating logics to equational theories. In the first subsection we will give the semantic perspective, leaving the syntactic viewpoint for the second part of this section. Underlying both approaches is one of the basic principles of algebraic logic, viz. that *formulas* of a logical language are viewed as *terms* of an algebraic language. In our case, this means that we are interested in the following algebraic language:

Definition 5.1 Let τ be a modal similarity type. The *corresponding algebraic similarity type* $\mathcal{F}_\tau$ contains as function symbols: all modal operators, and the boolean symbols $\neg$ (unary), $\vee$ (binary) and $\bot$ (constant). This means that for a set Φ of variables, the set $Ter_\tau(\Phi)$ of *τ-terms over Φ* is defined as follows:

1. the constant $\bot$, the constants from τ and the variables from Φ are the *basic* terms over Φ,
2. if s and t are terms over Φ, then so are $\neg s$, $s \vee t$ (and $s \wedge t$),
3. if Δ is a modal operator of rank n, and $s_1, \ldots, s_n$ are terms over Φ, then $\Delta(s_1, \ldots, s_n)$ is a term over Φ.

A τ-equation over Φ is a pair (s, t) of τ-terms over Φ, usually denoted as $s = t$.

Convention 5.2 Note that formally speaking, the modal and the algebraic similarity type do not coincide, since the latter also contains the boolean function symbols. In practice however, we will usually identify τ and $\mathcal{F}_\tau$.

We have taken the formulas-as-terms paradigm of algebraic logic quite literally: by our definitions, we have

$$Form(\tau, \Phi) = Ter_\tau(\Phi).$$

In many occasions however, we find it convenient to distinguish the modal and the algebraic perspective. Usually, we can make this distinction by using Greek lower case letters (ϕ, ψ, ...) for formulas, and Roman lower case (s, t, ...) for terms. In some cases, especially when we relate the algebraic and the logical viewpoint, we will use superscripts ($(\cdot)^f$ and $(\cdot)^t$ respectively), to denote whether we see a syntactic entity as a modal formula or as an algebraic term. To give two examples: ϕ^t is the formula/term ϕ, *viewed as an algebraic term*, while s^f is the term/formula s, *viewed as a modal formula*. Finally, adhering to usage in the theory of boolean algebras, we will frequently use the following symbols in algebraic terms:

$+$	for	$\vee$
$\cdot$	for	$\wedge$
$-$	for	$\neg$
0	for	$\bot$
1	for	$\top$.

5.1 Algebraizing modal semantics

We first turn to the *algebraic semantics* for the algebraic language just defined. In general, we could interpret τ-terms and τ-equations in *any* algebra of the appropriate similarity type. However, we are only interested in a particular *class* of algebras, namely the so-called *boolean algebras with operators*. Therefore, let us first introduce these algebras, and then discuss how they form a natural semantics for the algebraic language corresponding to modal logic.

Definition 5.3 Let τ be a modal similarity type. A *boolean algebra with τ-operators* is an algebra

$$\mathfrak{A} = (A, +, -, 0, f_\Delta)_{\Delta \in \tau}$$

such that $(A, +, -, 0)$ is a boolean algebra and every f_Δ is an *operator* of arity $\rho(\Delta)$; that is to say, f_Δ is an operation satisfying

Normality. $f_\Delta(a_1, \ldots, a_{\rho(\Delta)}) = 0$ whenever $a_i = 0$ for some i $(0 < i \leq \rho(\Delta))$.

Additivity. for all i (such that $0 < i \leq \rho(\Delta)$):

$$f_\Delta(a_1, \ldots, a_i + a_i', \ldots, a_{\rho(\Delta)}) = \\ f_\Delta(a_1, \ldots, a_i, \ldots, a_{\rho(\Delta)}) + f_\Delta(a_1, \ldots, a_i', \ldots, a_{\rho(\Delta)}).$$

If we abstract away from the particular modal similarity type τ, or if τ is known from the context, then we will speak of *boolean algebras with operators*, or *BAOs*, without reference to τ. Note that in the case of a unary operator f, the conditions of normality and additivity boil down to

$$\begin{aligned} f(0) &= 0 \\ f(x+y) &= fx + fy. \end{aligned}$$

Example 5.4 Consider the collection of binary relations over a given set U. These sets form a set algebra on which we can define the operations $|$ (composition), $(\cdot)^{-1}$ (inverse) and Id (the identity relation) as a binary, unary resp. nullary operation. It is easy to verify that these operations are actually operators — to give an example, we show that composition is additive in its second argument:

$(x, y) \in R \mid (S \cup T)$ iff
iff there is a z with $(x, z) \in R$ and $(z, y) \in S \cup T$
iff there is a z with $(x, z) \in R$ and $(z, y) \in S$ or $(z, y) \in T$
iff there is a z with $(x, z) \in R$ and $(z, y) \in S$,
or there is a z with $(x, z) \in R$ and $(z, y) \in T$
iff $(x, y) \in R \mid S$ or $(x, y) \in R \mid T$
iff $(x, y) \in R \mid S \cup R \mid T$.

The next example of a boolean algebra with operators plays such a central rôle in these notes, that it deserves a definition of its own.

Definition 5.5 Given an $n+1$-ary relation R on a set W, we define the n-ary operation m_R on subsets of W by

$$m_R(X_1, \ldots, X_n) = \{w \in W \mid \text{there are } w_1, \ldots, w_n \text{ such that } R_\Delta w w_1 \ldots w_n \text{ and } w_i \in X_i \text{ for all } i\}$$

Now let τ be a modal similarity type, and $\mathfrak{F} = (W, R_\Delta)_{\Delta \in \tau}$ a τ-frame. The *(full) complex algebra of* $\mathfrak{F}$, notation: $\mathfrak{Cm}\mathfrak{F}$ or $\mathfrak{F}^+$, is the extension of the power set algebra $\mathfrak{P}(W)$ with operations m_{R_Δ} for every operator Δ in τ. A *complex algebra* is a subalgebra of a full complex algebra.

Let K be a class of frames; we denote the class of full complex algebras of frames in K by **CmK**.

Note that for a *binary* relation R, the unary operation m_R yields, given a subset X of the universe, the set of all states which 'see' a state in X:

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

Proposition 5.6 *Let τ be a modal similarity type, and $\mathfrak{F} = (W, R_\Delta)_{\Delta\in\tau}$ a τ-frame. Then $\mathfrak{Cm}\mathfrak{F}$ is a boolean algebra with τ-operators.*

Proof. We have to show that operations of the form m_R are normal and additive. This rather easy proof is left to the reader. $\dashv$

Complex algebras are so to speak the concrete or real boolean algebras with operators, in the same way as set algebras are the concrete boolean algebras. Hence the obvious question is whether (similar to the boolean case) every boolean algebra with operators is isomorphic to a complex algebra. Theorem 6.6 in the next section will give a positive answer to this question; this result is fundamental to the duality theory of frames and boolean algebras with operators.

Let us now discuss the interpretation of τ-terms and τ-equations in (arbitrary) boolean algebras with τ-operators:

Definition 5.7 Let τ be a modal similarity type and Φ a set of variables. Let $\mathfrak{A} = (A, +, -, f_\Delta)_{\Delta\in\tau}$ be a boolean algebra with τ-operators. An *assignment* for Φ is a function $\theta : \Phi \to A$. We can extend θ uniquely to a map $\overline{\theta}$: $Ter_\tau(\Phi) \to A$ as follows:

$$\begin{aligned}
\overline{\theta}(p) &= \theta(p) \\
\overline{\theta}(\bot) &= 0 \\
\overline{\theta}(\neg s) &= -\overline{\theta}(s) \\
\overline{\theta}(s \vee t) &= \overline{\theta}(s) + \overline{\theta}(t) \\
\overline{\theta}(\Delta(s_1, \dots, s_n)) &= f_\Delta(\overline{\theta}(s_1), \dots, \overline{\theta}(s_n)).
\end{aligned}$$

Now let $s = t$ be a τ-equation. We say that $s = t$ is *true* in $\mathfrak{A}$, notation: $\mathfrak{A} \models s = t$, if for every assignment θ:

$$\overline{\theta}(s) = \overline{\theta}(t).$$

Occasionally, we will also use the concept of truth of a *modal formula* in a BAO. We say that a modal τ-formula ϕ is *true* in a boolean algebra with τ-operators $\mathfrak{A}$, notation: $\mathfrak{A} \models \phi$, if for all assignments θ,

$$\overline{\theta}(\phi^t) = 1,$$

(or, equivalently, if the equation $\phi^t = 1$ holds in $\mathfrak{A}$).

If $\mathfrak{A}$ is a *complex* algebra $\mathfrak{F}^+$, then we have (for the standard modal similarity type):

$$\overline{\theta}(\Diamond\phi) = m_{R_\Diamond}(\overline{\theta}(\phi)). \tag{5.16}$$

This is the key observation in the proof of the following Proposition which shows that the algebraic semantics can be seen as a generalization of the frame semantics:

Proposition 5.8 *Let τ be a modal similarity type, $\mathfrak{F}$ a τ-frame, ϕ a τ-formula and s and t τ-terms. Then*

$$\begin{aligned} \mathfrak{F} \models \phi \quad &\textit{iff} \quad \mathfrak{F}^+ \models \phi^t = 1 \\ \mathfrak{F}^+ \models s = t \quad &\textit{iff} \quad \mathfrak{F} \models s^f \leftrightarrow t^f. \end{aligned}$$

Proof. We will only prove the first part of the Proposition. Let $\mathfrak{F}$ be a frame on which the formula ϕ is valid. To show that $\mathfrak{F}^+ \models \phi^t = 1$, consider an arbitrary assignment θ of variables from Φ to elements of the algebra $\mathfrak{F}^+$. Since elements of $\mathfrak{F}^+$ are *subsets* of the power set POW(W) of the universe W of $\mathfrak{F}$, this means that θ is in fact nothing but a *valuation*. By induction on the complexity of a formula ψ, we will show that for all $w \in W$:

$$(\mathfrak{F}, \theta), w \Vdash \psi \text{ iff } w \in \overline{\theta}\psi. \tag{5.17}$$

The only interesting case in the proof of (5.17) is the modal case of the inductive step. For simplicity, we assume that ψ is of the form $\Diamond\chi$. We have the following equivalences:

$$\begin{aligned} (\mathfrak{F}, \theta), w \Vdash \Diamond\chi \quad &\text{iff} \quad \text{there is a } v \text{ such that } R_\Diamond wv \text{ and } (\mathfrak{F}, \theta), w \Vdash \chi \quad (IH) \\ &\text{iff} \quad \text{there is a } v \text{ such that } R_\Diamond wv \text{ and } v \in \overline{\theta}(\chi) \\ &\text{iff} \quad w \in m_R(\overline{\theta}(\chi)) \\ &\text{iff} \quad w \in \overline{\theta}(\Diamond\chi) \qquad (5.16) \end{aligned}$$

This proves (5.17).

Now since $\overline{\theta}(1) = W$, it follows from (5.17) that

$$\overline{\theta}(\phi) = \overline{\theta}(1) \text{ iff } \mathfrak{F}, \overline{\theta} \Vdash \phi. \tag{5.18}$$

But the right-hand side of (5.18) follows from the assumption that $\mathfrak{F} \models \phi$. Hence we have that $\overline{\theta}(\phi) = \overline{\theta}(1)$, implying that $\mathfrak{F}^+ \models \phi^t = 1$. $\dashv$

Proposition 5.8 justifies the identification of the *modal theory* of a class K of frames with the *equational theory* of the class **Cm**K of complex algebras of frames in K.

Definition 5.9 Let τ be a modal similarity type. For a set Σ of τ-formulas, we define Σ^{equ} as the set of corresponding equations:

$$\Sigma^{equ} = \{\phi^t = 1 \mid \phi \in \Sigma\}.$$

Conversely, for a set E of equations, we define the set E^{for} of corresponding modal formulas as

$$E^{for} = \{s^f \leftrightarrow t^f \mid s = t \in E\}.$$

It follows immediately from Proposition 5.8 that for any frame $\mathfrak{F}$:

$$\mathfrak{F} \models \Sigma \quad \text{iff} \quad \mathfrak{F}^+ \models \Sigma^{equ} \tag{5.19}$$

$$\mathfrak{F}^+ \models E \quad \text{iff} \quad \mathfrak{F} \models E^{for}. \tag{5.20}$$

This explains the following definition:

Definition 5.10 Let τ be a modal similarity type. For a set Σ of τ-formulas, we define V_Σ as the variety of boolean algebras with τ-algebras where the set of equations Σ^{equ} is valid.

5.2 Algebraizing modal axiomatics
We now turn to the second, axiomatic approach to the connection between modal logic and boolean algebras with operators. This approach is based on the observation that we can build an algebra on top of the set of formulas in such a way that for any normal modal logic, the relation between two formulas of *provable equivalence* is a congruence relation.

Definition 5.11 Let τ be an algebraic similarity type, and Φ a set of propositional variables. The formula algebra of τ over Φ is the algebra $\mathfrak{Form}(\tau,\Phi) = (Form(\tau,\Phi), I(\Delta))_{\Delta\in\tau}$ where each operator Δ is interpreted as the operation $I(\Delta)$ given by

$$I(\Delta)(t_1,\ldots,t_n) = \Delta(t_1,\ldots,t_n).$$

From the algebraic perspective, the formula algebra $\mathfrak{Form}(\tau,\Phi)$ is nothing but the *absolutely free algebra* or term algebra generated by the set Φ. This perspective on $\mathfrak{Form}(\tau,\Phi)$ as constituting an $\mathcal{F}_\tau$*-algebra* is very useful. For instance, the reader is advised to check that the extension $\overline{\theta}$ of an assignment θ (mapping variables of Φ to elements of an algebra $\mathfrak{A}$), is in fact the unique *homomorphism*: $\mathfrak{Form}(\tau,\Phi) \to \mathfrak{A}$ extending θ. Thus the modal valuations on a frame $\mathfrak{F}$ are the homomorphisms from $\mathfrak{Form}(\tau,\Phi)$ to the complex algebra $\mathfrak{Cm}\mathfrak{F}$.

Definition 5.12 Let τ be a modal similarity type, Φ a set of propositional variables and Λ a normal modal τ-logic. We define $\equiv_\Lambda$ as a binary relation between τ-formulas (in Φ) by

$$\phi \equiv_\Lambda \psi \text{ iff } \vdash_\Lambda \phi \leftrightarrow \psi.$$

The following proposition is fundamental to the syntactic approach to algebraizing modal logic.

Proposition 5.13 *Let τ be a modal similarity type, Φ a set of propositional variables and Λ a normal modal τ-logic. Then $\equiv_\Lambda$ is a congruence relation on $\mathfrak{Form}(\tau,\Phi)$.*

Proof. We confine ourselves to proving the Proposition for the standard modal similarity type only. It is not very difficult to show that $\equiv_\Lambda$ is an equivalence relation, so we leave this part of the proof to the reader. To show that $\equiv_\Lambda$ is a *congruence* relation on the formula algebra, we have to make clear that $\equiv_\Lambda$ has the following properties:

$$\begin{array}{rll}\phi_0 \equiv_\Lambda \psi_0 \text{ and } \phi_1 \equiv_\Lambda \psi_1 & \text{imply} & \phi_0 \vee \phi_1 \equiv_\Lambda \phi_0 \vee \psi_1 \\ \phi \equiv_\Lambda \psi & \text{implies} & \neg\phi \equiv_\Lambda \neg\psi \\ \phi \equiv_\Lambda \psi & \text{implies} & \Diamond\phi \equiv_\Lambda \Diamond\psi.\end{array} \tag{5.21}$$

Let us prove the last statement as an example: assume that $\phi \equiv_\Lambda \psi$, i.e., that $\vdash_\Lambda \phi \to \psi$. By some propositional reasoning it follows that $\vdash_\Lambda \neg\phi \to \neg\psi$. By an application of the rule of Universal Generalization we find $\vdash_\Lambda \Box(\neg\phi \to \neg\psi)$, and hence, using the K-axiom and Modus Ponens, $\vdash_\Lambda \Box\neg\phi \to \Box\neg\psi$. After some further propositional manipulations, we obtain $\vdash_\Lambda \neg\Box\neg\psi \to \neg\Box\neg\phi$, which gives $\vdash_\Lambda \Diamond\psi \to \Diamond\phi$. Likewise, we prove that $\vdash_\Lambda \Diamond\phi \to \Diamond\psi$. Finally, it follows from $\vdash_\Lambda \Diamond\phi \leftrightarrow \Diamond\psi$ that $\Diamond\phi \equiv_\Lambda \Diamond\psi$. $\dashv$

By (5.21) (or a generalization of it, in case of a polyadic modal operator), the following are correct definitions of functions on the set $Form(\tau_0,\Phi)/{\equiv_\Lambda}$ of equivalence classes under $\equiv_\Lambda$:

$$\begin{array}{rcl} [\phi]+[\psi] & := & [\phi\vee\psi] \\ -[\phi] & := & [\neg\phi] \\ f_\Delta[\phi_1,\ldots,\phi_n] & := & [\Delta(\phi_1,\ldots,\phi_n)] \end{array} \tag{5.22}$$

The last clause of this definition boils down to the following for unary diamonds:

$$f_\Diamond[\phi] := [\Diamond\phi].$$

Given a normal modal logic Λ, we now define the *Lindenbaum-Tarski algebra of* Λ to be the *quotient algebra* of the formula algebra over the congruence relation $\equiv_\Lambda$:

Definition 5.14 Let τ be a modal similarity type, Φ a set of propositional variables and Λ a normal, modal τ-logic in this language. The Lindenbaum-Tarski algebra of Λ over the set of generators Φ is the structure

$$\mathfrak{L}_\Lambda(\Phi) := (Form(\tau,\Phi)/\equiv, +, -, f_\Diamond),$$

where the operations are defined as in (5.22).

These Lindenbaum-Tarski algebras are the basic tools in the syntactic approach towards algebraizing modal logic. They also form our second main example of boolean algebras with operators. Let us prove additivity of $f_\Diamond$. We have to show that

$$f_\Diamond(a+b) = f_\Diamond a + f_\Diamond b, \tag{5.23}$$

for arbitrary elements a and b of $\mathfrak{L}_\Lambda(\Phi)$. Let a and b be such elements; by definition there are formulas ϕ and ψ such that $a=[\phi]$ and $b=[\psi]$. Then

$$f_\Diamond(a+b) = f_\Diamond([\phi]+[\psi]) = f_\Diamond([\phi\vee\psi]) = [\Diamond(\phi\vee\psi)]$$

while

$$f_\Diamond a + f_\Diamond b = f_\Diamond([\phi]) + f_\Diamond([\psi]) = [\Diamond\phi]+[\Diamond\psi] = [\Diamond\phi\vee\Diamond\psi]$$

Finally, an easy Λ-deduction shows that

$$\vdash_\Lambda \Diamond(\phi\vee\psi) \leftrightarrow (\Diamond\phi\vee\Diamond\psi).$$

Therefore, $[\Diamond(\phi\vee\psi)] = [\Diamond\phi\vee\Diamond\psi]$.

From the algebraic perspective, the Lindenbaum-Tarski algebra of a normal modal logic Λ is nothing but the *term model* of the equational theory Λ^{equ}. We can then apply Birkhoff's completeness theorem for equational logic to obtain a general completeness result for modal logic with respect to varieties of boolean algebras with operators. To make these claims more precise, we need to compare modal and equational derivation systems. We assume familiarity with some standard derivation system of equational logic, for instance with rules of reflexivity, symmetry, transitivity, congruence and replacement.

Definition 5.15 Let Σ be a set of modal formulas in some modal similarity type τ. We say that a τ-equation $s=t$ is (equationally) derivable from Σ^{bao}, notation: $\Sigma^{bao}\vdash s=t$, if the equation $s=t$ is derivable, using some standard derivation system of equational logic, from the following set of equations:

$$\Sigma^{equ} \cup BA \cup NORM \cup ADD.$$

Here BA is some standard set of equations axiomatizing boolean algebras, and $NORM$ and ADD are sets of equations forcing every operator from τ to be normal and additive.

Proposition 5.16 *Let Σ be a set of modal formulas in some modal similarity type τ. Then for all τ-formulas ϕ:*

$$\vdash_{\mathbf{K}\Sigma} \phi \textit{ iff } \Sigma^{bao} \vdash \phi^t = 1$$

Proof. Both directions of this proof go by an induction on the length of derivations. We only prove the direction from right to left, for which we load the induction hypothesis to:

$$\Sigma^{bao} \vdash s = t \implies \vdash_{\mathbf{K}\Sigma} s^f \leftrightarrow t^f. \tag{5.24}$$

Of course, for a proper proof of (5.24) we would have to give details of our 'standard derivation system of equational logic'. We confine ourselves to a few examples of standard equational derivation rules:

Transitivity. Assume that we obtained $\Sigma^{bao} \vdash s = u$ from $\Sigma^{bao} \vdash s = t$ and $\Sigma^{bao} \vdash t = u$. By the inductive hypothesis, $\vdash_{\mathbf{K}\Sigma} s^f \leftrightarrow t^f$ and $\vdash_{\mathbf{K}\Sigma} t^f \leftrightarrow u^f$. It follows by an easy propositional derivation that this implies $\vdash_{\mathbf{K}\Sigma} s^f \leftrightarrow u^f$.

Replacement. Let us give one typical example here: assume that we obtained $\Sigma^{bao} \vdash \Diamond s = \Diamond t$ from $\Sigma^{bao} \vdash s = t$. The inductive hypothesis is that

$$\vdash_{\mathbf{K}\Sigma} s^f \leftrightarrow t^f, \tag{5.25}$$

from which we get $\vdash_{\mathbf{K}\Sigma} \neg t^f \to \neg s^f$ by propositional reasoning. Now the rule of universal generalization yields $\vdash_{\mathbf{K}\Sigma} \Box(\neg t^f \to \neg s^f)$. From this, we derive $\vdash_{\mathbf{K}\Sigma} \Box\neg t^f \to \Box\neg s^f$ by the K-axiom and modus ponens. Then by propositional reasoning we get $\vdash_{\mathbf{K}\Sigma} \neg\Box\neg s^f \to \neg\Box\neg t^f$. Likewise, we find $\vdash_{\mathbf{K}\Sigma} \neg\Box\neg t^f \to \neg\Box\neg s^f$, so by a propositional derivation we find the desired[2]

$$\vdash_{\mathbf{K}\Sigma} \neg\Box\neg t^f \leftrightarrow \neg\Box\neg s^f. \quad \dashv \tag{5.26}$$

Theorem 5.17 *Let τ be a modal similarity type, and Σ a set of τ-formulas. Then $\mathfrak{L}_{\mathbf{K}\Sigma}(\Phi)$ is a free algebra for the variety V_Σ over the set Φ of generators.*

Proof. We leave it to the reader to verify that $\mathfrak{L}_{\mathbf{K}\Sigma}(\Phi)$ is in V_Σ, i.e., that it is a boolean algebra with operators in which the set Σ^{equ} of equations corresponding to the modal formulas in Σ is valid. To show that $\mathfrak{L}_{\mathbf{K}\Sigma}(\Phi)$ is *free* in V_Σ, we use the well-known fact (which can be found in any standard proof of Birkhoff's completeness result for equational logic) that the algebra $\mathfrak{Form}(\tau, \Phi)/\!\sim_\Sigma$ is free for V_Σ, where $\sim_\Sigma$ is the relation given by

[2]There is a technical problem involved here: since we are dealing with *syntax*: the formula in (5.26) is not the same as $\Diamond t^f \leftrightarrow \Diamond s^f$. In fact, the latter formula can *not* be derived in $\mathbf{K}\Sigma$, since in our set-up, *boxes*, not diamonds are the basic operators in the language of modal derivation systems. The symmetry can easily be restored by reconsidering the choice of basic operators on either the modal or the algebraic side, and modifying the derivation systems accordingly (cf. also Remark 4.5).

$$s \sim_\Sigma t \text{ iff } \Sigma^{bao} \vdash s = t.$$

Now it is immediate by Proposition 5.16 that for all terms s, t:

$$s \sim_\Sigma t \text{ iff } s^f \equiv_{\mathbf{K}\Sigma} t^f.$$

But this means that the Lindenbaum-Tarski algebra for $\mathbf{K}\Sigma$ is *identical* to $\mathfrak{Form}(\tau, \Phi)/{\sim_\Sigma}$ and thus also free for V_Σ. ⊣

The following Theorem states that modal logics are always complete with respect to the variety of boolean algebras with operators where their axioms are valid. Note that this is in sharp contrast to the relational semantics, where in general, modal logics are not complete with respect to the class of frames that they define.

Theorem 5.18 *Let τ be a modal similarity type, and Σ a set of modal τ-formulas. Then $\mathbf{K}\Sigma$ is sound and complete with respect to V_Σ, i.e., for all formulas ϕ we have*

$$\vdash_{\mathbf{K}\Sigma} \phi \text{ iff } \mathsf{V}_\Sigma \models \phi.$$

Proof. This Theorem is an immediate corollary of the previous one. For the completeness direction, assume that $\mathsf{V}_\Sigma \models \phi$, i.e., $\mathsf{V}_\Sigma \models \phi^t = 1$. Let Φ be a set of propositional variables containing all variables occurring in ϕ. Since $\mathfrak{L}_{\mathbf{K}\Sigma}(\Phi)$ is in V_Σ by Theorem 5.17, we find that $\mathfrak{L}_{\mathbf{K}\Sigma}(\Phi) \models \phi^t = 1$. Now consider the assignment $\iota : Form(\tau, \Phi) \to Form(\tau, \Phi)/{\equiv_{\mathbf{K}\Sigma}}$ given by

$$\iota(p) = [p].$$

It is easy to show that for all formulas ψ:

$$\iota(\psi^t) = [\psi].$$

Thus we find that $\bar{\iota}(\phi^t) = [\phi]$ and $\bar{\iota}(1) = [\top]$. From $\mathfrak{L}_{\mathbf{K}\Sigma}(\Phi) \models \phi^t = 1$ we infer that $\bar{\iota}(\phi^t) = \bar{\iota}(1)$, so we may conclude that

$$[\phi] = [\top].$$

Hence ϕ and $\top$ are provably equivalent in $\mathbf{K}\Sigma$, so $\vdash_{\mathbf{K}\Sigma} \phi$. ⊣

Let us finish this section with discussing the fundamental rôle that the Lindenbaum-Tarski algebra plays in the frame completeness theory of modal logic[3]. The basic idea is that we can prove completeness for many classes of logics by showing that the Lindenbaum-Tarski algebra of the logic can be represented as (i.e., is isomorphic to) an appropriate complex algebra. The slogan here is 'completeness via representation'. To be more precise, let us look at the basic modal logic $\mathbf{K}$ of the standard modal similarity type, and show that:

[3]For lack of space we will only consider *weak* completeness here. A proper treatment of the concept of strong completeness would involve an interesting but lengthy discussion on the various ways of defining and algebraizing the modal consequence relation $\Sigma \models \phi$ and the derivability relation $\Sigma \vdash \phi$. For details we refer the reader to [4].

if every Lindenbaum-Tarski algebra of **K** is representable as a complex algebra, (5.27)
then **K** is complete with respect to the class of all frames.

So, assume that $\mathfrak{L}_K(\Phi)$ is representable as a complex algebra. In order to prove completeness of **K** with respect to the class of all frames, assume that the formula ϕ is valid in every frame. By Proposition 5.16 this implies that $\mathfrak{F}^+ \models \phi^t = 1$, for every frame $\mathfrak{F}$. But then the equation $\phi^t = 1$ is valid in the Lindenbaum-Tarski algebra as well, since by assumption, this algebra can be embedded in some full complex algebra. Now (the proof of) Theorem 5.18 shows that this implies that $\vdash_{\mathbf{K}} \phi$. This proves (5.27).

So, we are left with the question whether Lindenbaum-Tarski algebras can be represented as complex algebras. Recall that in the previous subsection, we already hinted that *every* boolean algebra with operators is isomorphic to a complex algebra. This fundamental result in the algebraization of modal logic, which is due to Jónsson and Tarski [17], will be discussed and proved in the next section.

6. Basic duality

In this section, which really forms the heart of these lecture notes, we will show how to link up boolean algebras with operators with frames. We have already seen how to obtain a BAO from a frame (the complex algebra of the frame); now we will explain how, conversely, one can associate a frame with every boolean algebra with operators, namely, the ultrafilter frame of the algebra. We will prove the fundamental result that every boolean algebra with operators can be embedded in its canonical embedding algebra, which is nothing but the complex algebra of the ultrafilter frame of the original algebra. We will prove the fundamental result due to Jónsson and Tarski, that every BAO can be embedded in its canonical embedding algebra. We will use the above-mentioned construction to define yet another frame construction, viz. that of the ultrafilter extension of a frame. Finally, we will show how to extend this duality between algebras and frames from the level of structures to the level of *morphisms* between frames on the one hand and algebras on the other.

6.1 Ultrafilter frames and canonical embedding algebras

Suppose that we want to embed a BAO $\mathfrak{A}$ in a complex algebra. Obviously, the first question to ask is what the underlying *frame* of the complex algebra will be. In order to make our notation a bit simpler, let us assume for the moment that we are working in a similarity type with just one unary modality, and that $\mathfrak{A} = (A, +, -, 0, f)$ is a boolean algebra with one unary operator f. We have to find a universe W and a binary relation Q on W such that $\mathfrak{A}$ can be embedded in the complex algebra of the frame (W, Q). The easy half of the theorem is already given by Stone's Representation Theorem, namely: we know how to embed the boolean part of $\mathfrak{A}$ in the power set algebra of the set of *ultrafilters* of $\mathfrak{A}$.

Definition 6.1 Let $\mathfrak{A} = (A, +, -, 0)$ be a boolean algebra. A subset D of A has the *finite intersection property* or *f.i.p.* if for every finite subset $D_0 = \{a_1, \ldots, a_n\}$ of D we have $a_1 \cdot \ldots \cdot a_n \neq 0$.

A subset $F \subseteq A$ is a *filter* if it satisfies

(F0) $1 \in F$
(F1) F is closed under intersection, i.e., if $a, b \in F$ then $a \cdot b \in F$
(F2) F is upward closed, i.e., if $a \in F$ and $a \leq b$ then $b \in F$.

A filter is *proper* if it does not contain the smallest element 0 of $\mathfrak{A}$.

A filter F is an *ultrafilter* if it is proper, and satisfies

(F4) for all $a \in A$, either $a \in F$ or $-a \in F$.

We denote by $Uf\mathfrak{A}$ the set of all ultrafilters over $\mathfrak{A}$.

In many occasions, one is searching for an ultrafilter satisfying some specific properties. To prove the existence of such an ultrafilter, the following is often a good strategy: first, define a set D such that any ultrafilter which is a superset of D has the required properties, and second, show that D has the finite intersection property. This strategy is motivated by the facts that every set enjoying the f.i.p. is contained in a proper filter, and that every proper filter is contained in an ultrafilter. For future reference, we state the following Ultrafilter Theorem:

Proposition 6.2 (Ultrafilter Theorem) *Let $\mathfrak{A}$ be a boolean algebra, and F be a proper filter of $\mathfrak{A}$. Then there is an ultrafilter u such that $F \subseteq u$. As a corollary, every subset of A having the finite intersection property can be extended to an ultrafilter.*

In order to embed our BAO $\mathfrak{A} = (A, +, -, 0, f)$ in the complex algebra of a frame $\mathfrak{F} = (W, Q)$, it seems a natural choice to take W as the set of *ultrafilters* of (the boolean part of) $\mathfrak{A}$. Recall that in Stone's Representation Theorem, the embedding function $r : A \to \text{POW}(Uf\mathfrak{A})$ is defined by

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

For the definition of Q, we view the elements of the algebra as *propositions*, and imagine that the representation map $r(a)$ will be the set of states where a is *true* according to some valuation. Then if we read $f(a)$ as $\Diamond a$, we find that a state u should be in $r(fa)$ if and only if there is a v with Quv and $u \in r(a)$. So, in order to decide whether Quv should hold for two arbitrary states (ultrafilters) u and v, we should look at all the propositions a holding at v (i.e., all elements $a \in v$) and check whether fa holds at u (i.e., whether $fa \in u$). Putting it more formally, the natural, 'canonical' choice for Q seems to be the relation Q_f given by

$$Q_f uv \text{ iff } fa \in u \text{ for all } a \in v. \tag{6.28}$$

In the general case, we obtain the following definition:

Definition 6.3 Let $\mathfrak{A} = (A, +, -, 0, f_\Delta)_{\Delta \in \tau}$ be a boolean algebra with operators. The $n+1$-ary relation Q_f on the set $Uf\mathfrak{A}$ of ultrafilters of $\mathfrak{A}$ is given by

$$Q_f u u_1 \dots u_n \text{ iff } f(a_1, \dots, a_n) \in u \text{ for all } a_1 \in u_1, \dots, a_n \in u_n \tag{6.29}$$

The frame $(Uf\mathfrak{A}, Q_{f_\Delta})_{\Delta \in \tau}$ is called the *ultrafilter frame* of $\mathfrak{A}$, notation[4]: $\mathfrak{Cs}\mathfrak{A}$ or $\mathfrak{A}_+$. The complex algebra $\mathfrak{Cm}\mathfrak{Cs}\mathfrak{A} = (\mathfrak{A}_+)^+$ is called the *(canonical) embedding algebra* of $\mathfrak{A}$, notation: $\mathfrak{Em}\mathfrak{A}$.

[4]The notation $\mathfrak{Cs}\mathfrak{A}$ stems from an alternative name of this frame: *canonical extension*.

For later reference, we state the following proposition, which shows that we could have given an alternative but equivalent definition of the relation Q_f.

Proposition 6.4 *Let f be an n-ary operator on the boolean algebra $\mathfrak{A}$, and u, u_1, ..., u_n an n-tuple of ultrafilters of $\mathfrak{A}$. Then*

$$Q_f u u_1 \dots u_n \text{ iff } -f(-a_1, \dots, -a_n) \in u \text{ implies that for some } i\text{: } a_i \in u_i. \tag{6.30}$$

Proof. By some elementary manipulations on ultrafilters. ⊣

In fact, we have already encountered a frame which is very much like an ultrafilter frame, namely the *canonical* frame of a normal modal logic (cf. Definition 4.17). For, the worlds of the canonical frame are the maximal consistent theories of the logic, and an ultrafilter is nothing but an abstraction of a maximal consistent set. Being a bit more formal, we can show that the canonical frame of a logic is *isomorphic* to the ultrafilter frame of its Lindenbaum-Tarski algebra:

Theorem 6.5 *Let τ be a modal similarity type, and Λ a normal modal τ-logic. Then*

$$\mathfrak{F}^{\Lambda} \cong \mathfrak{L}_{\Lambda}(\Phi).$$

Proof. We leave it to the reader to show that the function θ defined by

$$\theta(\Gamma) = \{[\phi] \mid \phi \in \Gamma\}$$

mapping maximal Λ-consistent sets Γ to the set of equivalence classes of their members, is the required isomorphism between $\mathfrak{F}^{\Lambda}$ and $\mathfrak{L}_{\Lambda}(\Phi)$. ⊣

A second and very important example of an ultrafilter frame arises in the case where the boolean algebra with operators is itself the full complex algebra of some frame. This case is treated in the second part of this section.

We now return to perhaps the most fundamental theorem underlying the algebraization of modal logic, namely the Jónsson-Tarski Theorem which states that every boolean algebra with operators is embeddable in the full complex algebra of its ultrafilter frame:

Theorem 6.6 *Let τ be a modal similarity type, and $\mathfrak{A} = (A, +, -, 0, f_\Delta)_{\Delta \in \tau}$ a boolean algebra with τ-operators. Then the representation function $r : A \to \mathrm{POW}(Uf\mathfrak{A})$ given by*

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

is an embedding of $\mathfrak{A}$ into $\mathfrak{Em}\mathfrak{A}$.

As we already mentioned in the previous section, Theorem 6.6 plays a crucial rôle in the completeness theory of modal logic. However, there are some limitations to its usefulness here — let us illustrate this point by the example of the normal modal logic **K4**, which is axiomatized by the axiom

$$\Box p \to \Box\Box p \tag{4}$$

We know from Theorem 4.21 that **K4** is complete with respect to the class of transitive frames. Now suppose that we want to use the same strategy to prove this result as we did in the end of the previous section with respect to the completeness result of **K** for the class of *all* frames. We would then have to prove that the Lindenbaum-Tarski algebras of **K4** are embeddable in full complex algebras of *transitive* frames. Recall from Section 3 that the formula (4) *characterizes* the transitive frames, so that in our proposed completeness proof, we would have to show that (4) is valid in the ultrafilter frame $(\mathfrak{L}_{\mathbf{K4}}(\Phi))_+$ of $\mathfrak{L}_{\mathbf{K4}}(\Phi)$, or equivalently, that $((\mathfrak{L}_{\mathbf{K4}}(\Phi))_+)^+$ belongs to the variety V_4. Note that it follows from Theorem 5.17 that $\mathfrak{L}_{\mathbf{K4}}(\Phi)$ belongs to V_4.

The discussion above explains the relevance of the question which varieties of BAOs are closed under taking canonical embedding algebras.

Definition 6.7 Let τ be a modal similarity type, and C a class of boolean algebras with τ-operators. C is *canonical* if it closed under taking canonical embedding algebras, i.e., if for all algebras $\mathfrak{A}$, $\mathfrak{Em}\mathfrak{A}$ is in C if $\mathfrak{A}$ is in C.

In the original paper [17], Jónsson and Tarski proved (by algebraic and topological means) that classes of BAOs axiomatized by special kinds of (quasi-)equations are canonical. This important result can be derived as a corollary to our Theorem 7.4 below.

Before we set out to prove Theorem 6.6, we want to compare the two notions of canonicity we have now, viz. the logical one of Definition 4.20 and the algebraic one defined above. Using Theorem 5.17, we show that these two concepts are closely related.

Proposition 6.8 *Let τ be a modal similarity type, and Σ a set of τ-formulas. If V_Σ is a canonical variety, then Σ is canonical.*

Proof. Let V be the variety V_Σ, and assume that V is canonical. By Theorem 5.17, the Lindenbaum-Tarski algebra $\mathfrak{L}_{\mathbf{K\Sigma}}$ is in V; then by assumption, $\mathfrak{Em}\mathfrak{L}_{\mathbf{K\Sigma}}$ is in V. However, from Theorem 6.5 it follows that this algebra is isomorphic to the complex algebra of the canonical frame of $\mathbf{K\Sigma}$:

$$\mathfrak{Em}\mathfrak{L}_{\mathbf{K\Sigma}} = ((\mathfrak{L}_{\mathbf{K\Sigma}})_+)^+ \cong (\mathfrak{F}^{\mathbf{K\Sigma}})^+.$$

Now the fact that $(\mathfrak{F}^{\mathbf{K\Sigma}})^+$ is in V means that $(\mathfrak{F}^{\mathbf{K\Sigma}})^+ \models \Sigma^{equ}$; it follows from Proposition 5.8 that $\mathfrak{F}^{\mathbf{K\Sigma}} \models \Sigma$. But this means that Σ is canonical. $\dashv$

It is an obvious question whether the converse of Proposition 6.8 holds as well, i.e., whether a variety V_Σ is canonical if Σ is a canonical set of modal formulas. Note that canonicity of Σ only implies that one *particular* boolean algebra with operators has its embedding algebra in V_Σ, viz. the Lindenbaum-Tarski algebra over a *countably* infinite number of generators. In fact, we are facing an open problem here:

Open Problem 1 *Let τ be a modal similarity type, and Σ a canonical set of τ-formulas. Is V_Σ a canonical variety?*

Equivalently, suppose that E is a set of equations such that for all countable *boolean algebras with τ-operators we have*

$$\mathfrak{A} \models E \implies \mathfrak{Em}\mathfrak{A} \models E.$$

Is V_E a canonical variety?

Finally then, we give the proof of the Jónsson-Tarski Theorem:

Proof of Theorem 6.6. In order to make our notation a bit simpler, let us assume that we are working in a similarity type with just one n-ary modality, and that $\mathfrak{A} = (A, +, -, 0, f)$ is a boolean algebra with one n-ary operator f. By Stone's representation Theorem, the map $r : A \mapsto \text{POW}(Uf(\mathfrak{A}))$ given by

$$r(x) = \{u \in Uf(A) \mid x \in u\}$$

is a boolean embedding. So, it suffices to show that r is also a *modal* homomorphism, i.e., that

$$r(f(a_1, \ldots, a_n)) = m_{Q_f}(r(a_1), \ldots, r(a_n)). \tag{6.31}$$

We will first prove (6.31) for unary f. In other words, we have to prove that

$$r(fa) = m_{Q_f}(r(a)). \tag{6.32}$$

We start with the inclusion from right to left: assume $u \in m_{Q_f}(r(a))$. Then by definition of m, there is an ultrafilter u_1 with $u_1 \in r(a)$ (i.e. $a \in u_1$) and $Q_f uu_1$. By definition of Q_f this implies $f(a) \in u$, or $u \in r(f(a))$.

For the other inclusion, let u be an ultrafilter in $r(f(a))$, i.e. $f(a) \in u$. To prove that $u \in m_{Q_f}(r(a))$, it suffices to find an ultrafilter u_1 such that $Q_f uu_1$ and $u_1 \in r(a)$, or $a \in u_1$. The basic idea of the proof is that we first pick out the elements of A, apart from a, that *necessarily* have to be in u_1. These elements are given by the condition $Q_f uu_1$. By Proposition 6.4 we have that for every $-f(-y) \in u$, y has to be in u_1; therefore, we define

$$F := \{y \in A \mid -f(-y) \in u\}.$$

We will now show that there is an ultrafilter $u_1 \supseteq F$ containing a. First, an easy proof (using the additivity of f), shows that F is closed under intersection. Second, we prove that

$$F' := \{a \cdot y \mid y \in F\}$$

has the finite intersection property. As F is closed under intersection, it is sufficient to show that $a \cdot y \neq 0$ for $y \in F$. To arrive at a contradiction, suppose that $a \cdot y = 0$. Then $a \leq -y$, so by monotonicity[5] of f, $f(a) \leq f(-y)$; therefore, $f(-y) \in u$, contradicting $y \in F$.

Now by the Ultrafilter Theorem 6.2 there is an ultrafilter $u_1 \supseteq F'$. Note that $a \in u_1$, as $1 \in F$. Finally, $Q_f uu_1$ holds by definition of F: if $-f(-y) \in u$ then $y \in F \subseteq u_1$.

(∗) We will now prove (6.31) for arbitrary $n \geq 1$.

The proof is by induction on the arity n of f.

Note that the base step has already been proved above. So, assume that the inductive hypothesis holds for n. Let f be a normal and additive function of rank $n+1$, and suppose that $a_1, \ldots, a_{n+1}$ are elements of $\mathfrak{A}$ such that $f(a_1, \ldots, a_{n+1} \in u$. We have to find ultrafilters $u_1, \ldots, u_{n+1}$ of $\mathfrak{A}$ such that (i) $a_i \in u_i$ for all i with $1 \leq i \leq n+1$, and (ii) $Q_f uu_1 \ldots u_{n+1}$. Our strategy will be to have the induction hypothesis take care of $u_1, \ldots, u_n$ and then to search for u_{n+1}.

Let f' be the function $A^n \to A$ given by

[5] An operation g on a boolean algebra is *monotonic* if $a \leq b$ implies $fa \leq fb$. Operators are monotonic, because of the following: if $a \leq b$, then $a + b = b$, so $fa + fb = f(a+b) = fb$; but then $fa \leq fb$.

$$f'(x_1, \ldots, x_n) = f(x_1, \ldots, x_n, a_{n+1}).$$

It is easy to see that f' is normal and additive, so we may apply the induction hypothesis. This yields ultrafilters $u_1, \ldots, u_n$ such that

$$a_i \in u_i \text{ for all } i \text{ with } 1 \leq i \leq n. \tag{6.33}$$

and

$$f(x_1, \ldots, x_n, a_{n+1}) \in u, \text{ whenever } x_i \in u_i \text{ for every } 1 \leq i \leq n. \tag{6.34}$$

Now we set out to define an ultrafilter u_{n+1} such that $a_{n+1} \in u_{n+1}$ and $Q_f uu_1 \ldots u_{n+1}$. This second condition can be rewritten as follows (we abbreviate '$x_1 \in u_1, \ldots x_n \in u_n$' by '$\bar{x} \in \bar{u}$'):

$$\begin{array}{ll} Q_f uu_1 \ldots u_{n+1} & \text{iff} \\ \text{iff} & \text{for all } \bar{x}, y\text{: if } \bar{x} \in \bar{u}, \text{ then } y \in u_{n+1} \text{ implies } f(\bar{x}, y) \in u \\ \text{iff} & \text{for all } \bar{x}, y\text{: if } \bar{x} \in \bar{u}, \text{ then } f(\bar{x}, y) \notin u \text{ implies } y \notin u_{n+1} \\ \text{iff} & \text{for all } \bar{x}, y\text{: if } \bar{x} \in \bar{u}, \text{ then } -f(\bar{x}, y) \in u \text{ implies } -y \in u_{n+1} \\ \text{iff} & \text{for all } \bar{x}, z\text{: if } \bar{x} \in \bar{u}, \text{ then } -f(\bar{x}, -z) \in u \text{ implies } z \in u_{n+1}. \end{array}$$

This provides us with a minimal set of elements that u_{n+1} should contain; put

$$F := \{z \in A \mid (\exists \bar{x} \in \bar{u}) \ - f(\bar{x}, z) \in u\}.$$

If $-f(\bar{x}, -z) \in u$, we say that $\bar{x}$ *drive* z *into* F. We now take the second condition into account as well, defining

$$G := \{a_{n+1} \cdot z \mid z \in F\}.$$

Our aim is to prove the existence of an ultrafilter u_{n+1} containing G. It will be clear that this is sufficient to prove the theorem (note that $a_{n+1} \in G$ as $1 \in F$). In order to apply the Ultrafilter Theorem 6.2, we will show that G has the finite intersection property. We first need the following fact:

$$F \text{ is closed under intersection.} \tag{6.35}$$

Let z', z'' be in F; assume that z' and z'' are driven into F by $\bar{x}'$ and $\bar{x}''$, respectively. We will now see that $\bar{x} := (x'_1 \cdot x''_1, \ldots, x'_n \cdot x''_n)$ drives $z := z' \cdot z''$ into F, i.e., that $-f(\bar{x}, -z) \in u$.

Since f is monotonic, we have $f(\bar{x}, -z') \leq f(\bar{x}', -z')$, and hence we find that $-f(\bar{x}', -z') \leq -f(\bar{x}, -z')$. As u is upward closed and $-f(\bar{x}', -z') \in u$ by our 'driving assumption', this gives $-f(\bar{x}, -z') \in u$. In the same way we find $-f(\bar{x}, -z'') \in u$. Now

$$\begin{array}{rcl} f(\bar{x}, -z) & = & f(\bar{x}, -(z' \cdot z'')) \\ & = & f(\bar{x}, (-z') + (-z'')) \\ & = & f(\bar{x}, -z') + f(\bar{x}, -z''), \end{array}$$

whence

$$-f(\bar{x}, -z) = [-f(\bar{x}, -z')] \cdot [-f(\bar{x}, -z'')].$$

Therefore, $-f(\bar{x}, -z) \in u$, since u is closed under intersection. This proves (6.35).

We can now finish the proof and show that indeed

$$G \text{ has the finite intersection property.} \tag{6.36}$$

Let $a_{n+1} \cdot z_1, \ldots, a_{n+1} \cdot z_k$ be arbitrary elements of G, with every z_j in F. We have to show that

$$(a_{n+1} \cdot z_1) \cdot \ldots \cdot (a_{n+1} \cdot z_k) \neq 0.$$

First, observe that

$$(a_{n+1} \cdot z_1) \cdot \ldots \cdot (a_{n+1} \cdot z_k) = a_{n+1} \cdot (z_1 \cdot \ldots \cdot z_k).$$

As each z_j is in F, we find that $z_1 \cdot \ldots \cdot z_k \in F$ by (6.35).

Therefore, it suffices to show that

$$\text{for all } z \in F\text{: } a_{n+1} \cdot z = 0. \tag{6.37}$$

To prove (6.37), we reason by contraposition: suppose that $z \in F$ and $a_{n+1} \cdot z = 0$. Let $\bar{x} \in \bar{u}$ be a sequence that drives z into F, i.e., $-f(\bar{x}, -z) \in u$. From $a_{n+1} \cdot z = 0$ it follows that $a_{n+1} \leq -z$, so by monotonicity of f we get

$$-f(\bar{x}, -z) \leq -f(\bar{x}, a_{n+1}).$$

But then $-f(\bar{x}, a_{n+1}) \in u$, which contradicts (6.34). This proves (6.37) and with that (6.36) and Theorem 6.6. $\dashv$

6.2 Ultrafilter extensions

In the second part of this section we look at a special kind of ultrafilter frame, viz. the ultrafilter frame of a full complex algebra.

Definition 6.9 Let τ be a modal similarity type, and $\mathfrak{F} = (W, R_\Delta)_{\Delta \in \tau}$ a τ-frame. The *ultrafilter extension* $\mathfrak{ue}\,\mathfrak{F}$ of $\mathfrak{F}$ is defined as the ultrafilter frame of the complex algebra of $\mathfrak{F}$, i.e.,

$$\mathfrak{ue}\,\mathfrak{F} = (\mathfrak{F}^+)_+.$$

Putting it more directly (we now consider the standard similarity type), the ultrafilter extension of a frame $\mathfrak{F} = (W, R)$ is given as the frame

$$(Uf\mathrm{POW}(W), R^{ue}),$$

where $Uf\mathrm{POW}(W)$ is the set of ultrafilters of the power set algebra of W, and $R^{ue} = Q_{m_R}uv$ holds of two ultrafilters u and v, if $m_R(X) \in u$ for every $X \subseteq W$ such that $X \in v$.

The ultrafilter extension of a structure (model or frame) can be seen as a kind of *completion* of the original structure. To see this, we first need the following definition:

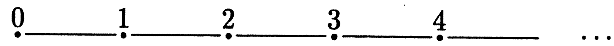
Definition 6.10 Let w be state in a frame $\mathfrak{F}$. We denote by u_w the *principal ultrafilter generated by* w, i.e., the set $u_w = \{X \subseteq W \mid w \in X\}$.

Note that any subset of a frame can, in principle, be viewed as (the extension of) a *proposition*. A filter over the universe of the frame can thus be seen as a *theory*, in fact as a closed theory, since filters are closed under intersection (conjunction) and also upward closed. A proper filter is then a *consistent* theory — it does not contain the empty set (falsum). In this perspective, an ultrafilter is a *complete* theory, or as we will call it, a *state of affairs*: it tells us of each proposition (subset of the universe) whether it holds (is a member of the ultrafilter) or not. Now in the original frame, not every state of affairs is 'realized' (in the sense that there is a world satisfying precisely all propositions belonging to the state of affairs), only the principal ultrafilters can be found. The ultrafilter extension of a frame provides realizations of *every* state of affairs, by simply adding them as points of the universe. So, now we have to define the relations holding between the states of the ultrafilter extension. In other words, we have to decide when to put an $n+1$-tuple of states of affairs in the relation R^{ue}. It is a natural suggestion to have $R^{ue}u_0u_1\dots u_n$ if u_0 'sees' the n-tuple $u_1,\dots,u_n$. To make this intuition precise, we interpret the condition as follows: whenever $X_1,\dots,X_n$ are propositions of $u_1,\dots,u_n$ respectively, then u_0 'sees' this combination, i.e. the proposition (subset of the universe) $m_\Delta(X_1,\dots,X_n)$ is a member of u_0.

Note that if one identifies each state w of a frame $\mathfrak{F}$ with the principal ultrafilter u_w, one can easily see that any frame $\mathfrak{F}$ is (isomorphic to) a *submodel* (in general not a *generated*) submodel of its ultrafilter extension. For, we have the following equivalences (in the basic modal similarity type):

$$\begin{array}{lll} Rwv & \text{iff} & w \in m_\Diamond(X) \text{ for all } X \subseteq W \text{ such that } v \in X \\ & \text{iff} & m_\Diamond(X) \in u_w \text{ for all } X \subseteq W \text{ such that } X \in u_v \\ & \text{iff} & R^{ue}u_wu_v. \end{array} \tag{6.38}$$

Example 6.11 Consider the frame $\mathfrak{N} = (\omega, <)$ (the ordering of the natural numbers):

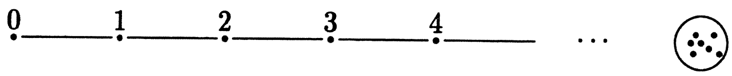


We will describe (and depict) the ultrafilter extension of $\mathfrak{N}$. Our first observation concerns the set of ultrafilters over *any* infinite set. It is an easy observation that there are two kinds of such ultrafilters: principal ones and co-finite ones, i.e., ultrafilters containing all co-finite sets and only infinite sets. (For, if an ultrafilter u contains a finite set, then it must contain a singleton, whence u is principal.) By the observation (6.38), inside $\mathfrak{ue}\,\mathfrak{N}$ the principal ultrafilters form an isomorphic copy of the frame $\mathfrak{N}$. So it suffices to show where the co-finite ultrafilters are situated. The key fact here is that

for any pair of ultrafilters u, u': if u' is co-finite, then $R^{ue}uu'$.

To prove this claim, let u' be a co-finite ultrafilter, and $X \in u'$. As X is infinite, for any $n \in \omega$ there is an m such that $n < m$ and $m \in X$. This shows that $m_\Diamond(X) = \omega$. But ω is an element of every ultrafilter u.

This shows that the ultrafilter extension of $\mathfrak{N}$ consists of a copy of $\mathfrak{N}$, followed by a big cluster consisting of all co-finite ultrafilters, viz.



Proposition 6.12 *Let τ be a modal similarity type, $\mathfrak{F}$ a τ-frame and ϕ a τ-formula. Then $\mathfrak{F} \models \phi$ if $\mathfrak{ue}\,\mathfrak{F} \models \phi$.*

Proof. This anti-preservation result of modal frame validity under taking ultrafilter extensions is an immediate corollary of Proposition 5.6 and the fact that the validity of equations is preserved under taking subalgebras:

$$\begin{array}{rrl} \mathfrak{ue}\,\mathfrak{F} \models \phi & \text{iff} & (\mathfrak{ue}\,\mathfrak{F})^+ \models \phi^t = 1 \\ & \text{iff} & ((\mathfrak{F}^+)_+)^+ \models \phi^t = 1 \\ & \text{only if} & \mathfrak{F}^+ \models \phi^t = 1 \\ & \text{iff} & \mathfrak{F} \models \phi \end{array}$$

Here the important step is the implication, which is justified by the fact that $((\mathfrak{F}^+)_+)^+$ is the *embedding* algebra of the complex algebra $\mathfrak{F}^+$. ⊣

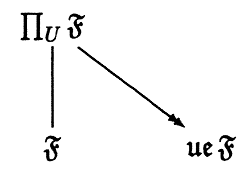
This proposition can and will be used to show that some frame properties are not modally definable.

Example 6.13 Working in the basic modal similarity type, we consider the property that every state has a reflexive successor, in the first-order frame language: $\forall x \exists y\,(Rxy \wedge Ryy)$. We claim that this property is not modally definable, although it is preserved under taking disjoint unions, generated subframes *and* bounded morphic images. To verify the claim, the reader is asked to look at the frame in Example 6.11. It is easy to see that every state of $\mathfrak{ue}\,\mathfrak{F}$ has a reflexive successor — take any non-principal ultrafilter. But $\mathfrak{F}$ itself clearly does not satisfy the property, as $\mathfrak{F}$ has *no* reflexive states.

Now suppose that the property were modally definable, let's say by the set of formulas Δ. Then we would have $\mathfrak{ue}\,\mathfrak{F} \models \Delta$, but $\mathfrak{F} \not\models \Delta$, a clear violation of Proposition 6.12.

Note the direction of the preservation result in Proposition 6.12. It states that modal validity is *anti*-preserved under taking ultrafilter extensions. This naturally raises the question whether the other direction holds as well, i.e., whether $\mathfrak{F} \models \phi$ implies $\mathfrak{ue}\,\mathfrak{F} \models \phi$. For a partial answer to this question, we need the following theorem, which is due to van Benthem, building on ideas from Fine:

Theorem 6.14 *Let τ be a modal similarity type, and $\mathfrak{F}$ a τ-frame. Then $\mathfrak{F}$ has an ultrapower $\prod_U \mathfrak{F}$ such that $\prod_U \mathfrak{F} \twoheadrightarrow \mathfrak{ue}\,\mathfrak{F}$.*



Proof. For a proof of this theorem, which essentially uses the notion of ω-saturation, we refer the reader to [2], or to [4]. ⊣

Corollary 6.15 *Let τ be a modal similarity type, and ϕ a τ-formula. having a first-order correspondent on the frame level. Then validity of ϕ is preserved under taking ultrafilter extensions.*

Proof. Suppose that there is a set Δ of first-order formulas such that for all frames $\mathfrak{F}$:

$$\mathfrak{F} \models \Delta \text{ iff } \mathfrak{F} \models \phi.$$

Let $\mathfrak{F}$ be a frame such that $\mathfrak{F} \models \phi$. It follows that $\mathfrak{F} \models \Delta$. If $\prod_U \mathfrak{F}$ is the ultrapower of $\mathfrak{F}$ such that $\mathfrak{F} \twoheadrightarrow \mathfrak{ue}\,\mathfrak{F}$, then $\prod_U \mathfrak{F} \models \Delta$, since validity of first-order formulas is preserved under taking ultrapowers. However, this implies that $\prod_U \mathfrak{F} \models \phi$, so by the preservation result of modal validity under taking bounded morphic images, we obtain that $\mathfrak{ue}\,\mathfrak{F} \models \phi$. ⊣

However, perhaps an even more important fact concerning Theorem 6.14 is that now we have sufficient testing material to find out whether a class of frames is modally definable — that is to say, if we confine ourselves to *first-order definable* classes of frames. In section 7 we will give a more precise formulation and a proof of this result by Goldblatt and Thomasson.

6.3 Basic duality theory

In this section we will show how the constructions of algebras from frames and frames from algebras can be extended to *morphisms* between frames or between algebras. As an important application of these 'lifting' constructions, we can link the operations on frame classes, viz. of taking bounded morphic images, generated subframes and disjoint unions, to the operations on classes of algebras of taking respectively subalgebras, homomorphic images and products. These links are formulated concisely in the Theorems 6.17 and 6.18, in which we use the following definitions:

Definition 6.16 Let τ be a modal similarity type, $\mathfrak{F}$ and $\mathfrak{G}$ two τ-frames, and $\mathfrak{A}$ and $\mathfrak{B}$ two boolean algebras with τ-operators. We recall (define, respectively) the following notation for relations between these structures:

$\mathfrak{F} \rightarrowtail \mathfrak{G}$	for	$\mathfrak{F}$ is a generated substructure of $\mathfrak{G}$
$\mathfrak{F} \twoheadrightarrow \mathfrak{G}$	for	$\mathfrak{G}$ is a bounded morphic image of $\mathfrak{F}$
$\mathfrak{A} \rightarrowtail \mathfrak{B}$	for	$\mathfrak{A}$ is a subalgebra of $\mathfrak{B}$
$\mathfrak{A} \twoheadrightarrow \mathfrak{B}$	for	$\mathfrak{B}$ is a homomorphic image of $\mathfrak{A}$.

Theorem 6.17 *Let τ be a modal similarity type, $\mathfrak{F}$ and $\mathfrak{G}$ two τ-frames, and $\mathfrak{A}$ and $\mathfrak{B}$ two boolean algebras with τ-operators. Then*

1. *If $\mathfrak{F} \rightarrowtail \mathfrak{G}$, then $\mathfrak{G}^+ \twoheadrightarrow \mathfrak{F}^+$.*
2. *If $\mathfrak{F} \twoheadrightarrow \mathfrak{G}$, then $\mathfrak{G}^+ \rightarrowtail \mathfrak{F}^+$.*
3. *If $\mathfrak{A} \rightarrowtail \mathfrak{B}$, then $\mathfrak{B}_+ \twoheadrightarrow \mathfrak{A}_+$.*
4. *If $\mathfrak{A} \twoheadrightarrow \mathfrak{B}$, then $\mathfrak{B}_+ \rightarrowtail \mathfrak{A}_+$.*

Proof. This Theorem follows immediately from the Propositions 6.21 and 6.22 below. ⊣

Theorem 6.18 *Let τ be a modal similarity type, and $\mathfrak{F}_i, i \in I$ a family of τ-frames; then*

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

Proof. We define a map η from the power set of the disjoint union $\biguplus_{i\in I} W_i$ to the carrier $\prod_{i\in I} \mathrm{POW}(W_i)$ of the product of the family of complex algebras $(\mathfrak{F}_i^+)_{i\in I}$.

Let X be a subset of $\biguplus_{i\in I} W_i$; $\eta(X)$ will have to be an element of the set $\prod_{i\in I} \mathrm{POW}(W_i)$. Note that elements of the set $\prod_{i\in I} \mathrm{POW}(W_i)$ are sequences σ such that $\sigma(i) \in \mathrm{POW}(W_i)$. So it suffices to say what the i-th element of the sequence $\eta(X)$ is:

$$\eta(X)(i) = X \cap W_i.$$

We leave it as an exercise to show that η is an isomorphism. ⊣

In order to prove Theorem 6.17, the reader is advised to recall the definition of a bounded morphism between two frames (Definition 2.21). We also need some terminology for morphisms between boolean algebras with operators:

Definition 6.19 Let $\mathfrak{A}$ and $\mathfrak{A}'$ be two boolean algebras with operators of the same similarity type, and let $\eta : A \to A'$ be a function. We say that η is a *boolean homomorphism* if η is a homomorphism from $(A, +, -, 0)$ to $(A', +, -, 0)$. We call η a *modal homomorphism* if η satisfies, for all modal operators Δ:

$$\eta(f_\Delta(a_1, \ldots, a_{\rho(\Delta)})) = f'_\Delta(\eta a_1, \ldots, \eta a_{\rho(\Delta)}).$$

Finally, η is a *(BAO-)homomorphism* if it is both a boolean and a modal homomorphism.

In the following definition, the construction of *dual*[6] or *lifted* morphisms is given.

Definition 6.20 Suppose that θ is a map from W to W'; then its *dual* $\theta^+ : \mathrm{POW}(W') \mapsto \mathrm{POW}(W)$ is defined as:

$$\theta^+(X') = \{u \in W \mid \theta(u) \in X'\}.$$

In the other direction, let $\mathfrak{A}$ and $\mathfrak{A}'$ be two BAOs, and $\eta : \mathfrak{A} \mapsto \mathfrak{A}'$ be a map from A to A''; then its *dual* is given as the following map from ultrafilters of $\mathfrak{A}'$ to subsets of A:

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

The following propositions assert that the duals of bounded morphisms are BAO-homomorphisms themselves:

Proposition 6.21 *Let $\mathfrak{F}, \mathfrak{F}'$ be frames, and $\theta : W \mapsto W'$ a map.*

1. *θ^+ is a boolean homomorphism.*
2. *If θ has (zig), then $m_R(\theta^+(Y'_1), \ldots, \theta^+(Y'_n)) \subseteq \theta^+ m_{R'}(Y'_1, \ldots, Y'_n)$.*
3. *If θ has (zag), then $m_R(\theta^+(Y'_1), \ldots, \theta^+(Y'_n)) \supseteq \theta^+ m_{R'}(Y'_1, \ldots, Y'_n)$.*
4. *If θ is a bounded morphism, then θ^+ is a BAO-homomorphism from $\mathfrak{Cm}\mathfrak{F}$ to $\mathfrak{Cm}\mathfrak{F}'$.*
5. *If θ is injective, then θ^+ is surjective.*

[6]Note that the word 'dual' is used here *not* in the sense of $\Diamond$ being the dual of $\Box$.

6. If θ is surjective, then θ^+ is injective.

Proof. For notational convenience, we assume that τ has only one modal operator, so that we can write $\mathfrak{F} = (W, R)$.

1. As an example, we treat complementation:

$$\begin{array}{lll} x \in \theta^+(-X') & \text{iff} & \theta x \in -X' \\ & \text{iff} & \theta x \notin X' \\ & \text{iff} & x \notin \theta^+(X'). \end{array}$$

From this it follows immediately that $\theta^+(-X') = -\theta^+(X')$.

2. Assume that θ has (zig). Then we have

$$\begin{array}{ll} x \in & m_R(\theta^+(Y_1'), \ldots, \theta^+(Y_n')) \\ \Longrightarrow & \text{there are } y_1, \ldots, y_n \text{ such that } \theta y_i \in Y_i' \text{ and } Rxy_1 \ldots y_n \\ \Longrightarrow & \text{there are } y_1, \ldots, y_n \text{ such that } \theta y_i \in Y_i' \text{ and } R'\theta x \theta y_1 \ldots \theta y_n \\ \Longrightarrow & \theta x \in m_{R'}(Y_1', \ldots, Y_n') \\ \Longrightarrow & x \in \theta^+ m_{R'}(Y_1', \ldots, Y_n'). \end{array}$$

3. Now suppose $x \in \theta^+ m_{R'}(Y_1', \ldots, Y_n')$, then $\theta x \in m_{R'}(Y_1', \ldots, Y_n')$. So there are $y_1', \ldots, y_n'$ in W' with $y_i' \in Y_i'$ and $R'\theta x y_1' \ldots y_n'$. As θ has (zag), there are $y_1, \ldots, y_n \in W$ with $\theta(y_i) = y_i'$ for all i, and $Rxy_1 \ldots y_n$. But then $y_i \in \theta^+ Y_i'$ for every i, so $x \in m_R(\theta^+(Y_1'), \ldots, \theta^+(Y_n'))$.

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

5. Assume that θ is injective, and let X be a subset of W. We have to find a subset X' of W' such that $\theta^+(X') = X$. Define

$$\theta[X] := \{\theta x \in W' \mid x \in W\}.$$

We claim that this set has the desired properties. It is immediate that $X \subseteq \theta^+(\theta[X])$. For the other direction, let x be an element of $\theta^+(\theta[X])$. Then by definition, $\theta(x) \in \theta[X]$, so there is a $y \in X$ such that $\theta(x) = \theta(y)$. Now by injectivity of θ, $x = y$. So $x \in X$.

6. Assume that θ is surjective, and let X' and Y' be distinct subsets of W'. Without loss of generality we may assume that there is an x' such that $x' \in X'$ and $x' \notin Y'$. As θ is surjective, there is an x in W such that $\theta(x) = x'$. So $x \in \theta^+(X')$, but $x \notin \theta^+(Y')$. So $\theta(X') \neq \theta(Y')$, whence θ^+ is injective. $\dashv$

In the other direction, i.e., from algebras to relational structures, we find that the duals of BAO-homomorphisms are bounded morphisms:

Proposition 6.22 *Let $\mathfrak{A}, \mathfrak{A}'$ be boolean algebras with operators, and η a map from A to A'.*

1. *If η is a boolean homomorphism, then η_+ maps ultrafilters to ultrafilters.*

2. *If $f'(\eta(a_1), \ldots, \eta(a_n)) \leq \eta(f(a_1, \ldots, a_n)$, then η_+ has (zig).*

3. *If $f'(\eta(a_1), \ldots, \eta(a_n)) \geq \eta(f(a_1, \ldots, a_n)$ and η is a boolean homomorphism, then η_+ has (zag).*

4. *If η is a BAO-homomorphism, then η is a bounded morphism from $\mathfrak{A}'_+$ to $\mathfrak{A}_+$.*

5. *If η is an injective boolean homomorphism, then $\eta_+ : Uf\mathfrak{A}' \to Uf\mathfrak{A}$ is surjective.*

6. *If η is an surjective boolean homomorphism, then $\eta_+ : Uf\mathfrak{A}' \to Uf\mathfrak{A}$ is injective.*

Proof. Again, without loss of generality we assume that τ has only one modal operator, so that we can write $\mathfrak{A} = (A, +, -, 0, f)$.

1. This part is left as an exercise.

2. Suppose that $Q_{f'}u'u'_1 \dots u'_n$ holds between some ultrafilters $u', u'_1, \dots, u'_n$ of $\mathfrak{A}'$. To show that $\mathfrak{A}_+ \models Q_f \eta_+ u' \eta_+ u'_1 \dots \eta_+ u'_n$, let $a_1, \dots, a_n$ be arbitrary elements of $\eta_+ u'_1, \dots, \eta_+ u'_n$ respectively. Then, by definition of η_+, $\eta a_i \in u'_i$, so $Q'_f u' u'_1 \dots u'_n$ gives $f'(\eta a_1, \dots, \eta a_n) \in u'$. Now the assumption yields $\eta f(a_1, \dots, a_n) \in u'$, as ultrafilters are upward closed. But then $f(a_1, \dots, a_n) \in \eta_+ u'$, which is what we wanted.

3. This part is left as an exercise.

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

5. Assume that η is injective, and let u be an ultrafilter of $\mathfrak{A}$. We want to follow the same strategy as in Proposition 6.21(5), and define

$$\eta[\iota] := \{\eta a \mid a \in u\}.$$

The difference with the earlier situation is that here, $\eta_+(\eta[u])$ is not well-defined, unless $\eta[u]$ is an ultrafilter.

We will first show that $\eta[u]$ is a *proper filter* of $\mathfrak{A}'$. For $(F1)$, $1 \in u$, so $\eta(1) = 1 \in \eta[u]$. For $(F2)$, assume $a', b' \in \eta[u]$. Then there are a, b in A such that $\eta a = a'$ and $\eta b = b'$. It follows that $\eta(a \cdot b) = \eta(a) \cdot \eta(b) = a' \cdot b' \in \eta[u]$, so $\eta[u]$ is closed under intersection. Likewise, one can show that η is upwards closed. Finally, suppose that $0' \in \eta[u]$. Then $0' = \eta(a)$ for some $a \in u$; as $0' = \eta(0)$, injectivity of η gives that $0 = a$, and hence, $0 \in u$. But then u is not an ultrafilter.

By the Ultrafilter Theorem 6.2, $\eta[u]$ can be extended to an ultrafilter u'. We claim that $u = \eta_+(u')$. First let a be in u, then $\eta a \in \eta[u] \subseteq u'$, so $a \in \eta_+(u')$. This shows that $u \subseteq \eta_+(u')$. For the other inclusion, it suffices to show that $a \notin \eta_+(u')$ if $a \notin u$; we reason as follows:

$$\begin{array}{rcl} a \notin u & \Longrightarrow & -a \in u \\ & \Longrightarrow & -\eta(a) = \eta(-a) \in \eta[u] \\ & \Longrightarrow & -\eta(a) \in u' \\ & \Longrightarrow & \eta(a) \notin u' \\ & \Longrightarrow & a \notin \eta_+(u') \end{array}$$

6. Similar to Proposition 6.21, part 6. $\dashv$

7. Applications of duality

In this last section we put some threads together and show how we can use the duality between frames and algebras to give very short proofs for some major theorems in the theory of modal logics.

Our first example shows that all the results given in Proposition 2.22 on the preservation of *modal* validity under certain frame operations fall out as simple consequences of the preservation results concerning *equational* validity under taking subalgebras, homomorphic images and products of algebras.

Proposition 7.1 *Let τ be a modal similarity type, ϕ a τ-formula and $\mathfrak{F}$ a τ-frame. Then*

1. *If $\mathfrak{G}$ is a bounded morphic image of $\mathfrak{F}$, then $\mathfrak{G} \models \phi$ if $\mathfrak{F} \models \phi$.*
2. *If $\mathfrak{G}$ is a generated subframe of $\mathfrak{F}$, then $\mathfrak{G} \models \phi$ if $\mathfrak{F} \models \phi$.*
3. *If $\mathfrak{F}$ is the disjoint union of a family $\{\mathfrak{F}_i \mid i \in I\}$, then $\mathfrak{F} \models \phi$ if for every $i \in I$, $\mathfrak{F}_i \models \phi$.*

Proof. We only prove the first part of the Proposition, leaving the other parts as an exercise to the reader.

Assume that $\mathfrak{F} \twoheadrightarrow \mathfrak{G}$, and $\mathfrak{F} \models \phi$. By Proposition 5.8, we have $\mathfrak{F}^+ \models \phi^t = 1$, and by Theorem 6.17, $\mathfrak{G}^+$ is a *subalgebra* of $\mathfrak{F}^+$. So by the fact that equational validity is preserved under taking subalgebras, we obtain that $\phi^t = 1$ holds in $\mathfrak{G}^+$. But then Proposition 5.8 implies that $\mathfrak{G} \models \phi$. $\dashv$

Our second example is a very simple proof of the Goldblatt-Thomason Theorem, which gives a precise structural characterization of the first-order definable classes of frames which are modally definable. Once we have set up the basic duality framework, this theorem is a more or less immediate corollary of Birkhoff's Theorem identifying equational classes and varieties:

Theorem 7.2 *Let τ be a modal similarity type. A first-order definable class K of τ-frames is modally definable if and only if it is closed under taking bounded morphic images, generated subframes, disjoint unions and it reflects*[7] *ultrafilter extensions.*

Proof. Let K be a class of frames satisfying the closure conditions given in the theorem. It suffices to show that any frame $\mathfrak{F}$ validating the modal theory of K, is itself a member of K.

Let $\mathfrak{F}$ be such a frame. It is not difficult to show that Proposition 5.8 implies that $\mathfrak{F}^+$ is a model for the equational theory of the class **Cm**K. It follows by Birkhoff's Theorem (identifying varieties and equational classes) that $\mathfrak{F}^+$ is in the variety generated by **Cm**K, so $\mathfrak{F}^+$ is in **HSPCm**K. In other words, there is a family $(\mathfrak{G}_i)_{i\in I}$ of frames in K, and there are boolean algebras with operators $\mathfrak{A}$ and $\mathfrak{B}$ such that

1. $\mathfrak{B}$ is the product $\prod_{i\in I} \mathfrak{G}_i^+$ of the complex algebras of the $\mathfrak{G}_i$,
2. $\mathfrak{A}$ is a subalgebra of $\mathfrak{B}$, and
3. $\mathfrak{F}^+$ is a homomorphic image of $\mathfrak{A}$.

By Theorem 6.18, $\mathfrak{B}$ is isomorphic to the complex algebra of the disjoint union $\mathfrak{G}$ of the family $(\mathfrak{G}_i)_{i\in I}$:

$$\mathfrak{B} \cong \mathfrak{G}^+ = (\biguplus_{i\in I} \mathfrak{G}_i)^+.$$

As K is closed under taking disjoint unions, $\mathfrak{G}$ is in K.

Now we have the following picture:

[7] This is to say that $\mathfrak{F}$ is a member of K whenever $\mathfrak{ue}\,\mathfrak{F}$ is.

$$\mathfrak{F}^+ \twoheadleftarrow \mathfrak{A} \rightarrowtail \mathfrak{G}^+.$$

By Theorem 6.17 it follows that

$$(\mathfrak{F}^+)_+ \rightarrowtail \mathfrak{A}_+ \twoheadleftarrow (\mathfrak{G}^+)_+.$$

Since K is closed under ultrapowers, Theorem 6.14 gives that $(\mathfrak{G}^+)_+ = \mathfrak{ue}\,\mathfrak{G}$ is in K. As K is closed under taking bounded morphic images and generated subframes, it follows that $\mathfrak{A}_+$ and $\mathfrak{ue}\,\mathfrak{F} = (\mathfrak{F}^+)_+$ (in that order) are in K. But then $\mathfrak{F}$ itself is also a member of K, since K reflects ultrafilter extensions. ⊣

As an important application, we return to the concept of canonicity. Here we will prove an important result and mention an intriguing open problem, both having to do with the relation between canonical varieties and first-order definable classes of frames. First we need a definition:

Definition 7.3 Let τ be modal similarity type, and K be a class of τ-frames. The variety generated by K, notation: $\mathbf{V}_{\mathsf{K}}$, is the class **HSPCm**K.

Theorem 7.4 *Let τ be modal similarity type, and* K *be a class of τ-frames which is closed under ultraproducts. Then the variety $\mathbf{V}_{\mathsf{K}}$ is canonical.*

Proof. Assume that the class K of τ-frames is closed under taking ultraproducts. We will first prove that the class **HSCm**K is canonical. Let $\mathfrak{A}$ be in this class, i.e., assume that there is a frame $\mathfrak{F}$ in K and an algebra $\mathfrak{B}$ such that

$$\mathfrak{A} \twoheadleftarrow \mathfrak{B} \rightarrowtail \mathfrak{F}^+.$$

It follows from Theorem 6.17 that

$$\mathfrak{Em}\mathfrak{A} \twoheadleftarrow \mathfrak{Em}\mathfrak{B} \rightarrowtail \mathfrak{Em}\mathfrak{F}^+ = (\mathfrak{ue}\,\mathfrak{F})^+. \tag{7.39}$$

From Theorem 6.14 we know that $\mathfrak{ue}\,\mathfrak{F}$ is the bounded morphic image of some ultrapower $\mathfrak{G}$ of $\mathfrak{F}$. Note that $\mathfrak{G}$ is in K, by our assumption. Now Theorem 6.17 gives

$$(\mathfrak{ue}\,\mathfrak{F})^+ \rightarrowtail \mathfrak{G}^+. \tag{7.40}$$

Since $\mathfrak{G}^=$ is in **Cm**K, (7.39) and (7.40) together imply that $\mathfrak{A}$ is in **HSCm**K. Hence this class is canonical.

To prove that the *variety* generated by K is canonical, we have to do some extra work, and we need one additional lemma:

Fact 7.5 Let τ be a modal similarity type, and K a class of τ-frames. Suppose that $\mathfrak{G}$ is an ultrapower of the disjoint union $\biguplus_{i\in I}\mathfrak{F}_i$, where $\{\mathfrak{F}_i \mid i \in I\}$ is a family of frames in K. Then $\mathfrak{G}$ is a bounded morphic image of a disjoint union of ultraproducts of frames in K.

For a *proof* of this Fact, which is essentially a result of Goldblatt, we refer to [15] or [4].

Now assume that $\mathfrak{A}$ is $\mathbf{V}_{\mathsf{K}} = $ **HSPCm**K. In other words, there are a family $\{\mathfrak{F}_i \mid i \in I\}$ of frames in K and an algebra $\mathfrak{B}$ such that

$$\mathfrak{A} \twoheadleftarrow \mathfrak{B} \rightarrowtail \prod_{i\in I} \mathfrak{F}_i^+.$$

In order to prove that $\mathfrak{Em}\mathfrak{A}$ is in $\mathbf{V}_{\mathsf{K}}$, it suffices to show that $\mathfrak{Em}(\prod_{i\in I} \mathfrak{F}_i^+)$ is in **SPCmK** — the remainder of the proof is as before. Let $\mathfrak{F}$ be the frame $\biguplus_{i\in I} \mathfrak{F}_i$, then by Theorem 6.18, $\mathfrak{F}^+ \cong \prod_{i\in I} \mathfrak{F}_i^+$. Hence, by Theorem 6.17:

$$\mathfrak{Em}(\prod_{i\in I} \mathfrak{F}_i^+) \cong ((\mathfrak{F}^+)_+)^+ = (\mathfrak{ue}\,\mathfrak{F})^+. \tag{7.41}$$

By Theorem 6.14, there is an ultrapower $\mathfrak{G}$ of $\mathfrak{F}$ such that $\mathfrak{G} \twoheadrightarrow \mathfrak{ue}\,\mathfrak{F}$. Now we apply our *Fact*, yielding a frame $\mathfrak{H}$ such that (i) $\mathfrak{H}$ is a disjoint union of ultraproducts of frames in K and (ii) $\mathfrak{H} \twoheadrightarrow \mathfrak{G}$. Putting these observations together in a picture, we now have:

$$\mathfrak{ue}\,\mathfrak{F} \twoheadleftarrow \mathfrak{G} \twoheadleftarrow \mathfrak{H},$$

Hence, by Theorem 6.17:

$$(\mathfrak{ue}\,\mathfrak{F})^+ \rightarrowtail \mathfrak{G}^+ \rightarrowtail \mathfrak{H}^+. \tag{7.42}$$

Note that $\mathfrak{H}$ is a disjoint union of K-frames, since K is closed under taking ultraproducts. This implies that $\mathfrak{H}^+$ is in **PCmK**. But then it follows from (7.41) and (7.42) that $\mathfrak{Em}(\prod_{i\in I} \mathfrak{F}_i^+)$ is in **SPCmK**, which is what we needed. ⊣

Example 7.6 Consider the modal similarity type $\{\circ, \otimes, \delta\}$ of arrow logic, where $\circ$ is binary, $\otimes$ is unary and δ is a constant (cf. Example 1.8). The standard interpretation of this language is given by the so-called *squares*: a square is a frame $\mathfrak{F} = (W, C, F, I)$ where for some base set U:

$$\begin{array}{rcl} W & = & U \times U \\ C((u,v),(w,x),(y,z)) & \text{iff} & u = w \ \text{ and } \ v = z \ \text{ and } \ x = y \\ F((u,v),(w,x)) & \text{iff} & u = x \ \text{ and } \ v = w \\ I(u,v) & \text{iff} & u = v \end{array}$$

It is not very difficult to show that the class SQ of (isomorphic copies of) squares allows a first-order definition (in the frame language with predicates C, F and I), cf. [26] for a proof. Therefore, Theorem 7.4 implies that the variety generated by SQ is canonical. This variety is well-known in the literature on algebraic logic as the variety RRA of Representable Relation Algebras (cf. [20]).

Rephrased in terminology from modal logic, Theorem 7.4 boils down to the following result, originally due to Fine (for the standard similarity type):

Corollary 7.7 *Let τ be modal similarity type, and* K *be a class of τ-frames which is closed under ultraproducts. Then the modal theory of* K *is a canonical logic.*

We now return to the (simple) Sahlqvist formulas of Section 3. There it was proved that every Sahlqvist formula corresponds to a first-order frame condition. This correspondence result can now be used to show that all Sahlqvist logics are canonical. This result has tremendous consequences in the completeness theory of modal logic, since it is a general theorem applying to many important frame classes.

Theorem 7.8 *Let τ be modal similarity type, and Σ a set of Sahlqvist axioms. Then*

1. *V_Σ is a canonical variety.*
2. *$\mathbf{K\Sigma}$ is a canonical logic and hence complete with respect to the class of frames defined by Σ (or, equivalently, by its first-order correspondents).*

Proof. Let Σ be a set of Sahlqvist axioms. By Theorem 3.18, Σ defines an *first-order definable* class K of τ-frames; it follows from Birkhoff's Theorem (identifying varieties and equational classes) that $\mathsf{V}_\Sigma = \mathbf{V}_{\mathsf{K}}$. As K is closed under taking ultraproducts, this variety is canonical. So, Proposition 6.8 implies that the logic $\mathbf{K\Sigma}$ is a canonical logic. ⊣

Note that Theorem 4.21 is an immediate corollary of the result above.

Let us finish this section with stating the question whether the obvious converse to Theorem 7.4 holds as well. This is in fact the foremost Open Problem in this area:

Open Problem 2 *Let τ be modal similarity type, and V a canonical variety of boolean algebras with τ-operators. Is there a class K of τ-frames, closed under taking ultraproducts, such that V is generated by K?*

Part IV: Notes

We conclude this paper with a few comments on topics on the interface of modal logic and boolean algebras with operators that we didn't cover, and with some suggestions for further reading.

A very important paper with implications for modal logic is Jónsson and Tarski's paper of 1951 [17]. If it had been widely read when it was published, the history of modal logic might have looked different. This paper extends the Stone duality for boolean algebras (see, for example, [21]) to boolean algebras with operators. Moreover, the paper shows that to get a full duality between modal frames and modal algebras one needs to consider so-called *general frames*. In the case of the standard modal language, these are structures of the form

$$(W, R, \mathcal{V})$$

where $\mathcal{V}$ is a collection of subsets of W that is closed under complementation, intersection and the operator m_R from Section 5. What we then get is a categorial duality between modal algebras and general frames. We refer the reader to Goldblatt's [12] for further details on the duality between frames and boolean algebras with operators.

In these notes we exploited the connection between frames and algebras mainly to get completeness results. But the connection can exploited to transfer many more results between logics and algebras. Here are a few examples of pairs of corresponding (modal) logical and algebraic properties:

Finite axiomatizability	Generating a finitely axiomatizable quasivariety
Finite frame property	Finite algebra property
Consequence compactness	Closure under ultrapowers
Craig Interpolation	Strong Amalgamation
Deduction theorem	Equationally definable principal congruences.

We refer the reader to Németi [20], and Blok and Pigozzi [6] for further details on algebraic counterparts of logical properties. Both papers also present valuable discussions about the question which logics have an algebraic counterpart.

Another important issue on the borderline between modal logic and boolean algebras with operators is the lattice of modal logics. The observation that the lattice of normal extensions of **K** is dually isomorphic to the lattice of subvarieties of the variety of modal algebras has been used by Blok [5] to invoke powerful results from the theory of lattices of varieties. As a consequence, one can get strong results on incompleteness, or on tabular and pretabular logics.

Finally, in these notes we concentrated on normal modal logics whose semantics employ arbitrary relations. As was briefly pointed out in the introduction, over the past decade or so there has been considerable interest in looking at modal languages that escape this format. On the one hand people have looked at richer languages that explore special relations, or at languages with non-normal modal operators. As an example of the former, one modal operator that has been considered is the *universal modality* A whose truth definition reads: $x \Vdash A\phi$ iff for all y: $y \Vdash \phi$: the algebraic semantics for this modal language is explored in Goranko and Passy [16]. As another example, the binary modal operator *Until* whose truth definition reads

$$x \Vdash \mathit{Until}(\phi, \psi) \text{ iff } \exists y\,(Rxy \wedge y \Vdash \phi \wedge \forall z\,(Rxz \wedge Rzy \rightarrow z \Vdash \psi)),$$

doesn't have a decent algebraic semantics as far as we know. On the other hand, weaker modal languages, and in particular, languages that aren't based on boolean algebras, but on alternative structures such as distributive lattices or Heyting algebras, have also been studied extensively; we refer the reader to Goldblatt's [14] for details.

Acknowledgment. Maarten de Rijke would like to acknowledge the financial support of the Netherlands Organization for Scientific Research (NWO), project NF 102/62-356 'Structural and Semantic Parallels in Natural Languages and Programming Languages'.

REFERENCES

1. H. Andréka, J.D. Monk and I. Németi (eds). *Algebraic Logic.* North-Holland, Amsterdam, 1991.
2. J. van Benthem. *Modal Logic and Classical Logic.* Bibliopolis, Naples, 1983.
3. P. Blackburn, C. Gardent and W. Meyer Viol. Talking About Trees. In: *Proc 6th EACL*, 1993.
4. P. Blackburn, M. de Rijke and Y. Venema. *Modal Logic. A Textbook.* Unpublished manuscript. CWI, Amsterdam, 1994.
5. W. Blok. The lattice of modal algebras: an algebraic investigation. *Theoria* 39:101–124, 1980.
6. W. Blok and D. Pigozzi. Abstract algebraic logic. This volume.
7. S. Burris and H.P. Sankappanavar, *Universal Algebra*, Springer Verlag, Berlin, 1981.
8. C. Chang and J. Keisler. *Model Theory.* North-Holland, Amsterdam, 1973.

9. J.N. Crossley (ed). *Algebra and Logic*. Springer Lecture Notes in Mathematics **450**, Berlin, 1975.

10. K. Fine. Some connections between elementary logic and modal logic. In: [18], pp. 15–31.

11. K. Fine. An incomplete logic containing **S4**. *Theoria* 40:23–29, 1974.

12. R. Goldblatt. Metamathematics of modal logic. *Reports on Mathematical Logic* 6:41–77, 1976, and 7:21–52, 1976.

13. R.I. Goldblatt and S.K. Thomason. Axiomatic classes in propositional modal logic. In: [9], pp. 163–173.

14. R. Goldblatt. Varieties of complex algebras. *Annals of Pure and Applied Logic* 44:173–242, 1989.

15. R. Goldblatt. On closure on canonical embedding algebras. In: [1], pp. 217–229.

16. V. Goranko and S. Passy. Using the universal modality. *Journal of Logic and Computation* 2:5–30, 1992.

17. B. Jónsson and A. Tarksi. Boolean algebras with operators. *American Journal of Mathematics* 78:891–939, 1951, and 74:127–162, 1952.

18. S. Kanger (ed.). *Proc. of the Third Scandinavian Logic Symposium Uppsala 1973.* North-Holland, Amsterdam, 1975.

19. Z. Manna and A. Pnueli. *The Temporal Logic of Reactive and Concurrent Systems. Vol. 1 Specification.* Springer-Verlag, New York, 1992.

20. I. Németi. Algebraizations of quantifier logics, an introductory overview. *Studia Logica* 50:485–569, 1991.

21. H. Rasiowa and R. Sikorski. *The Mathematics of Metamathematics*. Państwowe Wydawnictwo Naukowe, Warszawa, 1963.

22. H. Sahlqvist. Completeness and correspondence in the first and second order semantics for modal logic. In: [18], pp. 110–143.

23. K. Segerberg. Modal logics with linear alternative relations. *Theoria* 36:301–322, 1970.

24. C. Smoryński. *Self-Reference and Modal Logic.* Springer-Verlag, New York, 1985.

25. S.K. Thomason. Semantic analysis of tense logics. *Journal of Symbolic Logic* 37:716–720, 1972.

26. Y. Venema. *Many-Dimensional Modal Logic.* PhD thesis. ILLC, University of Amsterdam, 1991.