Annals of Pure and Applied Logic 37 (1988) 249-296
North-Holland

TOPOLOGY AND DUALITY IN MODAL LOGIC

Giovanni SAMBIN
Università di Padova, Seminario Matematico, Via Belzoni 7, 35131 Padova, Italy

Virginia VACCARO
Università di Napoli, Dipartimento di Matematica e Applicazioni, Via Mezzocannone 8, 80134 Napoli, Italy

Communicated by D. van Dalen
Received 15 September 1985; revised 26 June 1986

Contents

Foreword

This paper is the result of putting together a few ingredients, namely modal algebras, general frames, some topology, a little category theory, and mixing them up, at the light of available literature, with the flavouring of some opinions, technical or not, of the authors. If the recipe is good, only the reader can judge, after tasting, but here we can try to describe the outlook.

The aim is to discuss with some detail the connections between the algebraic approach, based on modal algebras, and the relational approach, based on frames, to the semantics of propositional normal modal logic. The study of such connections has been considered, by J. van Benthem [6], one of the three pillars of modal wisdom, and called by him duality theory (the other two being completeness and correspondence theory).

A posteriori, the first and fundamental result in duality theory is Jónsson–Tarski representation theorem for modal algebras [19], which was substantially improved by Halmos [20], who implicitly introduced categories. However, after

Kripke's fundamental work [23], modal algebras almost disappeared (the only exception is Lemmon [24]), until the discovery a few years later, that is roughly twelve years ago, of the incompleteness of Kripke's relational semantics. Then a wider notion of frame was introduced by S.K. Thomason [33], which was probably inspired by modal algebras and which does not suffer incompleteness, and a global mathematical study of modal logic was undertaken (see the collective review [9]). In particular, duality theory came into existence (see [11], [15], [26] and [35]).

The main novelty here is that we add a topology on any frame and extend the functors to the category of all frames. The technical outcome is a new categorical adjunction. More generally, beside a unified exposition of duality theory, a deeper understanding of the subject is thus achieved. Some old notions on frames, for instance that of p-morphism (here called contraction), acquire a new more natural definition, while some new ones, for instance that of weak subframe, are called into existence in a natural way. Also, applications to modal logic are not lacking. Some new results (for instance, the fact that any frame is equivalent to a refined frame also with respect to consequence) and simpler proofs of important theorems (see [30], [31] and Section III 4 below) have already been obtained, some other are in preparation; we are certain that the interested reader will find many more (our suggestions and hints to possible new roads are given in the footnotes).

Chapter I is a re-view of modal semantics, together with a discussion of ideas on which duality theory is based and their interplay with completeness theory. Chapter II contains the technical development of duality theory, with complete proofs. It can be read independently as a piece of pure mathematics.[1] Chapter III is an example of how duality theory can be applied. Using weak subframes, we obtain as corollaries both a simple description of the structure of, and the standard results (by Goldblatt and Thomason [16] and by van Benthem [3]) on, modal axiomatic classes of frames.

Even if we tried to keep an easily accessible language, and sometime gave proofs also of standard results, Stone's representation theorem for boolean algebras (an excellent exposition is in Chapter 1 of [8]) and very little of universal algebra, topology, category theory and modal logic are assumed to be known. In any case, standard references will be [10], [17] and [25].

We have chosen to suppress usual headings, like theorem, proof, remark, definition etc., in the hope that this can help avoiding fragmentation of the text. For the same reason, we have put little effort in separating new from known results and in giving credits, also because many of them have been deeply revised. Some important results have been given a name. When a word is italicized, its definition, sometimes implicit, can be found nearby (of course, this does *not* mean that all italicized words have been defined).

This paper is the first published outcome of a long though discontinuous work, which I, G.S., began a few years ago at the suggestion of Roberto Magari. My gratitude and indebtness to him certainly go far beyond the formality of any acknowledgement; but since this is one, I want to recall that some of the ideas used in Chapter II are due to him. I also like to thank personally Wolfgang Rautenberg, who encouraged and helped me at an early stage of the work and who let a first draft become public, [31]. V.V. joined in the project soon after and worked in particular on the topics now in Chapter III. Finally, we thank Johan van Benthem, Per Martin-Löf, Mario Servi and Silvio Valentini (they know why), and impersonally all the authors cited in this foreword (they also will know why, after reading).

CHAPTER I. Basic ideas and results

Introduction

This chapter is an overview on algebraic and relational semantics from the point of view of duality theory. Motivations and, where possible, intuitive explanations are given, leaving the mathematical unfolding for the remaining chapters.

We first briefly review notation and terminology. We use the propositional modal language L_M, containing connectives $\vee$, &, $\neg$ and $\Box$, the symbol $\bot$ (falsum) and propositional variables p, $q, \ldots$ The set of formulae FL_M in L_M is defined inductively as usual, including $\bot$ as an atomic formula. φ, $\psi, \ldots$ will be formulae. $\Diamond\varphi$ and $\varphi \rightarrow \psi$ are defined as usual by $\neg\Box\neg\varphi$ and $\neg\varphi \vee \psi$ respectively. A logic is here a set of formulae L containing classical tautologies and closed under the rules of Modus Ponens and Substitution

$$\text{SR:} \quad \frac{\varphi(p)}{\varphi(\psi)}.$$

We deal only with normal logics, i.e., logics containing the 'normality formula'

$$\text{NF:} \quad \Box(p \rightarrow q) \rightarrow (\Box p \rightarrow \Box q)$$

and closed under the necessitation (or 'normality') rule

$$\text{NR:} \quad \frac{\varphi}{\Box\varphi},$$

even if much of the algebraic approach extends easily to a wider class of logics. So, from now on, a logic is understood to be normal (and modal). As usual, K denotes the minimal logic.

Finally, we will use $\varphi \in L$ and $\vdash_L$ or $L \vdash \varphi$ as synonyms, thus assuming that $\vdash_L$ denotes an axiomatic system in which exactly all formulae of L are derivable.

1. Modal algebras

Respecting historical development, we start with algebraic semantics.

Given a logic L, we define an equivalence relation $\equiv_L$ on FL_M by identifying two formulae which cannot be distinguished by $\vdash_L$, i.e., by putting:

$$\varphi \equiv_L \psi \quad \text{iff} \quad L \vdash \varphi \leftrightarrow \psi.$$

The quotient set $FL_M/\equiv_L$ can then be enriched with boolean operations $+$, $\cdot$, v, 0 and 1 (sum, product, complementation, zero and one respectively) which, assuming $[\varphi]_L$ denotes $\{\psi \in L_M : \psi \equiv_L \varphi\}$, are defined by:

$$[\varphi]_L + [\psi]_L = [\varphi \vee \psi]_L,$$
$$[\varphi]_L \cdot [\psi]_L = [\varphi \,\&\, \psi]_L,$$
$$v[\varphi]_L = [\neg\varphi]_L,$$
$$0 = [\bot]_L, \qquad 1 = [\neg\bot]_L.$$

Such a definition is justified by the rule of replacement of equivalents for classical logic. Usual properties of classical propositional calculus are then expressed in the fact that $(FL_M/\equiv_L, +, \cdot, v, 0, 1)$ is a boolean algebra.

To express algebraically also the modal part of L, we can treat the modal connective $\Box$ similarly. So we define a unary operation τ on $FL_M/\equiv_L$ by:

$$\tau[\varphi]_L = [\Box\varphi]_L$$

which is a good definition when L is closed under the rule of replacement of equivalents

$$\text{RE:} \quad \frac{\varphi \leftrightarrow \psi}{\Box\varphi \leftrightarrow \Box\psi}.$$

This is certainly true when L is normal. The resulting algebra $A_L = (FL_M/\equiv_L, +, \cdot, v, 0, 1, \tau)$ is called the Lindenbaum algebra of the (modal) logic L. We will soon see that A_L offers all what can be said on L algebraically, in particular the completeness of algebraic semantics. Let us then introduce algebraic models of L. The easiest way is to look at them as models, in the classical model-theoretic sense, for a specific first-order theory related to L. So let L_{MA} be a first-order language with equality $=$, with no other predicate symbol and with function symbols $+$, $\cdot$, v, 0, 1 and τ (we use the same symbols for symbols and their interpretation). A structure adequate for L_{MA} is an algebra of similarity type $(2, 2, 1, 0, 0, 1)$. For any term t of L_{MA}, t^A is its interpretation in an algebra A (of the correct type) and for any assignment $\bar{a} = (a_0, a_1, \ldots)$ of variables in A, $t^A(\bar{a})$ is the value of t^A calculated in $\bar{a}$ (but we often omit the superscript). Then, as in the usual tarskian definition of truth, we say that the formula $t = u$ (t, u terms) is true in A on the assignment $\bar{a}$ if $t^A(\bar{a}) = u^A(\bar{a})$ and that it is valid in A if it is true on every assignment.

So far, the only difference from usual model theory is purely linguistic: as a habit, atomic formulae $t = u$ are called equations and equations true in A are called identities of A. The real difference is instead the underlying idea: we let every *formula* φ of the modal language L_M correspond biunivocally to a *term* φ^t of the language L_{MA}, and thus interpret φ in an algebra A through the aid of φ^t. Such correspondence is the most natural one, namely, φ^t is obtained from φ by replacing propositional variables $p_0, p_1, \ldots$ with individual variables $x_0, x_1, \ldots$ of L_{MA}, connectives $\vee$, &, $\neg$ with boolean operations $+$, $\cdot$, v respectively, formulae $\bot$, $\neg\bot$ with constants 0, 1, and the modal operator $\Box$ with the unary function τ. Now first note that the same sequence $\bar{a}$ of elements of an algebra A can be thought as an assignment in A both to propositional variables and to individual variables. So we can define the value of a formula φ in A on the assignment $\bar{a}$ to be the value of the corresponding term φ^t in A on $\bar{a}$, and say that φ is *true in A on the assignment* $\bar{a}$ if its value is 1, i.e., if $\varphi^t(\bar{a}) = 1$ holds in A. Accordingly, φ is said to be *valid in A* if $\varphi^t = 1$ is an identity of A. Finally, A is said to be an algebraic model of L, or briefly an *L-modal algebra*, if every formula of L is valid in A. In other words, we transform L into the set of equations $L^e = \{\varphi^t = 1 : \varphi \in L\}$ and we say that A is a model of L if it is a (classical) model of L^e. This can also be expressed by the inclusion $L^e \subseteq \mathrm{Id}(A)$, if $\mathrm{Id}(A)$ denotes the set of identities of A.

Let us give a closer look at L-modal algebras. For any algebra A, let $L(A)$ be the set of formulae valid in A; $L(A)$ is sometimes called the *logic* of A. Note that $L(A)^e = \mathrm{Id}^1(A)$, where $\mathrm{Id}^1(A)$ is the set of equations of the form $t = 1$, for some term t. (Note however that all equations in $\mathrm{Id}(A)$ are derivable from those in $\mathrm{Id}^1(A)$, since an equation $t = u$ is equivalent, in a boolean algebra, to $t \leftrightarrow u = 1$, where we put, as usual, $t \rightarrow u = vt + u$ and $t \leftrightarrow u = (t \rightarrow u) \cdot (u \rightarrow t)$.) It is easy to see that $L(A)$ contains all classical tautologies iff A is a boolean algebra and that $L(A)$ is a logic iff, moreover, $\tau 1 = 1$ and $\tau(x \rightarrow y) \rightarrow (\tau x \rightarrow \tau y) = 1$ are identities of A (to see the latter, use $\tau 1 = 1$ and properties of $=$). Therefore, since for a boolean algebra the pair of equations $\tau 1 = 1$ and $\tau(x \rightarrow y) \rightarrow (\tau x \rightarrow \tau y) = 1$ is equivalent to the pair $\tau 1 = 1$ and $\tau(x \cdot y) = \tau x \cdot \tau y$ (which is easily proved with the aid of some tautologies), we see that our definition of K-modal algebras, which are called *modal algebras* for short, is equivalent to the more standard one:

Definition. A modal algebra is a pair $A = (A, \tau)$ where A is a boolean algebra and τ is a unary operation on A such that $\tau 1 = 1$ and $\tau(x \cdot y) = \tau x \cdot \tau y$ are identities.

(To save words, here and in the whole paper, we follow the convention of denoting by $A, B, C, \ldots$ both the domain (or universe, carrier, ...) of a boolean algebra and the boolean algebra itself.) It is now obvious that for instance $S4$-modal algebras, usually called closure or topological boolean algebras, can be defined as modal algebras in which the equations $\tau x \rightarrow x = 1$ and $\tau x \rightarrow \tau\tau x = 1$

hold, GL-modal algebras, called diagonalizable algebras, the equation $\tau(\tau x \to x) \to \tau x = 1$, etc. (but note that often an order $\leq$ is defined by putting $x \leq y$ iff $x \cdot y = x$ and therefore the equation $\tau x \to \tau\tau x = 1$, for instance, is equivalently expressed by $\tau x \leq \tau\tau x$).

By its very definition, the class of L-modal algebras is an equational class (cf. [17, pp. 152 and 171]); we call it $\mathrm{MA}(L)$. Actually, what we have seen above shows also that the lattice of equational classes of modal algebras is (anti-) isomorphic to the lattice of logics. The logic corresponding to a given equational class of algebras K is $L(K)$, the set of formulae valid in every algebra of K; note that here too $L(K)^e = \mathrm{Id}^1(K)$, where $\mathrm{Id}^1(K)$ has the now obvious meaning.

Completeness of algebraic semantics is now the statement[2]:

(1) for every logic L and every formula φ, $L \vdash \varphi$ iff φ is valid in every L-modal algebra

which is equivalent to a more algebraic version

(2) for every logic L, $L^e = \mathrm{Id}^1(\mathrm{MA}(L))$.

This is quite easily achieved, as promised, through A_L. Since $L^e \subseteq \mathrm{Id}^1(\mathrm{MA}(L))$ by the definition of $\mathrm{MA}(L)$, (2) is proved once we show that

(3) for every logic L, $L^e = \mathrm{Id}^1(A_L)$.

In fact, one inclusion, $L^e \subseteq \mathrm{Id}^1(A_L)$, tells that A_L itself is an L-modal algebra, from which $\mathrm{Id}^1(\mathrm{MA}(L)) \subseteq \mathrm{Id}^1(A_L)$, while the other inclusion closes the chain, thus obtaining, beside (2), also the important by-product $\mathrm{Id}^1(A_L) = \mathrm{Id}^1(\mathrm{MA}(L))$.

Proof of (3). By definition of L^e, it is enough to prove $\varphi \in L$ iff $\varphi^t = 1 \in \mathrm{Id}^1(A_L)$. Note however that to prove $\varphi^t = 1$ to be an identity of A_L we have to show that $\varphi^t(\bar{a}) = 1$ for every assignment $\bar{a}$ in A_L, while what we know from the definition of A_L is that

(4) $\varphi \in L$ iff $[\varphi]_L = 1$ holds in A_L.

Actually, we can easily prove by induction, using the definition of operations of A_L, that $[\varphi]_L = \varphi^t(([p_0]_L, [p_1]_L, \ldots))$ and therefore $\varphi \in L$ iff $\varphi^t = 1$ is true on the assignment $([p_0]_L, [p_1]_L, \ldots)$. More generally, we can also prove that

(5) for all formulae $\psi_0, \psi_1, \ldots, \psi_n$, $[\varphi(\psi_0, \ldots, \psi_n)]_L = \varphi^t(([\psi_0]_L, [\psi_1]_L, \ldots))$

where $\varphi(\psi_0, \ldots, \psi_n)$ is the result of substituting $\psi_0, \ldots, \psi_n$ for $p_0, \ldots, p_n$ respectively in φ. Now the point is that L is closed under the rule SR, so that, if all propositional variables of φ are among $p_0, \ldots, p_n$, $\varphi \in L$ iff $\varphi(\psi_0, \ldots, \psi_n) \in L$ for all formulae $\psi_0, \ldots, \psi_n$. But then, since (4) holds, (3) is proved.

From the above proof of (3) we can also obtain an algebraic characterization of A_L. In fact, let A be any L-modal algebra and let f be any function from the set

$\{[p_i]_L : i \in N\}$ to A. Then, since $\{[p_i]_L : i \in N\}$ generates the whole A_L, putting $h([\varphi]_L) = (\varphi^t)^A((f[p_0]_L, f[p_1]_L, \ldots))$ defines a function h from the whole A_L to A; obviously, h extends f and moreover, since (5) holds, h is a homomorphism (to see this, use (5) with $\varphi = p_0 \vee p_1, \ldots, \tau p_0$). This property of A_L is exactly the definition of the fact that A_L is a free algebra, with free generators $[p_0]_L, [p_1]_L, \ldots$, over the equational class MA(L) (cf. [17, p. 162]).

Also, since the free algebras over the same class, with a set of free generators of a given cardinality, are all isomorphic, we can say that A_L is *the* free algebra over MA(L) on ω generators. Usually the notation $F_{MA(L)}(\omega)$ is used for such free algebra, but since we will use F for frames, we can here use the notation $F_L(\omega)^*$, the $*$ having a precise meaning which will be clear in Section 3.

2. Kripke frames

Relational semantics is based on the notion of *Kripke frame,* that is, a pair (X, r) where X is a set, usually considered as a universe of possible worlds, and r is a binary relation on X, usually considered as the relation of accessibility between worlds. To obtain an interpretation of modal formulae in (X, r), one must first assign to each world x the set of atomic formulae $p_0, p_1, \ldots$ which are assumed to be accepted by x as true (and it is understood that no world accepts $\perp$). Any such assignment, here called *valuation* and denoted by V, is then extended to all formulae by requiring that:

(i) The theory of a world x, i.e., the set T_x of all formulae accepted by x, must preserve the usual classical truth conditions for connectives (so that, for instance, $\varphi \vee \psi \in T_x$ iff $\varphi \in T_x$ or $\psi \in T_x$, but also $\neg\varphi \in T_x$ iff $\varphi \notin T_x$, for formulae φ, ψ) and must be closed under Modus Ponens.

(ii) A world x must accept a formula $\Box\varphi$ iff any world accessible from x accepts φ, that is, $\Box\varphi \in T_x$ iff $\varphi \in T_y$ for every y such that xry.

Such requirements have a unique solution, namely the usual inductive definition of the relation $x \Vdash_V \varphi$, to be read "x accepts φ on the valuation V". From a technical point of view, V is then a function which associates with each world x the set of formulae $T_x = \{\varphi \in FL_M : x \Vdash_V \varphi\}$. A triple (X, r, V) is called a *Kripke model* based on (X, r) and a formula φ is said to be *true in* it, written $(X, r, V) \vDash \varphi$, if $x \Vdash_V \varphi$ for every $x \in X$. And φ is said to be *valid in* (X, r), written $(X, r) \vDash \varphi$, if φ is true in every model based on (X, r).

It is well known that twelve years ago the hope of proving completeness of such semantics, which is expressed by the statement

(1) for every logic L, $\vdash_L \varphi$ iff φ is valid in every Kripke frame in which every formula of L is valid

has been shown to be badly founded by K. Fine [12] and S.K. Thomason [33]. In fact (1) fails for uncountably many logics, and some of them are quite simple (cf. [5], [8]).

The next problem was to characterize complete logics, i.e., logics satisfying (1), and this hope too now seems to be unreachable, as the work of J. van Benthem has shown. Waiting for a solution, Kripke frames are still used, both because they are simple and because they provide with a sensible interpretation of modal logic[3]. Their inadequacy is probably due to the lack of clarity of the notion of power set, which is implicitly used to define valuations; or, at least, changing that notion completeness is gained, as we now see. It is well known that this happens also for classical second-order logic.

3. Frames and their completeness

The concept of (general) frame arises naturally when looking at a Kripke frame not as a universe of worlds, each with its theory, but as a field of possible values, as we now explain.

Let (X, r) be a Kripke frame, V a valuation on it and $\Vdash_V$ the relation, generated by V, binding points with formulae (let us use the word point, instead of world, for elements of a frame). What we did in the preceding section was to think of $\Vdash_V$ as a collection of theories of formulae, each theory being associated with a point. Here we suggest to think of the same V and $\Vdash_V$ as a collection of sets of points, each set containing the points which accept a formula. Technically, the given V is here a function from $\{p_0, p_1, \ldots\}$ to $P(X)$, and it is extended to a function, still denoted by V, taking each formula into the set $V(\varphi) = \{x \in X : x \Vdash \varphi\}$. $V(\varphi)$ is called the *value* of φ in (X, r) under the valuation V and $T_V = \{V(\varphi) : \varphi \in FL_M\}$ is called the *field of possible values* of (X, r, V). When no valuation is given, we might say that the field of possible values is the whole $P(X)$. Of course, T_V and $P(X)$ are closed under the set-theoretic operations of union $\cup$, intersection $\cap$, complementation $-$, which is exactly what is needed to be able to find the value of compound formulae with principal signs $\vee$, $\&$, $\neg$ respectively, once the values of the components are given. Now we also want to be able to find the value of $\Box\varphi$, once the value of φ is given. What we need is then an operation, call it r^*, satisfying $r^*(V(\varphi)) = V(\Box\varphi)$ for every φ and V, which amounts to

(1) for every $C \in P(X)$, $r^*C = \{x \in X : \text{for every } y, xry \text{ implies } y \in C\}$.

So we take (1) as the definition of r^* and add it to the boolean algebra $P(X)$ (here and in the sequel, we do not indicate the usual boolean operations): what we obtain is a modal algebra, since obviously $V(\Box\neg\bot) = V(\neg\bot) = X$ and $V(\Box(\varphi \,\&\, \psi)) = V(\Box\varphi \,\&\, \Box\psi)$ for all formulae φ, ψ and every valuation V. For the same reasons, (T_V, r^*) too is a modal algebra, subalgebra of $(P(X), r^*)$. We can now generalize both situations by considering Kripke frames together with a field of possible values, which, as we have seen, must be a modal algebra.

Definition. A *frame* is a triple $F = (X, r, T)$ where (X, r) is a Kripke frame and T is a field of subsets of X closed under the operation r^* defined by (1).

T is called the field (of possible values) of F. To save words, we assume from now on that F denotes the frame (X, r, T) and G the frame (Y, s, U). The value of a formula in a frame F is obtained as before, except that only valuations in the field of F are considered. We also keep the same notation. So a Kripke frame (X, r) can be identified with the frame $(X, r, P(X))$. By what we have seen above, it is immediate that

(2) for every frame F, (T, r^*) is a modal algebra.

We call it the *dual* of F and denote it by F^*.

F and F^* are strongly tied together or, better, are two technical ways of looking at the same thing, namely valuations. In fact, we can easily see, or prove by induction, that

(3) for every formula φ, every frame F and every valuation V on F,
$V(\varphi) = \{x \in X : x \Vdash_V \varphi\} = (\varphi^t)^{F^*}((V(p_0), V(p_1), \ldots))$

i.e., the value of φ in the frame F coincides with the value of φ in the modal algebra F^* (recall that the same V can be seen both as a valuation on F and an assignment on F^*)[4].

An immediate consequence of (3) is that for every frame F, F and F^* validate the same formulae; in other words, putting $LF = \{\varphi \in FL_M : F \vDash \varphi\}$ (the *logic* of F) and recalling that $LF^* = \{\varphi \in FL_M : \varphi^t = 1 \in \mathrm{Id}(F^*)\}$,

(4) for every frame F, $LF = LF^*$.

The completeness of the semantics given by frames is now at hand: it is enough to construct a single frame à la Henkin. However, since such a construction will be used repeatedly in the sequel, we analyse it in some detail. We can isolate two preliminary steps:

(i) Construction of the model $M_L = (X_L, r_L, V_L)$ where: X_L is the set of all maximal consistent sets of formulae containing L; for every $S, T \in X_L$, Sr_LT iff for every formula φ, $\Box\varphi \in S$ implies $\varphi \in T$; $V_L(p_i) = \{S \in X_L : p_i \in S\}$.

M_L is called the canonical model and (X_L, r_L) the *canonical Kripke frame for* L.

(ii) Proof, by induction, of:

(5) for every formula φ and every $S \in X_L$, $S \Vdash_V \varphi$ implies $\varphi \in S$.

Every step of the induction is straightforward except the inductive step for $\Box$, where we need

(6) $\Box\varphi \in S$ iff for every $T \in X_L$, Sr_LT implies $\varphi \in T$.

To prove the non-trivial direction (from right to left), assume $\Box\varphi \notin S$. Then

$\varphi \notin \Box^{-1}S = \{\psi : \Box\psi \in S\}$ and, since $\Box^{-1}S$ is closed under Modus Ponens, there is a maximal set T' containing $\Box^{-1}S$ but not φ. This means $Sr_L T'$, but also $\varphi \notin T$, as we wanted.

From (5) we immediately derive

(7) $L \vdash \varphi$ iff $M_L \vDash \varphi$

since $\varphi \in L$ iff $\varphi \in S$ for every $S \in X_L$. (7) is sometimes called 'the fundamental theorem (of modal logic)'; it expresses completeness of the semantics given by models.

Also completeness for frames is now easily derivable. Let us say that a frame F is *a frame for L* if $F \vDash L$ (that is, $F \vDash \varphi$ for every $\varphi \in L$) or, equivalently, if $L \subseteq LF$. Then we want to prove

(8) (Completeness theorem) for every logic L, $L \vdash \varphi$ iff $F \vDash \varphi$ for every frame F for L.

As we did at the beginning of this section, we consider the frame $F_L = (X_L, r_L, T_L)$ generated by V_L over (X_L, r_L); in other words the field of possible values T_L is the collection $\{V_L(\varphi) : \varphi \in FL_M\}$ of all values actually taken by formulae on V_L. F_L is here called the *universal* (general) *frame for L*. Note that the canonical (Kripke) frame is obtained from F_L simply by dropping T_L.

In analogy with the case of modal algebras and following standard proofs of completeness for equational logic, we obtain (8) as a corollary of

(9) for every logic L, $L = L(F_L)$

which is obtained from (7) almost exactly as (3) of Section 1 was obtained from (4). Here again closure of $L(M_L)$, which is equal to L by (7), under the substitution rule is essential. In fact, note that for any valuation V on T_L, there is a sequence of formulae $\psi_1, \psi_2, \ldots$ such that $V(p_i) = V_L(\psi_i)$ for every i (a proof by induction is straightforward); so $V(\varphi(p_1, p_2, \ldots)) = V_L(\varphi(\psi_1, \psi_2, \ldots))$ for every formula φ. But then $\varphi \in L$ iff $\varphi(\psi_1, \psi_2, \ldots) \in L$ iff for each sequence of formulae $\psi_1, \psi_2, \ldots, V_L(\varphi(\psi_1, \psi_2, \ldots)) = X_L$ iff for every valuation V on T_L, $V(\varphi(p_1, p_2, \ldots) = X_L$ iff $F_L \vDash \varphi$.[5]

As for the free algebra A_L, the construction of F_L gives some good suggestions. One of these is the following: since a maximally consistent set of formulae can not separate two formulae φ, ψ if $\varphi \equiv_L \psi$, we can identify it with an ultrafilter of A_L (which we do also as far as notation is concerned). Then, simply translating the construction of F_L in algebraic language, we define $Sr_L T$ (S, T ultrafilters) to hold iff for every $a \in A_L$, $\tau a \in S$ implies $a \in T$ and take T_L to be the field of all subsets of the form $\{S \in X_L : a \in S\}$, when $a \in A_L$. What do we achieve? The circle, or rather the diagram, is closed, since we can easily prove:

(10) the dual (T_L, r_L^*) of the frame (X_L, r_L, T_L) is isomorphic to A_L.

The proof of (10) will follow the lines of that for (9) above, and thus no wonder that we can obtain completeness of frames (8) directly out of it: since $L = L(A_L)$ and $A_L \cong F_L^*$, $L = L(F_L^*)$ and hence also $L = L(F_L)$ by (4).

More important, however is another idea: why don't we repeat the same construction starting from an arbitrary modal algebra, instead of A_L? Actually, this is what we are going to do in the next section.

4. From modal algebras to frames

As we said above, the aim of this section is to provide a construction which shows that any modal algebra A is (isomorphic to) the modal algebra dual of some frame. We will obtain this by constructing a frame A_*, called the dual of A, such that $A \cong (A_*)^*$. Such a construction may appear more natural if we look at it backwards, that is postulate that we already know, given A, how to construct A_* and examine how it could be.

Recall that $(A_*)^*$, the dual of A_*, is simply the field of possible values over A_*, together with the operation corresponding to $\Box$. So the first step is to think of elements of A as possible values, and τ the additional operation. We then have to fill in with points (of A_*) every element of A. But how can we 'create' points of A_*? Here is the crucial point of the construction.

Note that, in any frame F, with each point x we can associate, in analogy with the complete theory of formulae T_x in Section 2, a complete theory of possible values, namely $U_x = \{C \in T : x \in C\}$. In mathematical words, U_x is an ultrafilter of the boolean algebra T. So with each point (still to be 'created') of A_* is associated an ultrafilter of the boolean algebra A. Now the idea is simply to reverse this, that is define points of A_* to *be* the ultrafilters of A. So $U(A) = \{S : S$ is an ultrafilter of $A\}$ is the domain of the frame A_*. It is then clear, after the above heuristic discussion, that an ultrafilter S, point of A_*, will belong to the possible value a (or better, to the possible value in $(A_*)^*$ corresponding to a) if the theory S holds a true, i.e., if $a \in S$. Therefore the isomorphism between A and $(A_*)^*$ must be the function $\beta : a \mapsto \{S \in U(A) : a \in S\}$. Completing the construction and checking that β is in fact an isomorphism is now easier. When is a theory T accessible from another theory S? Since we want S to hold the value τa iff all T's accessible from S hold the value a, again reversing things we choose the maximal relation compatible with this, namely the relation τ_* defined by

$$S \tau_* T \quad \text{iff} \quad \text{for every } a \in A, \quad \tau a \in S \text{ implies } a \in T.$$

And finally, as we said, the field of values will be A itself, but in a disguised form now: the place of an element a is taken by the set $\beta a = \{S \in U(A) : a \in S\}$ of all complete theories holding a (alias ultrafilters containing a).

Summing up, the *dual* A_* of A is the structure $(U(A), \tau_*, \beta A)$, where $\beta A = \{\beta a : a \in A\}$.[6] Proving that A_* is a frame is not trivial; actually, the fact that

A is closed under the operation $(\tau_*)^*$ will be proved by showing that

(1) $(\tau_*)^*\beta a = \beta(\tau a)$ for each a in A

which is also the only new step required to reach our aim, namely

(2) (Jónsson–Tarski representation theorem) every modal algebra A is isomorphic to its bidual $(A_*)^*$.

In fact, the reader who knows Stone's representation theorem for boolean algebras will have already noticed that, if $A = (A, \tau)$ with A a boolean algebra, then $(U(A), \beta A)$ is the Stone space dual of A, and βA is a boolean algebra isomorphic to A, β being the isomorphism. As we already remarked, a proof of (1) is very similar to the proof of (6) of Section 3.

An immediate consequence of (2) is that $LA = L(A_*)^*$ and therefore, by (3.4), also

(3) for every modal algebra A, $LA = LA_*$.

So, for any modal algebra there is an equivalent frame, and conversely. In particular

(4) for every frame F, $LF = L(F^*)_*$

so that any frame is equivalent to its bidual. Note however that we have not derived (4) from an analogue of (2) for frames, simply because it is *not* true that any frame F is isomorphic to its bidual $(F^*)_*$. This is due to the fact that bidual frames have a rather rich structure, which will be described in Chapter II. The frames with such a structure, that is isomorphic to the bidual of some frame, have been called descriptive by Goldblatt. They are the only frames for which an analogue of (2) can be proved, namely

(5) a frame F is isomorphic to its bidual $(F^*)_*$ iff F is descriptive

that is, iff F itself is isomorphic to the bidual of some frame. This follows easily from $F^* \cong ((F^*)_*)^*$, which is an instance of (2).

So, from a mathematical point of view, if we want a duality, in the sense of the category theory, to hold between modal algebras and frames, we must restrict to descriptive frames. From the point of view of logic, such a restriction is harmless as long as we are interested only in questions of completeness, in view of (3) or (4) above. However, it is philosophically debatable if such a restriction is justified. Moreover, we will show (Section III.2) that as soon as we extend our interest from validity of formulae to semantical consequence, descriptive frames are no longer enough. Finally, when one is working concretely with frames, it is much simpler to use all of them without bothering if they are descriptive or not. This is why we have chosen to keep on considering the class of all frames, also when looking at them as a category.

CHAPTER II. Duality theory

Introduction

We are now ready to begin, following the ideas discussed so far, and thus saving comments, the technical development of duality theory. The main known result, explicitly stated in [15] for the first time, is that the category **Mal**, of modal algebras and homomorphisms, and the category **DFra**, of descriptive frames and contractions (alias p-morphisms), are dual to each other (precise definitions are given below). A similar duality theorem is well known for boolean algebras (cf. [21]). In this case, the dual category is that of boolean spaces and continuous functions, where a boolean space is a topological space which is compact, Hausdorff and with a base of open and closed sets. A frame $A_* = (U(A), \tau_*, \beta A)$, dual of the modal algebra A, is such a space when we forget τ_*, and βA is the base for its topology. The idea here is to extend this to *all* frames, that is, to take the field T of a frame $F = (X, r, T)$ as the base of a topology on X. We thus will have a category **Fra** of all frames and suitable morphisms (weak contractions), of which **DFra** is a full subcategory, and functors between **Mal** and the whole **Fra**. We will show that such functors form an adjunction between **Fra** and **Mal** whose restriction to **DFra** will give the desired duality. Following an idea of Halmos, we think of modal algebras and frames as particular arrows in two 'bigger' categories. We can thus prove a general result (basic adjunction) which includes all the above as particular cases.

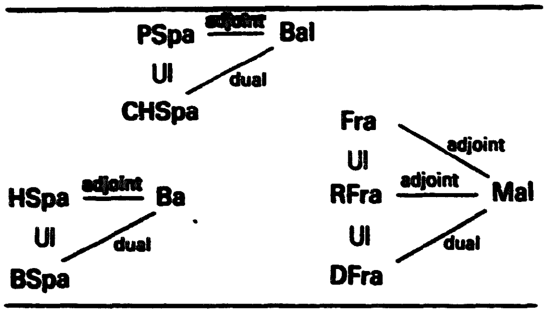
Two tables may help the reader: Table 1 informally summarizes the definitions of the categories to be introduced while Table 2 indicates the various categorical connections to be established.

The exposition will be detailed enough to avoid references to other sources and repetitions in later chapters (but the reader with little interest in adjunctions can skip Section 3 and most of Section 4, and instead follow the instructions given in footnote 7). Of course also in this purely technical chapter the reader will often see the relevance to modal semantics of some mathematical results, even if we do not explicitly mention it.

Table 1. Categories (in order of appearance).

Bal = boolean algebras A + hemimorphisms τ ($\tau 1 = 1$, $\tau(a \cdot b) = \tau a \cdot \tau b$)
Mal = modal algebras A (i.e., $A \xrightarrow{\tau} A$ in **Bal**) + homomorphisms h
Spa = spaces X + continuous relations ($r^{-1}U \subseteq T$)
Fra = frames F (i.e., $X \xrightarrow{r} X$ in **Spa**) + weak contractions c ($\overline{crx} = \overline{scx}$)
PSpa = spaces + point-closed continuous relations with composition $\star$
Ba = boolean algebras + homomorphisms
HSpa = zero-dimensional Hausdorff spaces + continuous functions
BSpa = boolean (i.e., compact Hausdorff) spaces + continuous functions
CHSpa = boolean spaces + point-closed continuous relations
RFra = refined frames (i.e., $X \xrightarrow{r} X$ with r point-closed, X Hausdorff) + weak contractions
DFra = descriptive frames (i.e., $X \xrightarrow{r} X$ in **CHSpa**) + contractions c ($crx = scx$)

Table 2



1. Introducing categories and topology

To simplify matters, we follow an idea of Halmos [20] and think of the category **Mal**, of modal algebras and homomorphisms, as obtained from a bigger auxiliary category **Bal**, of boolean algebras and hemimorphisms, where a *hemimorphism* from A to B is a function τ satisfying

$$\tau 1_A = 1_B$$

and

$$\text{for every } a, b \in A, \quad \tau(a \cdot_A b) = \tau a \cdot_B \tau b.$$

It is then clear that any modal algebra A is nothing but a pair (A, τ) where A is an object of **Bal** and τ a hemimorphism from A into itself. So A can be identified with a diagram of the form $A \xrightarrow{\tau} A$ in **Bal** (a rigorous definition of this identification is possible, but apparently useless, with the aid of the category of morphisms of **Bal**). Thus the notion of hemimorphism will allow us to treat at the same time operators τ on boolean algebras and homomorphisms between them.

We now want to do the same for frames, that is consider a frame $F = (X, r, T)$ as an arrow $(X, T) \xrightarrow{r} (X, T)$ in a bigger category. Thus objects will simply be the pairs (X, T), where X is a set and T is a subalgebra of $P(X)$; we call them *spaces*. A morphism from (X, T) to (Y, U) will be any relation $r \subseteq X \times Y$ satisfying the condition imposed on accessibility relations, but on any pair of spaces. Namely, for every $D \subseteq Y$ we put

$$r^*D = \langle x \in X : \text{for every } y \in Y, xry \text{ implies } y \in D \rangle$$

and say that r is a morphism from (X, T) to (Y, U) if

(1) for every $D \in U$, $\quad r^*D \in T$.

Just like that of hemimorphism, this definition will permit to treat at the same time accessibility relations on frames and morphisms between them.

Any space (X, T) is here meant to be endowed with the topology generated by taking T as a base for open subsets. Since T is closed under complements, T is then also a base for closed subsets and any set in T is both closed and open, alias *clopen*. Note however that T need not contain all clopen subsets of X (a typical example is when T is the family of finite and cofinite subsets of X, in which case all subsets of X are clopen), unless X is compact:

(2) if (X, T) is compact, then T coincides with the family of all clopen subsets of X.

Proof. Let D be clopen in X. Then, since D is open, $D = \bigcup C$ for some family $C \subseteq T$, and therefore, since D is closed and hence compact, $D = \bigcup C'$ for some finite subfamily $C' \subseteq C$. But then $D \in T$, because T is closed under finite unions.

So in general two different spaces may coincide if regarded as topological spaces and hence we can not forget the base; nevertheless, from now on we will write X for (X, T) and Y for (Y, U).

One should keep in mind that a morphism r from X to Y is not necessarily a function from X to Y; instead, we can think of it either as a function from $P(X)$ to $P(Y)$ by putting

(3) for every $C \subseteq X$, $rC = \{y \in Y : xry \text{ for some } x \in C\}$

or as a function from X to $P(Y)$, where the image of $x \in X$ under r is the *set* rx, short for $r\{x\}$ (and note that the case in which rx is empty is not excluded). This allows us to rewrite the definition of r^* in a simpler form:

(4) for every $D \subseteq Y$, $r^*D = \{x \in X : rx \subseteq D\}$.

The relation r^{-1}, defined by $yr^{-1}x$ iff xry, is usually called the inverse of r. However, some of the properties of inverse functions carry over to r^* rather than r^{-1}. For instance, since $rC = \bigcup_{x \in C} rx$, we immediately have:

(5) for every $C \subseteq X$ and $D \subseteq Y$, $rC \subseteq D$ iff $C \subseteq r^*D$.

In categorical terms, (5) says that r and r^* are adjoint, when considered as functions between $P(X)$ and $P(Y)$. Taking $D = rC$ and $C = r^*D$, (5) gives respectively

(6) (i) for every $C \subseteq X$, $C \subseteq r^*rC$,
(ii) for every $D \subseteq Y$, $rr^*D \subseteq D$.

Actually, when r is a function, that is rx is a singleton for every $x \in X$, the definition of r^* boils down to the usual definition of inverse. In fact in this case $rx \subseteq D$ iff $rx \cap D \neq \emptyset$, and hence $x \in r^*D$ iff $x \in r^{-1}D$. So

(7) if r is a function, then $r^* = r^{-1}$.

In general, a similar argument only shows how r^* and r^{-1} are connected: $x \in r^* - D$ iff $rx \cap D = \emptyset$ iff $x \notin r^{-1}D$, that is,

(8) for every $D \subseteq Y$, $r^* - D = -r^{-1}D$.

Of course, from (8) we have $-r^{-1} - D = r^*D$ and $r^{-1}D = -r^* - D$ for every

$D \subseteq Y$, which shows that r^{-1} bears the same relation to $\Diamond$ as r^* to $\Box$. Another consequence of (8) is that, since T and U are closed under complements, the condition (1) is equivalent to

(9) for every $D \in U$, $r^{-1}D \in T$.

Since obviously

(10) for every family $\{D_i : i \in I\}$ of subsets of Y, $r^{-1}(\bigcup_{i \in I} D_i) = \bigcup_{i \in I} r^{-1} D_i$

(9) becomes the usual definition of continuity as soon as r is a function. Thus a morphism may, and will, be called a *continuous relation*. Actually, extending to relations the familiar terminology for functions, we also say that r is closed when rC is closed whenever C is closed; and similarly for r^{-1}, open, clopen, etc. So a continuous relation always has an open inverse. Note, however, that unfortunately r^{-1} may be open without r being continuous even if r is a function (consider any space (X, T) in which T does not coincide with the family C of all clopen subsets, and the identity function $(X, T) \to (X, C)$). On the other hand, r may be continuous without r^{-1} being clopen (we omit counterexamples, which however are not too difficult). All what we can say, up to now, is that r^{-1} is open (closed) iff r^* is closed (open), by (8), and hence that r^{-1} is clopen iff r^* is clopen.

The composition of continuous relations $\circ$ is the usual set theoretic composition of relations, but note that we write $s \circ r$ for $\{(x, z) : xry \text{ and } ysz, \text{ for some } y\}$ since we want the equality $(s \circ r)C = s(rC)$ to hold, for every C. It is immediate to check that the composition of two continuous relations is still continuous, and that $\{(x, x) : x \in X\}$ is the identity morphism on X. Hence spaces and continuous relations form the category we were looking for, and we call it **Spa**. However, though fairly natural, this definition has to be modified a little if we want to obtain an adjunction with **Bal**, as we will see in the next section.

Any frame $F = (X, r, T)$ will be identified with the pair $((X, T), r)$, where (X, T) is a space and r is a continuous relation from X to X. So objects of **Fra** are diagrams of the form $X \xrightarrow{r} X$ in **Spa**.

The definition of morphisms in **Fra** is a bit less immediate and can be grasped better after the introduction of functors between **Spa** and **Bal**. We can follow two different lines of thought. The first is to adopt the general more traditional pattern of defining morphisms as *functions* which preserve the structure of objects. We then obtain the notion of contraction: given two frames F and G, a function c from X to Y is called a *contraction* (following the terminology of Rautenberg [28]) if it satisfies both

(11) for every $D \in U$, $c^{-1}D \in T$

and

(12) for every $x \in X$, $crx = scx$.

Of course a function satisfying (11) is continuous, but remind that the converse is in general not true. Condition (12) might puzzle some readers, but is simply a way of saying, using our conventions by which crx denotes $c(r\{x\})$, that the two relations $s \circ c$ and $c \circ r$ are equal. Now c, r, s are morphisms in **Spa** and hence $s \circ c = c \circ r$ is exactly what categorists usually express by saying that the diagram is commutative. So a contraction c from F to G is just a continuous relation from X

$$\begin{array}{ccc} X & \xrightarrow{c} & Y \\ {\scriptstyle r}\downarrow & & \downarrow{\scriptstyle s} \\ X & \xrightarrow{c} & Y \end{array}$$

to Y which is a function and makes the above diagram commute. Note that this parallels the characterization of a homomorphism h from A to B, A, B modal algebras, as a hemimorphism from A to B which is a boolean homomorphism and makes the following diagram commutative:

$$\begin{array}{ccc} A & \xrightarrow{h} & B \\ {\scriptstyle \tau_A}\downarrow & & \downarrow{\scriptstyle \tau_B} \\ A & \xrightarrow{h} & B \end{array}$$

A closer look shows that (12) is only a new dress for a well known requirement. In fact, simply by writing out the meaning of $crx = scx$, we see that for every $x \in X$, $crx \subseteq scx$ is equivalent to

(13) for every $y \in X$, xry implies $(cx)s(cy)$

and that $scx \subseteq crx$ is equivalent to

(14) for every $z \in Y$, if $(cx)sz$ then for some $y \in X$, xry and $cy = z$.

So (12) is equivalent to (13) and (14) together, which are traditionally used to define p-morphisms. Note that (12) is equivalent also to

(15) for every $z \in Y$, $r^{-1}c^{-1}z = c^{-1}s^{-1}z$

because $r^{-1} \circ c^{-1} = (c \circ r)^{-1}$ and similarly for s.

The second approach is to impose on a continuous relation $c: X \rightarrow Y$ the minimal conditions in order to obtain that its image c^* under the functor $(\ )^*$, defined in the next section, is a homomorphism between modal algebras. We will see that c^* is a boolean homomorphism iff c is a function and (11) holds, while c^* preserves the additional operator iff

(16) for every $D \in U$, $r^*c^*D = c^*s^*D$.

Since $c^* = c^{-1}$ whenever c is a function, and because of (8), (16) is equivalent to

(17) for every $D \in U$, $r^{-1}c^{-1}D = c^{-1}s^{-1}D$.

So we call *weak contraction* any function c satisfying (11) and (17). The name is due to the fact that

(18) every contraction is also a weak contraction

which is true because obviously (15) implies (17). Of course, the two notions coincide if only Kripke frames are considered (and, as we will show in Section 4, this holds also for descriptive frames). This probably explains why the notion of weak contraction has never appeared before (even if one could have reached it simply noting that (17) rather than (15) is used to prove the 'p-morphism lemma', i.e., to show that c preserves validity of modal formulae, cf. Section III.1 below). On the other hand, we believe that the results to follow justify our choice of taking weak contractions, rather than contractions, as morphisms in the category **Fra**.

2. The functors

The definition of functors is quite natural, and follows the ideas presented in Chapter I. In particular, the functor $(\ )^*$ from **Spa** to **Bal** gives no problems. For every space X, we let X^* be the field T of X; so X^* is a boolean algebra by definition. The image under $(\ )^*$ of a continuous relation $r:X\to Y$ is the function $r^*:P(Y)\to P(X)$ defined in the preceding section. By the definition of continuous relations, r^* maps U into T, and it is routine to check, using (1.4), that $r^*Y=X$ and $r^*(C\cap D)=r^*C\cap r^*D$ for every $C, D\in U$. So

(1) for every $r:X\to Y$, r^* is a hemimorphism from Y^* to X^*.

Finally, given two morphisms $r:X\to Y$ and $s:Y\to Z$, for every $C\in Z^*$ and $x\in X$, $x\in r^*s^*C$ iff $rx\subseteq s^*C$ iff, by (1.5), $srx\subseteq C$ iff $x\in(s\circ r)^*C$; that is, $(s\circ r)^*=r^*\circ s^*$. We thus have

(2) $(\ )^*$ is a contravariant functor from **Spa** to **Bal**.

The functor $(\ )_*$ from **Bal** to **Spa** is obtained as an elaboration on Stone's representation theorem for boolean algebras. Let us recall that $U(A)$ is the set of all ultrafilters of a boolean algebra A, and that $\beta_A:A\to P(U(A))$, defined by

$$\text{for every } a\in A,\quad \beta_A a=\{S\in U(A):a\in S\}$$

is an isomorphism from A to the subalgebra $\beta A=\{\beta_A a:a\in A\}$ of $P(U(A))$. The space $(U(A),\beta A)$ is usually called the *Stone space* of A; it is compact and Hausdorff and hence, having by definition a base of clopen subsets, it is a boolean space. (This is all we need of Stone's representation theory; the reader can easily reconstruct the proofs or consult any standard reference, like [8] or [21].)

The image A_* of a boolean algebra A under $(\ )_*$ is thus its Stone space $(U(A),\beta A)$, which clearly is a space. The image of a hemimorphism τ from A to

B is the relation $\tau_* \subseteq B_* \times A_*$ defined by

(4) for every $S \in U(B)$ and $T \in U(A)$, $S\tau_* T$ iff for every $a \in A$, $\tau a \in S$ implies $a \in T$.

Note that if, like many authors, we had used σ as primitive, here defined by $\sigma = \nu\tau\nu$, the definition of τ_* would have been

(5) $S\tau_* T$ iff for every $a \in A$, $a \in T$ implies $\sigma a \in S$

which instead is here an easy consequence of (4).

Showing that τ_* is indeed a continuous relation is not immediate at all. Actually, it follows only from

(6) for every $\tau: A \to B$ and $a \in A$, $(\tau_*)^* \beta_A a = \beta_B \tau a$

which is essentially the key step to obtain both the Jónsson–Tarski representation theorem I.4.2 and the fundamental theorem (cf. I.3.6).

Proof. First note that, putting $\tau^{-1}S = \{b \in A : \tau b \in S\}$, we obtain that $S\tau_* T$ iff $\tau^{-1}S \subseteq T$. By the definition of $(\ )^*$ and β, $S \in (\tau_*)^* \beta_A a$ iff $\tau_* S \subseteq \beta_A a$ and $S \in \beta\tau a$ iff $\tau a \in S$. So only

(7) $\tau a \in S$ iff $\tau_* S \subseteq \beta_A a$

is left to be proved. First assume $\tau a \in S$ and let $T \in \tau_* S$; then $\tau^{-1}S \subseteq T$ and hence $a \in T$, that is $T \in \beta_A a$. Conversely, assume $\tau a \notin S$. We will show that there is an ultrafilter which belongs to $\tau_* S$ but not to βa. It is easy to check that $\tau^{-1}S$ is always a filter, because τ is a hemimorphism. Under the assumption $\tau a \in S$, we can also show that $\tau^{-1}S \cup \{\nu a\}$ has the finite intersection property. In fact, suppose that $b_1 \cdot \ldots \cdot b_n \cdot \nu a = 0$ for some $b_1, \ldots, b_n \in \tau^{-1}S$; then $b_1 \cdot \ldots \cdot b_n \leq a$, and thus $a \in \tau^{-1}S$, against the assumption. So, by the ultrafilter theorem, there is an ultrafilter T' extending $\tau^{-1}S \cup \{\nu a\}$ which means, as we wanted, both $\tau^{-1}S \subseteq T'$, that is $T' \in \tau_* S$, and $a \notin T'$, that is $T \notin \beta a$.

We now still have to prove that $(\ )_*$ preserves composition, which also is not trivial. To prove it, we call topology on the stage. Recall that for any space X, the closure $\bar{D}$ of a subset D of X is the intersection of all clopen subsets $C \in \mathbf{T}$ containing D, because $\mathbf{T}$ is also a base for closed subsets; so $\bar{D} = \bigcap \{C \in \mathbf{T} : D \subseteq C\}$ (and this is all we need on the closure operator). This implies that, for any $r: X \to Y$, $\overline{rx} = \bigcap \{D \in \mathbf{U} : x \in r^* D\}$ and hence rx is closed, that is $rx = \overline{rx}$, iff $rx = \{y \in Y : \text{for every } D \in \mathbf{U}, x \in r^* D \text{ implies } y \in D\}$. In other words,

(8) for every $x \in X$, the following are equivalent:
 (i) rx is closed;
 (ii) xry iff for every $D \in \mathbf{U}$, $x \in r^* D$ implies $y \in D$

(and again, if one prefers r^{-1} to r^*, he will use

$$xry \text{ iff for every } D \in \mathbf{U},\ y \in D \text{ implies } x \in r^{-1}D$$

which is equivalent to (ii) because of (1.5)).

Let us say that r is *point-closed* if rx is closed for every $x \in X$. The point is that, for any hemimorphism $\tau: A \to B$, τ_* satisfies (ii) above. In fact, by definition, $S\tau_* T$ iff for every $a \in A$, $\tau a \in S$ implies $a \in T$. Observe that, by (3), $a \in T$ iff $T \in \beta a$, and hence also $\tau a \in S$ iff $S \in \beta\tau a$ iff $S \in (\tau_*)^*\beta a$, by (6). So $S\tau_* T$ iff for every $\beta a \in \beta A$, $S \in (\tau_*)^*\beta a$ implies $T \in \beta a$, which is the claim since βA is a base for A_*. Therefore, by (8),

(9) for any hemimorphism τ, τ_* is a point-closed continuous relation.

(The reader aware of the correspondence between filters on A and closed subsets of A_* can get a cheaper proof of (9): since $\tau^{-1}S$ is a filter, the set $\{T \in U(A): \tau^{-1}S \subseteq T\} = \tau_* S$ is closed.)

We still need two lemmas on point-closed relations. For any $r: X \to Y$, let us define the *pointwise closure* of r as the minimal point-closed relation $\bar{r}$ containing r; so $\bar{r}x = \overline{rx}$ for every $x \in X$, that is $x\bar{r}y$ iff $y \in \overline{rx}$. For every $D \in U$, $rx \subseteq D$ iff $\overline{rx} \subseteq D$, and hence $r^*D = \bar{r}^*D$. So

(10) for every $r, s: X \to Y$, if $\bar{r} = \bar{s}$ then $r^* = s^*$

that is, r^* does not determine r univocally. However this is true if we require r to be point-closed, since by (8) we have:

(11) if $r, s: X \to Y$ are point-closed and $r^*D = s^*D$ for every $D \in U$, then $r = s$.

The second lemma is an extension to relations of a standard result on continuous functions on compact spaces (cf. [10, p. 104, Theorem 9], but note that the structure of the proof has to be changed and that we do not need the assumption that the co-domain is Hausdorff). Our proof is quite different from that in [20, p. 165], and involves less notions and assumptions.

(12) if $r: X \to Y$ is continuous and point-closed and X is compact, then r is closed.

Proof. Suppose that D is a closed subset of X. We prove that rD is closed by showing that for every $y \notin rD$ there exists a clopen subset C of Y containing rD but not y. So, let $y \notin rD$. For every $x \in D$, $y \notin rx$ and hence, since rx is closed, there is a clopen subset C_x containing rx but not y. So $x \in r^*C_x$ and $D \subseteq \bigcup_{x \in D} r^*C_x$. Since D is closed and every r^*C_x is open because r is continuous, by the compactness of X there exists a finite subset $D' \subseteq D$ such that $D \subseteq \bigcup_{x \in D'} r^*C_x$. But then also

$$rD \subseteq r\Big(\bigcup_{x \in D'} r^*C_x\Big) = \bigcup_{x \in D'} rr^*C_x \subseteq \bigcup_{x \in D'} C_x,$$

the last inclusion being true by (1.6). Since $\bigcup_{x \in D'} C_x$ is itself clopen and does not contain y, it is the clopen subset C we wanted and the proof is complete.

The use of (12) is, for the moment, merely to show that, given $\tau: A \to B$ and

$\tau': B \to C$, $\tau_* \circ \tau'_*$ is point-closed. In fact, τ_* and τ'_* are point-closed by (9), and hence also closed by (12), because C_* and B_* are compact. So $\tau_* \circ \tau'_*$ too is closed, and a fortiori point-closed. We can then finally show that $\tau_* \circ \tau'_* = (\tau' \circ \tau)_*$. Since both members are pointclosed, by (11) it is enough to show that for every $a \in A$, $((\tau' \circ \tau)_*)^* \beta a = (\tau_* \circ \tau'_*)^* \beta a$. This follows by repeated use of (6) and the fact that $(\ )^*$ preserves composition:

$$((\tau' \circ \tau)_*)^* \beta a = \beta \tau' \tau a = (\tau'_*)^* \beta \tau a = ((\tau'_*)^* \circ (\tau_*)^*) \beta a = (\tau_* \circ \tau'_*)^* \beta a.$$

We thus have completed the proof of

(13) $(\ )_*$ is a contravariant functor from **Bal** to **Spa**.

We have actually proved that we have a little more than two functors, namely

(14) β is a natural isomorphism from the identity functor $\mathrm{Id}_{\mathbf{Bal}}$ into $((\ )_*)^*$

which is just what (6) says, together with the fact that $\beta_A : A \to \beta A = (A_*)^*$ is an isomorphism for each boolean algebra A.

It should be clear that there is no reasonable similar natural isomorphism in **Spa**, because $(X^*)_*$ is a boolean space, whatever space X is. However, for every space X we can define a function $\gamma_X : X \to (X^*)_*$ by putting as usual

(15) for every $x \in X$, $\quad \gamma_X x = \{C \in \boldsymbol{T} : x \in C\}$

and we can show that γ is 'almost' the inverse of β. First note that, by definition, both β and γ reverse the membership relation, in the sense that

(16) (i) for every $a \in A$ and $S \in U(A)$, $a \in S$ iff $S \in \beta a$;
(ii) for every $x \in X$ and $C \in \boldsymbol{T}$, $x \in C$ iff $C \in \gamma x$.

So, denoting by $1_A, 1_X, \ldots$ the identity morphisms of $A, X, \ldots$, we can easily prove

(17) (Triangular identities)
(i) for every boolean algebra A, $\quad (\beta_A)_* \circ \gamma_{A_*} = 1_{A_*}$;
(ii) for every space X, $\quad (\gamma_X)^* \circ \beta_{X^*} = 1_{X^*}$.

Proof. (i) For every $a \in A$ and $S \in U(A)$, using (16) we obtain $a \in S$ iff $S \in \beta a$ iff $\beta a \in \gamma S$ iff $a \in \beta^{-1}(\gamma S)$, and it is clear that, since β is a homomorphism, $\beta_* = \beta^{-1}$ (cf. (4.2) below). So $S = \beta_* \gamma S$.

(ii) for every $x \in X$ and $C \in \boldsymbol{T}$, $x \in C$ iff $C \in \gamma x$ iff $\gamma x \in \beta C$ iff $\gamma x \subseteq \beta C$ (remind that γx, the image of x under the relation γ, is a set) iff $x \in \gamma^* \beta C$. So $C = \gamma^* \beta C$.

We will make essential use of triangular identities in the next sections. Here we derive from them some properties of γ. Since the base of $(X^*)_*$ is $\{\beta C : C \in \boldsymbol{T}\}$, from $C = \gamma^* \beta C$ we have in particular

(18) γ_x is continuous.

Since γ is a function and hence $\gamma^* = \gamma^{-1}$ by (1.7), applying γ to both sides of

$C = \gamma^*\beta C$ we obtain

(19) for every $C \in T$, $\gamma C = \beta C \cap \gamma X$

(where of course γC is just the image of the set C under γ, namely $\gamma C = \{\gamma x : x \in C\}$). So in particular γX intersects properly every nonempty open subset in the base of $(X^*)_*$, which means that

(20) $\gamma_X X$ is a dense subset of $(X^*)_*$

(this little fact, which passed unobserved by modal logicians, will be quite useful).

Applying γ^* to both sides of (19) we obtain $\gamma^*\gamma C = \gamma^*\beta C \cap \gamma^*\gamma X = C \cap X = C$, which is just another way of saying that, for every $C \in T$, $\gamma x \in \gamma C$ iff $x \in C$. This also holds for all closed subsets of X. In fact, for any subset D of X and every $C \in T$, $D \subseteq C$ implies $\gamma^*\gamma D \subseteq \gamma^*\gamma C = C$ and conversely $\gamma^*\gamma D \subseteq C$ implies $D \subseteq C$ because $D \subseteq \gamma^*\gamma D$. So, for every $C \in T$, $D \subseteq C$ iff $\gamma^*\gamma D \subseteq C$, from which

(21) for every closed subset D of X, $\gamma^*\gamma D = D$.

Some additional properties of γ are tied to the structure of X. Since all our spaces have a base of clopen subsets, X is Hausdorff iff the closure of a point is the point itself. That is, X is Hausdorff iff

(22) if for every $C \in T$, $x \in C$ iff $y \in C$, then $x = y$

holds for every $x, y \in X$. But then the definition of γ says that (22) is equivalent to: if $\gamma x = \gamma y$ then $x = y$. Therefore

(23) γ_X is one–one iff X is Hausdorff.

Proving that

(24) γ_X is onto iff X is compact iff γ_X is closed

is also easy. Recall that a space X is compact iff every ultrafilter of neighbourhoods converges (cf. [10, p. 106]), that is iff

(25) for every $S \in U(T)$, $\bigcap S \neq \emptyset$.

Assuming that γ_X is onto, (25) is immediate: for every $S \in U(T)$ there exists $x \in X$ such that $S = \gamma x$, and hence $x \in \bigcap S$. Now assume X is compact. Then γ_x is closed because it is a continuous function with compact domain (cf. [10, p. 106] or just apply (12) above). Finally, if γ_X is closed, then in particular γX is closed and hence $\gamma X = U(T)$ by (20).

Finally, it is not difficult to see that γ behaves well with respect to relations:

(26) for every $r : X \to Y$, $x\bar{r}y$ iff $\gamma x (r^*)_* \gamma y$.

In fact, by the definition of $(r^*)_*$, $\gamma x (r^*)_* \gamma y$ iff for every $D \in U$, $r^*D \in \gamma x$ implies $D \in \gamma y$. By (16(ii)) this is equivalent to: for every $D \in U$, $x \in r^*D$ implies $y \in D$. And hence by (8) also to $x\bar{r}y$.[7]

3. Basic adjunction

The aim of this section is to prove the most general result we could find connecting boolean and modal algebras on one side with spaces and frames on the other. This takes the form of an adjunction between **Bal** and **PSpa**, a category of spaces defined below which is strictly related to **Spa**. From such an adjunction, all other similar result will follow either as particular cases or corollaries.

We use a few lines to recall the categorical notions we need. Let C, D be any categories, and let $A, B, \ldots$ be objects of D, and $X, Y, \ldots$ be objects of C. We say that two contravariant functors $(\ )_*: C \to D$ and $(\ )^*: D \to C$ form an *adjoint pair* (on the right)[8] if:

(1) for every A and X, there is a bijection

$\varphi_{X,A}: \mathrm{Hom}_C(X, A_*) \to \mathrm{Hom}_D(A, X^*)$;

(2) φ is natural in the variables A, X, that is, for every $f: X \to A_*$, $h: Y \to X$ and $k: B \to A$, the following hold:

$\varphi(f \circ h) = h^* \circ \varphi f, \qquad \varphi(k_* \circ f) = \varphi f \circ k.$

It is easily seen (it is just an exercise to reverse almost all arrows and compositions in the proof of Theorem 2.v in [25, p. 81]) that

(3) the functors $(\ )_*$ and $(\ )^*$ form an adjoint pair iff there exist two natural transformations, $\gamma: \mathrm{Id}_C \to ((\ )^*)_*$ and $\beta: \mathrm{Id}_D \to ((\ )_*)^*$, satisfying the triangular identities:

for every A, $\quad (\beta_A)_* \circ \gamma_{A_*} = 1_{A_*}$,

for every X, $\quad (\gamma_X)^* \circ \beta_{X^*} = 1_{X^*}$.

Sketch of proof. Assume that (1) and (2) hold. Then for every A put $\beta_A = \varphi(1_{A_*})$ and for every X put $\gamma_X = \varphi^{-1}(1_{X^*})$. Let $k: B \to A$. Then from $k_* \circ 1_{A_*} = 1_{B_*} \circ k_*$ using (2) we obtain $\varphi(1_{A_*}) \circ k = (k_*)^* \circ \varphi(1_{B_*})$, i.e., β is a natural transformation. Now from $\gamma_X = 1_{(X^*)_*} \circ \gamma_X$ using (2) we obtain

$$1_{X^*} = \varphi(\gamma_X) = \varphi(1_{(X^*)_*} \circ \gamma_X) = (\gamma_X)^* \circ \varphi(1_{(X^*)_*}) = (\gamma_X)^* \circ \beta_{X^*}.$$

Quite similarly, the other triangular identity and the fact that γ is a natural transformation follow from the fact that, since φ is bijective, also φ^{-1} is natural. Note that φ can be completely described using γ: from $f = 1_{A_*} \circ f$ we obtain $\varphi f = f^* \circ \varphi(1_{A_*})$, that is $\varphi f = f^* \circ \beta_A$. Also, $\varphi^{-1}(g) = g_* \circ \gamma$.

Conversely, define φ by $\varphi f = f^* \circ \beta_A$. To prove that φ is bijective, we show that $\theta g = g_* \circ \gamma$ (where $g: A \to X^*$) is its inverse. In fact,

$\theta\varphi f = (f^* \circ \beta)_* \circ \gamma$	by the definition of φ and θ
$= \beta_* \circ (f^*)_* \circ \gamma$	because $(\ )_*$ is a contravariant functor
$= \beta_* \circ \gamma \circ f$	because γ is a natural transformation
$= f$	by triangular identities.

Quite similarly, we can show $\varphi\theta g = g$. Finally,

$$\varphi(f \circ h) = (f \circ h)^* \circ \beta = h^* \circ f^* \circ \beta = h^* \circ \varphi f \quad \text{and}$$
$$\varphi(k_* \circ f) = f^* \circ (k_*)^* \circ \beta = f^* \circ \beta \circ k = \varphi(f) \circ k$$

show that φ is natural.

One must be very careful at this point: while it is true that β and γ, as defined in the preceding section, satisfy the triangular identities (see 2.17), it is *not* true that γ is a natural transformation from $\mathrm{Id}_{\mathrm{Spa}}$ to $((\ )^*)_*$. In fact, it is easy to find a space X and a relation $r: X \to X$ such that $\gamma \circ r \neq (r^*)_* \circ \gamma$. Let X be the natural numbers N, $T = P(X)$ and r the usual order $\leqslant$. Then for every $C \subseteq N$, r^*C is the maximal end segment contained in C, and so $T \in (r^*)_* \gamma x$ whenever T is a nonprincipal ultrafilter. On the other hand, by the definition of γ, $\gamma r x$ contains only principal ultrafilters.

This is quite unfortunate, because it compels us to consider another category of spaces. The idea is to identify all relations with the same image under $(\ )^*$ or, which is equivalent by (2.11), to take only point-closed continuous relations as morphisms. But then another problem arises, namely that the composition of point-closed relations is not, in general, point-closed (the reader can easily find counterexamples). Before giving up , however, we invent a new composition of morphisms $\star$, simply by defining $s \star r$ to be the minimal point-closed relation containing $s \circ r$. So

(4) for every point-closed continuous relations $r: X \to Y$ and $s: Y \to Z$, we put $s \star r = \overline{s \circ r}$.

The trouble now is to show that $\star$ is a good composition. Let us show it step by step.

(5) $\star$ is associative.

Proof. Assume $r: X \to Y$, $s: Y \to Z$ and $t: Z \to W$ are point-closed continuous relations. Let C be any clopen in W^*. Then

$$\begin{aligned} (t \star (s \star r))^* C &= (t \circ (s \star r))^* C && \text{because of (2.10) and (4)} \\ &= ((s \star r)^* \circ t^*) C && \text{because } (\ )^* \text{ is a functor} \\ &= ((s \circ r)^* \circ t^*) C && \text{again by (2.10)} \\ &= r^* s^* t^* C. \end{aligned}$$

Quite similarly, we obtain also $((t \star s) \star r)^* C = r^* s^* t^* C$ and hence we can apply (2.11) to obtain the claim $t \star (s \star r) = (t \star s) \star r$ since both members are point-closed by definition.

Since the identity morphism $1_X: X \to X$ must be point-closed, it is quite natural to define it as $\{(x, y): y \in \overline{\{x\}}\}$, alias $1_X: x \mapsto \overline{\{x\}}$. However, the proof that 1_X is in fact the identity morphism of X is a bit tedious, and can be jumped with no harm by the reader who sees no problems in always restricting to Hausdorff spaces. In fact, in this case 1_X is simply the identity function on X, since every point is closed.

(6) for every space X, the relation $1_X : x \mapsto \overline{\{x\}}$ is the identity morphism on X with respect to composition $\star$.

Proof. For any point-closed $s : Z \to X$ and every $z \in Z$, $(1_X \star s)z = \overline{1_X sz}$ by definition of $\star$, but since sz is closed, $1_X sz = sz$ and hence $\overline{1_X sz} = sz$. So $1_X \star s = s$.

Now assume $r : X \to Y$ is point-closed. First note that r can not distinguish points of X with equal closure, i.e.

(7) if $r : X \to Y$ is point-closed and $\overline{\{x\}} = \overline{\{y\}}$, then $rx = ry$.

In fact, from $\overline{\{x\}} = \overline{\{y\}}$ we have that for every $D \in U$, $x \in r^*D$ iff $y \in r^*D$, that is $rx \subseteq D$ iff $ry \subseteq D$. Since rx and ry are closed, this means $rx = ry$. From (7) we obtain $r\overline{\{x\}} = rx$ and hence

$$(r \star 1_X)x = \overline{(r \circ 1_X)x} = \overline{(r \circ 1_X)x} = \overline{r\overline{\{x\}}} = \overline{rx} = rx,$$

that is $r \star 1_X = r$ as we wanted.

Summing up, we have shown that

(8) taking spaces as objects and continuous point-closed relations as morphisms, with composition $\star$ defined in (4), gives a category, called **PSpa**.

Note that the functor $(\ )^*$ of the preceding section is also a functor from **PSpa** to **Bal** (recall that $r^* = \bar{r}^*$ for every relation r); similarly, $(\ )_*$ is a functor from **Bal** to **PSpa** (because of (2.9) and because $\tau'_* \circ \tau_* = \tau'_* \star \tau_*$, since $\tau'_* \circ \tau_*$ is point-closed). So we can finally start to go downhill towards our aim, which is proving

(9) (Basic adjunction) the functors $(\ )_* : \mathbf{Bal} \to \mathbf{PSpa}$ and $(\ )^* : \mathbf{PSpa} \to \mathbf{Bal}$ form an adjoint pair.

After theorem (3), (2.14) and triangular identities (2.17), the only fact left to be proved is that γ is a natural transformation. We deduce it from a more general lemma, which will be essential in the next section.

(10) for every point-closed relations r, s, c, d, the diagram

$$\begin{array}{ccc} X & \xrightarrow{r} & Y \\ {\scriptstyle c}\downarrow & & \downarrow{\scriptstyle d} \\ Z & \xrightarrow{s} & W \end{array}$$

commutes in PSpa, that is, $\overline{scx} = \overline{drx}$ for every $x \in X$, iff the diagram

$$\begin{array}{ccc} X^* & \xleftarrow{r^*} & Y^* \\ {\scriptstyle c^*}\uparrow & & \uparrow{\scriptstyle d^*} \\ Z^* & \xleftarrow{s^*} & W^* \end{array}$$

commutes in Bal, that is, $r^*d^*C = c^*s^*C$ for every $C \in W^*$.

Proof. By the definition of closure, $\overline{scx} = \overline{drx}$ iff for every $C \in W^*$, $scx \subseteq C$ iff $drx \subseteq C$. But using (1.5) we see that $scx \subseteq C$ iff $x \in c^*s^*C$, and similarly $drx \subseteq C$ iff $x \in r^*d^*C$, from which the claim.

We then also have

(11) γ is a natural transformation from $\mathrm{Id}_{\mathrm{PSpa}}$ into $((\ )^*)_*$.

Proof. First note that $\gamma : X \to (X^*)_*$ is a morphism in PSpa: it is point-closed simply because it is a function and every point in $(X^*)_*$ is closed, and it is a continuous relation by (2.18). Now to obtain the claim we have to show that

(12) for every $x \in X$ and $r : X \to Y$, $\overline{\gamma r x} = \overline{(r^*)_* \gamma x}$.

By (10), it is enough to show that $r^*\gamma^*\beta D = \gamma^*((r^*)_*)^*\beta D$ for every $\beta D \in U$. This can easily be obtained using triangular identities and (2.6) applied to r^*:

$$r^*\gamma^*\beta D = r^*D = \gamma^*\beta r^*D = \gamma^*((r^*)_*)^*\beta D.$$

4. Modal duality and other corollaries

As we promised in Section 3, modal duality as well as some other similar dualities or adjunctions, are easy corollaries of basic adjunction.

We first see what happens if we restrict to the usual case, in which morphisms $r : X \to Y$ and $\tau : A \to B$ are functions and homomorphisms respectively. Assume that $r : X \to Y$ is a function. Then by (1.7) $r^* = r^{-1}$ and hence, by (1.8), $r^* - D = -r^*D$ for every $D \in U$. So

(1) if $r : X \to Y$ is a function, then $r^* : Y^* \to X^*$ is a boolean homomorphism.

On the other side, assume that $\tau : A \to B$ is a homomorphism. Since τ is a hemimorphism, $\tau^{-1}S$ is a filter for every $S \in U(B)$ (see the proof of (2.6)); but from $\nu\tau a = \tau\nu a$ for every $a \in A$, we have also $a \notin \tau^{-1}S$ iff $\tau a \notin S$ iff $\nu\tau a = \tau\nu a \in S$ iff $\nu a \in \tau^{-1}S$, which means that $\tau^{-1}S$ is an ultrafilter. Hence $\tau_* S$, which is equal to $\{T \in U(A) : \tau^{-1}S \subseteq T\}$ by definition, is a singleton for every $S \in U(B)$, and hence

(2) if $\tau : A \to B$ is a homomorphism, then $\tau_* : B_* \to A_*$ is a function.

As a first little corollary of basic adjunction, we can prove also the converse of

(1) and (2) (but also a direct proof is possible, cf. [20, p. 57]):

(3) (i) τ is a homomorphism iff τ_* is a function;
(ii) under the assumption that Y is Hausdorff, $r:X \to Y$ is a function iff r^* is a homomorphism.

Proof. (i) Assume τ_* is a function. Then $(\tau_*)^*$ is a homomorphism by (1); so $(\tau_*)^* \circ \beta = \beta \circ \tau$ is a homomorphism and, β being an isomorphism, τ must be a homomorphism.

(ii) Assume r^* is a homomorphism. Then $(r^*)_*$ is a function by (2), and hence also $(r^*)_* \circ \gamma$ is a function, that is $(r^*)_* \gamma x$ is a singleton for every $x \in X$. But since Y is Hausdorff, $(r^*)_* \gamma x = \overline{(r^*)_* \gamma x}$ and hence, by (3.12), $\gamma r x$ is a singleton. Finally, rx is a singleton, because γ is one-one by (2.20).

Ignoring hemimorphisms which are not also homomorphisms, we obtain **Ba**, the usual category of boolean algebras and homomorphisms, as a subcategory of **Bal**. Similarly, we do ignore continuous relations which are not functions, but we do not obtain only continuous functions, because our definitions depend on bases (cf. the remarks following (1.10)). So, disregarding useless generalities, we also restrict to spaces (X, T) where T coincides with the family $C(X)$ of all clopen subsets of X, which we call *zero-dimensional* (note that usually a topological space is called zero-dimensional just in case $C(X)$ is a base and, in this sense, all our spaces are zero-dimensional; the difference is due, once more, to the fact that we consider the base T as part of the space). We then have:

(4) (Boolean adjunction) the category **Ba** is adjoint to the category **HSpa** of zero-dimensional Hausdorff spaces and continuous functions

Proof. The functors between **Bal** and **PSpa** continue to be functors here, by (3) and the fact that the composition of morphisms $\star$ in **PSpa** reduces to the usual composition of functions, because of the restriction to Hausdorff spaces. β and γ continue to be natural transformations, because β_A is always a homomorphism and γ_X is always a continuous function, and they obviously satisfy triangular identities. So we can apply (3.3).

An immediate corollary is the better known boolean duality. Here two categories are said to be *dual* of each other if one is equivalent to the opposite of the other, or, more directly, if they are equivalent via two contravariant functors (cf. [27, p. 18]). A duality is just a particular case of adjunction (in our sense), in which the units are natural isomorphisms. So boolean adjunction gives a duality simply by restricting to subcategories in which β and γ are isomorphisms. Since β_A is an isomorphism for each boolean algebra A, **Ba** is left unchanged. By (2.20), γ_X is an isomorphism iff the space X is compact, beside being Hausdorff, that is X is a boolean space. Also note that, by (1.2), every boolean space is zero-dimensional. We thus have:

(5) (Boolean duality) the categories **Ba** and **BSpa**, of boolean spaces and continuous functions, are dual to each other.

Actually, the same argument can be applied to basic adjunction, and we obtain:

(6) (Halmos' duality) the categories **Bal** and **CHSpa**, of boolean spaces and continuous point-closed relations, are dual to each other.

Continuous point-closed relations between boolean spaces were called boolean relations by Halmos [20]. Note that **CHSpa**, as a subcategory of **PSpa**, inherits the fancy composition ★, which however coincides with ∘ on boolean relations by (2.12) (and Halmos had to prove a similar lemma simply to show that **CHSpa** is indeed a category).

We have now finally arrived at modal duality. Recall that a frame $F=(X, r, T)$ is here identified with a morphism $(X, T)\xrightarrow{r}(X, T)$ in **Spa**. We have seen however that no adjunction exists between **Bal** and **Spa**, having β and γ as units. We are thus led to consider, as we did for spaces, another category of frames, in which objects are morphisms $X\xrightarrow{r}X$ in **PSpa**. Moreover, in order that any weak contraction $c:F\rightarrow G$ shall be point-closed and hence a morphism in **PSpa**, we also have to restrict to Hausdorff spaces. Frames (X, r, T) with r point-closed and (X, T) Hausdorff have already been considered in the literature, under the name of *refined* frames (cf. Section III.2 below). The category of refined frames and weak contractions is here called **RFra**.

Let us now give a second look at weak contractions. As a corollary of (3.10), we have

(7) a function $c:F\rightarrow G$, where F, G are arbitrary frames, is a weak contraction iff $c^{-1}(U)\subseteq T$ and $\overline{crx}=\overline{scx}$ for every $x\in X$.

We will often use this characterization from now on, even without explicit mention. Restricting to **RFra** it has an even sharper form: a function $c:F\rightarrow G$, with F, G refined, is a weak contraction iff c is a morphism in **PSpa** which makes the diagram

$$\begin{array}{ccc} X & \xrightarrow{r} & X \\ {\scriptstyle c}\downarrow & & \downarrow{\scriptstyle c} \\ Y & \xrightarrow{s} & Y \end{array}$$

commute in **PSpa**. Moreover, note that by (7) every closed weak contraction in **RFra** is always a contraction, since in this case from $\overline{crx}=\overline{scx}$ we also have $crx=scx$; this by the way gives support to our claim that weak contractions are more basic than contractions, the latter corresponding to closed continuous functions in topology.

Now modal duality is only a matter of putting together what we already know. The functors $(\ )^*$ and $(\ )_*$ between **Bal** and **PSpa** immediately yield functors between **Mal** and **RFra**, which we denote by the same symbols. Of course, if $F=((X, T), r)$, we put $F^*=((X, T)^*, r^*)$, which clearly is a modal algebra; similarly, if $A=(A, \tau)$ we put $A_*=(A_*, \tau_*)$, which is a (compact) refined frame since A_* is a boolean space and τ_* is point-closed by (2.9). We already know how the functors act on weak contractions and modal homomorphisms, since they are particular cases of continuous relations and hemimorphisms respectively. Moreover, (3.10) and (3) tell that

(8) $c:F\to G$ is a weak contraction iff $c^*:G^*\to F^*$ is a modal homomorphism.

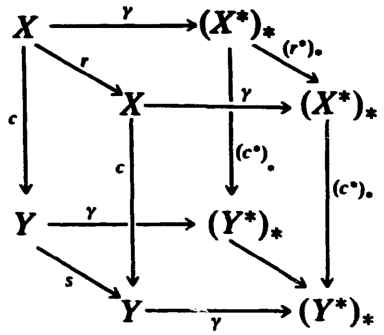
Using this and the fact that β is a natural isomorphism in **Bal**, we also have

(9) $h:A\to B$ is a homomorphism iff $h_*:B_*\to A_*$ is a weak contraction.

In particular, $(\ )_*$ and $(\ )^*$ are indeed functors between **Mal** and **RFra**, and we can finally prove:

(10) (Modal adjunction) the functors $(\ )_*$ and $(\ )^*$ between **Mal** and **RFra** form an adjoint pair.

In fact, assume that $c:F\to G$ is a weak contraction. Then all the information we need is contained in the following commutative cube of **PSpa**:



The assumption that $c:F\to G$ is a weak contraction tells that the left face is commutative, and hence also its image under $((\ )^*)_*$, which is the right face. The remaining four faces are commutative because γ is a natural transformation in **PSpa**; but then, by the remark following (7), the top (bottom) face shows that $\gamma:F\to(F^*)_*$ $(\gamma:G\to(G^*)_*)$ is a weak contraction, that is a morphism in **RFra** (actually, one of the aims of the introduction of weak contractions and of ★, as well as much of the work up to now, was just to reach this apparently innocuous result). What do the front and back faces express? To see it, let us compress the cube by juxtaposing them. We obtain the commutative square

$$\begin{array}{ccc} F & \xrightarrow{\gamma} & (F^*)_* \\ {\scriptstyle c}\downarrow & & \downarrow{\scriptstyle (c^*)_*} \\ G & \xrightarrow[\gamma]{} & (G^*)_* \end{array}$$

which tells that γ is a natural transformation from $\mathrm{Id}_{\mathbf{RFra}}$ into $((\ )^*)_*$.

Exactly similar (or actually, dual) is the proof of the fact that β is a natural isomorphism from $\mathrm{Id}_{\mathbf{Mal}}$ into $((\ )_*)^*$. So, since of course triangular identities continue to hold for β and γ (relations and hemimorphisms do not affect them), we can apply(3.3) and hence (10) is proved.

As for boolean and Halmos' duality, modal duality is obtained from (10) by restricting to the subcategory of RFra in which γ is an isomorphism. So let us first characterize isomorphisms. An isomorphism $c:F\to G$ must be a bijective continuous function with a continuous inverse, that is a (topological) homeomorphism. Conversely, if a weak contraction $c:F\to G$ is a homeomorphism, then from $\overline{crx}=\overline{scx}$ for all $x\in X$, we have in particular $\overline{crc^{-1}y}=\overline{scc^{-1}y}=\overline{sy}$ for all $y\in Y$, and hence also $\overline{c^{-1}sy}=\overline{rc^{-1}y}$, but still c^{-1} may fail to be a weak contraction (for example, consider $F=(N, r, P(N))$, $G=(N, r, T)$ where r is any relation, T is the family of finite and cofinite subsets of N and c is the identity function); to characterize isomorphisms we thus have to add the assumption that for every $C\in T$, $cC\in U$, or equivalently, that c^* is onto. Since γ^* is always onto, by (2.20) we have:

(11) $\gamma:F\to(F^*)_*$ is an isomorphism in RFra iff F is compact.

A compact refined frame is here called *descriptive* (and we will show in Section III.2 that this definition is equivalent to that of Goldblatt, who introduced the name). Restricting to descriptive frames, the distinction between weak contractions and contractions vanishes: a weak contraction $c:F\to G$, with F, G descriptive, is always also a contraction, because for every $x\in X$ from $\overline{crx}=\overline{scx}$ we obtain $crx=scx$, since c, r, s are point-closed and hence closed by (2.12). We thus have

(12) (Modal duality) the categories Mal and DFra, of descriptive frames and contractions, are dual to each other.

Unlike in the case of spaces, we can here also extend modal adjunction (10) to an adjunction between Mal and the whole of Fra, because, contrary to PSpa with respect to Spa, RFra *is* a subcategory of Fra, and actually of a rather nice kind:

(13) RFra is a reflective subcategory of Fra.

That is, there is a functor $R:\mathrm{Fra}\to\mathrm{RFra}$ which is left adjoint to the inclusion functor (note that here we follow [25, p. 89] word by word, since functors are covariant). To prove it, it is enough to show that for every frame F there is a refined frame F_R and a morphism $\rho_F:F\to F_R$ such that every morphism $c:F\to G$, with G refined, splits uniquely through ρ, that is $c=\bar{c}\circ\rho$ for a unique $\bar{c}:F_R\to G$ (cf. [25, p. 89]).

The most natural way to obtain a refined frame from any given frame F is to extend r into its pointwise closure $\bar{r}$ and identify points which are not separated by T. The latter aim is accomplished by the image of (X, T) under the morphism $\gamma:F\to(F^*)_*$. That is, we put $X_R=\{\gamma x:x\in X\}$ and $T_R=\{\gamma C:C\in T\}$. Now, defining a relation r_R on (X_R, T_R) by putting $\gamma x r_R \gamma y$ iff $x\bar{r}y$, we obtain that the triple (X_R, r_R, T_R) is exactly the substructure induced by $(F^*)_*$ on γX, that is

(14) for every $x \in X$, $(r^*)_*\gamma x \cap \gamma X = r_R\gamma x$.

In fact, by (2.26), $\gamma x(r^*)_*\gamma y$ is equivalent to $x\bar{r}y$, which is the definition of $\gamma x r_R \gamma y$.

So $F_R = (X_R, r_R, T_R)$ will be a refined frame, as soon as we prove that it is indeed a frame. We only have to show that T_R is closed under r_R^* (note that this is not true of every induced substructure), which immediately follows from

(15) for every $C \in T$, $\gamma r^*C = r_R^*\gamma C$.

To prove it, first note that

(16) for every $x \in X$, $r_R\gamma x = \gamma\overline{rx}$

because $\gamma y \in \gamma\overline{rx}$ iff $y \in \gamma^*\gamma\overline{rx}$, and $\gamma^*\gamma\overline{rx} = \overline{rx}$ by (2.21). Then, using (16) and (2.21) again, (14) follows easily: $\gamma x \in r_R^*\gamma C$ iff $r_R\gamma x = \gamma\overline{rx} \subseteq \gamma C$ iff $\overline{rx} \subseteq \gamma^*\gamma C = C$ iff $x \in r^*C$ iff $\gamma x \in \gamma r^*C$.

The next step is

(17) for every frame F, $\gamma : F \to F_R$ is a weak contraction

which by (7) is proved once we show that $\widetilde{\gamma rx} = r_R\gamma x$, where $\sim$ now denotes closure in F_R, to be distinguished from closure $\overline{\ \ }$ in $(F^*)_*$. But $\widetilde{\gamma rx} = \overline{\gamma rx} \cap \gamma X$ (cf. [10, p. 65]) and hence, since $\overline{\gamma rx} = (r^*)_*\gamma x$ because $\gamma : F \to (F^*)_*$ is a weak contraction, $\widetilde{\gamma rx} = (r^*)_*\gamma x \cap \gamma X = r_R\gamma x$ by (14), as we wanted.

To complete the proof of (13), let $c : F \to G$, with G refined, be any weak contraction. We want to show that there is only one weak contraction $\bar{c} : F_R \to G$ such that $\bar{c} \circ \gamma = c$. Of, course, since γ is onto F_R, the condition $\bar{c}\gamma x = cx$ uniquely defines $\bar{c}$, and so it only remains to show that $\bar{c}$ is a weak contraction. First, for every $D \in U$ we have that $\gamma x \in \bar{c}^{-1}D$ iff $\bar{c}\gamma x \in D$ iff $x \in c^{-1}D$ iff $\gamma x \in \gamma c^{-1}D$. So $\bar{c}^{-1}D = \gamma c^{-1}D \in T_R$. Finally, using the fact that $c\overline{rx} \subseteq \overline{crx}$ (cf. [10, p. 41]), one can easily derive that $\overline{c\overline{rx}} = \overline{crx}$. Hence from $\overline{crx} = \overline{scx}$ we have $\overline{s\bar{c}\gamma x} = \overline{scx} = \overline{crx} = \overline{c\overline{rx}} = \overline{\bar{c}\gamma\overline{rx}} = \overline{\bar{c}r_R\gamma x}$ (the last equality by (16)), which is what we need by (7).

Putting together (9) and (13), we obtain

(18) (Modal adjunction, extended) the categories Mal and Fra are adjoint

(the scrupulous reader can check that the composition of the natural bijections $\mathrm{Hom}(F, A_*) \cong \mathrm{Hom}(F_R, A_*)$ and $\mathrm{Hom}(F_R, A_*) \cong \mathrm{Hom}(A, (F_R)^*)$ is the natural bijection needed, observing that for every F, $F^* \cong (F_R)^*$).[9]

5. Frame constructions and duality

With the tools provided by modal duality, we here analyse (continuing the work in [11] and [15]) the usual frame constructions: subframes, images of contractions, disjoint union of frames. Actually, our notion of weak contraction suggests the introduction of a new notion, namely that of weak subframe, which will prove to be quite useful in the next chapter.

Giving all details here would mean boring the reader to death, and we thus assume more familiarity with categories than in the preceding sections. As we often did previously, we begin with the algebraic side:

(1) in the category Mal of modal algebras:
 (i) monomorphisms coincide with injective homomorphisms,
 (ii) subobjects coincide with subalgebras.

Of course, (ii) follows from (i), which is true because free modal algebras exist. On the other hand, the problem whether epimorphisms in Mal coincide with surjective homomorphisms remains open.

We can also describe quotients of modal algebras quite well. It is well known that the assignment $\theta \mapsto F_\theta = \{a \in A : a\,\theta\,1\}$ defines a biunivocal correspondence between congruences θ and filters F on a boolean algebra A. We say that a filter F on a modal algebra $A = (A, \tau)$ is closed under τ, briefly a *τ-filter*, if $a \in F$ implies $\tau a \in F$. We then have

(2) the lattice of congruences of a modal algebra A is isomorphic to the lattice of τ-filters of A.

Proof. It is enough to show that θ preserves τ iff F_θ is closed under τ. So assume that $a\,\theta\,b$ implies $\tau a\,\theta\,\tau b$. Then from $a \in F_\theta$, i.e. $a\,\theta\,1$, we have $\tau a\,\theta\,\tau 1$ and hence $\tau a \in F_\theta$ since $\tau 1 = 1$. Conversely, assume F_θ is closed under τ and let $a\,\theta\,b$. Then $a \leftrightarrow b \in F_\theta$ and hence $\tau(a \leftrightarrow b) \in F_\partial$, from which also $\tau a \leftrightarrow \tau b \in F_\theta$ because $\tau(a \leftrightarrow b) \leqslant \tau a \leftrightarrow \tau b$. So $\tau a\,\theta\,\tau b$.

Turning to frames, matters are not as simple. We first see the connections between special morphisms in the two categories. What directly follows from modal adjunction (cf. [27, p. 94], but still modulo the exercise of reversing arrows) is that the image under our functors of an epimorphism is a monomorphism. In addition, we can easily prove that our functors are faithful, namely

(3) (i) if c, d are weak contractions in RFra and $c^* = d^*$, then $c = d$;
 (ii) if h, g are homomorphisms and $h_* = g_*$, then $h = g$.

This holds simply because β and γ (in RFra) are monomorphisms. In fact, let $g_* = h_*$. Then $(g_*)^* = (h_*)^*$ and hence, since $(g_*)^*\beta = \beta g$, $\beta g = \beta h$, from which $g = h$ because β is a monomorphism. The proof of (i) is identical (also, we already proved it as (2.8)). So (cf. [27, p. 115]) the functors also 'co-reflect' monomorphisms, that is:

(4) (i) in RFra, c is an epimorphism iff c^* is a monomorphism;
 (ii) h is an epimorphism iff h_* is a monomorphism.

Proof. The two proofs are identical, and so let us prove (i). As mentioned above, one direction is just a consequence of modal adjunction, and to help the reader here is the argument. Assume c is epic and let $c^*g = c^*h$. Then $\varphi^{-1}g \circ c = \varphi^{-1}(c^*g) = \varphi^{-1}(c^*h) = \varphi^{-1}h \circ c$, from which $\varphi^{-1}g = \varphi^{-1}h$ since c is epic. So $g = h$ since φ is bijective. Conversely, assume c^* is monic and let $dc = ec$. Then $c^*d^* = c^*e^*$, from which $d^* = e^*$ since c^* is monic, and hence $d = e$ because $(\ )^*$ is faithful by (3(i)).

Since β is a natural isomorphism, h is a monomorphism iff $(h_*)^*$ is a monomorphism, and hence, by (4(i)) applied to h_*, also

(5) h is a monomorphism iff h_* is an epimorphism.

Note that this rests solely on the fact that β is a natural isomorphism. So, when also γ is a natural isomorphism, the dual statement holds, that is

(6) in **DFra**, c is a monomorphism iff c^* is an epimorphism.

Note that one direction of (6) holds more generally:

(7) in **RFra**, if c^* is an epimorphism, then c is a monomorphism.

This is true because γ is a monomorphism in **RFra**: if c^* is epic, then $(c^*)_*$ is monic by (4(ii)), and hence also $(c^*)_*\gamma = \gamma c$ is monic, from which the claim.

All of this may be amusing, but of little use, at least until we can characterize epimorphisms and monomorphisms in **Fra** more directly. Surely epimorphisms are not always surjective. In fact, $\gamma: \boldsymbol{F} \to (\boldsymbol{F}^*)_*$ is epic for every $\boldsymbol{F}$, by (4(i)) and the fact that γ^* is an isomorphism (by triangular identities, β is its inverse). However, we know by (2.24) that γ is onto only when $\boldsymbol{F}$ is compact. Still, γ is always 'almost' onto, in the sense that γX is dense in $(\boldsymbol{F}^*)_*$. We now see that the same is true for all epimorphisms:

(8) if $c: \boldsymbol{F} \to \boldsymbol{G}$ is an epimorphism in **RFra**, then cX is a dense subset of Y.

Proof. By (4(i)), which applies only to refined frames, c^* is monic, and hence one–one by (1). So, for every $D \in \boldsymbol{U}$ such that $D \neq \emptyset$, we have $c^*D \neq \emptyset = c^*\emptyset$. But then $cc^*D \cap cX \neq \emptyset$ and hence also $D \cap cX \neq \emptyset$, because $cc^*D \subseteq D$ by (1.6).

In particular, when $\boldsymbol{F}$ is compact and $\boldsymbol{G}$ is Hausdorff, c is point-closed and hence closed by (2.12). In this case $cX = \overline{cX} = Y$ and c is indeed onto. Thus

(9) in **DFra**, epimorphisms coincide with surjective contractions.

The problem whether monomorphisms in **DFra** coincide with injective contractions remains open; we can only give two partial results which together with (5) and (6), show that it is equivalent to the above mentioned problem about epimorphisms in **Mal**.

(10) if $c: \boldsymbol{F} \to \boldsymbol{G}$ is a weak contraction in **RFra** and c^* is onto, then c is one–one.

Proof. If $x, y \in X$ are distinct, there exists $C \in T$ which separates them, that is $x \in C$ and $y \notin C$. Since c^* is onto, there exists $D \in U$ such that $C = c^*D$. So $x \in c^*D$, that is $cx \in D$, and $y \notin c^*D$, that is $cy \notin D$. This means that $cx \neq cy$.

(11) in DFra, if $c: F \to G$ is one–one, then c^* is onto.

Proof. Since F and G are compact, c is closed and hence cX is compact. So for every $C \in T$, both cC and $c - C$ are closed, and hence clopen in the topology induced by U on cX. But then by (1.2) $cC = D \cap cX$ for some $D \in U$, which gives $C = c^*D$ since c is one–one.

We can now turn to frame constructions. Given a frame G and a subset X of Y, the triple (X, r, T) where $r = s \cap X^2$ and $T = \{D \cap X : D \in U\}$ is here called the *substructure induced* by G on X. Also, we say that X is an *s-hereditary* subset of G if $sX \subseteq X$, that is $y \in X$ whenever $x \in X$ and xsy. A frame F is traditionally said to be a (generated) subframe of a frame G if $X \subseteq Y$, F is the substructure induced by G on X and X is s-hereditary. All of this can be expressed through contractions. In fact, $r = s \cap X^2$ is equivalent to the requirement that $irx = six$ for every $x \in X$, where $i: X \to Y$ is the inclusion function, and $T = \{D \cap X : D \in U\}$ is equivalent to the requirement that i is continuous and i^* is onto, because $i^*D = D \cap X$ for every $D \subseteq Y$. So F is a *subframe* of G, written $F \subseteq G$, if $i: F \to G$ is a contraction and $i^*: G^* \to F^*$ is onto. It is now natural to widen this definition as follows: F is a *weak subframe* of G, written $F \subseteq_w G$, if $i: F \to G$ is a weak contraction and i^* is onto. More generally, we say that F is *embeddable* in G, written $F \hookrightarrow G$, if there exists a one–one weak contraction $c: F \to G$ such that c^* is onto. Trivially, every weak subframe of G is embeddable in G. Conversely, when F, G are refined, if $F \hookrightarrow G$ via c, then the substructure cF induced by G on cX is a weak subframe of G which is isomorphic to F (a detailed verification is left to the reader; observe that the assumption $c: F \to G$ must be used to be able to say that $cF \subseteq_w G$). In particular, since γ^* is always onto, we have

(12) for every frame F, $F_R \subseteq_w (F^*)_*$ and, when F is refined, γ is an embedding of F into $(F^*)_*$

which will be essential in the next chapter.

Our definitions trivially imply that

(13) if $F \hookrightarrow G$, in particular if $F \subseteq_w G$, then F^* is a homomorphic image of G^*.

Conversely, if $h: A \to B$ is onto, then $h_*: B_* \hookrightarrow A_*$. In fact, if h is onto then for any $S, T \in U(A)$, $h^{-1}(S) = h^{-1}(T)$ implies $S = T$, that is h_* is one–one, and $(h_*)^*$ is onto because $(h_*)^* \circ \beta = \beta \circ h$. Since weak contractions between descriptive frames are always contractions, we have

(14) if $h: A \to B$ is onto, then B_* is isomorphic to a subframe of A_*.

It is well known that the assignment $F \mapsto \bigcap \{\beta a : a \in F\} = C_F$ defines an isomorphism between boolean filters of A and closed subsets of $U(A)$. By using (2.6) and (1.5), one can easily show that a filter F is closed under τ iff $\tau_* C_F \subseteq C_F$, and hence

(15) τ-filters of a modal algebra A correspond biunivocally to τ_*-hereditary closed subsets of $U(A)$.

Note that a subframe F of a descriptive frame G is itself descriptive iff X is closed, because closed and compact subsets coincide in a compact space (cf. [10, pp. 102–103]). Therefore, combining (15) with (2) we obtain

(16) the lattice of congruences of a modal algebra A is anti-isomorphic to the lattice of descriptive subframes of A_*.

Similar results hold for weak contractions. By (1) and (4(i)), we immediately have

(17) if $c: F \to G$ is an epimorphism, then G^* is isomorphic to a subalgebra of F^*

and conversely, by (5) and (9),

(18) if A is a subalgebra of B and i is the inclusion, then $i^*: B_* \to A_*$ is a surjective contraction

However, a result corresponding to (16) is not immediate, because a characterization of quotient objects of a frame F, in terms of F itself, is not readily available. We now find it out. First recall that, given a space (X, T) and an equivalence relation θ on X, we can define the quotient space $(X/\theta, T_\theta)$ putting $T_\theta = \{D \subseteq X/\theta : c_\theta^{-1} D \in T\}$, where $c_\theta : x \mapsto [x]_\theta$ is the canonical mapping from X to X/θ (cf. [10, pp. 83–84]). It is known that, given an epimorphism $c: F \to G$, and putting $\theta(c) = \{(x, y) : cx = cy\}$, the quotient space $(X/\theta(c), T_{\theta(c)})$ is homeomorphic to (Y, U) iff c is closed. And c is closed iff the relation $\theta(c)$ is closed (cf. [10, pp.83–84]). So, let F be any frame and θ a closed equivalence relation on X. We want to add a relation r_θ to $(X/\theta, T_\theta)$ in such a way that $F/\theta = (X/\theta, r_\theta, T_\theta)$ is a frame and c_θ a weak contraction from F onto F/θ. The condition $c_\theta rx = r_\theta c_\theta x$ itself is met by a unique point-closed r_θ, and it is a good definition iff

(19) for all $x, y \in X$, $x \,\theta\, y$ implies $\theta rx = \theta ry$.

Thus a *congruence* on a frame F is a closed equivalence relation θ satisfying (19). Now it is not difficult to check that

(20) if θ is a congruence on F, then F/θ is a frame and $c_\theta : F \to F/\theta$ is a contraction

and conversely

(21) if $c: F \to G$ is a closed onto contraction, then $\theta(c)$ is a congruence on F and G is isomorphic to $F/\theta(c)$.

When F is descriptive, any epimorphism $c: F \to G$ is closed and onto by (9), and thus quotient objects correspond biunivocally to congruences on F. So, since by modal duality quotients objects of A_* correspond biunivocally to subobjects of A, we have

(22) the lattice of subalgebras of A is isomorphic to the lattice of congruences on A_*.

The isomorphism can easily be described. If F is descriptive and U is a subalgebra of T, then the corresponding congruence θ is defined by: $x\,\theta\,y$ iff for every $C \in U$, $x \in C$ iff $y \in C$.

Finally, dualizing the categorical definition of product of algebras, we obtain the definition of coproduct, that is disjoint union of frames. Therefore the *disjoint union* of a family $(F_i)_{i\in I}$ of pairwise disjoint frames, is the frame

$$\sum_{i\in I} F_i = \left(\bigcup_{i\in I} X_i, \bigcup_{i\in I} r_i, T\right)$$

where $C \in T$ iff $C \cap X_i \in T_i$ for every $i \in I$ (cf. [10, p. 72]). By modal duality we then have

(23) for each family $(F_i)_{i\in I}$ of frames, $(\sum_{i\in I} F_i)^*$ is isomorphic to $\prod_{i\in I} F_i^*$.

On the other hand, given a family $(A_i)_{i\in I}$ of modal algebras, the frame $\sum_{i\in I} A_{i*}$ is not homeomorphic to $(\prod_{i\in I} A_i)_*$, because the former is never compact when I is infinite. Rather, using the fact that the functor sending F to $(F^*)_*$ is a reflector (because of modal adjunction), and reflectors preserve coproducts, one can obtain

(24) for every family $(A_i)_{i\in I}$ of modal algebras, $((\sum_{i\in I} A_{i*})^*)_*$ is isomorphic to $(\prod_{i\in I} A_i)_*$

from which, for finite I,

$$\sum_{i\in I} A_{i*} \cong \left(\prod_{i\in I} A_i\right)_*.$$

CHAPTER III. Classes of frames

Introduction

The mathematical theory so far developed would be sterile if we could not apply it to problems usually encountered by modal logicians. To show the contrary, we have chosen a specific area, namely the study of classes of frames, and revised it with the aid of duality theory.

In particular, the notion of weak contraction permits to obtain easy proofs of the preservation of consequence under usual frame constructions and a proof of the fact that every frame is equivalent to a refined frame also with respect to consequence. Moreover, the notion of weak subframe, together with duality, is used to give a new simple description of the class of frames for a given logic and a purely frame-theoretic characterization of modal axiomatic classes of frames. Two wellknown theorems on classes of Kripke frames are obtained as corollaries.

We rely by now on the reader's confidence with the subject and thus will often justify a step in the proofs simply 'by modal duality', without explicit reference to specific results in Sections II.4 and II.5.

1. The logic of frame constructions

It is well known that frame and algebraic constructions preserve the validity of modal formulae. For algebras this is true since identities are preserved by homomorphic images, subalgebras and direct products. It is less known that the corresponding frame constructions preserve also semantical consequence. The notion of weak contraction is used here to give a complete and uniform proof of this fact.

Let φ be a formula and Γ a set of formulae. We say that φ is a *consequence of* Γ over the frame $\boldsymbol{F}$, written $\Gamma \models^{\boldsymbol{F}} \varphi$, if for every valuation V and every $x \in X$, $x \Vdash_V \Gamma$ (that is, $x \Vdash_V \psi$ for every $\psi \in \Gamma$) implies $x \Vdash_V \varphi$. In other words, putting $V(\Gamma) = \bigcap \langle V(\psi) : \psi \in \Gamma \rangle$,

(1) $\Gamma \models^{\boldsymbol{F}} \varphi$ iff for every valuation V, $V(\Gamma) \subseteq V(\varphi)$.

Of all the consequence relations considered in the literature (at least four), this is the strongest. To save words, we denote by $C\boldsymbol{F}$ the set $\{(\Gamma, \varphi) : \Gamma \models^{\boldsymbol{F}} \varphi\}$ and call it *the consequence of* $\boldsymbol{F}$ (in analogy with $L\boldsymbol{F}$, the logic of $\boldsymbol{F}$). We will say that two frames $\boldsymbol{F}$ and $\boldsymbol{G}$ are *strongly equivalent* if they have the same consequence, that is if $C\boldsymbol{F} = C\boldsymbol{G}$.

A technical lemma is the common part of all preservation results to follow. Let $c : \boldsymbol{F} \to \boldsymbol{G}$ be a weak contraction; for every valuation V on $\boldsymbol{G}$ we define a valuation $c^{-1}V$ on $\boldsymbol{F}$ by putting: $(c^{-1}V)(p) = c^{-1}V(p)$ for every propositional variable p. We then have:

(2) if $c : \boldsymbol{F} \to \boldsymbol{G}$ is a weak contraction, then for every valuation V on $\boldsymbol{G}$ and every formula φ, $(c^{-1}V)(\varphi) = c^{-1}V(\varphi)$.

Proof. The proof is by induction on the complexity of φ. The only interesting step is that for the modal operator, and here $\Diamond$ is easier to handle than $\Box$. We have

$c^{-1}V(\Diamond\varphi) = c^{-1}s^{-1}V(\varphi)$	by the definition of $V(\Diamond\varphi)$
$= r^{-1}c^{-1}V(\varphi)$	because c is a weak contraction
$= r^{-1}(c^{-1}V)(\varphi)$	by induction hypothesis
$= (c^{-1}V)(\Diamond\varphi)$	again by definition of $(c^{-1}V)(\Diamond\varphi)$.

This result can easily be extended to the case of a set of formulae because c^{-1} preserves intersections. Then we are ready to prove the next theorems.

(3) if $c:F\to G$ is a surjective weak contraction, then $CF\subseteq CG$.

Proof. Assume $\Gamma\models^F\varphi$ and let $y\in V(\Gamma)$. Then, since c is onto, there is $x\in X$ such that $y=cx$ and hence $x\in c^{-1}V(\Gamma)=(c^{-1}V)(\Gamma)$. But then, by (1), since $\Gamma\models^F\varphi$, $x\in(c^{-1}V(\varphi)=c^{-1}V(\varphi)$, that is $y\in V(\varphi)$. So, by (1), $\Gamma\models^G\varphi$.

(4) if F is a weak subframe of G, then $CG\subseteq CF$.

Proof. By definition of weak subframe, the inclusion map $i:F\to G$ is a weak contraction. Then, for every V on G, $i^{-1}V$ is a valuation on F. By (2) and the definition of i^{-1}, it follows that $(i^{-1}V)(\varphi)=V(\varphi)\cap X$. On the other hand, since $T=\{C\cap X:C\in U\}$, for every valuation V' on F there is V on G such that $V'(\varphi)=V(\varphi)\cap X$. So, let us assume that for every V on G, $V(\Gamma)\subseteq V(\varphi)$. Then $V(\Gamma)\cap X\subseteq V(\varphi)\cap X$, that is $(i^{-1}V)(\Gamma)\subseteq(i^{-1}V)(\varphi)$. So, by (1), $\Gamma\models^F\varphi$.

Obviously, (3) and (4) hold a fortiori for surjective contractions and subframes respectively. Moreover, simply by considering the case in which Γ is empty, we obtain the usual preservation results of validity of formulae.[10] It is worthwhile to note explicitly that, by (4), (3) and (II.5.12):

(5) for every frame F, $C(F^*)_*\subseteq CF$.

Even if $LF=L(F^*)_*$ holds for every frame, the inclusion in (5) is sometimes proper, as we will see in the next section.

Using (3) and (4) it is now immediate to prove that

(6) if $(F_i)_{i\in I}$ is a family of frames and $F=\sum_{i\in I}F_i$, then $CF=\bigcap_{i\in I}CF_i$.

In fact, since each F_i is a subframe of F, by (4), $C(F)\subseteq\bigcap C(F_i)$; conversely, since $(F_i)_{i\in I}$ is a family of pairwise disjoint frames, for every V on F, $V(\varphi)=\bigcup_{i\in I}V_i(\varphi)$ and $V(\Gamma)=\bigcup_{i\in I}V_i(\Gamma)$, where $V_i(\varphi)=V(\varphi)\cap X_i$. Assume that, for every $i\in I$, $V_i(\Gamma)\subseteq V_i(\varphi)$. Then $V(\Gamma)=\bigcup_{i\in I}V_i(\Gamma)\subseteq\bigcup_{i\in I}V_i(\varphi)=V(\varphi)$, so by (1), $\bigcap_{i\in I}C(F_i)\subseteq C(F)$.

We will later use also the following corollary of (3):

(7) if $c:F\to G$ is a weak contraction onto and $c^*:G^*\to F^*$ is onto, then $CF=CG$.

Proof. Since c^* is onto, that is $c^{-1}(U)=T$, every valuation on F is of the form $c^{-1}V$ for some valuation V on G. Therefore, by (1), it is enough to prove that, for every V on G, $V(\Gamma)\subseteq V(\varphi)$ iff $(c^{-1}V)(\Gamma)\subseteq(c^{-1}V)(\varphi)$. But this follows immediately from (2) and the assumption that c is onto.

2. Refined frames are enough

Several conditions on a frame have been considered in the literature, with the aim of restricting to a class of frames with enough structure to make their use simpler, but in the meantime wide enough to obtain completeness. We have already met descriptive frames and know that any frame F is equivalent to a descriptive frame, that is its bidual $(F^*)_*$. So one could restrict to the class of descriptive frames with no harm for completeness, and on the other hand with all the structure of the equational class of modal algebras offered on the tray by modal duality. However, this choice has not gained much consent, probably because infinite Kripke frames are not descriptive, and thus the original intuition is partly lost. Also, from a more technical point of view, it is not true that any frame is equivalent to a descriptive frame also with respect to consequence.

The aim of this section is to show, instead, that the choice of refined frames is the best compromise. Following S.K. Thomason, who introduced the notion in [33], a frame $F = (X, r, T)$ is usually said to be refined when

(1) xry iff $(\forall C \in T)(x \in r^*C \rightarrow y \in C)$

and

(2) $(\forall C \in T)(x \in C \leftrightarrow y \in C) \rightarrow x = y$

hold. Using topology we can save words and mental energy, and, by (II.2.8) and (II.2.23), say that F is refined if r is point-closed and F is Hausdorff (that is, the space (X, T) is Hausdorff). It is obvious that Kripke frames are refined and, conversely, all finite refined frames are Kripke frames. At this point it is also worthwhile to note that adding

(3) for each ultrafilter S on T, $\bigcap S = \{x\}$ for some $x \in X$

to (1) and (2), we obtain the original definition by Goldblatt [15] of descriptive frames. Now (3) is equivalent to compactness (see II.2.24) and hence a frame F satisfies (1)–(3) iff F is compact refined, that is, iff F is descriptive in our sense (cf. Section II.4).

In order to prove that the choice of refined frames is the best compromise, we begin with:

(4) any frame is strongly equivalent to a refined frame.

Actually, since we know how to construct the refinement F_R of a frame F (cf. Section II.4), (4) becomes

(5) for every frame F, $CF = C(F_R)$

which is quite easily proved using (1.7). In fact, by (II.4.17), γ is a weak contraction from F onto F_R, and obviously $\gamma^* : F_R^* \rightarrow F^*$ is onto.

We now want to show, with an example, that (4) can not be improved, in the sense that there are refined frames which are not strongly equivalent to a descriptive frame. An example is provided by the Kripke frame $F = (N, >)$, where N is the set of natural numbers and $>$ the usual greater than order. Since $V(\neg\Box^n\bot) = \{m : m > n\}$ for each V on F, putting $\Gamma = \{\neg\Box^n\bot : n \in N\}$ we obtain

$V(\Gamma)=\emptyset$ for each V and hence $\Gamma\models^F\bot$. On the other hand, suppose $\Gamma\models^G\bot$ holds for a descriptive frame G. Since $V(\neg\Box^n\bot)$ does not depend on V, from $V(\Gamma)=\emptyset$ we obtain, by compactness, that there exist $n_0,\ldots,n_k$ such that, for every V,

$$V(\neg\Box^{n_0}\bot\ \&\cdots\&\neg\Box^{n_k}\bot)=\emptyset.$$

But then $\bot$ is a consequence of $\varphi=\neg\Box^{n_0}\bot\ \&\cdots\&\neg\Box^{n_k}\bot$ on G, while, for every V on G, $V(\varphi)=\{m:m>\max(n_0,\ldots,n_k)\}$. So F and G are not strongly equivalent.

Now that our choice is made, we want to support it with something more. As it is known, for any logic L, the class MA(L) of L-modal algebras, which is an equational class, can be described as the class of homomorphic images of some free L-modal algebra $F_L(\alpha)^*$. By duality this is immediately transferred to the class of descriptive frames for L, and any descriptive frame becomes (isomorphic to) a subframe of a universal frame $F_L(\alpha)$, for some α. Of course, the same is not true for all frames, but weak subframes enable us to improve the situation by showing that all refined frames can somehow be embedded in a universal frame. Let us give a precise sense to this by saying that a frame F is *embeddable* in G if F is isomorphic to a weak subframe of G. Then by (II.5.12) we have:

(6) any refined frame F is embeddable in $(F^*)_*$.

Now let F be any refined frame for L; for some ordinal α, F^* is a homomorphic image of $F_L(\alpha)^*$ and hence $(F^*)_*$ is (isomorphic to) a subframe of $F_L(\alpha)$. So, by (6), since the composition of embeddings is an embedding,

(7) (Structure theorem) any refined frame for L is embeddable in the universal frame $F_L(\alpha)$, for some α.

In other words, the class of all refined frames for L can with no damage be described as formed by all weak subframes of all universal frames. This explains our choice of the name universal.

3. Modal axiomatic classes

The aim of this section is to characterize modal axiomatic classes of frames in terms of closure under specific frame constructions. The usual approach, which we also follow, is based on the idea of transferring, through modal duality, Birkhoff's theorem from modal algebras to frames. So, we certainly need closure under subframes, contractions and disjoint unions (dual of subalgebras, homomorphic images and direct products respectively). In addition, for example, the well known theorems by Goldblatt–Thomason [16] and van Benthem [3] about classes of Kripke frames, require closure under new constructions, namely that of state of affairs (SA-based) frames and that of ultrafilter extensions, respectively.

Since we here fix our attention on classes of refined frames, for which what we introduced so far (including our notion of weak subframe) is sufficient, we now put these results aside and later prove them as corollaries. This is possible since the above rather ad hoc constructions can easily be described in terms of our definitions and modal duality.

We begin by recalling some definition and notations. A class of frames K is called *modal axiomatic* if $K = \{\boldsymbol{F} : \boldsymbol{F} \models \Gamma\}$ for some set of modal formulae. We will use $\mathrm{RFr}(\Gamma)$ to denote the class of refined frames in which Γ is valid. In the other direction, for any class of frames K, we put $L(K) = \bigcap \{L\boldsymbol{F} : \boldsymbol{F} \in K\}$. The operators $\mathrm{RFr}(-)$ and $L(-)$ behave like their correspondent in classical model theory; for instance, any class K of refined frames is contained in $\mathrm{RFr}(L(K))$, which actually is the minimal modal axiomatic class containing K. So

(1) for any class K of refined frames, K is modal axiomatic iff $K = \mathrm{RFr}(L(K))$.

More typical here is the link between classes of frames and classes of modal algebras. For any class of frames K, we put $K^* = \{\boldsymbol{A} : \boldsymbol{A} \cong \boldsymbol{F}^* \text{ for some } \boldsymbol{F} \in K\}$; note that K^* is by definition closed under isomorphisms. Of course, $L(K) = L(K^*)$ because $L\boldsymbol{F} = L\boldsymbol{F}^*$ for every frame $\boldsymbol{F}$. Recall that for any set of formulae Γ, $\mathrm{MA}(\Gamma)$ is the equational class of modal algebras in which Γ is valid, that is $\boldsymbol{A} \in \mathrm{MA}(\Gamma)$ iff $\Gamma \subseteq L\boldsymbol{A}$. So, if K is any class of frames, $\boldsymbol{F}^* \in \mathrm{RFr}(L(K))^*$ iff $L(K) \subseteq L\boldsymbol{F}$ iff $\boldsymbol{F}^* \in \mathrm{MA}(L(K))$, and hence

(2) for every class K of frames, $\mathrm{RFr}(L(K))^* = \mathrm{MA}(L(K))$

since both classes are closed under isomorphisms. In particular, since $K = \mathrm{RFr}(L(K))$ iff K is modal axiomatic,

(3) if K is a modal axiomatic class, then K^* is an equational class.

Under which conditions on K can we prove the converse? If K^* is equational, then $K^* = \mathrm{MA}(L(K))$ because of Birkhoff's theorem [17, p. 171] and $L(K) = L(K^*)$. Of course, if $\boldsymbol{F} \in K$, then $\boldsymbol{F} \models L(K)$. Conversely, assume $\boldsymbol{F}$ is any refined frame such that $\boldsymbol{F} \models L(K)$; then $\boldsymbol{F}^* \in K^*$ and hence, by the definition of K^*, there is $\boldsymbol{G} \in K$ such that $\boldsymbol{G}^* \cong \boldsymbol{F}^*$. At this point, to be able to conclude that $\boldsymbol{F} \in K$, as we wish, it is enough that K is closed under biduals, isomorphisms and weak subframes. In fact, under such assumptions, $\boldsymbol{G} \in K$ implies $(\boldsymbol{G}^*)_* \in K$, hence $(\boldsymbol{F}^*)_* \in K$ because $(\boldsymbol{G}^*)_* \cong (\boldsymbol{F}^*)_*$, and finally $\boldsymbol{F} \in K$ by (2.6) because $\boldsymbol{F}$ is refined. Let us give a number to this partial result:

(4) let K be a class of refined frames closed under biduals, isomorphisms and weak subframes and assume K^* is equational; then K is modal axiomatic.

Note that the notion of weak subframe, or embedding, is exactly what we need to express the fact that

(5) for every frame $\boldsymbol{F}$, $(\boldsymbol{F}^*)_* \in K$ implies $\boldsymbol{F} \in K$

which is the crucial step to obtain (4). (5) could previously be obtained only by requiring the complement of K to be closed under biduals, that is: $F \notin K$ implies $(F^*)_* \notin K$. Though not very natural, such condition, together with closure under isomorphisms and biduals, can be used to obtain (4) in case K contains also frames which are not refined. With some caution, also the results below can be extended to all frames in a similar way.

We now want to get rid of the assumption in (4) that K^* is equational and substitute it with more direct assumptions on K. The idea once again is to use modal duality and transfer to frames known results of universal algebra. It is well known that K^* is equational iff $H(K^*) \subseteq K^*$, $S(K^*) \subseteq K^*$ and $P(K^*) \subseteq K^*$, where H, S and P are the usual operators forming all homomorphic images, subalgebras and direct products respectively (this fact is actually taken as the definition itself in [17, p. 152]). We then have:

(6) let K be a class of refined frames closed under isomorphisms and biduals; then
 (i) if K is closed under subframes, then $H(K^*) \subseteq K^*$;
 (ii) if K is closed under contractions, then $S(K^*) \subseteq K^*$;
 (iii) if K is closed under disjoint unions, then $P(K^*) \subseteq K^*$.

Proof. After modal duality, all the proofs are based on the fact that $A \in K^*$ iff $A_* \in K$, which holds since K is closed under isomorphisms and biduals. Thus we give only the proof of (iii). Let $(A_i)_{i \in I}$ be a family of modal algebras in K^*. Then for every $i \in I$, $A_{i*} \in K$ and, since K is closed under disjoint unions, $\sum_{i \in I} A_{i*} \in K$. By the definition of K^*, $(\sum_{i \in I} A_{i*})^* \in K^*$ and hence $\prod_{i \in I} A_i \in K^*$ by (II.5.24).

By (6) above, K^* is equational whenever K is closed under subframes, contractions and disjoint unions, beside biduals. Therefore, to obtain that K is modal axiomatic it is enough to substitute in (4) the assumption that K^* is equational with closure of K under contractions. On the other hand, by the results of the preceding section, every class which is modal axiomatic is, of course, closed under all such constructions. We thus have proved:

(7) a class K of refined frames is modal axiomatic iff K is closed under biduals, weak subframes, contractions and disjoint unions.

Results analogous to (4) and (7) for classes of descriptive frames are now an easy corollary (cf. [15, Section 12]).

The above characterization allows us to recognize a modal axiomatic class once we have it already, but does not give a method to construct it. We conclude this section with two such methods. The first has a given logic L as starting point, and is immediate: putting (2.7) and (1) together, we have

(8) a class K of refined frames is modal axiomatic iff K consists of all universal frames $F_{L(K)}(\alpha)$ together with all (isomorphic copies of) their weak subframes.

In other words, (8) says that, up to isomorphisms, RFr(L) is exactly the class of all universal frames for L and their weak subframes.

The second method shows how RFr($L(K)$), the minimal modal axiomatic class containing K, can be obtained from K by suitably applying some frame constructions. It is obtained by dualizing the description of the minimal equational class of modal algebras containing K^* as HSP(K^*). To save many words in the sequel, let us introduce some operators acting on classes of frames. And to save many problems, from now on we confine ourselves to refined frames. So, for any class of frames K, $W(K)$ is the class of subframes of some $G \in K$, $C(K)$ is the class of images of contractions from some $G \in K$, and $U(K)$ is the class of all disjoint unions of frames in K. In addition, we need also the operators B and W_s, where $B(K) = \{(F^*)_* : F \in K\}$ and $W_s(K) = \{F : F \subseteq_w G$ for some $G \in K\}$. It is easy to see that, for all operators O introduced, $O(K)$ is closed under isomorphisms if K is. We thus often omit to mention isomorphisms. We then have:

(9) for any class K of refined frames, RFr($L(K)$) $= W_sCBU(K)$; that is, the minimal modal axiomatic class containing K is formed by weak subframes of the image under a contraction of the bidual of a disjoint union of frames in K.

Proof. If $F \in W_sCBU(K)$, then $LK \subseteq LF$ because all the operators indicated preserve validity, and therefore $F \in$ RFr($L(K)$). Conversely, let $F \in$ RFr($L(K)$). Since F is isomorphic to a weak subframe of $(F^*)_*$, it is enough to show that $(F^*)_* \in WCBU(K)$. Now, from $F \in$ RFr($L(K)$) we have $F \vDash L(K)$ and hence, by (2) and the equality HSP(K) $=$ MA($L(K)$), $F^* \in$ HSP(K^*). A picture is

$$\prod_{i \in I} G_i^* \supseteq B \twoheadrightarrow F^*$$

where $G_i \in K$ for every $i \in I$. After duality, this becomes

$$\left(\prod_{i \in I} G_i^*\right)_* \twoheadrightarrow B_* \supseteq (F^*)_*$$

that is, since $(\prod_{i \in I} G_i^*)_* \cong ((\sum_{i \in I} G_i)^*)_*$ by (II.5.23), $(F^*)_* \in WCBU(K)$ as we wanted.

We can now give a dual form also to the fact that the single free algebra on ω generators is enough to generate the whole equation class [17, p. 172]. Just recall that $L(F_{L(K)}(\omega)) = L(K)$ and apply (9) to $\{F_{L(K)}(\omega)\}$ to obtain:

(10) for any class K of refined frames, K is modal axiomatic iff $K = W_sCBU(F_{L(K)}(\omega))$.

4. Some corollaries on Kripke frames

Throughout this section K will denote a class of Kripke frames. As previously remarked, every Kripke frame is indeed refined, but, in spite of this, we can not apply the results of the preceding section without modifications (an example of this is (3.3)). So, we still have to characterize modal axiomatic classes of Kripke frames. However, after the results of Section 3, we can have a new approach to the problem. In fact, we can do without diagrams and equational classes of algebras and operate only with refined frames and frame constructions. Let $\mathrm{KFr}(L(K))$ denote the minimal modal axiomatic class of Kripke frames containing K. So,

(1) for any class K of Kripke frames, K is modal axiomatic iff $K = \mathrm{KFr}(L(K))$.

Since $F \in \mathrm{KFr}(L(K))$ iff F is a Kripke frame and $F \in \mathrm{RFr}(L(K))$, we immediately obtain from (3.8) and (3.9) respectively

(2) for any class K of Kripke frames, K is modal axiomatic iff K consists of all (isomorphic copies of) Kripke subframes of all universal frames $F_{L(K)}(\alpha)$

and

(3) for any class K of Kripke frames, $\mathrm{KFr}(L(K))$ is, up to isomorphisms, exactly the class formed by Kripke frames in $W_sCBU(K)$.

In other words, (2) and (3) say that $\mathrm{KFr}(L(K))$ is obtained by considering, at first, K as any class of refined frames, and then by getting rid of every refined frame which is not also a Kripke frame.

When K is closed under disjoint unions, (3) is easily turned into

(4) for any class K of Kripke frames, K is modal axiomatic iff K is closed under isomorphisms, disjoint unions and for every Kripke frame F, $F \in W_sCB(K)$ implies $F \in K$

and the theorems of Goldblatt and Thomason and of van Benthem are then obtained as corollaries. For every frame F, we define F^d to be the *discretization* of F, that is the Kripke, or discrete, frame $(X, r, P(X))$ underlying F. The definition of 'state of affairs' frame F based on a given Kripke frame G, briefly SA-based on G, introduced and somehow heuristically justified in [16], can then be expressed in our terms. We say that a refined frame F is *SA-based* on a Kripke frame G, if there is a (general) frame H such that $H^d = G$ and F is embeddable in $(H^*)_*$. Note that with every refined frame also all its isomorphic copies are taken into account. The following result explains why SA-based frames could be used to characterize modal axiomatic classes:

(5) for every refined frame F and every Kripke frame G, F is SA-based on G iff F is (isomorphic to) a frame in $W_sCB(G)$.

Proof. Note that, since G is discrete, the condition $H^d = G$ is equivalent to $H^* \in S(G^*)$. But it is easy to check that $H^* \in S(G^*)$ is equivalent, modulo some isomorphisms, to $(H^*)_* \in CB(G)$. Replacing this in the definition of SA-based frames, we obtain the claim.

Using duality, it is then very simple to prove also that F is SA-based on G iff $F^* \in HS(G^*)$, which was the key step in [16]. We can now immediately obtain:

(6) (Goldblatt–Thomason theorem) a class K of Kripke frames is modal axiomatic iff K is closed under isomorphisms, disjoint unions and SA-based constructions.

Proof. We know that any modal axiomatic class is closed under isomorphisms and disjoint unions. Since B, C, W_s preserve the validity of modal formulae, by (5) it is also closed under SA-based constructions. The converse immediately follows from (5) and (4) above.

In [3], van Benthem characterizes the class of Kripke frames modally definable by a canonical set of modal formulae using the notion of ultrafilter extension. The *ultrafilter extension* ue(F) of a Kripke frame F is, in our notation, simply the frame $((F^*)_*)^{\mathrm{d}}$. Before going down to the proof of van Benthem's theorem it is useful to note that:

(7) for any frame F and G,
(i) if $F \in C(G)$, then $F^{\mathrm{d}} \in C(G^{\mathrm{d}})$;
(ii) if $F \in W_s(G)$, then $F^{\mathrm{d}} \in W_s(G^{\mathrm{d}})$.

In particular, if F is a Kripke frame, $F^{\mathrm{d}} = F$ and hence, by (7(ii)), F is embeddable in ue(F). Therefore, by (1.4):

(8) if F is a Kripke frame, then $C(\mathrm{ue}(F)) \subseteq CF$.

Recall that a set Γ of modal formulae is said to be *canonical* if, for every descriptive frame F, $F \vDash \Gamma$ implies $F^{\mathrm{d}} \vDash \Gamma$. We then have:

(9) (van Benthem theorem) a class of Kripke frames K is of the form $\{F : F \vDash \Gamma\}$ for a canonical set Γ of modal formulae, iff K is closed under subframes, contractions and disjoint unions, while both K and its complement are closed under ultrafilter extensions.

Proof. Assume $K = \mathrm{KFr}(\Gamma)$ with Γ canonical. Obviously, K is then closed under usual frame constructions. Moreover, if $\mathrm{ue}(F) \vDash \Gamma$, then by (8) also $F \vDash \Gamma$ and so the complement of K is closed under ultrafilter extensions. Finally, if $F \vDash \Gamma$, then obviously $(F^*)_* \vDash \Gamma$ and hence also $\mathrm{ue}(F) \vDash \Gamma$ because Γ is canonical.

Conversely, let $F \in W_s CB(K)$, F a Kripke frame. Then there is a frame $G \in K$ such that, up to isomorphisms, $(F^*)_* \in WC((G^*)_*)$, and so, by (7), $\mathrm{ue}(F) \in WC(\mathrm{ue}(G))$. Then the closure conditions guarantee that $\mathrm{ue}(F) \in K$, from which also $F \in K$ since the complement of K is closed under ultrafilter extensions. Thus we can apply (4) and obtain that K is modal axiomatic. So $K = \mathrm{KFr}(L(K))$ and the proof is complete once we show that $L(K)$ is canonical. Let F be a descriptive frame such that $F \vDash L(K)$. Then by (3.9) $F \in WCB(K)$, that is $F \in WC((G^*)_*)$, for some $G \in K$. But then, by (7), we have $F^{\mathrm{d}} \in WC(\mathrm{ue}(G))$, which implies $F^{\mathrm{d}} \in K$, that is $F^{\mathrm{d}} \vDash L(K)$.

We conclude with still another application of duality, namely a proof of equivalence between the notion of canonicity above (called d-persistence in [15]) and that introduced by Fine in [13]. This solves a problem raised in [3] (and simplifies terminology). We here say that a logic is Fine-canonical (canonical in [13]) if L is valid in every canonical frame for L, that is if $F_L(\alpha)^d \models L$ for every α. We can always restrict to logics, rather than sets of formulae, because obviously Γ is canonical iff $L = L(\mathrm{Fr}(\Gamma))$ is canonical. We then have:

(10) for every logic L, L is canonical iff L is Fine-canonical.

Proof. If L is canonical, then from $F_L(\alpha) \models L$ we obtain $F_L(\alpha)^d \models L$, and hence L is also Fine-canonical. Conversely, assume L is Fine-canonical and let $F \models L$ for some descriptive frame F. Then F^* is a homomorphic image of $F_L(\alpha)^*$ for some α, and hence by duality $F \cong (F^*)_*$ is (isomorphic to) a subframe of $F_L(\alpha)$. But then by (7(i)) we also have $F^d \subseteq F_L(\alpha)^d$, and hence $F^d \models L$ follows from the assumption $F_L(\alpha)^d \models L$.[11]

Footnotes

[1] Of course, the choice of some definitions, in particular that of space, is due to the fact that we keep modal logic in mind. Forgetting modal logic, the content of Chapter II could be more generally expressed in the framework of frames (in the categorical sense, cf. [18]) and locales. On the other hand, our basic duality shows that a few results in that area (say in the second chapter of [18]) can be extended to the case in which continuous relations, rather than functions, are considered.

[2] One could at this point follow an alternative indirect path to obtain (1). In fact, classical completeness gives: $L^c \vdash \varphi^t = 1$ iff $\varphi^t = 1$ is an identity of every modal algebra. Thus (1) is proved if we establish that: $L^c \vdash \varphi^t = 1$ iff $L \vdash \varphi$. The implication from right to left is immediate. The converse is obtained as a simple application of normalization theorem for proofs in natural deduction. In fact, any normal derivation of $\varphi^t = 1$ from L^c is easily transformed in a proof of φ from L.

[3] We have some doubts about the conception of Kripke frames as universe of possible worlds, because, with such highly metaphysical assumptions it is not clear whether the interpretation can help to understand modal logic, as it should, or viceversa. Still we believe that Kripke frames are acceptable and actually help comprehension if each of them is conceived as a single world (population, environment, etc.) populated with thinking subjects (individuals, persons, etc.); in this case accessibility from x to y means that the subject x has access to the opinions (beliefs, dogmas, accepted truths, etc.) of y. Then x believes in $\Box\varphi$ iff all opinions x can know about φ say that φ is true. Note that in the present assumptions, an individual has an opinion on everything.

[4] Actually, we could at this point look at things the other way round and derive Kripke's definition of validity from the following: for each valuation V of variables in T, the valuation of formulae is defined to be the unique homomorphism from A_K to F^* extending the function $V' : [p_i]_K \mapsto V(p_i)$.

[5] Comparing the proofs of (9) and (1.3), the unprejudiced reader will note their similarity in structure, due to the requirement of closure under substitutions. This shows that the widespread opinion that "algebraic semantics is syntax in disguised form" is, to say the least, superficial. Actually, this suggests to study the possibility of a modified form of Henkin construction, where logics (i.e., including closure under substitutions), rather than maximally consistent sets of formulae, are used.

[6] Adopting the picture sketched in footnote 3, the construction of A_* might be described as follows. If A is taken as the field of possible values, then $U(A)$ is the idealized world with exactly one

individual x for each complete theory U_x (and note that also individuals with no principles are tolerated, alias non-principal ultrafilters), x has access to the opinions of y iff y agrees on the truth of what is necessary for x, and finally a value is identified with the class of individuals holding it.

[7] One could at this point directly obtain modal duality (4.12) (but not modal adjunction (4.10)), without passing through basic adjunction of Section 3. In fact, in the category **DFra**, of descriptive frames and contractions, γ is a natural isomorphism by (23), (24) and (26). We leave this as an exercise; the reward is that one can skip Section 3 and Section 4, except (1), (2), the definition of functors preceding (8), and (12).

[8] The definition we adopted can be found in [14, p. 81]. In fact, *contrary* to McLane, we believe that dismissing *contra*variant functors in favour of opposite categories would here prevent us from the fun (!?) of interchanging points with sets when passing from a category to the other, like shown in Section 2.

[9] From our treatment of modal duality one can easily derive some related results. For instance, a duality between Kripke frames and complete atomic modal algebras, as proved in [35], is obtained modifying the definition of $(\ )_*$ by always restricting to principal ultrafilters. Moreover, with little adjustments it has been extended to tense logic by Paola Unterholzner [36]. Also, it is possible to adapt it to (propositional) dynamic logic (cf. [22], where however only objects are considered).

[10] The reader will see that we never use the full strength of the preservation results just proved. However, we have a suggestion at least, namely to use them to characterize classes of frames in which a given consequence, rather than a logic, is valid.

[11] A modal formula φ is said to be *natural* if for any refined frame F, $F \vDash \varphi$ implies $F^d \vDash \varphi$ (cf. [33] and [13]). Thus every natural formula is canonical. After reading a preprint of this paper, J. van Benthem has obtained an extension of his theorem (9) above to sets of natural formulae (cf. [7]).

References

[1] J.L. Bell and A.B. Slomson, Models and Ultraproducts; An Introduction (North-Holland, Amsterdam, 1969).

[2] J.F.A.K. van Benthem, Modal correspondence theory, Doctoral dissertation, University of Amsterdam, 1976.

[3] J.F.A.K. van Benthem, Canonical modal logics and ultrafilter extensions, J. Symbolic Logic 44 (1979) 1–8.

[4] J.F.A.K. van Benthem, Some kinds of modal completeness, Studia Logica 39 (1980) 125–141.

[5] J.F.A.K. van Benthem, Two simple incomplete modal logics, Theoria 44 (1978) 25–37.

[6] J.F.A.K. van Benthem, Modal correspondence theory, in: D. Gabbay and F. Guenthner, eds., Handbook of Philosophical Logic, Vol. II, (Reidel, Dordrecht, 1984) 167–247.

[7] J.F.A.K. van Benthem, Notes on modal definability, Report 86-11, Dept. of Math., University of Amsterdam, 1986.

[8] G. Boolos and G. Sambin, An incomplete system of modal logic, J. Philos. Logic 14 (1985) 351–358.

[9] R.A. Bull, Collective review in the J. Symbolic Logic 47 (1982) 440–445.

[10] R. Engelking, Outline of General Topology (North-Holland, Amsterdam, 1968).

[11] L. Esakia, Topological Kripke models (Russian), Soviet Math. Dokl. 15 (1974) 147–151.

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

[13] K. Fine, Some connections between elementary and modal logic, in: S. Kanger, ed., Proceedings of the Third Scandinavian Logic Symposium (North-Holland, Amsterdam, 1975) 15–31.

[14] P. Freyd, Abelian Categories: An Introduction to the Theory of Functors (Harper and Row, New York, 1964).

[15] R.I. Goldblatt, Metamathematics of modal logic, Reports on Mathematical Logic 6 (1976) 41–77 and 7 (1977) 21–52.

[16] R.I. Goldblatt and S.K. Thomason, Axiomatic classes in propositional modal logic, in: J.N. Crossley, ed., Algebra and Logic (Springer, Berlin, 1975) 163–173.
[17] G. Grätzer, Universal Algebra (Van Nostrand, New York, 1968).
[18] P.T. Johnstone, Stone Spaces (Cambridge Univ. Press, Cambridge, 1982).
[19] B. Jónsson and A. Tarski, Boolean algebras with operators, Amer. J. Math. 73 (1951) 891–939.
[20] P.R. Halmos, Algebraic Logic (Chelsea, New York, 1962).
[21] P.R. Halmos, Lectures on Boolean algebras (Van Nostrand, New York, 1963).
[22] D. Kozen, On the duality of dynamic algebras and Kripke models, in: E. Engeler, ed., Logic of Programs, Proceedings 1979, Lecture Notes in Computer Science 125 (Springer, Berlin, 1981) 1–11.
[23] S. Kripke, Semantic analysis of modal logic, Z. Math. Logik Grundl. Math. 9 (1963) 67–96.
[24] E.J. Lemmon, Algebraic semantics for modal logics, J. Symbolic Logic 31 (1966) 46–65 and 191–218.
[25] S. Mac Lane, Categories for the Working Mathematician (Springer, Berlin, 1971).
[26] R. Magari, Representation and duality theory for diagonalizable algebras (the algebraization of theories which express Theor, IV), Studia Logica 34 (1975) 305–313.
[27] B. Pareigis, Categories and Functors (Academic Press, New York, 1970).
[28] W. Rautenberg, Klassische und nichtklassische Aussagenlogik (Vieweg, Braunschweig, 1979).
[29] H. Sahlqvist, Completeness and correspondence in the first and second order semantics for modal logic, in: S. Kanger, ed., Proceedings of the Third Scandinavian Logic Symposium (North-Holland, Amsterdam, 1975) 110–143.
[30] G. Sambin, A simpler proof of Sahlqvist's theorem on completeness of modal logics, Bull. Sect. Logic, Polish Acad. Science 9 (1980) 50–56.
[31] G. Sambin, Topology and categorical duality in the study of semantics for modal logics, Report n.104, Fachbereich Mathematik, Freie Universität Berlin, December 1979.
[32] K. Segerberg, An essay in classical modal logic (Uppsala, 1971).
[33] S.K. Thomason, Semantical analysis of tense logic, J. Symbolic Logic 37 (1972) 150–158.
[34] S.K. Thomason, An incompleteness theorem in modal logic, Theoria 40 (1974) 30–34.
[35] S.K. Thomason, Categories of frames for modal logic, J. Symbolic Logic 40 (1975) 439–442.
[36] P. Unterholzner, Algebraic and relational semantics for tense logics, Rend. Sem. Mat. Univ. Padova 65 (1981) 119–128.