4

Completeness

This chapter is about the completeness — and incompleteness — of normal modal logics. As we saw in Section 1.6, normal modal logics are collections of formulas satisfying certain simple closure conditions. They can be specified either syntactically or semantically, and this gives rise to the questions which dominate the chapter: Given a semantically specified logic, can we give it a syntactic characterization, and if so, how? And: Given a syntactically specified logic, can we give it a semantic characterization (and in particular, a characterization in terms of frames), and if so, how? To answer either type of question we need to know how to prove (soundness and) *completeness* theorems, and the bulk of the chapter is devoted to developing techniques for doing so.

The chapter has two major parts. The first, comprising the first four sections, is an introduction to basic completeness theory. It introduces canonical models, explains and applies the completeness-via-canonicity proof technique, discusses the Sahlqvist Completeness Theorem, and proves two fundamental limitative results. The material introduced in these sections (which are all on the basic track) is needed to follow the second part and the algebraic investigations of Chapter 5.

In the second part of the chapter we turn to the following question: what are we to do when canonicity fails? (As will become clear, canonicity failure is a fact of life for temporal logic, propositional dynamic logic, and other applied modal languages.) This part of the chapter is technique oriented: it introduces five important ways of dealing with such difficulties.

Chapter guide

Section 4.1: Preliminaries (Basic track). This section introduces the fundamental concepts: normal modal logics, soundness, and completeness.

Section 4.2: Canonical Models (Basic track). Canonical models are introduced, and the fundamental Canonical Model Theorem is proved.

Section 4.3: Applications (Basic track). This section discusses the key concept of canonicity, and uses completeness-via-canonicity arguments to put canonical models to work. We prove completeness results for a number of modal and temporal logics, and finish with a discussion of the *Sahlqvist Completeness Theorem*.

Section 4.4: Limitative Results (Basic track). We prove two fundamental limitative results: not all normal logics are canonical, and not all normal logics are characterized by some class of frames. This section concludes our introduction to basic completeness theory.

Section 4.5: Transforming the Canonical Model (Basic track). Often we need to build models with properties for which we lack a canonical formula. What are we to do in such cases? This section introduces one approach: use transformation methods to try and massage the 'faulty' canonical model into the required shape.

Section 4.6: Step-by-Step (Basic track). Sometimes we can cope with canonicity failure using the step-by-step method. This is a technique for building models with special properties inductively.

Section 4.7: Rules for the Undefinable (Basic track). Special proof rules (that in a certain sense manage to express undefinable properties of models and frames) sometimes allow us to construct special canonical models containing submodels with undefinable properties.

Section 4.8: Finitary Methods I (Basic track). We discuss a method for proving weak completeness results for non-compact logics: *finite* canonical models. We use such models to prove the completeness of propositional dynamic logic.

Section 4.9: Finitary Methods II (Advanced track). This section further explores finitary methods, this time the direct use of filtrations. We illustrate this with an analysis of the normal logics extending **S4.3**.

4.1 Preliminaries

In this section we introduce some of the fundamental concepts that we will use throughout the chapter. We begin by defining *modal logics* — these could be described as propositional logics in a modal language.

Throughout the chapter we assume we are working with a fixed countable language of proposition letters.

Definition 4.1 (Modal Logics) A *modal logic* Λ is a set of modal formulas that contains all propositional tautologies and is closed under *modus ponens* (that is, if $\phi \in \Lambda$ and $\phi \to \psi \in \Lambda$ then $\psi \in \Lambda$) and *uniform substitution* (that is, if ϕ belongs to Λ then so do all of its substitution instances). If $\phi \in \Lambda$ we say that ϕ is a *theorem* of Λ and write $\vdash_\Lambda \phi$; if not, we write $\nvdash_\Lambda \phi$. If Λ_1 and Λ_2 are modal logics such that $\Lambda_1 \subseteq \Lambda_2$, we say that Λ_2 is an *extension* of Λ_1. In what follows, we usually drop the word 'modal' and talk simply of 'logics'. $\dashv$

Note that modal logics contain all substitution instances of the propositional tautologies: for example, $\Diamond p \vee \neg\Diamond p$, belongs to every modal logic. Even though such substitution instances may contain occurrences of $\Diamond$ and $\Box$, we still call them *tautologies*. Clearly tautologies are valid in every class of models.

Example 4.2 (i) The collection of all formulas is a logic, the *inconsistent logic*.

(ii) If $\{\Lambda_i \mid i \in I\}$ is a collection of logics, then $\bigcap_{i\in I} \Lambda_i$ is a logic.

(iii) Define Λ_{S} to be $\{\phi \mid \mathfrak{G} \Vdash \phi$, for all structures $\mathfrak{G} \in \mathsf{S}\}$, where S is any class of frames or any class of general frames. Λ_{S} is a logic. If S is the singleton class $\{\mathfrak{G}\}$, we usually call this logic $\Lambda_{\mathfrak{G}}$, rather than $\Lambda_{\{\mathfrak{G}\}}$.

(iv) If M is a class of models, then Λ_{M} need *not* be a logic. Consider a model $\mathfrak{M}$ in which p is true at all nodes but q is not. Then $p \in \Lambda_{\mathfrak{M}}$, but $q \notin \Lambda_{\mathfrak{M}}$. But q is obtainable from p by uniform substitution. $\dashv$

It follows from Examples 4.2(i) and 4.2(ii) that there is a smallest logic containing any set of formulas Γ; we call this *the logic generated by* Γ. For example, the logic generated by the empty set contains all the tautologies and nothing else; we call it **PC** and it is a subset of every logic. This generative perspective is essentially *syntactic*. However, as Example 4.2(iii) shows, there is a natural *semantic* perspective on logics: both frames and general frames give rise to logics in an obvious way. Even the empty class of frames gives rise to a logic, namely the inconsistent logic. Finally, Example 4.2(iv) shows that models may fail to give rise to logics. This 'failure' is actually the behavior we should expect: as we discussed in Section 1.6, genuine logics arise at the level of *frames*, via the concept of *validity*.

Definition 4.3 Let $\psi_1, \ldots, \psi_n, \phi$ be modal formulas. We say that ϕ is *deducible in propositional calculus from assumptions* $\psi_1, \ldots, \psi_n$ if $(\psi_1 \wedge \cdots \wedge \psi_n) \to \phi$ is a tautology. $\dashv$

All logics are closed under deduction in propositional calculus: if ϕ is deducible in propositional calculus from assumptions $\psi_1, \ldots, \psi_n$, then $\vdash_\Lambda \psi_1, \ldots, \vdash_\Lambda \psi_n$ implies $\vdash_\Lambda \phi$.

Definition 4.4 If $\Gamma \cup \{\phi\}$ is a set of formulas then ϕ *is deducible in* Λ *from* Γ (or: ϕ *is* Λ*-deducible from* Γ) if $\vdash_\Lambda \phi$ or there are formulas $\psi_1,\ldots,\psi_n \in \Gamma$ such that

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

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

It is a simple exercise in propositional logic to check that a set of formulas Γ is Λ-inconsistent if and only if there is a formula ϕ such that $\Gamma \vdash_\Lambda \phi \wedge \neg\phi$ if and only if for all formulas ψ, $\Gamma \vdash_\Lambda \psi$. Moreover, Γ is Λ-consistent if and only if every finite subset of Γ is. (That is, our notion of deducibility has the *compactness* property.) From now on, when Λ is clear from context or irrelevant, we drop explicit references to it and talk simply of 'theorems', 'deducibility', 'consistency' and 'inconsistency', and use the notation $\vdash \phi$, $\Gamma \vdash \phi$, and so on.

The preceding definitions merely generalize basic ideas of propositional calculus to modal languages. Now we come to a genuinely *modal* concept: *normal modal logics*. These logics are the focus of this chapter's investigations. We initially restrict our discussion to the basic modal language; the full definition is given at the end of the section. As we discussed in Section 1.6, the following definition is essentially an abstraction from Hilbert-style approaches to modal proof theory.

Definition 4.5 A modal logic Λ is *normal* if it contains the formulas:

(K) $\Box(p \to q) \to (\Box p \to \Box q)$,
(Dual) $\Diamond p \leftrightarrow \neg\Box\neg p$,

and is closed under *generalization* (that is, if $\vdash_\Lambda \phi$ then $\vdash_\Lambda \Box\phi$). $\dashv$

Syntactic issues do not play a large role in this book; nonetheless, readers new to modal logic should study the following lemma and attempt Exercise 4.1.2.

Lemma 4.6 *For any normal logic Λ, if $\vdash_\Lambda \phi \leftrightarrow \psi$ then $\vdash_\Lambda \Diamond\phi \leftrightarrow \Diamond\psi$.*

Proof. Suppose $\vdash_\Lambda \phi \leftrightarrow \psi$. Then $\vdash_\Lambda \phi \to \psi$ and $\vdash_\Lambda \psi \to \phi$. If we can show that $\vdash_\Lambda \Diamond\phi \to \Diamond\psi$ and $\vdash_\Lambda \Diamond\psi \to \Diamond\phi$, the desired result follows. Now, as $\vdash_\Lambda \phi \to \psi$, we have $\vdash_\Lambda \neg\psi \to \neg\phi$, hence by generalization $\vdash_\Lambda \Box(\neg\psi \to \neg\phi)$. By uniform substitution into the K axiom we obtain $\vdash_\Lambda \Box(\neg\psi \to \neg\phi) \to (\Box\neg\psi \to \Box\neg\phi)$. It follows by modus ponens that $\vdash_\Lambda \Box\neg\psi \to \Box\neg\phi$. Therefore, $\vdash_\Lambda \neg\Box\neg\phi \to \neg\Box\neg\psi$, and two uses of Dual yield $\vdash_\Lambda \Diamond\phi \to \Diamond\psi$, as desired. As $\vdash \psi \to \phi$, an analogous argument shows that $\vdash_\Lambda \Diamond\psi \to \Diamond\phi$, and the result follows. $\dashv$

Remark 4.7 The above definition of normal logics (with or without Dual, depending on the choice of primitive operators) is probably the most popular way of stipulating what normal logics are. But it's not the only way. Here, for example, is a simple diamond-based formulation of the concept, which will be useful in our later algebraic work: a logic Λ is normal if it contains the axioms $\Diamond\bot \leftrightarrow \bot$ and $\Diamond(p \vee q) \leftrightarrow \Diamond p \vee \Diamond q$, and is closed under the following rule: $\vdash_\Lambda \phi \to \psi$ implies

$\vdash_\Lambda \Diamond\phi \to \Diamond\psi$. This formulation is equivalent to Definition 4.5, as the reader is asked to show in Exercise 4.1.2. $\dashv$

Example 4.8 (i) The inconsistent logic is a normal logic.

(ii) **PC** is not a normal logic.

(iii) If $\{\Lambda_i \mid i \in I\}$ is a collection of normal logics, then $\bigcap_{i\in I} \Lambda_i$ is a normal logic.

(iv) If F is any class of frames, then Λ_{F} is a normal logic.

(v) If G is any class of general frames, then Λ_{G} is a normal logic. (The reader is asked to prove this in Exercise 4.1.1.) $\dashv$

Examples 4.8(i) and 4.8(iii) guarantee that there is a smallest normal modal logic containing any set of formulas Γ. We call this the normal modal logic *generated* or *axiomatized* by Γ. The normal modal logic generated by the empty set is called $\mathbf{K}$, and it is the smallest (or minimal) normal modal logic: for any normal modal logic Λ, $\mathbf{K} \subseteq \Lambda$. If Γ is a non-empty set of formulas we usually denote the normal logic generated by Γ by $\mathbf{K\Gamma}$. Moreover, we often make use of Hilbert axiomatization terminology, referring to Γ as *axioms* of this logic, and say that the logic was generated using the *rules of proof* modus ponens, uniform substitution, and generalization. We justified this terminology in Section 1.6, and also asked the reader to prove that the logic $\mathbf{K\Gamma}$ consists of precisely those formulas that can be proved in a Hilbert-style derivation from the axioms in Γ using the standard modal proof rules (see Exercise 1.6.6).

Defining a logic by stating which formulas generate it (that is, extending the minimal normal logic $\mathbf{K}$ with certain axioms of interest) is the usual way of syntactically specifying normal logics. Much of this chapter explores such axiomatic extensions. Here are some of the better known axioms, together with their traditional names:

(4) $\Diamond\Diamond p \to \Diamond p$

(T) $p \to \Diamond p$

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

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

(.3) $\Diamond p \wedge \Diamond q \to \Diamond(p \wedge \Diamond q) \vee \Diamond(p \wedge q) \vee \Diamond(q \wedge \Diamond p)$

(L) $\Box(\Box p \to p) \to \Box p$

There is a convention for talking about the logics generated by such axioms: if $\mathrm{A}_1, \ldots, \mathrm{A}_n$ are axioms then $\mathbf{KA}_1 \ldots \mathbf{A}_n$ is the normal logic generated by $\mathrm{A}_1, \ldots, \mathrm{A}_n$. But irregularities abound. Many historical names are firmly entrenched, thus modal logicians talk of **T**, **S4**, **B**, and **S5** instead of **KT**, **KT4**, **KB** and **KT4B** respectively. Moreover, many axioms have multiple names. For example, the axiom we call L (for Löb) is also known as G (for Gödel) and W (for wellfounded); and

K	the class of all frames
K4	the class of transitive frames
T	the class of reflexive frames
B	the class of symmetric frames
KD	the class of right-unbounded frames
S4	the class of reflexive, transitive frames
S5	the class of frames whose relation is an equivalence relation
K4.3	the class of transitive frames with no branching to the right
S4.3	the class of reflexive, transitive frames with no branching to the right
KL	the class of finite transitive trees (*weak* completeness only)

Table 4.1. *Some Soundness and Completeness Results*

the axiom we call .3 has also been called H (for Hintikka). We adopt a fairly relaxed attitude towards naming logics, and use the familiar names as much as possible.

Now that we know what normal modal logics are, we are ready to introduce the two fundamental concepts linking the syntactic and semantic perspectives: *soundness* and *completeness*.

Definition 4.9 (Soundness) Let S be a class of frames (or models, or general frames). A normal modal logic Λ is *sound* with respect to S if $\Lambda \subseteq \Lambda_{\mathsf{S}}$. (Equivalently: Λ is *sound* with respect to S if for all formulas ϕ, and all structures $\mathfrak{S} \in \mathsf{S}$, $\vdash_\Lambda \phi$ implies $\mathfrak{S} \Vdash \phi$.) If Λ is sound with respect to S we say that S *is a class of frames* (or models, or general frames) *for* Λ. ⊣

Table 4.1 lists a number of well-known logics together with classes of frames for which they are sound. Recall that a *right-unboundedness* frame (W, R) is a frame such that $\forall x \exists y \, Rxy$. Also, a frame (W, R) satisfying $\forall x \forall y \forall z \, (Rxy \wedge Rxz \rightarrow (Ryz \vee y = z \vee Rzy))$ is said to have *no branching to the right*.

The *soundness* claims made in Table 4.1 (with the exception of the last one, which was shown in Example 3.9) are easily demonstrated. In all cases one shows that the axioms are valid, and that the three rules of proof (modus ponens, generalization, and uniform substitution) preserve validity on the class of frames in question. In fact, the proof rules preserve validity on *any* class of frames or general frames (see Exercise 4.1.1), so proving soundness boils down to checking the validity of the axioms. Soundness proofs are often routine, and when this is the case we rarely bother to explicitly state or prove them. But the concept of *completeness*, leads to the problems that will occupy us for the remainder of the chapter.

Definition 4.10 (Completeness) Let S be a class of frames (or models, or general frames). A logic Λ is *strongly complete* with respect to S if for any set of formulas

$\Gamma \cup \{\phi\}$, if $\Gamma \Vdash_{\mathsf{S}} \phi$ then $\Gamma \vdash_{\Lambda} \phi$. That is, if Γ semantically entails ϕ on S (recall Definition 1.35) then ϕ is Λ-deducible from Γ.

A logic Λ is *weakly complete* with respect to S if for any formula ϕ, if $\mathsf{S} \Vdash \phi$ then $\vdash_{\Lambda} \phi$. Λ is strongly complete (weakly complete) with respect to a single structure $\mathfrak{S}$ if Λ is strongly complete (weakly complete) with respect to $\{\mathfrak{S}\}$. ⊣

Note that weak completeness is the special case of strong completeness in which Γ is empty, thus strong completeness with respect to some class of structures implies weak completeness with respect to that same class. (The converse does *not* hold, as we will later see.) Note that the definition of weak completeness can be reformulated to parallel the definition of soundness: Λ is weakly complete with respect to S if $\Lambda_{\mathsf{S}} \subseteq \Lambda$. Thus, if we prove that a syntactically specified logic Λ is both sound and weakly complete with respect to some class of structures S, we have established a perfect match between the syntactical and semantical perspectives: $\Lambda = \Lambda_{\mathsf{S}}$. Given a semantically specified logic Λ_{S} (that is, the logic of some class of structures S of interest) we often want to find a simple collection of formulas Γ such that Λ_{S} is the logic generated by Γ; in such a case we sometimes say that Γ *axiomatizes* S.

Example 4.11 With the exception of **KL**, all the logics mentioned in Table 4.1 are strongly complete with respect to the corresponding classes of frames. However **KL** is only weakly complete with respect to the class of finite transitive trees. As we will learn in section 4.4, **KL** is not strongly complete with respect to this class of frames, or indeed with respect to any class of frames whatsoever. ⊣

These completeness results are among the best known in modal logic, and we will soon be able to prove them. Together with their soundness counterparts (given in Example 4.1), they constitute perspicuous semantic characterizations of important logics. **K4**, for example, is not just the logic obtained by enriching **K** with some particular axiom: it is precisely the set of formulas valid on all transitive frames. There is always something arbitrary about syntactic presentations; it is pleasant (and useful) to have these semantic characterizations at our disposal.

We make heavy use, usually without explicit comment, of the following result.

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

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

To conclude this section, we extend the definition of normal modal logics to arbitrary similarity types.

Definition 4.13 Assume we are working with a modal language of similarity type τ. A *modal logic* in this language is (as before) a set of formulas containing all tautologies that is closed under modus ponens and uniform substitution. A modal logic Λ is *normal* if for every operator $\triangledown$ it contains: the axiom $K^i_\triangledown$ (for all i such that $1 \leq i \leq \rho(\triangledown)$); the axiom Dual$_\triangledown$; and is closed under the generalization rules described below.

The required axioms are obvious polyadic analogs of the earlier K and Dual axioms:

(K$^i_\triangledown$) $$\triangledown(r_1, \ldots, p \to q, \ldots, r_{\rho(\triangledown)}) \to \to \big(\triangledown(r_1, \ldots, p, \ldots, r_{\rho(\triangledown)}) \to \triangledown(r_1, \ldots, q, \ldots, r_{\rho(\triangledown)})\big).$$

(Dual$_\triangledown$) $$\triangle(r_1, \ldots, r_{\rho(\triangledown)}) \leftrightarrow \neg\triangledown(\neg r_1, \ldots, \neg r_{\rho(\triangledown)}).$$

(Here $p, q, r_1, \ldots, r_{\rho(\triangledown)}$ are distinct propositional variables, and the occurrences K$^i_\triangledown$ of p and q occur in the i-th argument place of $\triangledown$.) Finally, for a polyadic operator $\triangledown$, generalization takes the following form:

$$\vdash_\Lambda \sigma \text{ implies } \vdash_\Lambda \triangledown(\bot, \ldots, \sigma, \ldots, \bot).$$

That is, an n-place operator $\triangledown$ is associated with n generalization rules, one for each of its n argument positions.

Note that these axioms and rules don't apply to *nullary* modalities. Nullary modalities are rather like propositional variables and — as far as the minimal logic is concerned — they don't give rise to any axioms or rules. ⊣

Definition 4.14 Let τ be a modal similarity type. Given a set of τ-formulas Γ, we define $\mathbf{K}_\tau\Gamma$, the normal modal logic *axiomatized* or *generated* by Γ, to be the smallest normal modal τ-logic containing all formulas in Γ. Formulas in Γ are called *axioms* of this logic, and Γ may be called an *axiomatization* of $\mathbf{K}_\tau\Gamma$. The normal modal logic generated by the empty set is denoted by $\mathbf{K}_\tau$. ⊣

Exercises for Section 4.1

4.1.1 Show that if G is any class of general frames, then Λ_{S} is a normal logic. (To prove this, you will have to show that the modal proof rules preserve validity on any general frame.)

4.1.2 First, show that the diamond-based definition of normal modal logics given in Remark 4.7 is equivalent to the box-based definition. Then, for languages of arbitrary similarity type, formulate a $\triangle$-based definition of normal modal logics, and prove it equivalent to the ∇-based one given in Definition 4.13.

4.1.3 Show that the set of all normal modal logics (in some fixed language) ordered by set theoretic inclusion forms a complete lattice. That is, prove that every family $\{\Lambda_i \mid i \in I\}$ of logics has both an infimum and a supremum. (An infimum is a logic Λ such that $\Lambda \subseteq \Lambda_i$ for all $i \in I$, and for any other logic Λ' that has this property, $\Lambda \subseteq \Lambda'$; the concept of a supremum is defined analogously, with '$\supseteq$' replacing '$\subseteq$'.)

4.1.4 Show that the normal logic generated by $\Box(p \wedge \Box p \to q) \vee \Box(q \wedge \Box q \to p)$ is sound with respect to the class of **K4.3** frames (see Table 4.1). Further, show that the normal modal logic generated by $\Box(\Box p \to q) \vee \Box(\Box q \to p)$ is *not* sound with respect to this class of frames, but that it is sound with respect to the class of **S4.3** frames.

4.2 Canonical Models

Completeness theorems are essentially model existence theorems — that is the content of Proposition 4.12. Given a normal logic Λ, we prove its strong completeness with respect to some class of structures by showing that every Λ-consistent set of formulas can be satisfied in some suitable model. Thus the fundamental question we need to address is: *how do we build (suitable) satisfying models*?

This section introduces the single most important answer: build models out of *maximal consistent sets of formulas*, and in particular, build *canonical models*. It is difficult to overstress the importance of this idea. In one form or another it underlies almost every modal completeness result the reader is likely to encounter. Moreover, as we will learn in Chapter 5, the idea has substantial algebraic content.

Definition 4.15 (Λ-MCSs) A set of formulas Γ is *maximal Λ-consistent* if Γ is Λ-consistent, and any set of formulas properly containing Γ is Λ-inconsistent. If Γ is a maximal Λ-consistent set of formulas then we say it is a Λ-MCS. $\dashv$

Why use MCSs in completeness proofs? To see this, first note that every point w in every model $\mathfrak{M}$ for a logic Λ is associated with a set of formulas, namely $\{\phi \mid \mathfrak{M}, w \Vdash \phi\}$. It is easy to check (and the reader should do so) that this set of formulas is actually a Λ-MCS. That is: if ϕ is true in some model for Λ, then ϕ belongs to a Λ-MCS. Second, if w is related to w' in some model $\mathfrak{M}$, then it is clear that the information embodied in the MCSs associated with w and w' is 'coherently related'. Thus our second observation is: models give rise to collections of coherently related MCSs.

The idea behind the canonical model construction is to try and turn these observations around: that is, to work backwards from collections of coherently related MCSs to the desired model. The goal is to prove a Truth Lemma which tells us that

'ϕ belongs to an MCS' is actually *equivalent* to 'ϕ is true in some model'. How will we do this? By building a special model — the *canonical model* — whose points are all MCSs of the logic of interest. We will pin down what it means for the information in MCSs to be 'coherently related', and use this notion to define the required accessibility relations. Crucially, we will be able to prove an Existence Lemma which states that there are enough coherently related MCSs to ensure the success of the construction, and this will enable us to prove the desired Truth Lemma.

To carry out this plan, we need to learn a little more about MCSs.

Proposition 4.16 (Properties of MCSs) *If Λ is a logic and Γ is a Λ-MCS then:*

(i) *Γ is closed under modus ponens: if ϕ, $\phi \to \psi \in \Gamma$, then $\psi \in \Gamma$;*
(ii) *$\Lambda \subseteq \Gamma$;*
(iii) *for all formulas ϕ: $\phi \in \Gamma$ or $\neg\phi \in \Gamma$;*
(iv) *for all formulas ϕ, ψ: $\phi \vee \psi \in \Gamma$ iff $\phi \in \Gamma$ or $\psi \in \Gamma$.*

Proof. Exercise 4.2.1. ⊣

As MCSs are to be our building blocks, it is vital that we have enough of them. In fact, any consistent set of formulas can be extended to a maximal consistent one.

Lemma 4.17 (Lindenbaum's Lemma) *If Σ is a Λ-consistent set of formulas then there is an Λ-MCS Σ^+ such that $\Sigma \subseteq \Sigma^+$.*

Proof. Let $\phi_0, \phi_1, \phi_2, \ldots$ be an enumeration of the formulas of our language. We define the set Σ^+ as the union of a chain of Λ-consistent sets as follows:

$$\begin{aligned} \Sigma_0 &= \Sigma \\ \Sigma_{n+1} &= \begin{cases} \Sigma_n \cup \{\phi_n\}, & \text{if this is } \Lambda\text{-consistent} \\ \Sigma_n \cup \{\neg\phi_n\}, & \text{otherwise} \end{cases} \\ \Sigma^+ &= \textstyle\bigcup_{n \geq 0} \Sigma_n. \end{aligned}$$

The proof of the following properties of Σ^+ is left as Exercise 4.2.2: (i) Σ_n is Λ-consistent, for all n; (ii) exactly one of ϕ and $\neg\phi$ is in Σ^+, for every formula ϕ; (iii) if $\Sigma^+ \vdash_\Lambda \phi$, then $\phi \in \Sigma^+$; and finally (iv) Σ^+ is a Λ-MCS. ⊣

We are now ready to build models out of MCSs, and in particular, to build the very special models known as canonical models. With the help of these structures we will be able to prove the Canonical Model Theorem, a universal completeness result for normal logics. We first define canonical models and prove this result for the basic modal language; at the end of the section we generalize our discussion to languages of arbitrary similarity type.

Definition 4.18 The canonical model $\mathfrak{M}^\Lambda$ for a normal modal logic Λ (in the basic language) is the triple $(W^\Lambda, R^\Lambda, V^\Lambda)$ where:

(i) W^Λ is the set of all Λ-MCSs;
(ii) R^Λ is the binary relation on W^Λ defined by $R^\Lambda wu$ if for all formulas ψ, $\psi \in u$ implies $\Diamond\psi \in w$. R^Λ is called the *canonical relation*.
(iii) V^Λ is the valuation defined by $V^\Lambda(p) = \{w \in W^\Lambda \mid p \in w\}$. V^Λ is called the *canonical* (or *natural*) *valuation*.

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

All three clauses deserve comment. First, the canonical valuation equates the truth of a propositional symbol at w with its membership in w. Our ultimate goal is to prove a Truth Lemma which will lift this 'truth = membership' equation to arbitrary formulas.

Second, note that the states of $\mathfrak{M}^\Lambda$ consist of *all* Λ-consistent MCSs. The significance of this is that, by Lindenbaum's Lemma, *any* Λ-consistent set of formulas is a subset of some point in $\mathfrak{M}^\Lambda$ — hence, by the Truth Lemma proved below, any Λ-consistent set of formulas is true at some point in this model. In short, the single structure $\mathfrak{M}^\Lambda$ is a 'universal model' for the logic Λ, which is why it's called 'canonical'.

Finally, consider the canonical relation: a state w is related to a state u precisely when for each formula ψ in u, w contains the information $\Diamond\psi$. Intuitively, this captures what we mean by MCSs being 'coherently related'. The reader should compare the present discussion with the account of ultrafilter extensions in Chapter 2 — in Chapter 5 we'll discuss a unifying framework. In the meantime, the following lemma shows that we're getting things right:

Lemma 4.19 *For any normal logic Λ, $R^\Lambda wv$ iff for all formulas ψ, $\Box\psi \in w$ implies $\psi \in v$.*

Proof. For the left to right direction, suppose $R^\Lambda wv$. Further suppose $\psi \notin v$. As v is an MCS, by Proposition 4.16 $\neg\psi \in v$. As $R^\Lambda wv$, $\Diamond\neg\psi \in w$. As w is consistent, $\neg\Diamond\neg\psi \notin w$. That is, $\Box\psi \notin w$ and we have established the contrapositive. We leave the right to left direction to the reader. ⊣

In fact, the definition of R^Λ is exactly what we require; all that remains to be checked is that enough 'coherently related' MCSs exist for our purposes.

Lemma 4.20 (Existence Lemma) *For any normal modal logic Λ and any state $w \in W^\Lambda$, if $\Diamond\phi \in w$ then there is a state $v \in W^\Lambda$ such that $R^\Lambda wv$ and $\phi \in v$.*

Proof. Suppose $\Diamond\phi \in w$. We will construct a state v such that $R^\Lambda wv$ and $\phi \in v$. Let v^- be $\{\phi\} \cup \{\psi \mid \Box\psi \in w\}$. Then v^- is consistent. For suppose not. Then there are $\psi_1, \ldots, \psi_n$ such that $\vdash_\Lambda (\psi_1 \wedge \cdots \wedge \psi_n) \to \neg\phi$, and it follows by an easy argument that $\vdash_\Lambda \Box(\psi_1 \wedge \cdots \wedge \psi_n) \to \Box\neg\phi$. As the reader should check, the formula $(\Box\psi_1 \wedge \cdots \wedge \Box\psi_n) \to \Box(\psi_1 \wedge \cdots \wedge \psi_n)$ is a theorem of every normal modal logic, hence by propositional calculus, $\vdash_\Lambda (\Box\psi_1 \wedge \cdots \wedge \Box\psi_n) \to \Box\neg\phi$. Now, $\Box\psi_1 \wedge \cdots \wedge \Box\psi_n \in w$ (for $\Box\psi_1, \ldots, \Box\psi_n \in w$, and w is an MCS) thus it follows that $\Box\neg\phi \in w$. Using Dual, it follows that $\neg\Diamond\phi \in w$. But this is impossible: w is an MCS containing $\Diamond\phi$. We conclude that v^- is consistent after all.

Let v be any MCS extending v^-; such extensions exist by Lindenbaum's Lemma. By construction $\phi \in v$. Furthermore, for all formulas ψ, $\Box\psi \in w$ implies $\psi \in v$. Hence by Lemma 4.19, $R^\Lambda wv$. ⊣

With this established, the rest is easy. First we lift the 'truth = membership' equation to arbitrary formulas:

Lemma 4.21 (Truth Lemma) *For any normal modal logic Λ and any formula ϕ, $\mathfrak{M}^\Lambda, w \Vdash \phi$ iff $\phi \in w$.*

Proof. By induction on the degree of ϕ. The base case follows from the definition of V^Λ. The boolean cases follow from Proposition 4.16. It remains to deal with the modalities. The left to right direction is more or less immediate from the definition of R^Λ:

$\mathfrak{M}^\Lambda, w \Vdash \Diamond\phi$	iff	$\exists v\,(R^\Lambda wv \ \wedge\ \mathfrak{M}^\Lambda, v \Vdash \phi)$	
	iff	$\exists v\,(R^\Lambda wv \ \wedge\ \phi \in v)$	(Induction Hypothesis)
	only if	$\Diamond\phi \in w$	(Definition R^Λ)

For the right to left direction, suppose $\Diamond\phi \in w$. By the equivalences above, it suffices to find an MCS v such that $R^\Lambda wv$ and $\phi \in v$ — and this is precisely what the Existence Lemma guarantees. ⊣

Theorem 4.22 (Canonical Model Theorem) *Any normal modal logic is strongly complete with respect to its canonical model.*

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

At first glance, the Canonical Model Theorem may seem rather abstract. It is a completeness result with respect to a class of *models*, not frames, and a rather abstract class at that. (That **K4** is complete with respect to the class of transitive frames is interesting; that it is complete with respect to the singleton class containing only its canonical model seems rather dull.) But appearances are misleading: canonical models are by far the most important tool used in the present chapter. For a start, the Canonical Model Theorem immediately yields the following result:

Theorem 4.23 **K** *is strongly complete with respect to the class of all frames.*

Proof. By Proposition 4.12, to prove this result it suffices to find, for any **K**-consistent set of formulas Γ, a model $\mathfrak{M}$ (based on any frame whatsoever) and a state w in $\mathfrak{M}$ such that $\mathfrak{M}, w \Vdash \Gamma$. This is easy: simply choose $\mathfrak{M}$ to be $(\mathfrak{F}^{\mathbf{K}}, V^{\mathbf{K}})$, the canonical model for **K**, and let Γ^+ be any **K**-MCS extending Γ. By the previous lemma, $(\mathfrak{F}^{\mathbf{K}}, V^{\mathbf{K}}), \Gamma^+ \Vdash \Gamma$. ⊣

More importantly, it is often easy to get useful information about the structure of canonical frames. For example, as we will learn in the next section, the canonical frame for **K4** is transitive — and this immediately yields the (more interesting) result that **K4** is complete with respect to the class of transitive frames. Even when a canonical model is not as cleanly structured as we would like, it still embodies a vast amount of information about its associated logic; one of the important themes pursued later in the chapter is how to make use of this information indirectly. Furthermore, canonical models are mathematically natural. As we will learn in Chapter 5, from an algebraic perspective canonical models are not abstract oddities: indeed, they are precisely the structures one is lead to by considering the ideas underlying the Stone Representation Theorem.

To conclude this section we sketch the generalizations required to extend the results obtained so far to languages of arbitrary similarity types.

Definition 4.24 Let τ be a modal similarity type, and Λ a normal modal logic in the language over τ. The *canonical model* $\mathfrak{M}^{\Lambda} = (W^{\Lambda}, R^{\Lambda}_{\triangle}, V^{\Lambda})_{\triangle \in \tau}$ for Λ has W^{Λ} and V^{Λ} as defined in Definition 4.18, while for an n-ary operator $\triangle \in \tau$ the relation $R^{\Lambda}_{\triangle} \subseteq (W^{\Lambda})^{n+1}$ is defined by $R^{\Lambda}_{\triangle} w u_1 \dots u_n$ if for all formulas $\psi_1 \in u_1$, $\dots, \psi_n \in u_n$ we have $\triangle(\psi_1, \dots, \psi_n) \in w$. ⊣

There is an analog of Lemma 4.19.

Lemma 4.25 *For any normal modal logic* Λ, $R^{\Lambda}_{\triangle} w u_1 \dots u_n$ *iff for all formulas* $\psi_1, \dots, \psi_n$, $\nabla(\psi_1, \dots, \psi_n) \in w$ *implies that for some* i *such that* $1 \leq i \leq n$, $\psi_i \in u_i$.

Proof. See Exercise 4.2.3. ⊣

Now for the crucial lemma — we must show that enough coherently related MCSs exist. This requires a more delicate approach than was needed for Lemma 4.20.

Lemma 4.26 (Existence Lemma) *Suppose* $\triangle(\psi_1, \dots, \psi_n) \in w$. *Then there are* $u_1, \dots, u_n$ *such that* $\psi_1 \in u_1, \dots, \psi_n \in u_n$ *and* $R^{\Lambda} w u_1 \dots u_n$.

Proof. The proof of Lemma 4.20 establishes the result for any unary operators in the language, so it only remains to prove the (trickier) case for modalities of higher arity. To keep matters simple, assume that $\triangle$ is binary; this illustrates the key new idea needed.

So, suppose $\triangle(\psi_1, \psi_2) \in w$. Let ϕ_0, ϕ_1, ... enumerate all formulas. We construct two sequences of sets of formulas

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

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

The key step in the inductive construction is

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

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

With this lemma established, the real work has been done. The Truth Lemma and the Canonical Model Theorem for general modal languages are now obvious analogs of Lemma 4.21 and Theorem 4.22. The reader is asked to state and prove them in Exercise 4.2.4.

Exercises for Section 4.2

4.2.1 Show that all MCSs have the properties stated in Proposition 4.16. In addition, show that if Σ and Γ are distinct MCSs, then there is at least one formula ϕ such that $\phi \in \Sigma$ and $\neg\phi \in \Gamma$.

4.2.2 Lindenbaum's Lemma is not fully proved in the text. Give proofs of the four claims made at the end of our proof sketch.

4.2.3 Prove Lemma 4.25. (This is a good way of getting to grips with the definition of normality for modal languages of arbitrary similarity type.)

4.2.4 State and prove the Truth Lemma and the Canonical Model Theorem for languages of arbitrary similarity type. Make sure you understand the special case for nullary modalities (recall that we have no special axioms or rules of proof for these).

4.3 Applications

In this section we put canonical models to work. First we show how to prove the frame completeness results noted in Example 4.11 using a simple and uniform method of argument. This leads us to isolate one of most important concepts of modal completeness theory: *canonicity*. We then switch to the basic temporal language and use similar arguments to prove two important temporal completeness results. We conclude with a statement of the *Sahlqvist Completeness Theorem*, which we will prove in Chapter 5.

Suppose we suspect that a normal modal logic Λ is strongly complete with respect to a class of frames F; how should we go about proving it? Actually, there is no infallible strategy. (Indeed, as we will learn in the following section, many normal modal logics are not complete with respect to any class of frames whatsoever.) Nonetheless, a very simple technique works in a large number of interesting cases: simply show that the canonical frame for Λ belongs to F. We call such proofs *completeness-via-canonicity* arguments, for reasons which will soon become clear. Let's consider some examples.

Theorem 4.27 *The logic* **K4** *is strongly complete with respect to the class of transitive frames.*

Proof. Given a **K4**-consistent set of formulas Γ, it suffices to find a model $(\mathfrak{F}, V)$ and a state w in this model such that (1) $(\mathfrak{F}, V), w \Vdash \Gamma$, and (2) $\mathfrak{F}$ is transitive. Let $(W^{\mathbf{K4}}, R^{\mathbf{K4}}, V^{\mathbf{K4}})$ be the canonical model for **K4** and let Γ^+ be any **K4**-MCS extending Γ. By Lemma 4.21, $(W^{\mathbf{K4}}, R^{\mathbf{K4}}, V^{\mathbf{K4}}), \Gamma^+ \Vdash \Gamma$ so step (1) is established. It remains to show that $(W^{\mathbf{K4}}, R^{\mathbf{K4}})$ is transitive. So suppose w, v and u are points in this frame such that $R^{\mathbf{K4}}wv$ and $R^{\mathbf{K4}}vu$. We wish to show that $R^{\mathbf{K4}}wu$. Suppose $\phi \in u$. As $R^{\mathbf{K4}}vu$, $\Diamond\phi \in v$, so as $R^{\mathbf{K4}}wv$, $\Diamond\Diamond\phi \in w$. But w is a **K4**-MCS, hence it contains $\Diamond\Diamond\phi \to \Diamond\phi$, thus by modus ponens it contains $\Diamond\phi$. Thus $R^{\mathbf{K4}}wu$. $\dashv$

In spite of its simplicity, the preceding result is well worth reflecting on. Two important observations should be made.

First, the proof actually establishes something more general than the theorem claims: namely, that the canonical frame of *any* normal logic Λ containing $\Diamond\Diamond p \to \Diamond p$ is transitive. The proof works because all MCSs in the canonical frame contain the 4 axiom; it follows that the canonical frame of any extension of **K4** is transitive, for all such extensions contain the 4 axiom.

Second, the result suggests that there may be a connection between the structure of canonical frames and the frame correspondences studied in Chapter 3. We know from our previous work that $\Diamond\Diamond p \to \Diamond p$ *defines* transitivity — and now we know that it imposes this property on canonical frames as well.

Theorem 4.28 **T**, **KB** *and* **KD** *are strongly complete with respect to the classes of reflexive frames, of symmetric frames, and of right-unbounded frames, respectively.*

Proof. For the first claim, it suffices to show that the canonical model for $\mathbf{T}$ is reflexive. Let w be a point in this model, and suppose $\phi \in w$. As w is a **T**-MCS, $\phi \to \Diamond\phi \in w$, thus by modus ponens, $\Diamond\phi \in w$. Thus $R^{\mathbf{T}}ww$.

For the second claim, it suffices to show that the canonical model for $\mathbf{KB}$ is symmetric. Let w and v be points in this model such that $R^{\mathbf{KB}}wv$, and suppose that $\phi \in w$. As w is a **KB**-MCS, $\phi \to \Box\Diamond\phi \in w$, thus by modus ponens $\Box\Diamond\phi \in w$. Hence by Lemma 4.19, $\Diamond\phi \in v$. But this means that $R^{\mathbf{KB}}vw$, as required.

For the third claim, it suffices to show that the canonical model for $\mathbf{KD}$ is right-unbounded. (This is slightly less obvious than the previous claims since it requires an existence proof.) Let w be any point in the canonical model for $\mathbf{KD}$. We must show that there exists a v in this model such that $R^{\mathbf{KD}}wv$. As w is a $\mathbf{KD}$-MCS it contains $\Box p \to \Diamond p$, thus by closure under uniform substitution it contains $\Box\top \to \Diamond\top$. Moreover, as $\top$ belongs to all normal modal logics, by generalization $\Box\top$ does too; so $\Box\top$ belongs to $\mathbf{KD}$, hence by modus ponens $\Diamond\top \in w$. Hence, by the Existence Lemma, w has an $R^{\mathbf{KD}}$ successor v. $\dashv$

Once again, these result hint at a link between definability and the structure of canonical frames: after all, T defines reflexivity, B defines symmetry, and D right unboundedness. And yet again, the proofs actually establish something more general than the theorem states: the canonical frame of *any* normal logic containing T is reflexive, the canonical frame of *any* normal logic containing B is symmetric, and the canonical frame of *any* normal logic containing D is right unbounded. This allows us to 'add together' our results. Here are two examples:

Theorem 4.29 **S4** *is strongly complete with respect to the class of reflexive, transitive frames.* **S5** *is strongly complete with respect to the class of frames whose relation is an equivalence relation.*

Proof. The proof of Theorem 4.27 shows that the canonical frame of *any* normal logic containing the 4 axiom is transitive, while the proof of the first clause of Theorem 4.28 shows that the canonical frame of *any* normal logic containing the T axiom is reflexive. As **S4** contains both axioms, its canonical frame has both properties, thus the completeness result for **S4** follows.

As **S5** contains both the 4 and the T axioms, it also has a reflexive, transitive canonical frame. As it also contains the B axiom (which by the proof of the second clause of Theorem 4.28 means that its canonical frame is symmetric), its canonical relation is an equivalence relation. The desired completeness result follows. $\dashv$

As these examples suggest, canonical models are an important tool for proving frame completeness results. Moreover, their utility evidently hinges on some sort of connection between the properties of canonical frames and the frame correspondences studied earlier. Let us introduce some terminology to describe this important phenomenon.

Definition 4.30 (Canonicity) A formula ϕ is *canonical* if, for any normal logic Λ, $\phi \in \Lambda$ implies that ϕ is valid on the canonical frame for Λ. A normal logic Λ is *canonical* if its canonical frame is a frame for Λ. (That is, Λ is canonical if for all ϕ such that $\vdash_\Lambda \phi$, ϕ is valid on the canonical frame for Λ.) $\dashv$

Clearly 4, T, B and D axioms are all canonical formulas. For example, any normal logic Λ containing the 4 axiom has a transitive canonical frame, and the 4 axiom is valid on transitive frames. Similarly, any modal logic containing the B axiom has a symmetric canonical frame, and the B axiom is valid on symmetric frames.

Moreover **K4**, **T**, **KB**, **KD**, **S4** and **S5** are all canonical logics. Our previous work has established that all the axioms involved are valid on the relevant canonical frames. But (see Exercise 4.1.1) modus ponens, uniform substitution, and generalization preserve frame validity. It follows that *every* formula in each of these logics is valid on that logic's canonical frame. In general, to show that $\mathbf{KA_1} \dots \mathbf{A_n}$ is a canonical logic it suffices to show that $A_1, \dots, A_n$ are canonical formulas.

Definition 4.31 (Canonicity for a Property) Let ϕ be a formula, and P be a property. If the canonical frame for any normal logic Λ containing ϕ has property P, and ϕ is valid on any class of frames with property P, then ϕ is *canonical for* P. For example, we say that the 4 axiom is canonical for transitivity, because the presence of 4 forces canonical frames to be transitive, and 4 is valid on all transitive frames. $\dashv$

Let us sum up the discussion so far. Many important frame completeness results can be proved straightforwardly using canonical models. The key idea in such proofs is to show that the relevant canonical frame has the required properties. Such proofs boil down to the following task: showing that the axioms of the logic are canonical for the properties we want (which is why we call them completeness-via-canonicity arguments).

Now for some rather different application of completeness-via-canonicity arguments. The theorems just proved were *syntactically* driven: we began with syntactically specified logics (for example **K4** and **T**) and showed that they could be semantically characterized as the logics of certain frame classes. Canonical models are clearly useful for such proofs — but how do they fare when proving *semantically* driven results? That is, suppose F is a class of frames we find interesting, and we have isolated a set of axioms which we hope generates Λ_{F}. Can completeness-via-canonicity arguments help establish their adequacy?

As such semantically driven questions are typical of temporal logic, let us switch to the basic temporal language. Recall from Example 1.14 that this language has two diamonds, F and P, whose respective duals are G and H. The F operator looks forward along the flow of time, and P looks backwards. Furthermore, recall from Example 1.25 that we are only interested in the frames for this language in which the relations corresponding to F and P are mutually converse. That is, a bidirectional frame is a triple $(W, \{R_P, R_F\})$ such that

$$R_P = \{(y, x) \mid (x, y) \in R_F\}.$$

Recall that by convention we present bidirectional frames as unimodal frames (T, R); in such presentations we understand that $R_F = R$ and $R_P = R^{\smile}$. The class of all bidirectional frames is denoted by F_t, and a bidirectional model is a model whose underlying frame belongs to F_t.

So, what is a temporal logic? As a first step towards answering this we define:

Definition 4.32 The *minimal temporal logic* Λ_{F_t} is $\{\phi \mid \mathsf{F}_t \Vdash \phi\}$. $\dashv$

That is, the minimal temporal logic contains precisely the formulas valid on all bidirectional frames. This is a semantic definition, and, given our interest in frames, a sensible one. But can we axiomatize Λ_{F_t}? That is, can we give Λ_{F_t} a simple *syntactic* characterization? First, note that Λ_{F_t} is *not* identical to the minimal normal logic in the basic temporal language. As we noted in Example 1.29(v), for any frame $\mathfrak{F} = (W, \{R_F, R_P\})$ we have that

$$\mathfrak{F} \Vdash (q \to HFq) \land (q \to GPq) \text{ iff } \mathfrak{F} \in \mathsf{F}_t.$$

The two conjuncts define the 'mutually converse' property enjoyed by R_F and R_P. Clearly, both conjuncts belong to Λ_{F_t}. Equally clearly, they do *not* belong to the minimal normal logic in the basic temporal language Nonetheless, although Λ_{F_t} is stronger, it is not much stronger: the only axioms we need to add are these converse-defining conjuncts.

Definition 4.33 A *normal temporal logic* Λ is a normal modal logic (in the basic temporal language) that contains $p \to GPp$ and $p \to HFp$ (the *converse axioms*). The smallest normal temporal logic is called $\mathbf{K}_t$. We usually call normal temporal logics *tense logics*.

Note that in the basic temporal language the K axioms are $G(p \to q) \to (Gp \to Gq)$ and $H(p \to p) \to (Hp \to Hp)$, and the Dual axioms are $Fp \leftrightarrow \neg G \neg p$ and $Pp \leftrightarrow \neg H \neg p$. Closure under generalization means that if $\vdash_\Lambda \phi$ then $\vdash_\Lambda G\phi$ and $\vdash_\Lambda H\phi$. $\dashv$

We want to show that $\mathbf{K}_t$ generates exactly the formulas in Λ_{F_t}. Soundness is immediate: clearly $\mathbf{K}_t \subseteq \Lambda_{\mathsf{F}_t}$. We show completeness using a canonicity argument. So, what are canonical models for tense logics? Nothing new: simply the following instance of Definition 4.24:

Definition 4.34 The *canonical model for a tense logic* Λ is the structure $\mathfrak{M}^\Lambda = (T^\Lambda, \{R_P^\Lambda, R_F^\Lambda\}, V^\Lambda)$ where:

(i) T^Λ is the set of all Λ-MCSs;
(ii) R^Λ_P is the binary relation on T^Λ defined by $R^\Lambda_P ts$ if for all formulas ϕ, $\phi \in s$ implies $P\phi \in t$.
(iii) R^Λ_F is the binary relation on T^Λ defined by $R^\Lambda_F ts$ if for all formulas ϕ, $\phi \in s$ implies $F\phi \in t$.
(iv) V^Λ is the valuation defined by $V^\Lambda(p) = \{t \in T^\Lambda \mid p \in t\}$. ⊣

We immediately inherit a number of results from the previous section, such as an Existence Lemma, a Truth Lemma, and a Canonical Model Theorem telling us that each tense logic is complete with respect to its canonical model. This is very useful — but it is not quite enough. We want to show that $\mathbf{K}_t$ generates all the *temporal* validities. None of the results just mentioned allow us to conclude this, and for a very obvious reason: we don't yet know whether canonical frames for tense logics are bidirectional frames! In fact they are, and this is where the converse axioms come into play. As the next lemma shows, these axioms are canonical; they force R^Λ_P and R^Λ_F to be mutually converse.

Lemma 4.35 *For any tense logic Λ, if $R^\Lambda_P ts$ then $R^\Lambda_F st$, and if $R^\Lambda_F ts$ then $R^\Lambda_P st$.*

Proof. Rather like the proof that B is canonical for symmetry (see Theorem 4.28 item (ii)). We leave it to the reader as Exercise 4.3.2. ⊣

Thus canonical frames of tense logics are bidirectional frames, so from now on we present them as pairs (T^Λ, R^Λ). Moreover, we now have the desired result:

Corollary 4.36 $\mathbf{K}_t$ *is strongly complete with respect to the class of all bidirectional frames, and* $\mathbf{K}_t = \Lambda_{\mathsf{F}_t}$.

Proof. $\mathbf{K}_t$ is strongly complete with respect to its canonical model. As we've just seen, this model is based on a *bidirectional* frame, so the strong frame completeness result follows. Strong completeness implies weak completeness, so $\Lambda_{\mathsf{F}_t} \subseteq \mathbf{K}_t$. The inclusion $\mathbf{K}_t \subseteq \Lambda_{\mathsf{F}_t}$ has already been noted. ⊣

With this basic result established, we are ready to start a semantically driven exploration of tense logic. That is, we can now attempt to capture the logics of 'time-like' classes of frames as axiomatic extensions of $\mathbf{K}_t$. Here we limit ourselves to the following question: how can the temporal logic of *dense unbounded weak total orders* be axiomatized? From the point of view of tense logic, this is an interesting problem: dense frames and totally ordered frames both play an important role in modeling temporal phenomena. Moreover, as we will see, there is an instructive problem that must be overcome if we build totally ordered models. This will give us a gentle initiation to the fundamental difficulty faced by semantically driven completeness results, a difficulty which we will explore in more detail later in the chapter.

Definition 4.37 A bidirectional frame (T, R) is *dense* if there is a point between any two related points ($\forall xy\,(Rxy \to \exists z\,(Rxz \wedge Rzy))$). It is *right-unbounded* if every point has a successor, *left-unbounded* if every point has a predecessor, and *unbounded* if it is both right and left unbounded. It is *trichotomous* if any two points are equal or are related one way or the other ($\forall xy\,(Rxy \vee x = y \vee Ryx)$), and a *weak total order* (or *weakly linear*) if it is both transitive and trichotomous. We call a frame with all these properties a DUWTO-frame. $\dashv$

Note that weakly linear frames are allowed to contain both reflexive and irreflexive points. Indeed, they are allowed to contain non-empty subsets S such that for all $s, s' \in S$, Rss'. Thus they do not fully model the idea of linearity. Linearity is better captured by the class of *strict* total orders, which are transitive, trichotomous and *irreflexive*. Building strictly totally ordered models is harder than building weakly totally ordered models; we examine the problem in detail later in the chapter.

Our first task is to select suitable axioms. Three of the choices are fairly obvious.

(4) $FFp \to Fp$
(D$_r$) $Gp \to Fp$
(D$_l$) $Hp \to Pp$

Note that $FFp \to Fp$ is simply the 4 axiom in tense logical notation. We know (by the proof of Theorem 4.27) that it is canonical for transitivity, hence choosing it as an axiom ensures the transitive canonical frame we want. Next, D$_r$ (a tense logical analog of the D axiom) is (by the proof of the third claim of Theorem 4.28) canonical for right-unboundedness. Similarly, its backward-looking companion $Hp \to Pp$ is canonical for *left*-unboundedness, so we obtain an unbounded canonical frame without difficulty.

What about density? Here we are in luck. The following formula is canonical for density:

(Den) $Fp \to FFp$

This is worth a lemma, since the proof is not trivial. (Note that density is a universal-existential property, rather than a universal property like transitivity or reflexivity. This means that proving canonicity requires establishing the *existence* of certain MCSs.)

Lemma 4.38 $Fp \to FFp$ *is canonical for density.*

Proof. Let Λ be any tense logic containing $Fp \to FFp$, let (T^Λ, R^Λ) be its canonical frame, and let t and t' be points in this frame such that $R^\Lambda tt'$. We have to show that there is a Λ-MCS s such that $R^\Lambda ts$ and $R^\Lambda st'$. If we could show that

$\{\phi \mid G\phi \in t\} \cup \{F\psi \mid \psi \in t'\}$ was consistent we would have the desired result (for by Lemma 4.35, any MCS extending this set would be a suitable choice for s).

So suppose for the sake of contradiction that this set is not consistent. Then, for some finite set of formulas $\phi_1, \ldots, \phi_m, \psi_1, \ldots, \psi_n$ from this set,

$$\vdash_\Lambda (\phi_1 \wedge \cdots \wedge \phi_m \wedge F\psi_1 \wedge \cdots \wedge F\psi_n) \to \bot \ .$$

Define $\widehat{\phi}$ to be $\phi_1 \wedge \cdots \wedge \phi_m$ and $\widehat{\psi}$ to be $\psi_1 \wedge \cdots \wedge \psi_n$. Note that $\widehat{\psi} \in t'$.

Now, $\vdash_\Lambda F\widehat{\psi} \to F\psi_1 \wedge \cdots \wedge F\psi_n$, hence $\vdash_\Lambda \widehat{\phi} \wedge F\widehat{\psi} \to \bot$, hence $\vdash_\Lambda \widehat{\phi} \to \neg F\widehat{\psi}$, and hence $\vdash_\Lambda G\widehat{\phi} \to G\neg F\widehat{\psi}$. Because $G\phi_1, \ldots, G\phi_m \in t$, we have that $G\widehat{\phi} \in t$ too, hence $G\neg F\widehat{\psi} \in t$, and hence $\neg G\neg F\widehat{\psi} \notin t$. That is, $FF\widehat{\psi} \notin t$. But this means that $F\widehat{\psi} \notin t$, as (by uniform substitution in Den) $F\widehat{\psi} \to FF\widehat{\psi} \in t$. But now we have a contradiction: as $\widehat{\psi} \in t'$ and $R^\Lambda tt'$, $F\widehat{\psi}$ must be in t. We conclude that $\{\phi \mid G\phi \in t\} \cup \{F\psi \mid \psi \in t'\}$ is consistent after all. (Note that this proof makes no use of the converse axioms, thus we have also proved that $\Diamond p \to \Diamond\Diamond p$ is canonical for density.) $\dashv$

So it only remains to ensure trichotomy — but here we encounter an instructive difficulty. Because modal (and temporal) validity is preserved under the formation of disjoint unions (see Proposition 3.14) no formula of tense logic defines trichotomy. Moreover, a little experimentation will convince the reader that canonical frames may have disjoint point generated subframes; such canonical frames are clearly not trichotomous. In short, to prove the desired completeness result we need to build a model with a property for which no modal formula is canonical. This is the problem we encounter time and time again when proving semantically driven results.

In the present case, a little lateral thinking leads to a solution. First, let us get rid of a possible preconception. Until now, we have always used the entire canonical model — but we do not need to do this. A point generated submodel suffices. More precisely, if $\mathfrak{M}^\Lambda, w \Vdash \Gamma$, then as modal satisfaction is preserved in generated submodels (see Proposition 2.6) $\mathfrak{S}, w \Vdash \Gamma$, where $\mathfrak{S}$ is the submodel of $\mathfrak{M}^\Lambda$ generated by w.

The observation is trivial, but its consequences are not. By restricting our attention to point-generated submodels, we increase the range of properties we can impose. In particular, we *can* impose trichotomy on point-generated submodels. We met the relevant axioms when working with the basic modal language. From our discussion of **S4.3** and **K4.3** (in particular, Exercise 4.3.3) we know that

$(.3_r)$ $\quad (Fp \wedge Fq) \to F(p \wedge Fq) \vee F(p \wedge q) \vee F(q \wedge Fp)$

is canonical for no-branching-to-the-right. Analogously

$(.3_l)$ $\quad (Pp \wedge Pq) \to P(p \wedge Pq) \vee P(p \wedge q) \vee P(q \wedge Pp).$

is canonical for no-branching-to-the-left. Call a frame with no branching to the left or right a *non-branching* frame.

Proposition 4.39 *Any trichotomous frame (T, R) is non-branching. Furthermore, if R is transitive and non-branching and $t \in T$, then the subframe of (T, R) generated by t is trichotomous.*

Proof. Trivial — though the reader should recall that when forming generated subframes for the basic temporal language, we generate on both the relation corresponding to F and that corresponding to P. That is, we generate both forwards and backwards along R. $\dashv$

In short, although no formula is canonical for trichotomy, there is a good 'approximation' to it (namely, the non-branching property) for which we do have a canonical formula (namely, the conjunction of $.3_l$ and $.3_r$). With this observed, the desired result is within reach.

Definition 4.40 Let $\mathbf{K}_t\mathbf{Q}$ be the smallest tense logic containing 4, D$_l$, D$_r$, Den, $.3_l$ and $.3_r$. $\dashv$

Theorem 4.41 $\mathbf{K}_t\mathbf{Q}$ *is strongly complete with respect to the class of* DUWTO-*frames.*

Proof. If Γ is $\mathbf{K}_t\mathbf{Q}$-consistent set of formulas, extend it to a $\mathbf{K}_t\mathbf{Q}$-MCS Γ^+. Let $\mathfrak{M}$ be the canonical model for $\mathbf{K}_t\mathbf{Q}$, and let $\mathfrak{S}$ be the submodel of $\mathfrak{M}$ generated by Γ^+. As we just noted, $\mathfrak{S}, \Gamma^+ \Vdash \Gamma$. Moreover, the frame underlying $\mathfrak{S}$ is a DUWTO-frame as required. First, as $\mathbf{K}_t\mathbf{Q}$ contains axioms that are canonical for transitivity, unboundedness, and density, $\mathfrak{M}$ has these properties; it is then not difficult to show that $\mathfrak{S}$ has them too. Moreover, as the conjunction of $.3_l$ and $.3_r$ is canonical for non-branching, $\mathfrak{M}$ is non-branching and $\mathfrak{S}$ trichotomous. $\dashv$

To conclude, two important remarks. First, the need to build models possessing properties for which no formula is canonical is the fundamental difficulty facing semantically driven results. In the present case, a simple idea enabled us to bypass the problem — but we won't always be so lucky and in the second part of the chapter we develop more sophisticated techniques for tackling the issue.

Second, the relationships between completeness, canonicity and correspondence are absolutely fundamental to the study of normal modal logics. These relationships are further discussed in the following section, and explored algebraically in Chapter 5, but let's immediately mention one of the most elegant positive results in the area: the *Sahlqvist Completeness Theorem*. In Chapter 3 we proved the Sahlqvist Correspondence Theorem: every Sahlqvist formula *defines* a first-order class of frames. Here's its completeness theoretic twin, which we will prove in Chapter 5:

Theorem 4.42 *Every Sahlqvist formula is canonical for the first-order property it defines. Hence, given a set of Sahlqvist axioms Σ, the logic $\mathbf{K\Sigma}$ is strongly complete with respect to the class of frames F_Σ (that is, the first-order class of frames defined by Σ).*

This is an extremely useful result. Most commonly encountered axioms in the basic modal language are Sahlqvist (the Löb and McKinsey formulas are the obvious exceptions) thus it provides an immediate answer to a host of completeness problems. Moreover, like the Sahlqvist Correspondence Theorem, the Sahlqvist Completeness Theorem applies to modal languages of *arbitrary* similarity type. Finally, the Theorem generalizes to a number of extended modal logics, most notably D-logic (which we introduce in Chapter 7). Note that Kracht's Theorem (see Chapter 2) can be viewed as a providing a sort of 'converse' to Sahlqvist's result, for it gives us a way of computing formulas that are canonical for certain first-order classes of frames.

Exercises for Section 4.3

4.3.1 Let 1.1 be the axiom $\Diamond p \to \Box p$. Show that $\mathbf{K1.1}$ is sound and strongly complete with respect to the class of all frames (W, R) such that R is a partial function.

4.3.2 Let Λ be a normal temporal logic containing the axioms $p \to GPp$ and $p \to HFp$. Show that if $R_P^\Lambda ts$ then $R_F^\Lambda st$, and if $R_F^\Lambda ts$ then $R_P^\Lambda st$.

4.3.3 Use canonical models to show that **K4.3** is strongly complete with respect to the class of frames that are transitive and have no branching to the right, and that **S4.3** is strongly complete with respect to the class of frames that are reflexive, transitive and have no branching to the right.

Then, by proving suitable completeness results (and making use of the soundness results proved in Exercise 4.1.4), show that the normal logic axiomatized by $\Box(p \land \Box p \to q) \lor \Box(q \land \Box q \to p)$ is **K4.3**. Further, show that the normal modal logic axiomatized by $\Box(\Box p \to q) \lor \Box(\Box q \to p)$ is **S4.3**. Try proving the equivalence of these logics syntactically.

4.3.4 Prove directly that $\Diamond\Box p \to \Box\Diamond p$ is canonical for the Church-Rosser property.

4.3.5 Let W5 be the formula $\Diamond\Box p \to (p \to \Box p)$, and let **S4W5** be the smallest normal logic extending **S4** that contains W5. Find a simple class of frames that characterizes this logic.

4.3.6 Show that **S5** is complete with respect to the the class of *globally related frames*, that is, those frames (W, R) such that $\forall w\, Rww$.

4.3.7 Consider a similarity type τ with one binary operator $\triangle$. For each of the following Sahlqvist formulas, first compute the (global) first-order correspondent. Then, give a *direct* proof that the modal formula is canonical for the corresponding first-order property.

(a) $p \triangle q \to q \triangle p$,
(b) $(p \triangle q) \triangle r \to p \triangle (q \triangle r)$,
(c) $((q \triangle \neg(p \triangle q)) \land p) \to \bot$.

4.4 Limitative Results

Although completeness-via-canonicity is a powerful method, it is not infallible. For a start, not every normal modal logic is canonical. Moreover, not every normal logic is the logic of some class of frames. In this section we prove both claims and discuss their impact on modal completeness theory.

We first demonstrate the existence of non-canonical logics. We will show that **KL**, the normal modal logic generated by the Löb axiom $\Box(\Box p \to p) \to \Box p$, is not canonical. We prove this by showing that **KL** is not sound and strongly complete with respect to any class of frames. Now, every *canonical* logic is sound and strongly complete with respect to some class of frames. (For suppose Λ is a canonical logic and Γ is a Λ-consistent set of formulas. By the Truth Lemma, Γ is satisfiable on $\mathfrak{F}^{\Lambda}$; as Λ is canonical, $\mathfrak{F}^{\Lambda}$ is a frame for Λ.) Hence if **KL** is not sound and strongly complete with respect to any class of frames, it cannot be canonical either.

Theorem 4.43 **KL** *is not sound and strongly complete with respect to any class of frames, and hence it is not canonical.*

Proof. Let Γ be $\{\Diamond q_1\} \cup \{\Box(q_i \to \Diamond q_{i+1}) \mid 1 \leq i \in \omega\}$. We will show that Γ is **KL**-consistent, and that no model based on a **KL**-frame can satisfy all formulas in Γ at a single point. The theorem follows immediately.

To show that Γ is consistent, it suffices to show that every finite subset Ψ of Γ is consistent. Given any such Ψ, for some natural number n there is a finite set Φ of the form $\{\Diamond q_1\} \cup \{\Box(q_i \to \Diamond q_{i+1}) \mid 1 \leq i < n\}$ such that $\Psi \subseteq \Phi \subset \Gamma$. We show that Φ, and hence Ψ, is consistent.

Let $\widehat{\Phi}$ be the conjunction of all the formulas in Φ. To show that $\widehat{\Phi}$ is **KL**-consistent, it suffices to show that it can be satisfied in a model based on a frame for **KL**, for this shows that $\neg\widehat{\Phi}$ is *not* valid on all frames for **KL**, and hence is *not* one of its theorems. Let $\mathfrak{F}$ be the frame consisting of $\{0, \ldots, n\}$ in their usual order; as this is a transitive, converse well-founded frame, by Example 3.9 it is a frame for **KL**. Let $\mathfrak{M}$ be any model based on $\mathfrak{F}$ such that for all $1 \leq i \leq n$, $V(q_i) = \{i\}$. Then $\mathfrak{M}, 0 \Vdash \widehat{\Phi}$ and $\widehat{\Phi}$ is **KL** consistent.

Next, suppose for the sake of a contradiction that **KL** is sound and strongly complete with respect to some class of frames F; note that as **KL** is not the inconsistent logic, F must be non-empty. Thus any **KL**-consistent set of formulas can be satisfied at some point in a model based on a frame in F. In particular, there is a model $\mathfrak{M}$ based on a frame in F and a point w in $\mathfrak{M}$ such that $\mathfrak{M}, w \Vdash \Gamma$. But this is impossible: because $\mathfrak{M}, w \Vdash \Gamma$, we can inductively define an infinite path through $\mathfrak{M}$ starting at w; however as $\mathfrak{M}$ is based on a frame for **KL** it cannot contain such infinite paths. Hence **KL** is not sound and strongly complete with respect to any class of frames, and so cannot be canonical. $\dashv$

Remark 4.44 A normal logic Λ is said to be *compact* when any Λ-consistent set Σ can be satisfied in a frame for Λ at a single point. So the above proof shows that **KL** is not compact. Note that a non-compact logic cannot be canonical, and cannot be sound and strongly complete with respect to any class of frames. We will see a similar compactness failure when we examine PDL in Section 4.8. ⊣

What are we to make of this result? The reader should *not* jump to the conclusion that it is impossible to characterize **KL** as the logic of some class of frames. Although no *strong* frame completeness result is possible, as we noted in Table 4.1 there is a elegant *weak* frame completeness result for **KL**, namely:

Theorem 4.45 **KL** *is weakly complete with respect to the class of all finite transitive trees.*

Proof. The proof uses the finitary methods studied later in the chapter. The reader is asked to prove it in Exercises 4.8.7 and 4.8.8. ⊣

Thus **KL** is the logic of all finite transitive trees — and there exist non-canonical but (weakly) complete normal logics. We conclude that, powerful though it is, the completeness-via-canonicity method cannot handle all interesting frame completeness results.

Let us turn to the second conjecture: are all normal logics weakly complete with respect to some class of frames? No: *incomplete* normal logics exist.

Definition 4.46 Let Λ be a normal modal logic. Λ is *(frame) complete* if there is a class of frames F such that $\Lambda = \Lambda_{\mathsf{F}}$, and (*frame*) *incomplete* otherwise. ⊣

We now demonstrate the existence of incomplete logics in the basic temporal language. The demonstration has three main steps. First, we introduce a tense logic called $\mathbf{K}_t$**Tho** and show that it is consistent. Second, we show that no frame for $\mathbf{K}_t$**Tho** can validate the McKinsey axiom (which in tense logical notation is $GF\phi \to FG\phi$). It is tempting to conclude that $\mathbf{K}_t$**ThoM**, the smallest tense logic containing both $\mathbf{K}_t$**Tho** and the McKinsey axiom, is the inconsistent logic. Surprisingly, this is *not* the case. $\mathbf{K}_t$**ThoM** is consistent — and hence is not the tense logic of any class of frames at all. We prove this in the third step with the help of general frames.

$\mathbf{K}_t$**Tho** is the tense logic generated by the following axioms:

(.3_r) $F(p \wedge q) \to F(p \wedge Fq) \vee F(p \wedge q) \vee F(p \wedge Fq)$
(D_r) $Gp \to Fp$
(L_l) $H(Hp \to p) \to Hp$

As we have already seen, the first two axioms are canonical for simple first-order conditions (no branching to the right, and right-unboundedness, respectively). The third axiom is simply the Löb axiom written in terms of the backward looking operator H; it is valid on precisely those frames that are transitive and contain no infinite descending paths. (Note that such frames cannot contain reflexive points.) Let $\mathbf{K}_t\mathbf{Tho}$ be the tense logic generated by these three axioms. As all three axioms are valid on the natural numbers, $\mathbf{K}_t\mathbf{Tho}$ is consistent. If (T, R) is a frame for $\mathbf{K}_t\mathbf{Tho}$ and $t \in T$, then $\{u \in T \mid Rtu\}$ is a right-unbounded strict total order.

Now for the second step. Let $\mathbf{K}_t\mathbf{ThoM}$ be the smallest tense logic containing $\mathbf{K}_t\mathbf{Tho}$ and the McKinsey axiom $GFp \to FGp$. What are the frames for this enriched logic? The answer is: none at all, or, to put it another way, $\mathbf{K}_t\mathbf{ThoM}$ defines the empty class of frames. To see this we need the concept of cofinality.

Definition 4.47 Let $(U, <)$ be a strict total order and $S \subseteq U$. S is *cofinal in* U if for every $u \in U$ there is an $s \in S$ such that $u < s$. $\dashv$

For example, both the even numbers and the odd numbers are cofinal in the natural numbers. Indeed, they are precisely the kind of cofinal subsets we will use in the work that follows: mutually complementary cofinal subsets.

Lemma 4.48 *Let* $\mathfrak{T}$ *be any frame for* $\mathbf{K}_t\mathbf{Tho}$. *Then* $\mathfrak{T} \not\Vdash GFp \to FGp$.

Proof. Let t be any point in $\mathfrak{T}$, let $U = \{u \in T \mid Rtu\}$, and let $<$ be the restriction of R to U. As $\mathfrak{T}$, validates all the $\mathbf{K}_t\mathbf{Tho}$ axioms, $(U, <)$ is a right-unbounded strict total order. Suppose we could show that there is a non-empty proper subset S of U such that both S and $U \backslash S$ are cofinal in U. Then the lemma would be proved, for we would merely need to define a valuation V on $\mathfrak{T}$ such that $V(p) = S$, and $(\mathfrak{T}, V), t \not\Vdash GFp \to FGp$.

Such subsets S of U exist by (3.18) in Chapter 3. For a more direct proof, take an ordinal κ that is larger than the size of U. By ordinal induction, we will define a sequence of pairs of sets $(R_\alpha, S_\alpha)_{\alpha \leq \kappa}$ such that $R_\kappa \cap S_\kappa = \varnothing$ and both R_κ and S_κ are cofinal. We can easily prove the lemma from this by taking $S = S_\kappa$. The definition is as follows:

(i) For $\alpha = 0$, take some points r_0 and s_0 in U such that $r_0 < s_0$ and define $R_0 = \{r_0\}$ and $S_0 = \{s_0\}$.

(ii) If α is a successor ordinal $\beta + 1$, then distinguish two cases:
   - (a) if R_β or S_β is cofinal, then define $R_\alpha = R_\beta$ and $S_\alpha = S_\beta$,
   - (b) if neither R_β nor S_β is cofinal, then take some upper bound r_β of S_β (that is, $r_\beta > s$ for all $s \in S_\beta$), take some s_β bigger than r_β and define $R_\alpha = R_\beta \cup \{r_\beta\}$ and $S_\alpha = S_\beta \cup \{s_\beta\}$

(iii) If α is a limit ordinal, then define $R_\alpha = \bigcup_{\beta<\alpha} R_\beta$ and $S_\alpha = \bigcup_{\beta<\alpha} S_\beta$.

It is easy to prove that $R_\alpha \cap S_\alpha = \varnothing$ for every ordinal $\alpha \leq \kappa$, so it remains to be shown that both R_κ and S_κ are cofinal. The key to this proof is the observation that if R_κ and S_κ were not cofinal, then the (implicitly defined) partial map $r : \kappa \to U$ would be total and injective (further proof details are left to the reader). This would contradict the assumption that κ exceeds the size of U. ⊣

We are ready for the final step. As **K**$_t$**ThoM** defines the empty class of frames, it is tempting to conclude that it is also complete with respect to this class; that is, that **K**$_t$**ThoM** is the inconsistent logic. However, this is not the case.

Theorem 4.49 **K**$_t$**ThoM** *is consistent and incomplete.*

Proof. Let $(\mathbb{N}, <)$ be the natural numbers in their usual order. Let A be the collection of finite and cofinite subsets of $\mathbb{N}$; we leave it to the reader to show that A is closed under boolean combinations and modal projections. Thus $(\mathbb{N}, <, A)$ is a general frame; we claim that it validates all the **K**$_t$**ThoM** axioms. Now, it certainly validates all the **K**$_t$**Tho** axioms, for these are already valid on the underlying frame. But what about M? As we noted in Example 1.34, $GFp \to FGp$ cannot be falsified under assignments mapping p to either a finite or a co-finite set. Hence all the axioms are valid and **K**$_t$**ThoM** must be consistent.

Now, by Lemma 4.48, **K**$_t$**ThoM** is not the logic of any non-empty class of frames. But as **K**$_t$**ThoM** is consistent, it's not the logic of the empty class of frames either. In short, it's not the logic of any class of frames whatsoever, and is incomplete. ⊣

Frame incompleteness results are not some easily fixed anomaly. As normal logics are sets of formulas closed under three rules of proof, the reader may be tempted to think that these rules are simply too weak. Perhaps there are yet-to-be-discovered rules which would strengthen our deductive apparatus sufficiently to overcome incompleteness? (Indeed, later in the chapter we introduce an additional proof rule, and it will turn out to be very useful.)

Nonetheless, no such strengthening of our deductive apparatus can eliminate frame incompleteness. Why is this? Ultimately it boils down to something we learned in Chapter 3: frame consequence is an essentially a second-order relation. Moreover, as we discussed in the Notes to Chapter 3, it is a very strong relation indeed: strong enough to simulate the standard second-order consequence relation. Frame incompleteness results reflect the fact that (over frames) modal logic is second order logic in disguise.

There are many incomplete logics. Indeed, if anything, incomplete logics are the norm. An analogy may be helpful. When differential calculus is first encountered, most students have rather naive ideas about functions and continuity; polynomials, and other simple functions familiar from basic physics, are taken to be typical of all real-valued functions. The awakening comes with the study of analysis. Here the student encounters such specimens as everywhere-continuous but nowhere-differentiable functions — and comes to see that the familiar functions are actually abnormally well-behaved. The situation is much the same in modal logic. The logics of interest to philosophers — logics such as **T**, **S4** and **S5** — were the first to be semantically characterized using frames. It is tempting to believe that such logics are typical, but they are actually fairly docile creatures; the lattice of normal logics contains far wilder inhabitants.

The significance of the incompleteness results depends on one's goals. Logicians interested in applications are likely to focus on certain intended classes of models, and completeness results for these classes. Beyond providing a salutary warning about the folly of jumping to hasty generalizations, incompleteness results are usually of little direct significance here. On the other hand, for those whose primary interest is syntactically driven completeness results, the results could hardly be more significant: they unambiguously show the inadequacy of frame-based classifications. Unsurprisingly, this has had considerable impact on the study of modal logic. For a start, it lead to a rebirth of interest in alternative tools — and in particular, to the renaissance of *algebraic semantics*, which we will study in Chapter 5. Moreover, it has lead modal logicians to study new types of questions. Let us consider some of the research themes that have emerged.

One response has been to look for general syntactic constraints on axioms which guarantee canonicity. The most elegant such result is the Sahlqvist Completeness Theorem, which we have already discussed. A second response has been to investigate the interplay between completeness, canonicity, and correspondence. Typical of the questions that can be posed is the following: *If $A_1, \ldots, A_n$ are axioms that define an elementary class of frames, is* $\mathbf{KA_1} \ldots \mathbf{A_n}$ *frame complete?* (In fact, the answer here is *no* — as the reader is asked to show in Exercise 4.4.3.) The most significant positive result that has emerged from this line of enquiry is the following:

Theorem 4.50 *If F is a first-order definable class of frames, then Λ_{F} is canonical.*

Again, we prove this in Chapter 5 using algebraic tools (see Theorem 5.56). Tantalizingly, at the time of writing the status of the converse was unknown: If a normal modal logic Λ is canonical, then there is a first-order definable class of frames F such that $\Lambda = \Lambda_{\mathsf{F}}$. This conjecture seems plausible, but neither proof nor counterexample has been found.

A third response has been to examine particular classes of normal modal logics more closely. The entire lattice may have undesirable properties — but many sub-regions are far better behaved. We will examine a particularly well-behaved sub-region (namely, the normal logics extending **S4.3**) in the final section of this chapter.

This concludes our survey of basic completeness theory. The next four sections (all of which are on the basic track) explore the following issue: how are we to prove completeness results when we need to build a model that has a property for which no formula is canonical? Some readers may prefer to skip this for now and go straight on to the following chapter. This discusses completeness, canonicity and correspondence from an *algebraic* perspective.

Exercises for Section 4.4

4.4.1 Recall that any normal modal logic that has the finite model property also has the finite frame property. What are the consequences of this for incomplete normal modal logics?

4.4.2 The logic **KvB** consists of all formulas valid on the general frame $\mathfrak{J}$. The domain J of $\mathfrak{J}$ is $\mathbb{N} \cup \{\omega, \omega+1\}$ (the set of natural numbers together with two further points), and R is defined by Rxy iff $x \neq \omega + 1$ and $y < x$ or $x = \omega + 1$ and $y = \omega$. (The frame (J, R) is shown in Figure 6.2 in Chapter 6.) A, the collection of subsets of J admissible in $\mathfrak{J}$, consists of all $X \subseteq J$ such that either X is finite and $\omega \notin X$, or X is co-finite and $\omega \in X$.

(a) Show that $\Box\Diamond(\top) \to \Box(\Box(\Box p \to p) \to p)$ is valid on $\mathfrak{J}$.
(b) Show that on any *frame* on which the previous formula is valid, $\Box\Diamond(\top) \to \Box(\bot)$ is valid too.
(c) Show that $\Box\Diamond(\top) \to \Box(\bot)$ is *not* valid on $\mathfrak{J}$.
(d) Conclude that **KvB** is incomplete.

4.4.3 Consider the formulas (T) $p \to \Diamond p$, (M) $\Box\Diamond p \to \Diamond\Box p$, (E) $\Diamond(\Diamond p \wedge \Box q) \to \Box(\Diamond p \vee \Box p)$ and (Q) $(\Diamond p \wedge \Box(p \to \Box p)) \to p$. Let Λ denote the normal modal logic axiomatized by these formulas.

(a) Prove that E corresponds to the following first-order formula: $\forall xy_1y_2\,((Rxy_1 \wedge Rxy_2) \to (\forall z\,(Ry_1z \to Ry_2z) \vee \forall z\,(Ry_2z \to Ry_1z)))$.
(b) Prove that within the class of frames validating both T and M, Q defines the frames satisfying the condition $R^{\smile} \subseteq R^*$ (that is, if Rst then there is finite path back from t to s).
(c) Prove that the conjunction of the four axioms defines the class of frames with a trivial accessibility relation — that is, $\mathrm{T} \wedge \mathrm{M} \wedge \mathrm{E} \wedge \mathrm{Q}$ corresponds to $\forall xy\,(Rxy \leftrightarrow x = y)$. (Hint: consider the effect of the McKinsey formula on the frames satisfying the condition $R^{\smile} \subseteq R^*$.)
(d) Consider the so-called *veiled recession frame* $(\mathbb{N}, R, A)$, where $\mathbb{N}$ is the set of natural numbers, Rmn holds iff $m \leq n+1$ and A is the collection of finite and co-finite subsets of $\mathbb{N}$. Show that all four axioms are valid on this general frame, but that the formula $p \to \Box p$ can be refuted.
(e) Conclude that Λ is incomplete, although it defines an elementary class of frames.
(f) Does this contradict Theorem 4.50?

4.4.4 Given a class K of frames, let $\Theta(\mathsf{K}) = \Lambda_{\mathsf{K}}$ denote the set $\{\phi \mid \mathfrak{F} \Vdash \phi$ for all $\mathfrak{F}$ in $\mathsf{K}\}$ and given a logic Λ, let $\mathsf{Fr}(\Lambda)$ denote the class of frames on which Λ is valid.

(a) Show that the operations Θ and Fr form a so-called *Galois connection*. That is, prove that for all classes K and logics Λ:

$$\Lambda \subseteq \Theta(\mathsf{K}) \text{ iff } \mathsf{K} \subseteq \mathsf{Fr}(\Lambda).$$

(b) What does it mean for a logic Λ if $\Lambda = \Theta(\mathsf{Fr}(\Lambda))$? (Give an example of a logic for which it does *not* hold.)
(c) What does it mean for a frame class K if $\mathsf{K} = \mathsf{Fr}(\Theta(\mathsf{K}))$? (Give an example of a frame class for which it does *not* hold.)

4.5 Transforming the Canonical Model

What is the modal logic of partial orders? And what is the tense logic of strict total orders? Such questions bring us face to face with the fundamental problem confronting semantically driven completeness results. Partial orders are *antisymmetric*, and strict total orders are *irreflexive*. No modal formula defines either property, and (as the reader probably suspects) no formula is canonical for them either. Thus, to answer either question, we need to build a model for which we lack a canonical formula — and hence we will need to expand our repertoire of model building techniques. This is the main goal of the present section and the three that follow.

In this section we explore a particularly natural strategy: transforming the canonical model. Although a canonical model may lack some desired properties, it does get a lot of things right. Perhaps it is possible to reshape it, transforming it into a model with all the desired properties? We have done this once already, though in a very simple way: in the completeness proof for $\mathbf{K}_t\mathbf{Q}$ (see Theorem 4.41 and surrounding discussion) we formed a point-generated submodel of the canonical model to ensure trichotomy. Here we will study two more sophisticated transformations — *unraveling* and *bulldozing* — and use them to answer the questions with which this section began.

It seems plausible that **S4** is the modal logic of partial orders: Theorem 4.29 tells us that **S4** is complete with respect to the class of reflexive transitive frames (that is, *preorders*) and there don't seem to be any modal formulas we could add to **S4** to reflect antisymmetry. Furthermore, it seems reasonable to hope that we could prove this using some sort of model transformation: as every **S4**-consistent set of formulas can be satisfied on a preorder, and as we know that modal languages are blind to antisymmetry (at least as far as frame definability is concerned) maybe we can find a way of transforming any satisfying preorder into a partial order without affecting satisfiability? (It's worth stressing that this informal line of argument is *not* a proof; it's intended solely to motivate the work that follows.)

A transformation called *unraveling* will enable us do this. Indeed, unraveling will let us prove the stronger result that **S4** is complete with respect to the class of *reflexive and transitive trees*. (This will be useful in Chapter 6 when we discuss decidability). We briefly discussed unraveling in Chapter 2, where we used it to show that modal logic has the tree property (see Proposition 2.15). Informally, given any model, unraveling builds a new model, whose points are *paths* of the

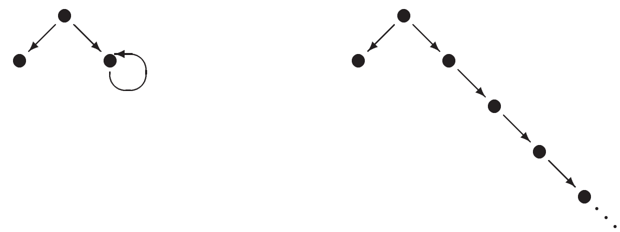


Fig. 4.1. A model and its unraveling

original model. That is, transition sequences in the original model are explicitly represented as states in the unraveled model. More precisely:

Definition 4.51 (Unraveling) Let (W, R) be a frame generated by some point $w \in W$. The *unraveling* of (W, R) around w is the frame $(\vec{W}, \vec{R})$ where:

(i) $\vec{W}$ is the set of all finite sequences $(w, w_1, \ldots, w_n)$ such that $w_1, \ldots, w_n \in W$ and $Rww_1, \ldots, Rw_{n-1}w_n$, and
(ii) If $\vec{s}_1, \vec{s}_2 \in \vec{W}$, then $\vec{R}\vec{s}_1\vec{s}_2$ if there is some $v \in W$ such that $\vec{s}_1 + (v) = \vec{s}_2$, where + denotes sequence concatenation.

If $\mathfrak{M} = (W, R, V)$ is a model and $(\vec{W}, \vec{R})$ is the unraveling of (W, R) around w, then we define the valuation $\vec{V}$ on $(\vec{W}, \vec{R})$ as follows:

$$\vec{V}(p) = \{(w, w_1, \ldots, w_n) \in \vec{W} \mid w_n \in V(p)\}$$

The model $\vec{\mathfrak{M}} = (\vec{W}, \vec{R}, \vec{V})$ is called the unraveling of $\mathfrak{M}$ around w. $\dashv$

A simple example is given in Figure 4.1. As this example suggests (and as the reader should check) unraveling any frame around a generating point w yields an *irreflexive*, *intransitive*, and *asymmetric* frame. Indeed, note that unraveled frames are *trees*: the root node is the sequence (w), and the relation $\vec{R}$ is just the familiar (immediate) successor (or daughter-of) relation on trees.

Lemma 4.52 *Let $\vec{\mathfrak{M}} = (\vec{W}, \vec{R}, \vec{V})$ be the unraveling of $\mathfrak{M} = (W, R, V)$ around w. Then (W, R) is a bounded morphic image of $(\vec{W}, \vec{R})$, and $\mathfrak{M}$ is a bounded morphic image of $\vec{\mathfrak{M}}$.*

Proof. Let $f : \vec{W} \to W$ be defined by $f(w, w_1, \ldots, w_n) = w_n$. It is easy to see that f is surjective, has the back and forth property, and that for any $\vec{s} \in \vec{W}$, $\vec{s}$ and $f(\vec{s})$ satisfy the same propositional variables. $\dashv$

A simple corollary is that *any* satisfiable set of formulas is satisfiable on a (irreflexive, intransitive, and asymmetric) tree: for if a set of formulas is satisfiable, it is satisfiable on a point-generated model (take the submodel generated by the satisfying point), hence by unraveling we have the result. It follows that **K** is (strongly) complete with respect to this class of models.

But our real interest is **S4**. How do we use unraveling to make the *partially ordered* models we require for the completeness result? In the most obvious way possible: we simply take the reflexive transitive closures of unraveled models. More precisely, suppose we unravel $\mathfrak{M}$ around some generating point w to obtain $(\vec{W}, \vec{R}, \vec{V})$. Now consider the model $\mathfrak{M}^* = (\vec{W}, R^*, \vec{V})$ where R^* is the reflexive transitive closure of $\vec{R}$. Trivially, $\mathfrak{M}^*$ is an **S4** model. Moreover, as $(\vec{W}, \vec{R})$ is a tree, $(\vec{W}, R^*)$ is an *antisymmetric* frame. Indeed, it is a *reflexive and transitive tree*, for R^* is simply the familiar dominates (or ancestor-of) relation on trees. So only one question remains: is $\mathfrak{M}$ a bounded morphic image of $\mathfrak{M}^*$? In general, *no*. But if the model $\mathfrak{M}$ we started with was itself reflexive and transitive, *yes*:

Lemma 4.53 *Let $\mathfrak{M} = (W, R, V)$ be a reflexive transitive model generated by some $w \in W$, and let $(\vec{W}, \vec{R}, \vec{V})$ be the unraveling of $\mathfrak{M}$ around w. Let R^* be the reflexive transitive closure of $\vec{R}$, and define $\mathfrak{M}^*$ to be $(\vec{W}, R^*, \vec{V})$. Then $\mathfrak{M}$ is a bounded morphic image of $\mathfrak{M}^*$.*

Proof. It is easy to see that the function f defined in Lemma 4.52 remains the required bounded morphism; as far as surjectivity, the back property, and the distribution of proposition letters are concerned, nothing has changed. We only have to check that taking the reflexive transitive closure of $\vec{R}$ does not harm the forth property. But, as R is itself reflexive and transitive, the forth property survives. $\dashv$

Theorem 4.54 **S4** *is strongly complete with respect to the class of partially ordered reflexive and transitive trees.*

Proof. If Σ is an **S4**-consistent set of formulas, and Σ^+ is an **S4**-MCS extending Σ, then $\mathfrak{M}^{\mathbf{S4}}, \Sigma^+ \Vdash \Sigma$. Moreover, as the **S4** axioms are canonical, $\mathfrak{M}^{\mathbf{S4}}$ is a reflexive transitive model. We now transform this model into the required partial order in two steps.

Step 1. Let $\mathfrak{M}^S$ be the submodel of $\mathfrak{M}^{\mathbf{S4}}$ generated by Σ^+. Clearly this is a reflexive, transitive, point-generated model such that $\mathfrak{M}^S, \Sigma^+ \Vdash \Sigma$.

Step 2. Let $\mathfrak{M}^* = (\vec{W}, R^*, \vec{V})$ be the reflexive transitive closure of the unraveling of $\mathfrak{M}^S$ around Σ^+.

By Lemma 4.53, $\mathfrak{M}^S$ is a bounded morphic image of $\mathfrak{M}^*$ under f, hence for all sequences $\vec{s} \in f^{-1}[\Sigma]$, we have $\mathfrak{M}^*, \vec{s} \Vdash \Sigma$, and by the surjectivity of f there is at least one such $\vec{s}$. Hence we have satisfied Σ on a reflexive and transitive tree. $\dashv$

The previous proof could be summed up as follows: we found a way to use the information in a canonical model *indirectly*. The canonical model for **S4** did not have the structure we wanted — nonetheless, we successfully tapped into the information it contained via a short sequence of bisimulations ($\mathfrak{M}^*$ had $\mathfrak{M}^S$ as a bounded morphic image, and $\mathfrak{M}^S$ was a generated submodel of $\mathfrak{M}^{\mathbf{S4}}$).

Unraveling is an intrinsically *global* transformation that can change a model's geometry drastically. This is in sharp contrast to the transformation we will now examine — *bulldozing* — which works locally, and (in spite of its name) rather more gently. We will use bulldozing to answer the second of the questions posed above. Recall that a *strict* total order (STO) is a relation that is transitive, trichotomous and *irreflexive*. The class of strict total orders contains such important structures as $(\mathbb{N}, <)$, $(\mathbb{Z}, <)$, $(\mathbb{Q}, <)$, and $(\mathbb{R}, <)$ (the natural numbers, the integers, the rationals and the reals in their usual order) and is widely used to model various temporal phenomena. What is its tense logic?

Once again, it is not hard to find a plausible candidate: $\mathbf{K}_t\mathbf{4.3}$, the tense logic generated by 4, $.3_l$ and $.3_r$, seems the only reasonable candidate. For a start, $\mathbf{K}_t\mathbf{4.3}$ is strongly complete with respect to the class of *weak* total orders. (To see this, observe that the axioms are canonical for transitivity and non-branching. Hence any point generated submodel $\mathfrak{M}^S$ of the canonical model is transitive and trichotomous, and the completeness result is immediate.) Moreover, there simply are no other plausible axioms — in particular, irreflexivity is not definable. Has this (somewhat dangerous) line of reasoning led to the right answer? Let us see.

If we could find a way of transforming weakly linear models into strictly linear models we would have the desired completeness result. Note that unraveling won't help — it would turn the weak total order into a tree, thus destroying trichotomy. If only we could find a method which replaced the undesirable parts of the model with some suitable STO, and left the good parts untouched: then trichotomy would not be affected, and we would have assembled the required strict total order. Bulldozing is a way of doing this. The first step is to pin down what the 'undesirable' parts of weak total orders are. The obvious response is 'reflexive points' — but while this isn't exactly wrong, it misses the crucial insight. The entities we really need to think about are *clusters*, introduced in Chapter 2. We repeat the definition:

Definition 4.55 Let (T, R) be a transitive frame. A *cluster* on (T, R) is a subset C of T that is a maximal equivalence relation under R. That is, the restriction of R to C is an equivalence relation, and this is *not* the case for any other subset D of T such that $C \subset D$. A cluster is *simple* if it consists of a single reflexive point, and *proper* if it contains more than one point. When we say that a model contains clusters, we mean that its underlying frame does. $\dashv$

The point is this: we should not think in terms of removing isolated reflexive points; rather, we should remove entire clusters at one stroke. (Intuitively, the information in a cluster is information that 'belongs together'.) Any transitive trichotomous frame can be thought of as a strictly totally ordered collection of clusters (cf. Exercise 1.1.1). If we could remove each cluster as a single chunk, and replace it with something equivalent, we would have performed a local model transformation.

So the key question is: what should we replace clusters with? Clearly some sort of STO — but how can we do this in a truth preserving way? Note that any cluster C, even a simple one, introduces an infinity of information recurrence in both the forward and backward directions: we can follow paths within C, moving forwards and backwards, for as long as we please. Thus, when we replace a cluster C with a STO, we must ensure that the STO duplicates all the information in C infinitely often, in both directions. Bulldozing does precisely this in a straightforward way. We simply impose a strict total order on the cluster (that is, we pick some path through the cluster that visits each point once and only once) and then lay out infinitely many copies of this path in both the forward and backward direction. We then replace the cluster by the infinite repetition of the chosen path. We have squashed the clusters down into infinitely long STOs — hence the name 'bulldozing'.

Theorem 4.56 $\mathbf{K}_t$**4.3** *is strongly complete with respect to the class of strict total orders.*

Proof. Let Σ be a $\mathbf{K}_t$**4.3**-consistent set of formulas; expand it to a $\mathbf{K}_t$**4.3**-MCS Σ^+. Let $\mathfrak{M} = (T, R, V)$ be the canonical model for $\mathbf{K}_t$**4.3**. By the canonicity of the axioms, $\mathfrak{M}$ is transitive and non-branching. Let $\mathfrak{M}^S = (S, R^S, V^S)$ be the submodel of $\mathfrak{M}$ generated by Σ^+; $\mathfrak{M}^S$ is a transitive and trichotomous model such that $\mathfrak{M}^S, \Sigma^+ \Vdash \Sigma$. But $\mathfrak{M}^S$ may contain clusters, which we will bulldoze away.

Step 1. Index the clusters in $\mathfrak{M}^S$ by some suitable set I.
Step 2. Define an arbitrary strict total order $<^i$ on each cluster C_i.
Step 3. Define $C_i^\flat$ to be $C_i \times \mathbb{Z}$. ($\mathbb{Z}$ is the set of integers.)
Step 4. Define B, the set underlying the bulldozed model, to be $S^- \cup \bigcup_{i \in I} C_i^\flat$, where S^- is the set $(S \setminus \bigcup_{i \in I} C_i)$ of points *not* belonging to any cluster.
Step 5. Define a mapping $\beta : B \to S$ by: $\beta(b) = b$, if $b \in S^-$; and $\beta(b) = s$, if $b = (s, z)$.
Step 6. Define an ordering $<^b$ on B by $b <^b b'$ iff

either ($b \in S^-$ or $b' \in S^-$) and $\beta(b) R^S \beta(b')$;
or $b = (s, z)$ and $b' = (s', z')$ and

either s and s' belong to distinct clusters and $\beta(b) R^S \beta(b')$;
or s and s' belong to the same cluster and $z <_{\mathbb{Z}} z'$ (where $<_{\mathbb{Z}}$ is the usual ordering on the integers);
or s and s' belong to the same cluster C_i and $z = z'$ and $s <^i s'$.

Step 7. Define a valuation V^b on $(B, <^b)$ by $b \in V(p)$ iff $\beta(b) \in V^S(p)$.
Step 8. Define $\mathfrak{M}^B$, *the bulldozed model*, to be $(B, <^b, V^b)$.

We now make the following claims:

Claim 1. *The mapping β is a surjective bounded morphism from $(B, <^b)$ to (S, R^S), and the model $\mathfrak{M}^S$ is a bounded morphic image of $\mathfrak{M}^B$ under β.*

Claim 2. *$(B, <^b)$ is a strict total order.*

Proving these claims is a matter of checking the definitions; we leave this to the reader as Exercise 4.5.5. With this done, the theorem is immediate. By Claim 1, for any $b \in \beta^{-1}(\Sigma^+)$ we have $\mathfrak{M}^B, b \Vdash \Sigma$, and since β is surjective, there is at least one such b. Thus $\mathfrak{B}$ is a model of Σ, and by Claim 2 it has the structure we want. $\dashv$

Although it works more locally, like unraveling, bulldozing is a way of using the information in canonical models *indirectly*. Indeed, like unraveling, it accesses the information in the relevant canonical model via a sequence of bisimulations: the final model $\mathfrak{M}^B$ had $\mathfrak{M}^S$ as a bounded morphic image, and $\mathfrak{M}^S$ in turn was a generated submodel of $\mathfrak{M}$.

Bulldozing is a flexible method. For example, we're not forced to define $C_i^\flat$ to be $C_i \times \mathbb{Z}$; any unbounded STO would do. Moreover, if we used a *reflexive* total order (for example $(\mathbb{Z}, \leq)$) instead, we could prove analogous completeness results for reflexive total orders; for example, the reader is asked to show in Exercise 4.5.6 that **S**$_t$**4.3** is the logic of this class of frames. Moreover, for modal languages, we only need to ensure infinite information repetition in the *forward* direction, so structures such as $(\mathbb{N}, <)$ and $(\mathbb{N}, \leq)$ suffice.

But there are more interesting variations. For example, instead of simply ordering the points in the cluster, one can *embed* the cluster in some suitable total order, and work with its embedded image instead. By embedding the clusters in a *dense* set, it is possible to build dense totally order ordered models. And by combining such ideas with other transformations (notably filtrations) the method can be used to prove many classic completeness results of modal and tense logics.

Model manipulation methods, and completeness proofs making use of them, abound. Further examples are mentioned in the Notes, but it is not remotely possible to be encyclopedic: such methods trade on specific insights into the geometry of relational structures, and this gives rise to a wide variety of variants and combinations. The reader should certainly be familiar with such methods — they are often simple to adapt to specific problems — but it is just as important to appreciate the general point that has emerged from our discussion: even if the canonical model is not quite what we need, it can still be extremely useful. The following section further explores this theme.

Exercises for Section 4.5

4.5.1 **K** is complete with respect to the class of irreflexive frames. Unraveling shows this, but there is a much simpler transformation proof. (Hint: given a model $\mathfrak{M}$, tinker with the disjoint union of $\mathfrak{M}$ with itself.)

4.5.2 Formulate the unraveling method for modal languages containing two diamonds. Then formulate the method in such a way that bidirectional frames unravel into bidirectional frames.

4.5.3 Consider a similarity type τ with one binary operator $\triangle$. Call a τ-frame $\mathfrak{F} = (W, T)$ *acyclic* if the binary relation $R = \{(s,t) \in W^2 \mid Tstu$ or $Tsut$ for some $u \in W\}$ is acyclic (that is to say, R^+ is irreflexive). Prove that the basic modal logic $\mathbf{K}_\tau$ is strongly sound and complete with respect to the class of acyclic frames.

4.5.4 Show that the canonical model for $\mathbf{K}_t\mathbf{Q}$ contains proper clusters.

4.5.5 Prove Claims 1 and 2 of Theorem 4.56.

4.5.6 Let $\mathbf{K}_t\mathbf{QT}$ be the smallest normal temporal logic containing both $\mathbf{K}_t\mathbf{Q}$ and $p \to Fp$. Show, using a light bulldozing argument, that $\mathbf{K}_t\mathbf{QT}$ is strongly complete with respect to the class of all dense unbounded reflexive total orders.

4.6 Step-by-step

Three main ideas underly the step-by-step method:

(i) Don't consider the entire canonical model to be the key ingredient of a completeness proof. Rather, think of *selections of* MCS*s from the canonical model* as the basic building blocks.

(ii) The standard way of proving completeness is by constructing a model for a consistent set of formulas. Take the term 'constructing' as literally as possible: break it down into a sequence of steps.

(iii) Putting the first two observations together, think of the construction of a model as the stepwise selection of the needed MCSs. More precisely, think of the model construction process as approaching a limit via a sequence of ever better approximations, using local configurations of the canonical model to make improvements at each step of the construction.

The method gives us enormous control over the models we build, and even at this stage it's easy to see why. First, we do not have to worry about unpleasant features of the canonical model (such as clusters) since we only work with selections of the information that canonical structures contain. Furthermore, as we select our information one step at a time, we obtain an iron grip on what ends up in the model.

To illustrate the method's potential, we use it to prove that the logic $\mathbf{K}_t\mathbf{Q}$ defined in Definition 4.40 is strongly complete with respect to $(\mathbb{Q}, <)$. In what follows, consistency means $\mathbf{K}_t\mathbf{Q}$-consistency, and $\mathfrak{M}^c$ $(= (T^c, R^c, V^c)$ is this logic's canonical model. Furthermore we fix a maximal consistent set Σ; the goal of our proof is to construct a model $\mathfrak{M} = (T, <, V)$ for Σ such that $(T, <)$ is an ordering which is isomorphic to $(\mathbb{Q}, <)$. At each step of the construction we will be dealing with an approximation of $\mathfrak{M}$ consisting of a strictly ordered finite set of points (that will ultimately end up) in T and for each of these, the set of all formulas that we want to be the point's modal type (that is, the set of formulas holding at the point).

Definition 4.57 A *network* is a triple $\mathcal{N} = (N, R, \nu)$ such that R is a binary relation on the set N, and ν is a labeling function mapping each point in N to a maximal consistent set. $\dashv$

We are not interested in networks that are blatantly faulty as approximations of our desired model. For example, we want R to be a strict total ordering. Moreover, whenever a formula ψ is in the label set of a point s, then $F\psi$ should be in $\nu(t)$ for any t with Rts. Such requirements lead to the following definition.

Definition 4.58 A network $\mathcal{N} = (N, <, \nu)$ is *coherent* if it satisfies:

(C1) $<$ is a strict total ordering,
(C2) $\nu(s)R^c\nu(t)$ for all $s, t \in N$ such that $s < t$.

A *network for* Σ is a network such that Σ is the label set of some node. $\dashv$

C1 and C2 are the minimal requirements for a network to be useful to us; note that both requirements are *universal*. (C2 is equivalent to the requirement that if $s < t$ then $F\phi \in \nu(s)$ for all $\phi \in \nu(t)$ and $P\phi \in \nu(t)$ for all $\phi \in \nu(s)$.) But if a network is to really resemble a model, it must also satisfy certain *existential* requirements.

Definition 4.59 A network $\mathcal{N} = (N, <, \nu)$ is *saturated* if it satisfies:

(S1) R is unbounded to the left and to the right,
(S2) R is dense,
(S3) $\mathcal{N}$ is modally saturated. That is, we demand that (F) if $F\psi \in \nu(s)$ for some $s \in N$, then there is some $t \in N$ such that Rst and $\psi \in \nu(t)$, and (P) if $P\psi \in \nu(s)$ for some $s \in N$, then there is some $t \in N$ such that Rts and $\psi \in \nu(t)$.

A network is *perfect* if it is both coherent and saturated. $\dashv$

We want networks to give rise to models. Let's now check that we have imposed sufficiently many criteria on networks to achieve this.

Definition 4.60 Let $\mathcal{N} = (N, R, \nu)$ be a network. The frame $\mathfrak{F}_{\mathcal{N}} = (N, R)$ the *underlying frame* of $\mathcal{N}$. The *induced valuation* $V_{\mathcal{N}}$ on $\mathfrak{F}$ is defined by $V_{\mathcal{N}}(p) = \{s \in N \mid p \in \nu(s)\}$. The structure $\mathfrak{I}_{\mathcal{N}} = (\mathfrak{F}_{\mathcal{N}}, V_{\mathcal{N}})$ is the *induced model*. $\dashv$

The following lemma shows that our definition of perfection is the right one.

Lemma 4.61 (Truth Lemma) *Let $\mathcal{N}$ be a countably infinite perfect network. Then for all formulas ψ, and all nodes s in N,*

$$\mathfrak{I}_{\mathcal{N}}, s \Vdash \psi \text{ iff } \psi \in \nu(s).$$

Moreover, $\mathfrak{F}_{\mathcal{N}}$ is isomorphic to the ordering of the rational numbers.

Proof. The first part of the proof is by induction on the degree of ψ. The base case is clear from the definition of the induced valuation, and the steps for the booleans are straightforward. As for the modal operators, the coherency of $\mathcal{N}$ drives the left to right implication through, and saturation takes care of the other direction.

Finally, the underlying frame of a perfect network must be a dense, unbounded, strict total ordering. Hence, if it is countably infinite, it must be isomorphic to $(\mathbb{Q}, <)$ by Cantor's Theorem. (Readers unfamiliar with this theorem should try to prove this classic result from first principles. The standard proof builds up the isomorphism using a step-by-step argument!) $\dashv$

It follows from Lemma 4.61 that we have reduced the task of finding a model for our MCS Σ to the quest for a countable, perfect network for Σ. And now we arrive at the heart of the step-by-step method: the crucial idea is that each witness to the imperfection of a coherent network can be removed, one step at a time. Such witnesses will be called *defects*. There are three kinds of defect: each corresponds to a violation of a saturation condition.

Definition 4.62 Let $\mathcal{N} = (N, R, \nu)$ be a network. An S1-defect of $\mathcal{N}$ consists of a node $s \in N$ that has no successor, or no predecessor; an S2-defect is a pair (s, t) of nodes for which there is no intermediate point. An S3-defect consists of (F) a node s and a formula $F\psi \in \nu(s)$ for which there is no t in N such that Rst and $\psi \in \nu(t)$, or (P) a node s and a formula $P\psi \in \nu(s)$ for which there is no t in N such that Rts and $\psi \in \nu(t)$. $\dashv$

Now we need to say more what it to repair a defect. To make this precise, we need the notion of one network *extending* another.

Definition 4.63 Let $\mathcal{N}_0$ and $\mathcal{N}_1$ be two networks. We say that $\mathcal{N}_1$ *extends* $\mathcal{N}_0$ if $\mathfrak{F}_{\mathcal{N}_0}$ is a subframe of $\mathfrak{F}_{\mathcal{N}_1}$ and ν_0 agrees with ν_1 on N_0. $\dashv$

The key lemma of this (or for that matter, any) step-by-step proof states that any defect of a finite coherent network can be repaired. More precisely:

Lemma 4.64 (Repair Lemma) *For any defect of a finite, coherent network $\mathcal{N}$ there is a finite, coherent $\mathcal{N}' \rhd \mathcal{N}$ lacking this defect.*

Proof. Let $\mathcal{N} = (N, <, \nu)$ be a finite, coherent network and assume that $\mathcal{N}$ has some defect. We prove the Lemma by showing that all three types of defect can be removed.

S1-defects.
These are left as an exercise to the reader.

S2-defects.
Assume that there are nodes s and t in N for which there is no intermediate point.

How should we repair this defect? The basic idea is simple: just throw in a new point between s and t, and find an appropriate label for it. This can be done easily, since it follows by coherence of $\mathcal{N}$ that $\nu(s)R^c\nu(t)$, and by canonicity of the density axiom that there is some MCS Γ such that $\nu(s)R^c\Gamma R^c\nu(t)$. Hence, take some *new* node u (new in the sense that $u \notin N$) and define $\mathcal{N}' = (N', <', \nu')$ by

$$\begin{aligned} N' &:= N \cup \{u\}, \\ <' &:= < \cup \{(x,u) \mid x \leq s\} \cup \{(u,x) \mid t \leq x\}, \\ \nu' &:= \nu \cup \{(u, \Gamma)\}. \end{aligned}$$

It is clear that $\mathcal{N}'$ is a network that does not suffer from the old defect. But is $\mathcal{N}'$ coherent? Condition C1 is almost immediate by the definition, so we concentrate on C2. Let x and y be two arbitrary nodes in $\mathcal{N}'$ such that $x <' y$; we have to check that $\nu(x)R^c\nu(y)$. Now, as $<'$ is irreflexive, x and y are distinct. Moreover, there can only be a problem if one of the nodes is the new point u; assume that $y = u$ (the other case is similar). If $y = s$ then we have $\nu'(y)R^c\nu'(u)$ by our assumption on Γ, so suppose that $y \neq s$. By definition of $<'$ and the fact that there are no old nodes between s and t, this means that $y < s$, so by the coherency of $\mathcal{N}$ we have that $\nu(y)R^c\nu(s)$. Hence, it follows by the transitivity of R^c that $\nu(y)R^c\Gamma$; but then it is immediate by the definition of ν' that $\nu'(y)R^c\nu'(u)$.

S3-defects.
We only treat the P-defects; the case for F-defects follows by symmetry. Assume that there is a node s in N and a formula $P\psi$ in $\nu(s)$ for which there is no t in N such that $t < s$ and $\psi \in \nu(t)$.

Again, the basic strategy is simple: we insert a new point s' into the network (before s!) and choose an adequate label for it; this has to be a maximal consistent set containing ψ and preceding $\nu(s)$ in the preorder R^c. But *where* should s' be inserted? If we are not careful we will destroy the coherency of $\mathcal{N}$. The following maneuver (which takes advantage of the fact that $\mathfrak{F}_{\mathcal{N}}$ is a *finite* STO) overcomes the difficulty.

Let m be the unique point in N such that (1) $(m, P\psi)$ is an S3-defect in $\mathcal{N}$, and (2) for all $w < m$, $(w, P\psi)$ is *not* a defect. Such an m must exist (it is either s

itself, or one of the finitely many points preceding s) and, as we will see, we can repair $(m, P\psi)$ without problems by simply inserting the new point s' immediately before m. Repairing this minimal defect automatically repairs the defect $(s, P\psi)$.

Choose some new point s' (that is, $s' \notin S$) and let Ψ be an MCS containing ψ such that $\Psi R^c \nu(m)$; such a Ψ exists by the Existence Lemma for normal logics. Define $\mathcal{N}' = (N', <', \nu')$ as follows.

$$\begin{array}{rcl} N' & := & N \cup \{s'\} \\ <' & := & < \cup \{(x, s') \mid x < m\} \cup \{(s', x) \mid m \leq x\} \\ \nu' & := & f \cup \{(s', \Psi)\} \end{array}$$

Observe that $\mathfrak{F}_{\mathcal{N}'}$ is a strict total order, and that $\mathcal{N}'$ does *not* contain the defect $(s, P\psi)$. It only remains to ensure that $\mathcal{N}'$ satisfies the second coherency condition.

Consider two nodes $x, y \in N'$ such that $x <' y$. Again, the only cases worth checking are when either x or y is the new point s'. If we have $x = s'$ we are in a similar situation as in the case of S2-defects, so we do not go into details here.

Hence, assume that $y = s'$. By construction $\nu(s') = \Psi R^c \nu(m)$, and by the coherency of $\mathcal{N}$, $\nu(x) R^c \nu(m)$. But R^c is the canonical relation for $\mathbf{K}_t\mathbf{Q}$ — a relation with no branching to the left — hence either $\Psi R^c \nu(x)$, $\Psi = \nu(x)$ or $\nu(x) R^c \Psi$. We claim that the first two options are impossible. For, if $\Psi R^c \nu(x)$ then $\psi \in \Psi$ would imply that $P\psi \in \nu(x)$ and this contradicts the minimality of m; and if $\Psi = \nu(x)$, then $\psi \in \nu(x)$ would mean that $(s, P\psi)$ was not a defect in the first place! We conclude that $\nu(u) R^c \Psi$, which establishes coherence. $\dashv$

With both the Truth Lemma for Induced Models and the Repair Lemma at our disposal, we can prove the desired strong completeness result. The idea is straightforward. We start with a singleton network and extend it step-by-step to larger (but finite) networks by repeated use of the Repair Lemma. We obtain the required perfect network by taking the union of our sequence of networks.

Theorem 4.65 $\mathbf{K}_t\mathbf{Q}$ *is strongly complete with respect to* $(\mathbb{Q}, <)$.

Proof. Choose some set $S = \{s_i \mid i \in \omega\}$ (we will use its elements to build the required frame) and enumerate the set of potential defects (that is, the union of the sets S, $S \times S$ and $S \times \{F, P\} \times Form$). Given a consistent set of formulas Σ, expand it to an MCS Σ_0. Let $\mathcal{N}_0$ be the network $(\{s_0\}, \varnothing, (s_0, \Sigma_0))$. Trivially, $\mathcal{N}_0$ is a finite, coherent network for Σ_0.

Let $n \geq 0$ and suppose $\mathcal{N}_n$ is a finite, coherent network. Let D be the defect of $\mathcal{N}_n$ that is minimal in our enumeration. Such a D exists, since any finite network must at least have S1- and S2-defects. Form $\mathcal{N}_{n+1}$ by repairing the defect D as described in the proof of the Repair Lemma. Observe that D will not be a defect of any network extending $\mathcal{N}_n$.

Let $\mathcal{N} = (N, <, \nu)$ be given by

$$N = \bigcup_{n\in\omega} N_n,\ < = \bigcup_{n\in\omega} <_n,\ \text{ and }\ \nu = \bigcup_{n\in\omega} \nu_n.$$

It is easy to see that $\mathfrak{F}_{\mathcal{N}}$ is a strict total order. Moreover, as we chose the points in $\mathcal{N}$ from a countably infinite set, $\mathcal{N}$ is countable.

It should be intuitively clear that $\mathcal{N}$ is perfect, but the actual proof has to take care of a subtlety. Suppose that $\mathcal{N}$ is not perfect; let D be the minimal (according to our enumeration) defect of $\mathcal{N}$, say $D = D_k$. By our construction, there must be an approximation $\mathcal{N}_i$ of $\mathcal{N}$ of which D is also a defect. Note that D need *not* be the minimal defect of $\mathcal{N}_i$ — this is the subtlety. Fortunately, there can be at most k defects that are more urgent, so D will be repaired before stage $k + i$ of the construction.

Finally, by the perfection of $\mathcal{N}$ it follows from Lemma 4.61 that the induced model $\mathfrak{I}_{\mathcal{N}}$ satisfies Σ at s_0. ⊣

The step-by-step method is one of the most versatile tools at the modal logician's disposal: a wide variety of results in modal and tense logic have been using this method, it is the tool of choice for many stronger modal systems such as Arrow Logic and Since-Until logic, and we will make use of step-by-step arguments when we discuss rules for the undefinable in the following section. We urge the reader to experiment with it. A good starting point is Exercise 4.6.1.

Exercises for Section 4.6

4.6.1 Consider a modal language with three diamonds $\Diamond_1$, $\Diamond_2$ and $\Diamond_3$. Give a complete axiomatization for the class of frames $\mathfrak{F} = (W, R_1, R_2, R_3)$ satisfying $R_3 = R_1 \cap R_2$.

4.6.2 Consider, for a modal language with two diamonds $\Diamond_0$ and $\Diamond_1$, the normal modal logic $(\mathbf{S5})^{\mathbf{2}}$ axiomatized by **S5** axioms for both diamonds, and the commutativity axiom $\Diamond_0\Diamond_1 p \leftrightarrow \Diamond_1\Diamond_0 p$. Prove that this logic is complete for the class of square frames. A square frame for this language is of the form $\mathfrak{F} = (W, R_0, R_1)$ where for some set U we have

$$\begin{array}{rcl} W & = & U^2, \\ R_i st & \text{iff} & s_i = t_i. \end{array}$$

Hint: take as approximations networks of the form (N, ν) where ν is a labeling mapping *pairs* over N to maximal consistent sets.

4.6.3 Consider a similarity type τ with one binary operator $\circ$, as in arrow logic. Call a τ-frame $\mathfrak{F} = (W, T)$ a *relativized square* if W is some collection of pairs over a base set U, and $T \subseteq W^3$ satisfies $Tstu$ iff $s_0 = t_0$, $t_1 = u_0$ and $s_1 = u_1$.

(a) Prove that the basic modal logic $\mathbf{K}_\tau$ is strongly sound and complete with respect to the class of relativized squares.
(b) Try to axiomatize the logic of the class of frames (W, R) in which W is as above, but T satisfies $Tstu$ iff $s_0 = t_1$, $t_0 = u$ and $u_0 = s_1$.

4.7 Rules for the Undefinable

In the previous two sections we proved semantically driven completeness results by using standard canonical models indirectly. The present section takes a rather different approach: we enrich the deductive system with a special proof rule, and consider a special (not necessarily generated) submodel of the canonical model for this new logic. The submodel that we study contains only special *distinguishing* (or *witnessing*) MCSs. The completeness proof shows that this new canonical model has all the good properties of the original, and that, in addition, it is already in the right shape. We will make use of ideas introduced in our discussion of the step-by-step method in the previous section (in particular, the concept of a defect).

The running example in this section will (again) be the tense logic of dense unbounded strict total orderings. Recall that the difficulty when working with this logic is that there is no axiom ensuring the irreflexivity of the canonical frame — we have all the other required properties: point generated submodels of the candidate logic $\mathbf{K}_t\mathbf{Q}$ are transitive, trichotomous, dense, and unbounded. Now, in previous sections we achieved irreflexivity indirectly: either we bulldozed away clusters, or we used the canonical model for $\mathbf{K}_t\mathbf{Q}$ to induce a model on a carefully constructed irreflexive frame. In this section we will construct a canonical frame that is transitive, non-branching, dense *and irreflexive* right from the start. Indeed, if we work with a countably infinite language, every point generated subframe of this canonical model will be countable, and hence (by Cantor's Theorem) isomorphic to $(\mathbb{Q}, <)$.

The starting point of the enterprise is that irreflexivity, although not definable in basic modal languages, *can* be characterized in an alternative sense:

> If a temporal formula ψ is satisfiable on an irreflexive frame, then for any proposition letter p not occurring in ψ, the conjunction $(\neg Pp \wedge p \wedge \neg Fp) \wedge \psi$ is also satisfiable on that frame.

For, if $\mathfrak{F}, V, s \Vdash \psi$, then $\mathfrak{F}, V', s \Vdash (\neg Pp \wedge p \wedge \neg Fp) \wedge \psi$, where V' is just like V except that it assigns the singleton $\{s\}$ to p. The condition that p does not occur in ψ is crucial here: it ensures that changing the set assigned to p does not affect the satisfaction of ψ.

Now, by taking the contrapositive of the above statement, we turn it into a proof rule:

(IRR) if $\vdash (\neg Pp \wedge p \wedge \neg Fp) \rightarrow \phi$ then $\vdash \phi$, provided p does not occur in ϕ.

We have just seen that this rule is sound on the class of irreflexive frames. Moreover, note that on the class of strict total orders the formula $(\neg P\phi \wedge \phi \wedge \neg F\phi)$ is true at some state s iff s is the *only* state where ϕ holds (we need trichotomy and transitivity to guarantee this). That is, the formula $\neg P\phi \wedge \phi \wedge \neg F\phi$ acts as a sort of 'name' for the satisfying point. Call this formula $name(\phi)$. Bearing these remarks in mind, let us now see how adding this rule is of any help in proving the desired completeness result.

Definition 4.66 The logic $\mathbf{K}_t\mathbf{Q}^+$ is obtained by adding to $\mathbf{K}_t\mathbf{Q}$ the irreflexivity rule IRR. In what follows, consistency means $\mathbf{K}_t\mathbf{Q}^+$-consistency, $\vdash \phi$ means that ϕ is provable in $\mathbf{K}_t\mathbf{Q}^+$, and so on. The canonical model for $\mathbf{K}_t\mathbf{Q}^+$ is denoted by $\mathfrak{M}^c$, the canonical relation by R^c. $\dashv$

The remainder of this section is devoted to proving completeness of the proof system $\mathbf{K}_t\mathbf{Q}^+$ with respect to $(\mathbb{Q}, <)$. Of course the *result* is not surprising: we have already seen that plain old $\mathbf{K}_t\mathbf{Q}$ is strongly complete with respect to $(\mathbb{Q}, <)$. It is the *method* that is important: rules such as IRR give us a way of forming more cleanly structured canonical models.

Our goal is to construct an irreflexive version of the canonical model for $\mathbf{K}_t\mathbf{Q}^+$. The basic idea is to work only with special *witnessing* MCSs:

Definition 4.67 A maximal consistent model is called witnessing if it contains a formula of the form $name(\phi)$. $\dashv$

Why are these witnessing MCSs so interesting? Well, suppose that we are dealing with a collection W of witnessing maximal consistent sets. This collection induces a model in the obvious way: the relation is just the canonical accessibility relation restricted to W and likewise for the valuation. Now suppose that we can prove a Truth Lemma for this model; that is, suppose we can show that 'truth and membership coincide' for formulas and MCSs. It is then immediate that the underlying relation of the model is irreflexive: $name(\phi) \in \Gamma$ implies $\phi \in \Gamma$ and $F\phi \notin \Gamma$.

This is all very well, but it is obvious that we cannot just throw away non-witnessing MCSs from the canonical model without paying a price. How can we be sure that we did not throw away too many MCSs? An examination of the standard canonical completeness proof reveals that there are two spots where claims are made concerning the existence of certain MCSs.

(i) There is the Existence Lemma, which is needed to prove the Truth Lemma. In our case, whenever the formula $F\phi$ is an element of one of our witnessing MCSs (Γ, say) then there must be a *witnessing* Δ such that $\Gamma R^c \Delta$ and $\phi \in \Delta$. But if Δ is witnessing, then there is some δ with $name(\delta) \in \Delta$; it follows from the definition of the canonical accessibility relation that $F(\phi \wedge name(\delta)) \in \Gamma$. This shows that it will not do to just take the witnessing MCSs: the Existence Lemma requires stronger saturation conditions on MCSs, namely that whenever $F\phi \in \Gamma$, then there is some δ such that $F(\phi \wedge name(\delta)) \in \Gamma$ too.

(ii) If there are axioms in the logic that are canonical for some property with existential import, how can we make sure that the trimmed down version of the canonical model still validates these properties? Examples are the formulas $\Diamond\Box p \to \Box\Diamond p$, or, in the present case, the density axiom. The point is that from the density of the standard canonical frame we may not infer that its subframe formed by witnessing MCSs is dense as well: why should there be a *witnessing* MCS between two witnessing MCSs?

These two kinds of problems will be taken care of in two different ways. We first deal with the Existence Lemma. To start with, let us see how sets of MCSs give rise to models — the alternative versions of the canonical model that we already mentioned.

Definition 4.68 Let W be a set of maximal consistent sets of formulas. Define $\mathfrak{M}^c|_W$ to be the submodel of the canonical model induced by W; that is, $\mathfrak{M}^c|_W = (W, R, V)$ where R is the relation R^c restricted to W, and V is the canonical relation restricted to W. ⊣

Obviously, we are only interested in such models for which we can prove a Truth Lemma. The following definition gives a sufficient condition for that.

Definition 4.69 A set W of maximal consistent sets is called *diamond saturated* if it satisfies the requirement that for each $\Sigma \in W$ and each formula $F\psi \in \Sigma$ there is a set $\Psi \in W$ such that $\Sigma R^c \Psi$ and $\psi \in \Psi$, and the analogous condition holds for past formulas. ⊣

Lemma 4.70 (Truth Lemma) *Let W be a diamond saturated set of maximal consistent sets of formulas. Then for any $\Gamma \in W$ and any formula ϕ:*

$$\mathfrak{M}^c|_W, \Gamma \Vdash \phi \text{ iff } \phi \in \Gamma.$$

Proof. Straightforward by a induction on ϕ. ⊣

Our goal is now to prove the existence of diamond saturated collections of witnessing MCSs.

Proposition 4.71 *Let ξ be some consistent formula. Then there is a countable, diamond saturated collection W of witnessing* MCSs *such that $\xi \in \Xi$ for some $\Xi \in W$.*

Proof. The basic idea of the proof is to define W step-by-step, in a sort of parallel Lindenbaum construction on graphs. During the construction we are dealing with finite approximations of W. At each stage, one of the shortcomings of the current approximation is taken care of; this can be done in such a way that the limit of the construction has no shortcomings at all. A finite approximation of W will consist of a finite graph together with a labeling which assigns a finite set of formulas to each node of the graph. We associate a formula with each of these finite labeled graphs, and require that this corresponding formula be consistent for each of the approximations. The first graph has no edges, and just one point of which the label set is the singleton $\{\xi\}$. The construction is such that the graph is growing in two senses: edges may be added to the graph, and formulas may be added to the label sets. (Some readers may find it helpful to think of this process as a rather abstract tableau construction.) All this is done to ensure that in the limit we are dealing with a (possibly infinite) labeled graph meeting the requirements that (1) the label set of each point is a MCS, (2) each label set contains a witness and (3) if a formula of the form $F\phi$ ($P\phi$) belongs to the label set of some node, then there is an edge connecting this node to another one containing ϕ in its label set. Finally, W is defined as the range of this infinite labeling function — note that the label function will not be required to be injective.

Now for the technical details. Approximations to W will be called *networks*: a network is a quadruple $\mathcal{N} = (N, E, d, \Lambda)$ such that (N, E) is a finite, undirected, connected and acyclic graph; d is a direction function mapping each edge (s, t) of the graph to either R or its converse $R^{\breve{}}$; and Λ is a label function mapping each node of N to a finite set of formulas.

As in our earlier example of a step-by-step construction, we first want to formulate coherence conditions on networks and define the notion of a defect of network with respect to its ideal, W. We start with a formulation of the coherence of a network. Since we are working in the basic temporal similarity type — that is, we have diamonds both for looking along R and along $R^{\breve{}}$ — there is an obvious way of describing the network, from each of its nodes. Let $\mathcal{N} = (N, E, d, \Lambda)$ be some network, and let s and t be two adjacent nodes of $\mathcal{N}$. We use the following notational conventions:

$$\langle st \rangle := \begin{cases} F & \text{if } d(s,t) = R, \\ P & \text{if } d(t,s) = R^{\breve{}}. \end{cases}$$

and let $E(s)$ denote the set of nodes adjacent to s. Finally, we let $\lambda(s)$ denote the conjunction $\bigwedge \Lambda(s)$. Define

$$\begin{array}{rcl} \Delta(\mathcal{N}, s) & := & \lambda(s) \wedge \bigwedge_{v \in E(s)} \langle sv \rangle \theta(\mathcal{N}, v, s), \\ \theta(\mathcal{N}, t, s) & := & \lambda(t) \wedge \bigwedge_{s \neq v \in E(s)} \langle tv \rangle \theta(\mathcal{N}, v, t). \end{array}$$

In words, $\Delta(\mathcal{N}, s)$ starts with a local description $\lambda(s)$ of s and then proceeds to its neighbors. For each neighbor v, $\Delta(\mathcal{N}, s)$ writes a future operator if $d(s, v) = R$ (and a past operator if $d(s, v) = R^{\breve{}}$) and then starts to describe the network after v by calling θ. $\theta(\mathcal{N}, v, s)$ first gives a local description $\lambda(v)$ of v, and then recursively proceeds to the neighbors of v — except for s. The omission of s, together with the finiteness and acyclicity of the graph, ensures that we end up with a finite formula.

The following claim shows that it does not really matter from which perspective we describe $\mathcal{N}$.

Lemma 4.72 *For any network $\mathcal{N}$ and any two nodes s, t in $\mathcal{N}$, $\Delta(\mathcal{N}, s)$ is consistent iff $\Delta(\mathcal{N}, t)$ is consistent.*

Proof. By the connectedness of $\mathcal{N}$ it is sufficient to prove the Lemma for adjacent s and t; the general case can be proved by a simple induction on the length of the path connecting the two nodes.

So suppose that s and t are adjacent; without loss of generality assume that $d(s,t) = R$. Since $\mathcal{N}$ is fixed it will not lead to confusion if we abbreviate $\Delta(\mathcal{N}, x)$ by $\Delta(x)$ and $\theta(\mathcal{N}, x, y)$ by $\theta(x, y)$. Then by definition, $\Delta(s)$ is given by

$$\begin{aligned} \Delta(s) &= \lambda(s) \wedge \bigwedge_{u \in E(s)} \langle su \rangle \theta(u, s) \\ &= \lambda(s) \wedge F\theta(t, s) \wedge \bigwedge_{t \neq u \in E(s)} \langle su \rangle \theta(u, s) \\ &= F\theta(t, s) \wedge \theta(s, t). \end{aligned}$$

Likewise, we can show that

$$\Delta(t) = \theta(t, s) \wedge P\theta(s, t).$$

But it is a general property of any logic extending $\mathbf{K}_t$ that for any two formulas α and β, $F\alpha \wedge \beta$ is consistent iff $\alpha \wedge P\beta$ is consistent. From this, the Lemma is immediate. $\dashv$

The upshot of Lemma 4.72 is a good definition of the coherence of a network: we will call a network $\mathcal{N}$ *coherent* if $\Delta(\mathcal{N}, s)$ is consistent for each of (equivalently: some of) its nodes s. However, being finite, our networks will never be perfect. What kinds of defects can they have?

A *defect* of a network is either (D1) a pair (s, ϕ) such that neither ϕ nor $\neg\phi$ belongs to $\Lambda(s)$; (D2) a pair $(s, F\phi)$ such that $F\phi \in \Lambda(s)$ while there is no witness for this (in the sense that $\phi \in \Lambda(t)$ for some node t with Est and $d(s,t) = R$); (D3) a similar pair $(s, P\phi)$; or (D4) a node s without a name; that is, $name(\phi) \in \Lambda(s)$ for no formula ϕ.

We will show that each kind of defect of a network can be repaired. For this we need some terminology. A network $\mathcal{N}'$ *extends* a network $\mathcal{N}$, if $N \subset N'$, while $E = E' \cap N \times N$, $d = d'|_N$ and $\Lambda(s) \subseteq \Lambda'(s)$ for each node s of $\mathcal{N}$.

Lemma 4.73 *For any defect of a finite, coherent network $\mathcal{N}$ there is a finite, coherent $\mathcal{N}' \rhd \mathcal{N}$ lacking this defect.*

Proof. Let $\mathcal{N} = (N, E, d, \Lambda)$ be a coherent network and assume that $\mathcal{N}$ has some defect. We will prove the Lemma by showing how to remove the various types of defect.

D1-defects.
Assume that there is a node s and a formula ϕ such that neither ϕ nor $\neg\phi$ belongs to Λ. Since the formula $\Delta(\mathcal{N}, s)$ is consistent, it follows that either $\Delta(\mathcal{N}, s) \wedge \phi$ or $\Delta(\mathcal{N}, s) \wedge \neg\phi$ is consistent; let $\pm\phi$ denote the formula such that $\Delta(\mathcal{N}, s) \wedge \pm\phi$ is consistent. Now define $\mathcal{N}'$ by $N' := N$, $E' := E$, $d' := d$, while Λ' is given by $\Lambda'(t) = \Lambda(t)$ for $t \neq s$ and

$$\Lambda(s) := \Lambda(s) \cup \{\pm\phi\}.$$

Clearly, $\mathcal{N}'$ is a finite network lacking the defect (s, ϕ). It is also obvious that $\Delta(\mathcal{N}', s)$ is the formula $\Delta(\mathcal{N}, s) \wedge \pm\phi$, so $\Delta(\mathcal{N}', s)$ is consistent, and hence, $\mathcal{N}'$ is coherent.

D2-defects.
Assume that there is a node s and a formula ϕ such that $F\phi \in \Lambda(s)$ while there is no witness for this. Take a *new* node t (that is, t does not belong to N) and define $\mathcal{N}'$ as follows.

$$\begin{aligned} N' &:= N \cup \{t\}, \\ E' &:= E \cup \{(s,t)\}, \\ d' &:= d \cup \{((s,t), R)\}, \\ \Lambda' &:= \Lambda \cup \{(u, \{\phi\})\}. \end{aligned}$$

It is obvious that $\mathcal{N}'$ extends $\mathcal{N}$ and that the defect has been repaired. Finally, it is clear by the definitions that $\Delta(\mathcal{N}', s) = \Delta(\mathcal{N}, s)$: the only information that the new node adds to the description is a conjunct $F\phi$ and by assumption this was already a member of $\Lambda(s)$, and thus a conjunct of $\lambda(s)$. Hence, the coherence of $\mathcal{N}'$ is an immediate consequence of the coherence of $\mathcal{N}$.

D3-defects.
Repaired analogously to D2-defects.

D4-defects.
These are repaired in the same way as D1-defects, using the fact that if $\Delta(\mathcal{N}, s)$ is consistent, then there is a propositional variable p that does not occur in any of the label sets. And here — at last — we use the IRR-rule to show that the formula $\Delta(\mathcal{N}, s) \wedge name(p)$ is consistent. $\dashv$

Finally, we return to the proof of Proposition 4.71. Assume that ξ is a consistent formula.

By a standard step-by-step construction we can define a sequence $(\mathcal{N}_i)_{i \in \mathbb{N}}$ of networks such that

(i) $\mathcal{N}_0$ is a one-node network with label set $\{\xi\}$,
(ii) $\mathcal{N}_j$ extends $\mathcal{N}_i$ whenever $i < j$,
(iii) For every defect of any network $\mathcal{N}_i$ there is a network $\mathcal{N}_j$ with $j > i$ lacking this defect.

Let N be the set $\bigcup_{i \in \mathbb{N}} N_i$; and for $s \in N$, define $\Lambda(s) = \bigcup_{i \in \mathbb{N}} \Lambda_i(s)$. We claim that for every $s \in N$, $\Lambda(s)$ is a witnessing MCS. We first show that for all formulas ϕ, either ϕ or $\neg\phi$ belongs to $\Lambda(s)$. Let $i \in \mathbb{N}$ be such that s is already in existence in $\mathcal{N}_i$; if neither ϕ nor $\neg\phi$ belongs to $\Lambda_i(s)$, this constitutes a defect of $\mathcal{N}_i$. Hence, by the construction there is some $j > i$ such that either ϕ or $\neg\phi$ belongs to $\Lambda_j(s)$. But then the same formula belongs to $\Lambda(s)$. In the same manner we can prove that every set $\Lambda(s)$ contains a name. Now assume that $\Lambda(s)$ is not consistent; then there are formulas $\phi_1, \ldots, \phi_n$ in $\Lambda(s)$ such that $\phi_1 \wedge \cdots \wedge \phi_n$ is inconsistent. By construction, there must be a $k \in \mathbb{N}$ such that each ϕ_i belongs already to $\Lambda_k(s)$. But this contradicts the consistency of $\Delta(\mathcal{N}_k, s)$ and hence, the coherency of $\mathcal{N}_k$.

Finally, define W as the range of Λ. The preceding paragraphs show that W is a collection of witnessing MCSs. By our definition of $\mathcal{N}_0$, it follows that ξ belongs to some MCS in W.

Now let $F\phi$ be some formula in $\Gamma \in W$. By definition, there is some $s \in N$ such that $\Gamma = \Lambda(s)$, and thus, some $i \in \mathbb{N}$ such that $F\phi \in \Lambda_i(s)$. By our construction there is some $j \geq i$ and some $t \in N_j$ such that $E_j st$ and $\phi \in \Lambda_j(t)$. It follows that $\phi \in \Lambda(t)$, so it remains to prove that $\Lambda(s) R^c \Lambda(t)$. In order to reach a contradiction, suppose otherwise. Then there is a formula $\psi \in \Lambda(t)$ such that $F\psi \notin \Lambda(s)$. Since $\Lambda(s)$ is a MCS, this implies that $\neg F\psi \in \Lambda(s)$. Now let $k \in \mathbb{N}$ be large enough that $\psi \in \Lambda_k(t)$ and $\neg F\psi \in \Lambda_k(s)$. From this it is immediate that $\Delta(\mathcal{N}_k, s)$ is inconsistent; this contradicts the coherency of $\mathcal{N}_k$. This proves that W is diamond saturated.

But then we have prove that W meets all requirements phrased in the Proposition. $\dashv$

This shows that we have more or less solved the first problem concerned with working in a trimmed down version of the canonical model: we have established that every consistent formula ξ can be satisfied in an *irreflexive* canonical-like model. Let's now think about the second kind of problem. Concretely, how can we prove that we have not destroyed the nice properties of the canonical frame by moving to a subframe? In particular, how can we ascertain *density*? We will see that here we will make good use of the special naming property of the formulas $name(\phi)$, namely that they can be used as identifiers of MCSs.

Lemma 4.74 *Let W be a diamond saturated collection of witnessing maximal consistent sets of formulas, and let $<$ denote the relation R^c restricted to W. Then the frame $(W, <)$ is a non-branching, unbounded, dense, strict ordering.*

Proof. Let W and $<$ be as in the statement of the lemma. Clearly, $(W, <)$ is a subframe of the canonical frame; hence, it inherits every *universal* property of $\mathfrak{T}$, such as transitivity or non-branching. Irreflexivity follows from the fact that $\Gamma R^c \Gamma$ for no witnessing Γ. This shows that $<$ is a non-branching, strict ordering of W.

Unboundedness is not a universal condition, but nevertheless follows rather easily: simply use the fact that the formulas $F\top$ and $P\top$ are theorems of the logic and hence, belong to every maximal consistent set. Unboundedness then follows by the diamond saturation of W.

The case of density is more difficult, and here's where names are genuinely useful. Assume that Γ and Δ are two MCSs such that $\Gamma < \Delta$. We have to find a MCS Θ *in* W that lies between Γ and Δ. Let δ be the formula such that $name(\delta) \in \Delta$. It follows from $\Gamma < \Delta$ that $Fname(\delta) \in \Gamma$, so using the density axiom, we find that $FFname(\delta) \in \Gamma$. From this we may infer the existence of a MCS $\Theta \in W$ with $\Gamma < \Theta$ and $Fname(\delta) \in \Theta$.

But is $\Theta < \Delta$? Note that since $<$ is non-branching to the right, we already know that $\Theta < \Delta$ or $\Theta = \Delta$ or $\Delta < \Theta$. But it clearly cannot be the case that $\Theta = \Delta$, since $F\delta \in \Theta$ and $\neg F\delta \in \Delta$. Neither is it possible that $\Delta < \Theta$, for suppose otherwise. It would follows from $F\delta \in \Theta$ that $FF\delta \in \Delta$, so by the transitivity axiom, $F\delta \in \Delta$; but this would contradict the fact that $\neg F\delta \in \Delta$. $\dashv$

We now have all the ingredients for the main theorem of this section:

Theorem 4.75 $\mathbf{K}_t\mathbf{Q}^+$ *is complete with respect to* $(\mathbb{Q}, <)$.

Proof. Given any consistent formula ξ, construct a countable, diamond saturated set W of witnessing MCSs for ξ, as in the proof of Proposition 4.71. By the Truth Lemma 4.70, ξ is satisfiable at some MCS Ξ in the model $\mathfrak{M}^c|_W$ induced by W; and by Lemma 4.74, this model is based on a non-branching, unbounded, dense, strict ordering. But then the subframe generated by Ξ is based on a countable, dense, unbounded, strict total order and hence, isomorphic to the ordering of the rationals. $\dashv$

How widely applicable are these ideas? Roughly speaking, the situation is as follows. The basic idea is widely applicable; various rules for the undefinable have been employed in many different modal languages, and for many different classes of models (we'll see further examples in Chapter 7). Moreover, the use of such rules can be fruitfully combined with other techniques, notably the step-by-step method (this combination sometimes succeeds when all else fails). Rules for the undefinable are fast becoming a standard item in the modal logicians' toolkit.

Nonetheless the method has its limitations, at least in the kinds of modal languages we have been considering so far. These limitations are centered on the problem of working with submodels of the original canonical model.

As we saw, the first problem — retaining sufficiently many MCSs for proving the Truth Lemma — has a fairly satisfactory solution. Two remarks are in order here.

(i) The method only works well when we are working in tense logic. In the proof of the 'multiple Lindenbaum Lemma', we crucially needed operators for looking in *both* directions in order to show that it does not matter from which perspective we describe a graph. If we have no access to the information of nodes lying 'behind', we are forced to add a countably infinite *family* of more and more complex rules, instead of one single irreflexivity rule.

But there are no problems in generalizing the proof of Lemma 4.71 to similarity types with more than one tense diamond and/or versatile polyadic operators. For example, in Exercise 4.7.3 is asked to use the method to prove completeness for the language of PDL with converse programs.

(ii) Observe that we only proved *weak* completeness for $\mathbf{K}_t\mathbf{Q}^+$. This is because our proof of Lemma 4.71 only works with finite networks. In the presence of names, however, it is possible to prove a stronger version of Lemma 4.71; the basic idea is that when a MCS Γ contains a name, other MCSs may have complete access to the information in Γ through the finite 'channel' of Γ's name. For details we refer to Exercise 4.7.2.

There is a second problem which seems to be more serious. Which properties of the canonical frame can we guarantee to hold on a trimmed down version? In general, very little. Obviously, universal properties of the canonical model hold in each of its submodels, and first-order properties that are the standard translation of closed modal formulas (such as $\forall x \exists y\, Rxy$) are valid in each subframe for which a Truth Lemma holds, but that is about it.

It is at this point where the names come in very handy. In fact, in order to prove the inheritance of universal-existential properties like density, the names seem to be really indispensable. *If*, on the other hand, we have names at our disposal, we can prove completeness results for a wide range of logics. Roughly speaking, in case the logic is a tense logic, we can show that every Sahlqvist formula is 'distinguishing-canonical'. The crucial observation is that the witnessing submodel of the canonical model is a *named* model.

Definition 4.76 Let τ be some modal similarity type. A τ-model $\mathfrak{M}$ is called *named* if for every state s in $\mathfrak{M}$ there is a formula ϕ such that s is the only point in $\mathfrak{M}$ satisfying ϕ. $\dashv$

Theorem 4.77 *Let τ be some modal similarity type, and suppose that $\mathfrak{M} = (\mathfrak{F}, V)$ is a named τ-model. Then for every very simple Sahlqvist formula σ:*

$$\mathfrak{M} \Vdash \sigma \textit{ iff } \mathfrak{F} \Vdash \sigma. \tag{4.1}$$

If, in addition, $\mathfrak{M}$ *is a versatile model for* τ*, then (4.1) holds for every Sahlqvist formula.*

Proof. Let $\mathfrak{M}$ be a named model. It was the aim of Exercise 1.4.7 to let the reader show that the collection

$$A := \{V(\phi) \mid \phi \text{ a formula }\}$$

is closed under the boolean and modal operations. Hence, the structure $\mathfrak{g} = (\mathfrak{F}, A)$ is a general frame. Since $\mathfrak{M}$ is named, A contains all singletons. The result then follows from Theorem 5.90 in Chapter 5 — for the second part of the Theorem Exercise 5.6.1 is needed as well. $\dashv$

The use of rules for the undefinable really comes into its own in some of the extended modal languages studied for Chapter 7. Two main paths have been explored, and we will discuss both. In the first, the *difference operator* is added to an orthodox modal language. It is then easy to state a rule for the undefinable (even if the underlying modal language does not contain converse operators) and (by extending the remarks just made) to prove a D-Sahlqvist theorem. In the second approach, atomic formulas called *nominals* and operators called *satisfaction operators* are added to an orthodox modal language. These additions make it straightforward to define simple rules for the undefinable (even if the underlying modal language does not contain converse operators) and to prove a general completeness result without making use of step-by-step arguments.

Exercises for Section 4.7

4.7.1 We are working in the basic modal similarity type. First, prove that a frame is intransitive ($\forall xyz\,(Rxy \wedge Ryz \rightarrow \neg Rxz)$) iff we can falsify the formula $\Box p \rightarrow \Diamond\Diamond p$ at every state of the frame.

Second, let $\mathbf{KB}'$ be the logic $\mathbf{K}$, extended with the symmetry axiom $p \rightarrow \Box\Diamond p$ and the rule

(ITR) if $\vdash (\Box p \wedge \Box\Box\neg p) \rightarrow \phi$ then $\vdash \phi$, provided p does not occur in ϕ,

Show that $\mathbf{KB}'$ is sound and complete with respect to the class of symmetric, intransitive frames.

4.7.2 Assume that we are working with the logic $\mathbf{K}_t\mathbf{Q}^+$. Show that for each consistent set Σ there is a diamond saturated set of MCSs W such that $\Sigma \subseteq \Xi$ for some $\Xi \in W$.

(Hint: use a construction analogous to the one employed in the proof of Proposition 4.71. Add an infinite set of *new* variables to the language and first prove that $\Sigma \cup \{name(p)\}$ is consistent for any new variable p. A network is now allowed to have one special node with an *infinite* label set, which should contain $\Sigma \cup \{name(p)\}$. A description of a network is now an infinite set of formulas.)

4.7.3 Assume that we extend the language of PDL with a *reverse* program constructor:

- if π is a program then so is π^{-1}.

The intended accessibility relation of π^{-1} is the converse relation of R_π. Let $\mathbf{PDL}_\omega$ be the axiom system of PDL (see Section 4.8), modulo the following changes:

(i) Add the converse axiom schemas $p \to [\pi]\langle\pi^{-1}\rangle p$ and $p \to [\pi^{-1}]\langle\pi\rangle p$,
(ii) Replace the Segerberg induction axiom with the following infinitary rule:

(ω–$*$) If $\vdash \phi \to [\pi^n]\psi$ for all $n \in \omega$, then $\vdash \phi \to [\pi^*]\psi$.

Prove that this logic is sound and complete with respect to the standard models.

4.8 Finitary Methods I

In this section we introduce finite canonical models. We use such models to prove weak completeness results for non-compact logics. We examine one of the best known examples — propositional dynamic logic — in detail. More precisely, we will axiomatize the validities regular (test free) propositional dynamic logic. Recall from Chapter 1 that this has a set of diamonds $\langle\pi\rangle$ indexed by a collection of programs Π. Π consists of a collection of basic programs, and the programs generated from them using the constructors $\cup$, ;, and $*$. A frame for this language is a transition system $\mathfrak{F} = (W, R_\pi)_{\pi\in\Pi}$, but we are only interested in *regular frames*, that is, frames such that for all programs π, π_1 and π_2:

$$\begin{aligned} R_{\pi_1\cup\pi_2} &= R_{\pi_1} \cup R_{\pi_2} \\ R_{\pi_1;\pi_2} &= R_{\pi_1}; R_{\pi_2} \\ R_{\pi^*} &= (R_\pi)^*. \end{aligned}$$

We say that a formula ϕ is a PDL-validity (written $\Vdash \phi$) if it is valid on all regular frames.

The collection of PDL-validities is not compact: consider the set

$$\Sigma \;=\; \{\langle a^*\rangle p, \neg p, \neg\langle a\rangle p, \neg\langle a\rangle\langle a\rangle p, \neg\langle a\rangle\langle a\rangle\langle a\rangle p, \ldots\}.$$

Any finite subset of Σ is satisfiable on a regular frame at a single point, but Σ itself is not. This compactness failure indicates that a *strong* completeness result will be out of reach (recall Remark 4.44) so our goal (as with **KL**) should be to prove a weak completeness result. It is is not too hard to come up with a candidate axiomatization. For a start, the first two regularity conditions given above can be axiomatized by Sahlqvist axioms. The last condition is more difficult, but even here we have something plausible: recall that in Example 3.10 we saw that this last condition is *defined* by the formula set

$$\Delta = \{(p \wedge [\pi^*](p \to [\pi]p)) \to [\pi^*]p, \langle\pi^*\rangle p \leftrightarrow (p \vee \langle\pi\rangle\langle\pi^*\rangle p) \mid \pi \in \Pi\}.$$

This suggests the following axiomatization.

Definition 4.78 A logic Λ in the language of propositional dynamic logic is a *normal propositional dynamic logic* if it contains every instance of the following axiom schemas:

(i) $[\pi](p \to q) \to ([\pi]p \to [\pi]q)$
(ii) $\langle\pi\rangle p \leftrightarrow \neg[\pi]\neg p$
(iii) $\langle\pi_1;\pi_2\rangle p \leftrightarrow \langle\pi_1\rangle\langle\pi_2\rangle p$
(iv) $\langle\pi_1 \cup \pi_2\rangle p \leftrightarrow \langle\pi_1\rangle p \vee \langle\pi_2\rangle p$
(v) $\langle\pi^*\rangle p \leftrightarrow (p \vee \langle\pi\rangle\langle\pi^*\rangle p)$
(vi) $[\pi^*](p \to [\pi]p) \to (p \to [\pi^*]p)$

and is closed under modus ponens, generalization ($\vdash_\Lambda \phi$ implies $\vdash_\Lambda [\pi]\phi$, for all programs π) and uniform substitution. We call the smallest normal propositional dynamic logic **PDL**. In this section, $\vdash \phi$ means that ϕ is a theorem of **PDL**, consistency means **PDL**-consistency, and so on. $\dashv$

As we've already remarked, axioms (iii) and (iv) are (conjunctions of) Sahlqvist axioms; they are canonical for the first two regularity conditions, respectively. Further, observe that Axiom (v) is a Sahlqvist formula as well; it is canonical for the condition $R_{\pi^*} = Id \cup R_\pi ; R_{\pi^*}$. Thus we've isolated the difficult part: axiom (vi), which we will call the *induction* axiom for obvious reasons, is the formula we need to think about if we are to understand how to cope with the canonicity failure. It is probably a good idea for the reader to attempt Exercise 4.8.1 right away.

Proving the soundness of **PDL** is straightforward (though the reader should (re-)check that the induction axiom really is valid on all regular frames). We will prove completeness with the help of *finite* canonical models. Our work falls into two parts. First we develop the needed background material: finitary versions of MCSs, Lindenbaum's Lemma, canonical models, and so on. Following this, we turn to the completeness proof proper.

Recall that a set of formulas Σ is closed under subformulas if for all $\phi \in \Sigma$, if ψ is a subformula of ϕ then $\psi \in \Sigma$.

Definition 4.79 (Fischer-Ladner Closure) Let X be a set of formulas. Then X is *Fischer-Ladner closed* if it is closed under subformulas and satisfies the following additional constraints:

(i) If $\langle\pi_1;\pi_2\rangle\phi \in X$ then $\langle\pi_1\rangle\langle\pi_2\rangle\phi \in X$
(ii) If $\langle\pi_1 \cup \pi_2\rangle\phi \in X$ then $\langle\pi_1\rangle\phi \vee \langle\pi_2\rangle\phi \in X$
(iii) If $\langle\pi^*\rangle\phi \in X$ then $\langle\pi\rangle\langle\pi^*\rangle\phi \in X$.

If Σ is any set of formulas then $\mathrm{FL}(\Sigma)$ (the *Fischer Ladner closure* of Σ) is the smallest set of formulas containing Σ that is Fischer Ladner closed.

Given a formula ϕ, we define $\sim\phi$ as the following formula:

$$\sim\phi = \begin{cases} \psi & \text{if } \phi \text{ is of the form } \neg\psi, \\ \neg\phi & \text{otherwise}. \end{cases}$$

A set of formulas X is *closed under single negations* if $\sim\phi$ belongs to X whenever $\phi \in X$.

We define $\neg\mathrm{FL}(\Sigma)$, the *closure of* Σ, as the smallest set containing Σ which is Fischer Ladner closed and closed under single negations. $\dashv$

It is convenient to talk as if $\sim\phi$ really is the negation of ϕ, and we often do so in what follows. The motivation of closing a set under *single* negations is simply to have a 'connective' that is just as good as negation, while keeping the set finite. (If we naively closed under ordinary negation, then any set would have an infinite closure.)

It is crucial to note that if Σ is finite, then so is its closure. Some reflection on the closure conditions will convince the reader that this is indeed the case, but it is not entirely trivial to give a precise proof. We leave this little combinatorial puzzle to the reader as Exercise 4.8.2.

We are now ready to define the generalization of the notion of a maximal consistent set that we will use in this section.

Definition 4.80 (Atoms) Let Σ be a set of formulas. A set of formulas A is an *atom* over Σ if it is a maximal consistent subset of $\neg\mathrm{FL}(\Sigma)$. That is, A is an atom over Σ if $A \subseteq \neg\mathrm{FL}(\Sigma)$, A is consistent, and if $A \subset B \subseteq \neg\mathrm{FL}(\Sigma)$ then B is inconsistent. $At(\Sigma)$ is the set of all atoms over Σ. $\dashv$

Lemma 4.81 *Let Σ be any set of formulas, and A any element of $At(\Sigma)$. Then:*

(i) *For all $\phi \in \neg FL(\Sigma)$: exactly one of ϕ and $\sim\phi$ is in A.*
(ii) *For all $\phi \vee \psi \in \neg FL(\Sigma)$: $\phi \vee \psi \in A$ iff $\phi \in A$ or $\psi \in A$.*
(iii) *For all $\langle \pi_1; \pi_2 \rangle \phi \in \neg FL(\Sigma)$: $\langle \pi_1; \pi_2 \rangle \phi \in A$ iff $\langle \pi_1 \rangle \langle \pi_2 \rangle \phi \in A$.*
(iv) *For all $\langle \pi_1 \cup \pi_2 \rangle \phi \in \neg FL(\Sigma)$: $\langle \pi_1 \cup \pi_2 \rangle \phi \in A$ iff $\langle \pi_1 \rangle \phi \in A$ or $\langle \pi_2 \rangle \phi \in A$.*
(v) *For all $\langle \pi^* \rangle \phi \in \neg FL(\Sigma)$: $\langle \pi^* \rangle \phi \in A$ iff $\phi \in A$ or $\langle \pi \rangle \langle \pi^* \rangle \phi \in A$.*

Proof. With the possible exception of the last item, obvious. $\dashv$

Atoms are a straightforward generalization of MCSs. Note, for example, that if we choose Σ to be the set of all formulas, then $At(\Sigma)$ is just the set of all MCSs. More generally, the following holds:

Lemma 4.82 *Let $\mathcal{M}$ be the set of all* MCS*s, and Σ any set of formulas. Then*

$$At(\Sigma) = \{\Gamma \cap \neg FL(\Sigma) \mid \Gamma \in \mathcal{M}\}.$$

Proof. Exercise 4.8.3. ⊣

Unsurprisingly, an analog of Lindenbaum's Lemma holds:

Lemma 4.83 *If $\phi \in \neg FL(\Sigma)$ and ϕ is consistent, then there is an $A \in At(\Sigma)$ such that $\phi \in A$.*

Proof. If Σ is infinite, the result is exactly Lindenbaum's Lemma, so let us turn to the more interesting finite case. There are two ways to prove this. We could simply apply Lindenbaum's Lemma: as ϕ is consistent, there is an MCS Γ that contains ϕ. Thus, by the previous lemma, $\Gamma \cap \neg\mathrm{FL}(\Sigma)$ is an atom containing ϕ.

But this is heavy handed: let's look for a finitary proof instead. Note that the information in an atom A can be represented by the single formula $\bigwedge_{\phi\in A}\phi$. We will write such conjunctions of atoms as $\widehat{A}$. Obviously $\widehat{A} \notin A$.

Using this notation, we construct the desired atom as follows. Enumerate the elements of $\neg\mathrm{FL}(\Sigma)$ as $\sigma_1, \ldots, \sigma_m$. Let A_1 be $\{\sigma_1\}$. Suppose that A_n has been defined, where $n < m$. We have that

$$\vdash \widehat{A}_n \leftrightarrow (\widehat{A}_n \wedge \sigma_{n+1}) \vee (\widehat{A}_n \wedge {\sim}\sigma_{n+1}),$$

as this is a propositional tautology, thus either $A_n \cup \{\sigma_{n+1}\}$ or $A_n \cup \{{\sim}\sigma_{n+1}\}$ is consistent. Let A_{n+1} be the consistent extension, and let A be A_m. Then A is an atom containing ϕ. ⊣

Note the technique: we forced a finite sequence of choices between σ and ${\sim}\sigma$. Actually, we did much the same thing in the proof of Lemma 4.26, the Existence Lemma for modal languages of arbitrary similarity type, and we'll soon have other occasions to use the idea.

Now that we have Lemma 4.83, it is time to define finite canonical models:

Definition 4.84 (Canonical Model over Σ) Let Σ be a finite set of formulas. The *canonical model over* Σ is the triple $(At(\Sigma), \{S^{\Sigma}_{\pi}\}_{\pi\in\Pi}, V^{\Sigma})$ where for all propositional variables p, $V^{\Sigma}(p) = \{A \in At(\Sigma) \mid p \in A\}$, and for all atoms $A, B \in At(\Sigma)$ and all programs π,

$$AS^{\Sigma}_{\pi}B \text{ if } \widehat{A} \wedge \langle\pi\rangle\widehat{B} \text{ is consistent}.$$

V^{Σ} is called the *canonical valuation*, and the S_{π} are called the *canonical relations*. We generally drop the Σ superscripts. ⊣

Although we have defined it purely finitarily, the canonical model over Σ is actually something very familiar: a filtration. Which filtration? Exercise 4.8.4 asks the reader to find out. Further, note that although some of the above discussion is specific to propositional dynamic logic (for example, the use of the Fischer Ladner closure) the basic ideas are applicable to any modal language. In Exercise 4.8.7 we ask the reader to apply such techniques to the logic **KL**.

But of course, the big question is: does this finite canonical model *work*? Given a consistent formula ϕ, we need to satisfy ϕ in a regular model. This gives two natural requirements on the canonical model: first, we need to prove some kind of Truth Lemma, and second, we want the model to be regular. The good news is that we can easily prove a Truth Lemma; the bad news is that we are unable to show regularity. This means that we cannot use the canonical model itself; rather, we will work with the canonical relations S_π for the atomic relations only, and define relations R_π for the other programs in a way that *forces* the model to be regular.

Definition 4.85 (Regular PDL-model over Σ) Let Σ be a set of formulas. For all basic programs a, define R_a^Σ to be S_a^Σ. For all complex programs, inductively define the **PDL**-relations R_π^Σ in the usual way using unions, compositions, and reflexive transitive closures. Finally, define $\mathfrak{R}$, the *regular* PDL-*model over* Σ to be $(At(\Sigma), \{R_\pi^\Sigma\}_{\pi \in \Pi}, V^\Sigma)$, where V^Σ is the canonical valuation. Again, we generally drop the Σ superscripts. $\dashv$

But of course, *now* the main question is, will be able to prove a Truth Lemma? Fortunately, we can prove the key element of this lemma, namely, an Existence Lemma (cf. Lemma 4.89 below). First the easy part. As the canonical relations S_a are identical to the **PDL**-relations R_a for all basic programs a, we have:

Lemma 4.86 (Existence Lemma for Basic Programs) *Let A be an atom, and a a basic program. Then for all formulas $\langle a \rangle \psi$ in $\neg$FL(Σ), $\langle a \rangle \psi \in A$ iff there is a $B \in At(\Sigma)$ such that AR_aB and $\psi \in B$.*

Proof. This can be proved by appealing to the standard Existence Lemma and then taking intersections (as in Lemma 4.83) — but it is more interesting to prove it finitarily. For the right to left direction, suppose there is a $B \in At(B)$ such that AR_aB and $\psi \in B$. As R_a and S_a are identical for basic programs, AS_aB, thus $\widehat{A} \wedge \langle a \rangle \widehat{B}$ is consistent. As ψ is one of the conjuncts in $\widehat{B}$, $\widehat{A} \wedge \langle a \rangle \psi$ is consistent. As $\langle a \rangle \psi$ is in $\neg$FL(Σ) it must also be in A, for A is an atom and hence *maximal* consistent in $\neg$FL(Σ).

For the left to right direction, suppose $\langle a \rangle \psi \in A$. We construct an appropriate atom B by forcing choices. Enumerate the formulas in $\neg$FL(Σ) as $\sigma_1, \dots, \sigma_m$. Define B_0 to be $\{\psi\}$. Suppose as an inductive hypothesis that B_n is defined such that $\widehat{A} \wedge \langle a \rangle \widehat{B_n}$ is consistent (where $0 \leq n < m$). We have

$$\vdash \langle a \rangle \widehat{B}_n \leftrightarrow \langle a \rangle ((\widehat{B}_n \wedge \sigma_{n+1}) \vee (\widehat{B}_n \wedge {\sim}\sigma_{n+1}))$$

thus

$$\vdash \langle a \rangle \widehat{B}_n \leftrightarrow (\langle a \rangle (\widehat{B}_n \wedge \sigma_{n+1}) \vee \langle a \rangle (\widehat{B}_n \wedge {\sim}\sigma_{n+1})).$$

Therefore either for $B' = B_n \cup \{\sigma_{n+1}\}$ or for $B' = B_n \cup \{\sim\sigma_{n+1}\}$ we have that $\widehat{A} \wedge \langle a \rangle \widehat{B'}$ is consistent. Choose B_{n+1} to be this consistent expansion, and let B be B_m. B is the atom we seek. $\dashv$

Now for the hard part. Axioms (v) and (vi) cannot enforce the desired identity between S_π and R_π. But good news is at hand. These axioms are very strong and manage to 'approximate' the desired behavior fairly well. In particular, they are strong enough to ensure that $S_\pi \subseteq R_\pi$ for arbitrary programs π. This inclusion will enable us to squeeze out a proof of the desired Existence Lemma. The following lemma is the crucial one.

Lemma 4.87 *For all programs* π, $S_{\pi^*} \subseteq (S_\pi)^*$.

Proof. We need to show that for all programs π, if $AS_{\pi^*}B$ then there is a finite sequence of atoms $C_0, \ldots, C_n$ such that $A = C_0 S_\pi C_1, \ldots, C_{n-1} S_\pi C_n = B$. Let $\mathcal{D}$ be the set of all atoms reachable from A by such a sequence. We will show that $B \in \mathcal{D}$.

Define δ to be $\bigvee_{D \in \mathcal{D}} \widehat{D}$. Note that $\delta \wedge \langle \pi \rangle \neg\delta$ is *inconsistent*, for suppose otherwise. Then $\delta \wedge \langle \pi \rangle \widehat{E}$ would be consistent for at least one atom E *not* in $\mathcal{D}$, which would mean that $\widehat{D} \wedge \langle \pi \rangle \widehat{E}$ was consistent for at least one $D \in \mathcal{D}$. But then by $DS_\pi E$, E could be reached from A in finitely many S_π steps, which would imply that $E \in \mathcal{D}$ — which it is not.

As $\delta \wedge \langle \pi \rangle \neg\delta$ is inconsistent, $\vdash \delta \to [\pi]\delta$, hence by generalization $\vdash [\pi^*](\delta \to [\pi]\delta)$. By axiom (vi), $\vdash \delta \to [\pi^*]\delta$. Now, as $AS_{\pi^*}A$, $\widehat{A}$ is one of the disjuncts in δ, thus $\vdash \widehat{A} \to \delta$ and hence $\vdash \widehat{A} \to [\pi^*]\delta$. As our initial assumption was that $\widehat{A} \wedge \langle \pi^* \rangle \widehat{B}$ is consistent, it follows that $\widehat{A} \wedge \langle \pi^* \rangle (\widehat{B} \wedge \delta)$ is consistent too. But this means that for one of the disjuncts $\widehat{D}$ of δ, $\widehat{B} \wedge \widehat{D}$ is consistent. As B and D are atoms, $B = D$ and hence $B \in \mathcal{D}$. $\dashv$

With the help of this lemma, it is straightforward to prove the desired inclusion:

Lemma 4.88 *For all programs* π, $S_\pi \subseteq R_\pi$.

Proof. Induction on the structure of π. The base case is immediate, for we defined R_a to be S_a for all basic programs a. So suppose $AS_{\pi_1;\pi_2}B$, that is, $\widehat{A} \wedge \langle \pi_1; \pi_2 \rangle \widehat{B}$ is consistent. By axiom (iii), $\widehat{A} \wedge \langle \pi_1 \rangle \langle \pi_2 \rangle \widehat{B}$ is consistent as well. Using a 'forcing choices' argument we can construct an atom C such that $\widehat{A} \wedge \langle \pi_1 \rangle \widehat{C}$ and $\widehat{C} \wedge \langle \pi_2 \rangle \widehat{B}$ are both consistent. But then, by the inductive hypothesis, $AR_\pi C$ and $CR_\pi B$. It follows that $AR_{\pi_1;\pi_2}B$, as required. A similar argument using axiom 4 shows that $S_{\pi_1 \cup \pi_2} \subseteq R_{\pi_1 \cup \pi_2}$.

The case for reflexive transitive closures follows from the previous lemma and the observation that $S_\pi \subseteq R_\pi$ implies $(S_\pi)^* \subseteq (R_\pi)^*$. $\dashv$

We can now prove an Existence Lemma for *arbitrary* programs.

Lemma 4.89 (Existence Lemma) *Let A be an atom and let $\langle\pi\rangle\psi$ be a formula in $\neg FL(\Sigma)$. Then $\langle\pi\rangle\psi \in A$ iff there is a B such that $AR_\pi B$ and $\psi \in B$.*

Proof. The left to right direction puts the crucial inclusion to work. Suppose $\langle\pi\rangle\psi \in A$. We can build an atom B such that $AS_\pi B$ by 'forcing choices' in the now familiar manner. But we have just proved that $S_\pi \subseteq R_\pi$, thus $AR_\pi B$ as well.

For the right to left direction we proceed by induction on the structure of π. The base case is just the Existence Lemma for basic programs, so suppose π has the form $\pi_1;\pi_2$, and further suppose that $AR_{\pi_1;\pi_2}B$ and $\psi \in B$. Thus there is an atom C such that $AR_{\pi_1}C$ and $CR_{\pi_2}B$ and $\psi \in B$. By the Fischer Ladner closure conditions, $\langle\pi_2\rangle\psi$ belongs to $\neg$FL(Σ), hence by the inductive hypothesis, $\langle\pi_2\rangle\psi \in C$. Similarly, as $\langle\pi_1\rangle\langle\pi_2\rangle\psi$ is in $\neg$FL(Σ), $\langle\pi_1\rangle\langle\pi_2\rangle\psi \in A$. Hence by Lemma 4.81, $\langle\pi_1;\pi_2\rangle\psi \in A$, as required.

We leave the case $\pi = \pi_1 \cup \pi_2$ to the reader and turn to the reflexive transitive closure: suppose π is of the form ρ^*. Assume that $AR_{\rho^*}B$ and $\psi \in B$. This means there is a finite sequence of atoms $C_0, \dots, C_n$ such that $A = C_0 R_\rho C_1, \dots, C_{n-1} R_\rho C_n = B$. By a subinduction on n we prove that $\langle\rho^*\rangle\psi \in C_i$ for all i; the required result for $A = C_0$ is then immediate.

Base case: $n = 0$. This means $A = B$. From axiom (v) we have that $\vdash \langle\rho^*\rangle\psi \leftrightarrow \psi \vee \langle\rho\rangle\langle\rho^*\rangle\psi$, and hence that $\vdash \psi \to \langle\rho^*\rangle\psi$. Thus $\langle\rho^*\rangle\psi \in A$.

Inductive step. Suppose the result holds for $n \leq k$, and that

$$A = C_0 R_\rho C_1, \dots, C_k R_\rho C_{k+1} = B.$$

By the inductive hypothesis, $\langle\rho^*\rangle\psi \in C_1$. Hence $\langle\rho\rangle\langle\rho^*\rangle\psi \in A$, for $\langle\rho\rangle\langle\rho^*\rangle\psi \in \neg$FL$(\Sigma)$. But $\vdash \langle\rho^*\rangle\psi \leftrightarrow \psi \vee \langle\rho\rangle\langle\rho^*\rangle\psi$. Hence $\langle\rho^*\rangle\psi \in A$.

This completes the subinduction, and establishes the required result for $\langle\rho^*\rangle$. It also completes the main induction and thus the proof of the lemma. $\dashv$

Lemma 4.90 (Truth Lemma) *Let $\mathfrak{R}$ be the regular* **PDL***-model over Σ. For all atoms A and all $\psi \in \neg FL(\Sigma)$, $\mathfrak{R}, A \Vdash \psi$ iff $\psi \in A$.*

Proof. Induction on the number of connectives. The base case follows from the definition of the canonical valuation over Σ. The boolean case follows from Lemma 4.81 on the properties of atoms. Finally, the Existence Lemma pushes through the step for the modalities in the usual way. $\dashv$

The weak completeness result for propositional dynamic logic follows.

Theorem 4.91 **PDL** *is weakly complete with respect to the class of all regular frames.*

Exercises for Section 4.8

4.8.1 Show that the induction axiom is not canonical.

4.8.2 Prove that for a finite set Σ, its closure set $\neg\mathrm{FL}(\Sigma)$ is finite as well.

4.8.3 Prove Lemma 4.82. That is, show that $At(\Sigma) = \{\Gamma \cap \neg\mathrm{FL}(\Sigma) \mid \Gamma \in \mathcal{M}\}$, where $\mathcal{M}$ is the set of all MCSs, and Σ is any set of formulas.

4.8.4 Show that the finite models defined in the **PDL** completeness proofs are isomorphic to certain filtrations.

4.8.5 Show that for any collection of formulas Σ, $\vdash \bigvee_{A \in At(\Sigma)} \widehat{A}$.

4.8.6 Extend the completeness proof in the text to PDL with tests. Once you have found an appropriate axiom governing tests, the main line of the argument follows that given in the text. However because test builds modalities from formulas you will need to think carefully about how to state and prove analogs of the key lemmas (such as Lemmas 4.87 and 4.88).

4.8.7 Use finite canonical models to show that **KL** is weakly complete with respect to the class of finite strict partial orders (that is, the class of finite irreflexive transitive frames). (Hint: given a formula ϕ, let Φ be the set of all ϕ's subformulas closed under single negations. Let the points in the finite canonical model be all the maximal **KL**-consistent subsets of Φ. For the relation R, define Rww' iff (1) for all $\Box\phi \in w$, $\Box\phi, \phi \in w'$ and (2) there is some $\Box\phi \in w'$ such that $\Box\phi \notin w$. Use the natural valuation. You will need to make use of the fact that $\vdash_{KL} \Box\Box\phi \to \Box\phi$; bonus points if you can figure out how to prove this yourself!)

4.8.8 Building on the previous result, show that **KL** is weakly complete for the class of finite transitive trees. (Hint: unravel.)

4.9 Finitary Methods II

As we remarked at the end of Section 4.4, although the incompleteness results show that frame-theoretic tools are incapable of analyzing the entire lattice of normal modal logics, they are capable of yielding a lot of information about some of its subregions. The normal logics extending **S4.3** are particularly well-behaved, and in this section we prove three results about them. First, we prove Bull's theorem: all such logics have the *finite frame property*. Next, we show that they are all *finitely axiomatizable*. Finally, we show that each of these logics has a *negative characterization in terms of finite sets of finite frames*, which will be important when we analyze their computational complexity in Chapter 6.

The logics extending **S4.3** are logics of frames that are rooted, transitive, and connected ($\forall xy\,(Rxy \vee Ryx)$)). To see this, recall that **S4.3** has as axioms 4, T, and .3. These formulas are canonical for transitivity, reflexivity, and no branching to the right, respectively. Hence any point-generated submodel of the canonical model

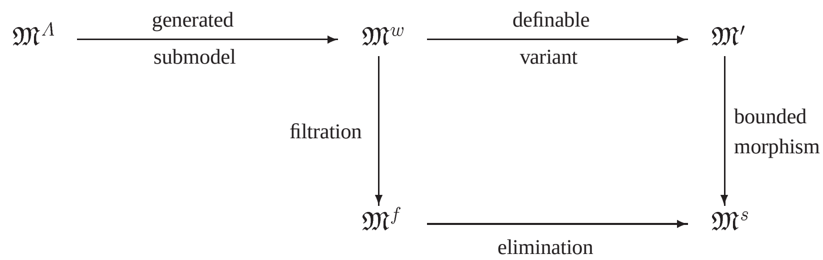


Fig. 4.2. The models we will construct, and their relationships

for these logics inherits all three properties, and will in addition be rooted and connected. Now, any connected model is reflexive. Thus *rootedness*, *transitivity*, and *connectedness* are the fundamental properties, and we will call any frame that has them an **S4.3** *frame*. Note that any **S4.3** frame can be viewed as a chain of *clusters* (see Definition 2.43), a perspective which will frequently be useful in what follows.

Bull's Theorem

Our first goal is to prove Bull's theorem: all extensions of **S4.3** have the finite frame property. In Definition 3.23 we defined the finite frame property as follows: Λ has the finite frame property with respect to a class of finite frames F if and only if $\mathsf{F} \Vdash \Lambda$, and for every formula ϕ such that $\phi \notin \Lambda$ there is some $\mathfrak{F} \in \mathsf{F}$ such that ϕ is falsifiable on $\mathfrak{F}$. Using the terminology introduced in this chapter, we can reformulate this more concisely as follows: Λ has the finite frame property if and only if there is a class of finite frames F such that $\Lambda = \Lambda_{\mathsf{F}}$. So, to prove Bull's Theorem, we need to show that if Λ extends **S4.3**, then any Λ-consistent formula ϕ is satisfiable in a finite model (W, R, V) such that $(W, R) \Vdash \Lambda$. In short, Bull's Theorem is essentially a general weak completeness result covering all logics extending **S4.3**.

But how are we to build the required models? By transforming the canonical model. Suppose ϕ is Λ-consistent. Let w be any Λ-MCS containing ϕ, and let $\mathfrak{M}^w = (W^w, R^w, V^w)$ be the submodel of $\mathfrak{M}^\Lambda$ generated by w. Then $\mathfrak{M}^w, w \Vdash \phi$, and (as just discussed) $\mathfrak{M}^w$ is based on an **S4.3** frame. We are going to transform $\mathfrak{M}^w$ into a finite model $\mathfrak{M}^s$ that satisfies ϕ and is based on an **S4.3** frame that validates Λ.

Figure 4.2 shows what is involved. We are going to transform $\mathfrak{M}^w$ in two distinct ways. One involves taking a filtration and eliminating certain points; this is the technical heart of the proof. The other involves defining a bounded morphism on a definable variant $\mathfrak{M}'$ of $\mathfrak{M}^w$; this part uses the results on definable variants and distinguishing models proved in Section 3.4. These transformations offer us two perspectives on the properties of $\mathfrak{M}^s$, and together yield enough information to prove the result.

And so to work. We first discuss the filtration/elimination transformation. Let Φ be the (finite) set consisting of all subformulas of $\Diamond\phi$, and let $\mathfrak{M}^f = (W^f, R^f, V^f)$ be the result of transitively filtrating $\mathfrak{M}^w$ through Φ. Recall that the relation R^f used in transitive filtrations is defined by $R^f|u||v|$ iff $\Diamond\psi \in v$ implies $\Diamond\psi \in u$, for all $\Diamond\psi \in \Phi$, and all $u, v \in W^w$; see Lemma 2.42. As Φ is finite, so is W^f. By the Filtration Theorem (Theorem 2.39) $\mathfrak{M}^f, |u| \Vdash \psi$ iff $\mathfrak{M}^w, u \Vdash \psi$, for all $\psi \in \Phi$, and all $u \in W^w$. Moreover, R^f is transitive, reflexive, and connected, and $|w|$ is a root of the filtration, thus $\mathfrak{M}^f$ is based on an **S4.3** frame. Hence the frame underlying $\mathfrak{M}^f$ is a finite chain of finite clusters.

Now for the key elimination step. We want to build a finite model based on a frame for Λ. Now, we don't know whether $\mathfrak{M}^w$ is based on such a frame, but we *do* know that $\mathfrak{M}^w \Vdash \Lambda$. If we could transfer the truth of Λ in $\mathfrak{M}^w$ to a finite *distinguishing* model, then by item (iii) of Lemma 3.27 we would have immediately have Bull's Theorem. Unfortunately, while $\mathfrak{M}^f$ is finite, and also (being a filtration) distinguishing, we have no guarantee that $\mathfrak{M}^f \Vdash \Lambda$. This reflects something discussed in Section 2.3: the natural map associated with a filtration need not be bounded morphism. It also brings us to the central idea of the proof: *eliminate all points in* $\mathfrak{M}^f$ *which prevent the natural map from being a bounded morphism*. Obviously, any model built from $\mathfrak{M}^f$ by eliminating points will be finite and distinguishing. So the crucial questions facing us are: which points should be eliminated? And how do we know that they can be thrown away without affecting the satisfiability of formulas in Φ?

Recall that the natural map associated with a filtration sends each point u in the original model to the equivalence class $|u|$ in the filtration. So if the natural map from the frame underlying $\mathfrak{M}^w$ to the frame underlying $\mathfrak{M}^f$ is *not* a bounded morphism, this means that for some $\beta, \alpha \in W^f$ we have that $R^f\beta\alpha$ but

$$\neg\forall v \in \beta\, \exists z\, (R^w vz \wedge z \in \alpha),$$

or equivalently, that $R^f\beta\alpha$ but

$$\exists v \in \beta\, \forall z\, (z \in \alpha \to \neg R^w vz).$$

This motivates the following definition:

Definition 4.92 Suppose β, $\alpha \in W^f$. We say that α is *subordinate* to β (α *sub* β) if there is a $v \in \beta$ such that for all $z \in \alpha$, it is *not* the case that $R^w vz$. $\dashv$

So: if $\mathfrak{M}^f$ is *not* a bounded morphic image of $\mathfrak{M}^w$ under the natural map, then there is some $\alpha \in W^f$ such that for some $\beta \in W^f$, $R^f\beta\alpha$ and α *sub* β. We must get rid of all such α; we will call them *eliminable* points. But to show that we can safely eliminate them, we need to understand the *sub* relation a little better.

Lemma 4.93 (i) *If α sub β, then there is a $v \in \beta$ such that for all $z \in \alpha$, $R^w zv$.*

(ii) *If α sub β then $R^f \alpha\beta$.*

(iii) *The sub relation is transitive and asymmetric.*

(iv) *Suppose $\alpha, \beta, \gamma \in W^f$ such that α sub γ and not α sub β. Then β sub γ.*

Proof. For item (i), note that by definition there is a $v \in \beta$ such that for all $z \in \alpha$, it is *not* the case that $R^w vz$. But R^w is a connected relation, hence for every $z \in \alpha$, $R^w zv$.

For item (ii), suppose α *sub* β. By item (i), this means that there is some element v of β, such that every element of α R^w-precedes v. Now if $\Diamond\psi \in \beta$, then $\mathfrak{M}^w, v \Vdash \Diamond\psi$. Hence (by the transitivity of R^w) for all $z \in \alpha$, $\mathfrak{M}^w, z \Vdash \Diamond\psi$ too. This means that $\Diamond\psi \in \alpha$, that is, $R^f \alpha\beta$. (It follows that if the natural map fails to be bounded morphism because of its behavior on the points β and α, then the eliminable point α belongs to the *same* cluster as β.)

Items (iii) and (iv) are left for the reader as Exercise 4.9.1. $\dashv$

We are now ready for the key result: we can safely get rid of all the eliminable points; there are enough *non*-eliminable points left to prove an Existence Lemma:

Lemma 4.94 (Existence Lemma) *Let $u \in W^w$ and suppose $\Diamond\psi \in u \cap \Phi$. Then there is a $|v| \in W^f$ such that $R^f|u||v|$, $\psi \in |v|$, and $|v|$ is* not *eliminable.*

Proof. Construct a maximal sequence $\alpha_0, \alpha_1, \ldots$ through W^f with the following properties:

(i) $\alpha_0 = |u|$.

(ii) If $i > 0$ and odd, then α_i is some $|v|$ such that $\psi \in v$, $R^f \alpha_{i-1}|v|$, and not $|v|$ *sub* α_{i-1}.

(iii) If $i > 0$ and even, then α_i is some $|v|$ such that $R^f|v|\alpha_{i-1}$ and α_{i-1} *sub* $|v|$.

Here's the basic idea. Think of this sequence as a series of moves through the model. We are given $\Diamond\psi$, and our goal is to find a R^f-related ψ-containing point that is not eliminable. So, on our first move (an *odd* move) we select an R^f-related ψ-containing point (we are guaranteed to find one, pretty much as in any Existence Lemma). If the point is *not*-eliminable we have found what we need and are finished. Unfortunately, the point may well be eliminable. If so, we make a second move (an *even* move) to another point *in the same cluster* — namely a point to which the first point we found is subordinate. We iterate the process, and eventually we will find what we are looking for. We now make this (extremely sketchy) outline precise.

Claim 1. *For every item $\alpha_i = |v|$ in the sequence, $\Diamond\psi \in v$.*

If $i = 0$, $\alpha_i = |u|$ and by assumption $\Diamond\psi \in u$. If $i > 0$ and odd, then $\psi \in |v|$ by construction, hence $\psi \in v$. As v is a Λ-MCS it contains $\psi \to \Diamond\psi$, thus $\Diamond\psi \in v$ also. Finally, if $i > 0$ and even, then as we have just seen, $\Diamond\psi \in \alpha_{i-1}$. By construction, $R^f|v|\alpha_{i-1}$ hence $\Diamond\psi \in |v|$ and hence $\Diamond\psi \in v$. This proves Claim 1.

Claim 2. *The sequence terminates.*

Suppose i is even. By property (iii), α_{i+1} *sub* α_{i+2} and by property (ii), it is not the case that α_{i+1} *sub* α_i. Hence by item 3 of Lemma 4.93, α_i *sub* α_{i+2}. By item (ii) of Lemma 4.93, *sub* is a transitive and asymmetric relation, thus each α_i, for i even, is distinct. As there are only finitely many elements in W^f, the sequence must terminate. This proves Claim 2.

Claim 3. *The sequence does not terminate on even i.*

Suppose i is even. We need to show that there is an $\alpha_{i+1} \in W^f$ such that $R^f\alpha_i\alpha_{i+1}$ and not α_{i+1} *sub* α_i. Let $\{\beta_1, \ldots, \beta_m\}$ be $\{\beta \in W^f \mid \beta$ *sub* $\alpha_i\}$. Then for each k $(1 \le k \le m)$ there is a $v_k \in \alpha_i$ such that not $R^w v_k z$, for all $z \in \beta_k$. Let v be one of these points v_k such that for all k, $R^w v_k v$, for $1 \le k \le m$. (It is always possible to choose such a v as R^w is connected.) As $\alpha_i = |v|$, by Claim 1 $\Diamond\psi \in v$. By the Existence Lemma for normal logics (Lemma 4.20), there is a $x \in W$ such that $\psi \in x$ and $R^w vx$. Moreover, not $|x|$ *sub* $|v|$. For suppose for the sake of a contradiction that $|x|$ *sub* $|v|$. Then $|x| = \beta_k$, for some $1 \le k \le m$, and hence not $R^w v_k x$. But $R^w v_k v$ and $R^w vx$, hence (by transitivity) $R^w v_k x$ — contradiction. We conclude that not $|x|$ *sub* $|v|$, hence (recalling that $|v| = \alpha_i$) we can always choose α_{i+1} to be $|x|$. This proves Claim 3.

We can now prove the result. By Claims 2 and 3, the sequence terminates on $\alpha_m = |v|$, for some odd number m. By construction, $\psi \in v$, hence $\psi \in |v|$. Since α_{m+1} does not exist, α_m is not eliminable. By construction, for all even i, $R^f\alpha_i\alpha_{i+1}$. By item (ii) of Lemma 4.93, for all odd i, $R^f\alpha_i\alpha_{i+1}$. Hence by the transitivity of R^f, $R^f|u||v|$, and we are through. $\dashv$

We now define the model $\mathfrak{M}^s$. Let W^s be the set of non-eliminable points in W^f. (Note that by the previous lemma there must be at least one such point, for $\Diamond\phi \in w \cap \Phi$.) Then $\mathfrak{M}^s = (W^s, R^s, V^s)$ is $\mathfrak{M}^f$ restricted to W^s. Hence $\mathfrak{M}^s$ is a finite distinguishing model, and (W^s, R^s) is an **S4.3** frame.

Lemma 4.95 $\mathfrak{M}^s$ *satisfies* ϕ.

Proof. First, we show by induction on the structure of ψ that for all $\psi \in \Phi$, and all $|u| \in W^s$, $\mathfrak{M}^s, |u| \Vdash \psi$ iff $\psi \in u$. The only interesting case concerns the modalities. So suppose $\Diamond\psi \in u$. By the previous lemma, there is some $|v|$ such that $R^f|u||v|$, $\psi \in |v|$, and $|v|$ is not eliminable. As $\psi \in |v|$, $\psi \in v$, hence by the inductive hypothesis, $\mathfrak{M}^s, |v| \Vdash \psi$, hence $\mathfrak{M}^s, |u| \Vdash \Diamond\psi$ as desired. The converse is straightforward; we leave it to the reader.

It follows that ϕ is satisfied somewhere in $\mathfrak{M}^s$. For, as $\Diamond\phi \in w \cap \Phi$, by Lemma 4.94 there is a non-eliminable $|u|$ such that $R^f|w||u|$ and $\phi \in |u|$. Hence $\phi \in u$, and $\mathfrak{M}^s, |u| \Vdash \phi$. $\dashv$

We are almost there. If we can show that $\mathfrak{M}^s \Vdash \Lambda$, then as $\mathfrak{M}^s$ is a finite distinguishing model, its frame validates Λ and we are through. Showing that $\mathfrak{M}^s \Vdash \Lambda$, will take us along the other path from $\mathfrak{M}^w$ to $\mathfrak{M}^s$ shown in Figure 4.2. That is, we will show that $\mathfrak{M}^s$ is a bounded morphic image of a definable variant $\mathfrak{M}'$ of $\mathfrak{M}^w$.

The required bounded morphism f is easy to describe: it agrees with the natural map on all *non*-eliminable points, and where the natural map sent a point w to a point that has been eliminated, $f(w)$ will be a point 'as close as possible' to the eliminated point. Let's make this precise. Enumerate the elements of W^s. Define $f : W^w \to W^s$ by

$$f(w) \quad = \quad \begin{cases} |w|, \text{ if } |w| \in W^s \\ \text{the first element in the enumeration which is an } R^s\text{-minimal} \\ \text{element of } \{\alpha \in W^s \mid R^s|w|\alpha\}, \text{ otherwise.} \end{cases}$$

As W^s is finite, the minimality requirement (which captures the 'as close as possible' idea) is well defined.

As we will show, f is a bounded morphism from (W^w, R^w) into (W^s, R^s). But we have no guarantee that f is a bounded morphism from the *model* $\mathfrak{M}^w$ to $\mathfrak{M}^s$, for while the underlying frame morphism is fine, we need to ensure that the valuations agree on propositional symbols. We fix this as follows. For any propositional symbol p, define $V'(p)$ to be $\{u \in W^w \mid f(u) \in V^s(p)\}$, and let $\mathfrak{M}'$ be (W^w, R^w, V'). That is, $\mathfrak{M}'$ is simply a variant of $\mathfrak{M}^w$ that agrees with $\mathfrak{M}^s$ under the mapping f. But it is not just any variant: as we will now see, it is a *definable* variant. It is time to pull all the threads together and prove the main result.

Theorem 4.96 (Bull's Theorem) *Every normal modal logic extending* $\mathbf{S4.3}$ *has the finite frame property.*

Proof. First we will show that $\mathfrak{M}'$ is a definable variant of $\mathfrak{M}^w$. If β is any of the equivalence classes that make up the filtration $\mathfrak{M}^f$, then $\beta \subseteq W^w$. Moreover, $\mathfrak{M}^w$ can define any such β: the defining formula $\hat{\beta}$ is simply a conjunction of all the formulas in some subset of Φ, the set we filtrated through. (Incidentally, we take the conjunction of the empty set to be $\bot$.) It follows that $\mathfrak{M}^w$ can define $V'(p)$ for any propositional symbol p. To see this, note that $V^s(p)$ is either the empty set or some finite collection of equivalence classes $\{\beta_1, \ldots, \beta_n\}$. In the former case, define δ_p to be $\bot$. In the latter case, define δ_p to be $\bigvee_{i \in n} \hat{\beta}_i$. Either way, δ_p defines $V'(p)$ in $\mathfrak{M}^w$, for $V'(p)$ is $\{u \in W^w \mid f(u) \in V^s(p)\}$. Thus $\mathfrak{M}'$ is a definable variant of $\mathfrak{M}^w$. (Note that this argument makes use of facts about all four models constructed in the course of the proof.)

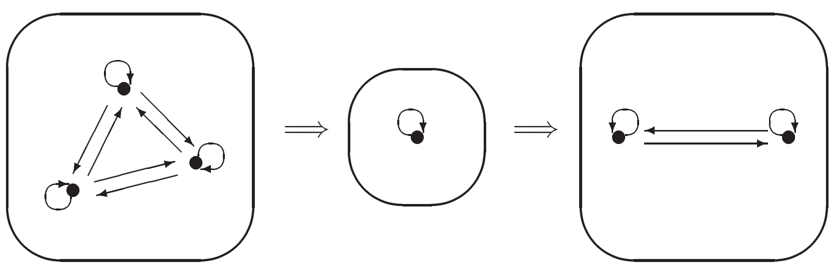
Next we claim that f is indeed a surjective bounded morphism from $\mathfrak{M}^s$ onto $\mathfrak{M}'$; we show here that it satisfies the back condition and leave the rest to the reader. Suppose $R^s f(u) f(v)$. As $f(v) \in W^s$, it is not eliminable, hence not $f(v)$ *sub* $f(u)$. But this means that every element in $f(u)$ R^w-precedes an element in $f(v)$, as required.

But now Bull's Theorem follows. If Λ is a normal modal logic extending **S4.3** and ϕ is a Λ-consistent formula, build $\mathfrak{M}^s$ as described above. By Lemma 4.95, $\mathfrak{M}^s$ satisfies ϕ. Moreover $\mathfrak{M}^s \Vdash \Lambda$. To see this, simply follow the upper left-to-right path through Figure 4.2. $\mathfrak{M}^\Lambda \Vdash \Lambda$, hence so does $\mathfrak{M}^w$, for it is a generated submodel of $\mathfrak{M}^\Lambda$. As $\mathfrak{M}'$ is a definable variant of $\mathfrak{M}^w$, by Lemma 3.25 item (iii), $\mathfrak{M}' \Vdash \Lambda$. Hence, as $\mathfrak{M}^s$ is a bounded morphic image of $\mathfrak{M}'$, it too validates Λ as required. But $\mathfrak{M}^s$ is a finite distinguishing model, hence, by Lemma 3.27 item (iii), its frame validates Λ and we are through. $\dashv$

Finite axiomatizability

We now show that every normal logic extending **S4.3** is *finitely axiomatizable*. (A logic Λ is finitely axiomatizable if there is a *finite* set of formulas Γ such that Λ is the logic generated by Γ.) The proof makes use of a special representation for finite **S4.3** frames.

Because every finite **S4.3** frame is a finite chain of finite clusters, any such frame can be represented as a list of positive integers: each positive integer in the list records the cardinality of the corresponding cluster. For example, the list $[3, 1, 2]$ represents the following frame:



Such representations will allow us to reduce the combinatorial heart of the following proofs to a standard result about lists. The following definition pins down the relationship between lists that will be important.

Definition 4.97 A *list* is a finite non-empty list of positive integers. A list **t** *contains* a list **s** if **t** has a sublist of the same length as **s**, each item of which is greater or equal than the corresponding item of **s**. A list **t** *covers* a list **s** if **t** contains **s** and the last item of **t** is greater than or equal to the last item of **s**. $\dashv$

For example, the list $[9, 40, 1, 9, 3]$ contains the list $[8, 2, 9]$, for it has $[9, 40, 9]$ as a sublist, but it does not cover this list. But $[9, 40, 1, 9, 10]$ covers $[8, 2, 9]$.

The modal relevance of list covering stems from the following lemma:

Lemma 4.98 *Let* $\mathfrak{F}$ *and* $\mathfrak{G}$ *be finite* **S4.3** *frames, and let* **f** *and* **g** *be their associated lists. Then* **f** *covers* **g** *iff there is a bounded morphism from* $\mathfrak{F}$ *onto* $\mathfrak{G}$.

Proof. Exercise 4.9.2. $\dashv$

In view of this result, the following well-known result can be viewed as asserting the existence of infinite sequences of bounded morphisms:

Theorem 4.99 (Kruskal's Theorem) *Every countably infinite sequence of lists* $\mathfrak{t}$ *contains an infinite subsequence* $\mathfrak{s}$ *such that for all lists* s_i *and* s_j *in* **s**, $i > j$ *implies* s_i *covers* s_j.

Proof. Let us call a (finite or infinite) subsequence $(t_i)_{i \in I}$ of a sequence of lists $\mathfrak{t} = (t_i)_{i \in \omega}$ a *chain in* $\mathfrak{t}$ if for all $i, j \in I$, t_j covers t_i whenever $j > i$. We assume familiarity with the notions of the head, the tail and the sum of a list. For instance, the head of $[8, 2, 9]$ is 8, its tail is $[2, 9]$ and its sum is 19. Call $\mathfrak{s}$ *smaller* than $\mathfrak{t}$ if the sum of $\mathfrak{s}$ is smaller than that of $\mathfrak{t}$.

In order to prove the lemma, we will show the following holds:

every countably infinite sequence of lists $\mathfrak{t}$ contains a chain of length 2. (4.2)

Assume that (4.2) does not hold; that is, there are countably infinite sequences without chains of length 2 as subsequences.

Without loss of generality we may assume that $\mathfrak{t}$ does not contain infinitely many lists of length 1. For otherwise, consider its subsequence $(t_i)_{i \in I}$ of these one-item lists. This subsequence may be identified with a sequence of natural numbers $(n_i)_{i \in \omega}$. But

any sequence $(n_i)_{i \in \omega}$ of natural numbers contains a subsequence $(n_i)_{i \in I}$ such that for all $i, j \in I$, $i < j$ implies $n_i \leq n_j$, (4.3)

as can easily be proved. But if $n_i \leq n_j$ then clearly t_j covers t_i. But then we may also assume that $\mathfrak{t}$ does not contain one-item lists at all: simply consider the sequence found by eliminating all one-item lists.

Let $\mathfrak{t}$ be a *minimal* such sequence. That is, $\mathfrak{t}$ is a sequence of more-item lists, $\mathfrak{t}$ has no 2-chains, and for all n, there are no more-item lists $t'_n, t'_{n+1}, \ldots$ such that t'_n is smaller than t_n, while the sequence $t_0, t_1, \ldots, t_{n-1}, t'_n, t'_{n+1}, \ldots$ has no 2-chains.

Now we arrive at the heart of the argument. Define $(n_i)_{i\in\omega}$ and $(u_i)_{i\in\omega}$ as the sequences of the heads and the tails of $\mathbf{t}$; that is, for each i, n_i is the head of t_i and u_i is the tail of t_i. By (4.3), there is a subsequence $(n_i)_{i\in I}$ such that $i < j$ implies $n_i \leq n_j$, whenever $i, j \in I$. Now consider the corresponding subsequence $(u_i)_{i\in I}$ of $\mathbf{u}$. We need the following result:

$$\text{any subsequence } (v_i)_{i\in\omega} \text{ of tails of } t \text{ contains a 2-chain.} \tag{4.4}$$

By the same argument as before, we may assume that $\mathbf{v}$ contains only more-item lists. Let k be the natural number such that v_0 is the tail of t_k, and consider the sequence $t_0, t_1, \ldots, t_{k-1}, v_0, v_1, \ldots$. Since v_0 is the tail of t_k and hence, *smaller* than t_k, it follows by the minimality of t_k that the mentioned sequence contains a 2-chain. But obviously this 2-chain can only occur in the $\mathbf{v}$-part of the sequence. This proves (4.4).

But if $\mathbf{u}$ contains a 2-chain, this means that there are two numbers i and j in I with $i < j$ and u_j covers u_i. Also, by definition of I, $n_i \leq n_j$. But then $t_i = [m_i] * u_i$ is covered by $t_j = [m_j] * u_j$. This proves (4.2).

Finally, it remains to prove the lemma from (4.2). Let $\mathbf{t}$ be an arbitrary countably infinite sequence of lists. By successive applications of (4.2), it follows that $\mathbf{t}$ contains infinitely many chains. We claim that one these chains is infinite. For if we suppose that there are only finite chains, we may consider the sequence $\mathbf{z}$ of last items of right-maximal finite chains in $\mathbf{t}$ (a chain is right-maximal if it can not be extended to the right). There must be infinitely many such right-maximal chains, so $\mathbf{z}$ is an infinite sequence. Hence, by yet another application of (4.2), $\mathbf{z}$ contains a chain of length 2. But then some chain was not right-maximal after all. $\dashv$

We now extract the consequences for logics extending **S4.3**:

Corollary 4.100 *There is no infinite sequence Λ_0, Λ_1, $\ldots$ of normal logics containing* $\mathbf{S4.3}$ *such that for all i, $\Lambda_i \subset \Lambda_{i+1}$.*

Proof. Suppose otherwise. Then for some infinite sequence of logics $\Lambda_0, \Lambda_1, \ldots$ extending **S4.3**, and for all natural numbers i, there is a formula ϕ_i such that $\phi_i \notin \Lambda_i$ and $\phi_i \in \Lambda_{i+1}$. So, by Bull's Theorem, for all natural numbers i there is a finite **S4.3** frame $\mathfrak{F}_i$ that validates Λ_i and does not satisfy ϕ_i. Let $\mathbf{t}$ be the infinite sequence of lists t_i associated with the frames $\mathfrak{F}_i$. By the Kruskal's Theorem, there exist natural numbers k and l, such that $k > l$ and t_k covers t_l. Hence by Lemma 4.98 there is a bounded morphism from $\mathfrak{F}_k$ onto $\mathfrak{F}_l$. It follows that $\mathfrak{F}_l \Vdash \phi_l$ and we have a contradiction. $\dashv$

Theorem 4.101 *Every normal modal logic extending* $\mathbf{S4.3}$ *is finitely axiomatizable.*

Proof. To arrive at a contradiction, we will assume that there does exist an extension Λ of **S4.3** that is not finitely axiomatizable. We will construct a infinite sequence $\Lambda_0 \subset \Lambda_1 \subset \cdots$ of extensions of **S4.3**, thus contradicting Corollary 4.100.

As Λ is not finitely axiomatizable, it must be a proper extension of **S4.3**. Let ϕ_0 be an arbitrary formula in $\Lambda \setminus \mathbf{S4.3}$, and define Λ_0 to be the logic generated by $\mathbf{S4.3} \cup \{\phi_0\}$. Then $\mathbf{S4.3} \subset \Lambda_0 \subset \Lambda$. The latter inclusion is strict because Λ is not finitely axiomatizable. Hence, there exists $\phi_1 \in \Lambda \setminus \Lambda_0$. Let Λ_1 be the logic generated by $\Lambda_0 \cup \{\phi_1\}$. Continuing in this fashion we find the required infinite sequence $\Lambda_0 \subset \Lambda_1 \subset \cdots$ of extensions of **S4.3**. ⊣

A negative characterization

We turn to the final task: showing that every normal logic extending **S4.3** has a negative characterization in terms of finite sets of finite frames. Once again, the proof makes use of the representation of **S4.3** frames as lists of positive integers.

First some terminology. A set of lists X is *flat* if for every two distinct lists in X, neither covers the other. In view of Lemma 4.98, the modal relevance of flatness is this: if two frames are associated with distinct lists belonging to a flat set, then neither frame is a bounded morphic image of the other.

Lemma 4.102 *All flat sets are finite. Furthermore, for any set of lists Y there is a maximal set X such that $X \subseteq Y$ and X is flat.*

Proof. Easy consequences of Kruskal's Theorem. ⊣

If X is a flat set of lists, then $C(X)$ is the set of lists covered by some list in X. Note that $C(X)$ is finite and that $X \subseteq C(X)$. If X is a set of lists, then $B(X)$ is the class of all finite **S4.3** frames $\mathfrak{F}$ such that there is a bounded morphism from $\mathfrak{F}$ onto some frame whose list is in X.

Theorem 4.103 *For every normal modal logic Λ extending* **S4.3** *there is a finite set* N *of finite* **S4.3** *frames with the following property: for any finite frame $\mathfrak{F}$, $\mathfrak{F} \Vdash \Lambda$ iff $\mathfrak{F}$ is an* **S4.3** *frame and there does not exist a bounded morphism from $\mathfrak{F}$ onto any frame in* N.

Proof. Let $\Lambda \supseteq$ **S4.3**, and let L' be the set of lists associated with finite **S4.3** frames which do not validate Λ. Let L be a maximal flat set such that $L \subseteq L'$. Note that $C(L) \subseteq L'$.

We claim that for any finite **S4.3** frame $\mathfrak{F}$, $\mathfrak{F} \Vdash \Lambda$ iff $\mathfrak{F} \notin B(C(L))$. The left to right implication is clear, for as no frame whose list belongs to $C(L)$ validates Λ, there cannot be a bounded morphism from $\mathfrak{F}$ onto any such frame. For the other direction, we show the contrapositive. Suppose that $\mathfrak{F} \not\Vdash \Lambda$. Let $\mathfrak{F}$'s list be $\mathbf{f}$. Then

$\mathbf{f} \in L'$. Now either $\mathbf{f} \in C(L)$ or $\mathbf{f} \in L' \setminus C(L)$. If $\mathbf{f} \in C(L)$, then the identity morphism on $\mathfrak{F}$ guarantees that $\mathfrak{F} \in B(C(L))$ as required. So suppose instead that $\mathbf{f} \in L' \setminus C(L)$. This means that $\mathbf{f} \notin L$ (as $L \subseteq C(L)$), hence as L is a *maximal* flat subset of L', **f** must cover some list **g** in L. Thus by Lemma 4.98, any **S4.3** frame $\mathfrak{G}$ whose list is **g** is a bounded morphic image of $\mathfrak{F}$, hence $\mathfrak{F} \in B(C(L))$ as required. This completes the proof of the claim.

We can now define the desired finite set N: for each $\mathbf{g} \in C(L)$, choose a frame whose list is **g**, and let N be the set of all our choices. ⊣

Exercises for Section 4.9

4.9.1 Show that the *sub* relation is transitive and asymmetric. Furthermore, show that if α *sub* γ and not α *sub* β, then β *sub* γ.

4.9.2 Prove Lemma 4.98. That is, let $\mathfrak{F}$ and $\mathfrak{G}$ be finite **S4.3** frames, and let **f** and **g** be their associated lists. Then show that **f** covers **g** iff there is a bounded morphism from $\mathfrak{F}$ onto $\mathfrak{G}$. (First hint: look at how we defined the bounded morphism used in the proof of Bull's theorem. Second hint: look at the statement (but not the proof!) of Lemma 6.39.)

4.9.3 Give a complete characterization of all the normal logics extending **S5**. Your answer should include axiomatizations for all such logics.

4.9.4 Let $\mathbf{K}_t\mathbf{4.3}$ be the smallest tense logic containing 4, T, $.3_l$ and $.3_r$. Show that there are tense logics extending $\mathbf{K}_t\mathbf{4.3}$ that do not have the finite frame property. (Hint: look at the tense logic obtained by adding the Grzegorczyk axiom in the operator F. Is the Grzegorczyk axiom in P satisfiable in a model for this logic? Is the Grzegorczyk axiom in P satisfiable in a *finite* model for this logic?)

4.10 Summary of Chapter 4

- ▶ *Completeness*: A logic Λ is *weakly* complete with respect to a class of structures S if every formula valid on S is a Λ-theorem. It is *strongly* complete with respect to S if whenever a set of premises entails a conclusion over S, then the conclusion is Λ-deducible from the premises.
- ▶ *Canonical Models and Frames*: Completeness theorems are essentially model existence theorems. The most important model building technique is the canonical model construction. The points of the underlying canonical frames are maximal consistent sets of formulas, and the relations and valuation are defined in terms of membership of formulas in such sets.
- ▶ *Canonicity* Many formulas are canonical for a property P. That is, they are valid on any frame with property P, and moreover, when used as axioms, they guarantee that the canonical frame has property P. When working with such formulas, it is possible to prove strong completeness results relatively straightforwardly.

- ▶ *Sahlqvist's Completeness Theorem*: Sahlqvist formulas not only define first-order properties of frames, each Sahlqvist formula is also canonical for the first-order property it defines. As a consequence, strong completeness is automatic for any logic that is axiomatized by axioms in Sahlqvist form.
- ▶ *Limitative Results*: The canonical model method is not universal: there are weakly complete logics whose axioms are not valid on any canonical frame. Indeed, no method is universal, for there are logics that are not sound and weakly complete with respect to any class of frames at all.
- ▶ *Unraveling and Bulldozing*: Often we need to build models with properties for which no modal formula is canonical. Sometimes this can be done by transforming the logic's canonical model so that it has the relevant properties. Unraveling and bulldozing are two useful transformation methods.
- ▶ *Step-by-step*: Instead of modifying canonical models directly, the step-by-step method builds models by selecting MCSs. Because it builds these selections inductively, it offers a great deal of control over the properties of the resulting model.
- ▶ *Rules for the Undefinable*: By enriching our deductive machinery with special proof rules, it is sometimes possible to construct canonical models that have the desired properties right from the start, thus avoiding the need to massage the (standard) canonical model into some desired shape.
- ▶ *Finitary Methods*: The canonical model method establishes *strong* completeness. Only *weak* completeness results are possible for for non-compact logics such as propositional dynamic logic, and finite canonical models (essentially filtrations of standard canonical models) are a natural tool for proving such results.
- ▶ *Logics extending* **S4.3**: Although the incompleteness results show that a frame based analysis of all normal logics is impossible, many subregions of the lattice of normal modal logics are better behaved. For example, the logics extending **S4.3** all have the finite frame property, are finitely axiomatizable, and have negative characterizations in terms of finite frames.

Notes

Modal completeness results can be proved using a variety of methods. Kripke's original modal proof systems (see [290, 291] were tableaux systems, and completeness proofs for tableaux typically don't make use of MCSs (Fitting [145] is a good introduction to modal tableaux methods). Completeness via normal form arguments have also proved useful. For example, Fine [139] uses normal forms to prove the completeness of the normal logic generated by the McKinsey axiom; this logic is not canonical (see Goldblatt [193]).

Nonetheless, most modal completeness theory revolves, directly or indirectly, around canonical models; pioneering papers include Makinson [314] (who uses a method tantalizingly close to the step-by-step construction to pick out generated subframes of canonical models) and Cresswell [97]. But the full power of canonical models and completeness-via-canonicity arguments did not emerge clearly till the work of Lemon and Scott [303]. Their monograph stated and proved the Canonical Model Theorem and used completeness-via-canonicity arguments to establish many important frame completeness results. One of their theorems was a general canonicity result for axioms of the form $\Diamond^k \Box^j p \to \Box^m \Diamond^n p$, where $k, j, m, n \geq 0$. Although not as general as Sahlqvist's [388] later result (Theorem 4.42), this covered most of the better known modal systems, and was impressive testimony to the generality of the canonical model method.

That **KL** is weakly complete with respect to the class of finite transitive trees is proved in Segerberg [396]. (Strictly speaking, Segerberg proved that **KL4** is complete with respect to the transitive trees, as it wasn't then known that 4 was derivable in **KL**; derivations of 4 were independently found by De Jongh, Kripke, and Sambin: see Boolos [67, page 11] and Hughes and Cresswell [241, page 150].) Segerberg first proves weak completeness with respect to the class of finite strict partial order (the result we asked the reader to prove in Exercise 4.8.7), however he does so by filtrating the canonical model for **KL**, whereas we asked the reader to use a finite canonical model argument. Of course, the two arguments are intimately related, but the finite canonical model argument (which we have taken from taken from Hughes and Cresswell [241, Theorem 8.4] is rather more direct. Segerberg then proves weak completeness with respect to finite trees by unraveling the resulting model (just as we asked the reader to do in Exercise 4.8.8).

The incomplete tense logic $\mathbf{K}_t$**ThoM** discussed in the text was the first known frame incomplete logic, and it's still one of the most elegant and natural examples. It can be found in Thomason [427], and the text follows Thomason's original incompleteness proof. Shortly afterward, both Fine [137] and Thomason [427] exhibited (rather complex) examples of incomplete logics in the the basic modal language. The (much simpler) incomplete logic **KvB** examined in Exercise 4.4.2 is due to van Benthem [38]; **KvB** is further examined in Cresswell [96]. In Exercise 4.4.3 we listed three formulas which jointly define a first-order class of frames, but which when used as axioms give rise to an incomplete normal logic; this example is due to van Benthem [36]. Both the original paper and the discussion in [42] are worth looking at. The logic of the veiled recession frame was first axiomatized by Blok [63]. It was also Blok [64, 65] who showed that incompleteness is the rule rather than the exception among modal logics.

Although filtration and unraveling had been used earlier to prove completeness results, the systematic use of transformation methods stems from the work of Segerberg [396]. Segerberg refined the filtration method, developed the bulldozing technique, and used them (together with other transformation) to prove many important completeness results, including characterizations of the tense logics of $(\mathbb{N}, <)$, $(\mathbb{Z}, <)$, $(\mathbb{Q}, <)$, $(\mathbb{R}, <)$ and their reflexive counterparts.

We do not know who first developed the modal step-by-step method. Certainly the idea of building models inductively is a natural one, and has long been used in both algebraic logic (see [237]) and set-theory (see [410]). One influential source for the method is the work of Burgess: for example, in [76] he uses it to prove completeness results in Since-Until logic (see also Xu [458] for some instructive step-by-step proofs for this language). Moreover, in [77], his survey article on tense logic, Burgess proves a number of completeness results for the basic modal language using the method. A set of lecture notes by De Jongh and Veltman [255] is the source of the popularity among Amsterdam logicians. Recent work on Arrow Logic uses the method (and the related mosaic method) heavily, often combined with the use of rules for the undefinable (see, for example, [326]). Step-by-step arguments are now widely used in a variety of guises.

Gabbay [158] is one of the earliest papers on rules for the undefinable, and one of the most influential (an interesting precursor is Burgess [75], in which these rules are used in the setting of branching time logic). Gabbay and Hodkinson [164] is an important paper which shows that such rules can take a particularly simple form in the basic temporal language. For rules in modal languages equipped with the D-operator, see de Rijke [104] and Venema [439]. For rules in modal languages with nominals, see Passy and Tinchev [362], Gargov and Goranko [171], Blackburn and Tzakova [61], and Blackburn [55].

The axiomatization of **PDL** given in the text is from Segerberg's 1977 abstract, (see [400]). But there was a gap in Segerberg's completeness proof, and by the time he had published a full corrected version (see [402]) very different proofs by Parikh [357] and Kozen and Parikh [279], had appeared. It seems that several other unpublished completeness proofs were also in circulation at this time: see Harel's survey of dynamic logic [215] for details. The proof in the text is based on lecture notes by Van Benthem and Meyer Viol [48].

Bull's Theorem was the first general result about the fine structure of the lattice of normal modal logics. Bull's original proof (in [72]) was algebraic; the model-theoretic proof given in the text is due to Fine [136]. A discussion of the relationship between the two proofs may be found in Bull and Segerberg [73]. Moreover, Goldblatt [183] presents Fine's proof from a rather different perspective, emphasizing a concept he calls 'clusters within clusters'; the reader will find it instructive to compare Goldblatt's presentation with the one in the text, which uses Fine's original argument. Fine's paper also contains the finite axiomatizability result for logics extending **S4.3** (Theorem 4.101) and the (negative) characterization in terms of finite sets of finite frames (Theorem 4.103), and the text follows Fine's original proofs here too.

The work of Bull and Fine initiated a (still flourishing) investigation into subregions of the lattice of normal modal logics. For example, the position of logics in the lattice characterized by a single structure is investigated in Maksimova [317], Esakia and Meskhi [132] and (using algebraic methods) Blok [65]. In [138] and [141], Fine adapts his methods to analyze the logics extending **K4.3** (the adaptation is technically demanding as not all these logics have the finite frame property). Moreover, the Berlin school has a long tradition in this area: see Rautenberg [374, 375, 376], Kracht [283, 285, 286], and Wolter [452]. More recently, the structure of the lattice of tense logics has received attention: see, for example, Kracht [281] and Wolter [450]. And Wolter [451] investigates the transfer of properties when the converse operator P is added to a logic (in the basic modal language) that extends **K4**, obtaining various axiomatizability and decidability results.

Work by Zakharyaschev has brought new ideas to bear. As we pointed out in the Notes to Chapter 3, in the 1960s (the early years following the introduction of relational semantics for modal logic) it was hoped that one could describe and understand *any* modal formula by imposing first-order conditions on its frames. But the incompleteness results, and the discovery of modal formulas that do not correspond to any first-order conditions, destroyed this hope. In a series of papers Zakharyaschev [462, 463, 464, 465] has studied an alternative, purely frame-theoretic approach to the classification of modal formulas. Given a modal (or intuitionistic) formula ϕ, one can effectively construct finite rooted frames $\mathfrak{F}_1, \ldots, \mathfrak{F}_n$ such that a general frame $\mathfrak{g}$ refutes ϕ iff there is a (not necessarily generated) subframe $\mathfrak{g}'$ of $\mathfrak{g}$ which satisfies certain natural conditions and which can be mapped to one of the $\mathfrak{F}_i$ by a bounded morphism. Conversely, with every finite rooted frame $\mathfrak{F}$ Zakharyaschev associates a *canonical* formula which can be refuted on a frame iff that frame contains a subframe (satisfying certain natural conditions) that can be mapped to $\mathfrak{F}$ by a bounded morphism. Like the search for first-order characterizations, the classification approach in terms of canonical formulas is not universal either. But its limitations are of a different kind: it only characterizes *transitive* general frames — but for every modal (and intuitionistic) formula. Zakharyaschev [459] is a very accessible survey of canonical formulas, with plenty of motivations, examples and definitions; technical details and discussions of the algebraic and logical background of canonical formulas are provided by Chagrov and Zakharyaschev [86, Chapter 9].