Intensional logics

de Jongh, D.H.J.; Veltman, F.J.M.M.

Publication date
1999

INTENSIONAL LOGICS

DICK DE JONGH & FRANK VELTMAN

1999

Contents

Chapter 1

Introduction to modal logic

This first chapter contains an introduction to modal logic. In section 1.1 the syntactic side of the matter is discussed, and in section 1.2 the subject is approached from a semantic point of view.

1.1 The syntactic approach to modal logic

Modal logic can be described briefly as the logic of necessity and possibility. The language of propositional modal logic differs from the language of standard propositional logic in that in addition to the usual truth functional connectives $\wedge$, $\vee$, $\rightarrow$, $\leftrightarrow$, $\neg$, $\perp$, and $\top$, it incorporates a unary sentence operator $\Box$ which is intended not to be truth-functional. For future convenience we right away define its dual $\Diamond$ as $\neg\Box\neg$.

For the time being we are interested in modal logic proper, in which $\Box\varphi$ is to be read as 'It is necessary that φ', and $\Diamond\varphi$ — consequently — as 'It is possible that φ'. In due course we will discuss other interpretations of $\Box$.

Modal logics are meant to be extensions of the classical propositional calculus. This means that every modal logic will at least validate all tautologies of the propositional calculus. In developing a proof system for modal logic, we can make use of the fact that being a theorem of the propositional calculus is a decidable property (that is: an effective procedure exists by which we can decide for arbitrary formulas whether they are theorems or not). This fortunate circumstance enables us to take each tautology separately as an axiom instead of having to give an axiomatization of the propositional calculus. In doing so we save ourselves time and trouble in making derivations, because we introduce any tautology directly in our derivations rather than having to derive it first from some axiomatic base. (Since there is no decision procedure for universal validity in first order predicate logic, this strategy cannot be followed in .)

Axioms for the modal logic **K**. For the reasons explained above all formulas with the form of a tautology of the propositional calculus are considered as axioms. These include formulas in which $\Box$ appears, as for example $(\Box p \wedge \Box q) \to \Box p$. In derivations we justify the use of a tautology under the name **Axiom 1**.

In addition we need axioms and rules for manipulating $\Box$. We add as **Axiom 2** all formulas of the form

$$\Box(\varphi \to \psi) \to (\Box\varphi \to \Box\psi)$$

As derivation rules we take **Modus Ponens**:

$$\frac{\varphi, \varphi \to \psi}{\psi}$$

and **Necessitation**:

$$\frac{\varphi}{\Box\varphi}$$

The resulting system is known as *the minimal normal modal logic* **K**. (In what way this logic is 'minimal' and 'normal' is best explained semantically).

What it means for a formula φ to be *derivable in* **K** (φ is *a theorem of* **K**) can now be defined in the standard fashion.

Definition 1 $\vdash_{\mathbf{K}} \varphi$ if and only if there is a sequence of formulas of which φ is the last such that every formula in the sequence is either an axiom of **K** or else is derived by means of one of the rules from formulas appearing higher up in the sequence.

The definition of $\Delta \vdash_{\mathbf{K}} \varphi$ (φ is derivable from the set Δ of premises) presents some difficulty. If we were to proceed as in the case of $\vdash_{\mathbf{K}} \varphi$, except that now formulas from Δ may be inserted at every stage of the derivation, we would get the following implausible consequence: the necessitation rule will make $\varphi \vdash \Box\varphi$ valid for every φ, but this means as much as 'Everything that is the case is necessarily so', something we would be loath to accept as logically valid.

We can avoid this by forbidding the necessitation rule to be applied after one of the premises has been used; this restriction only affects those parts of a proof in which a premise is 'really' made use of, for theorems can always be derived first.

In fact, our first attempt to define $\Delta \vdash \varphi$ was not guided by the principle 'thinking comes before acting'. This can be seen if we stop for a moment to consider what we are engaged in: we are trying to explicate the notion of necessity.

The reason we added the necessitation rule to our axiom system in the first place was that we are bound to accept as necessary every theorem of the logic in question. But premises can (and usually will) bring into play some contingent extra-logical facts about the domain of discussion and of course we do not want these to come out necessary, too.

In the definitions given above the '**K**' in $\vdash_{\mathbf{K}}$ is there to distinguish the notion of derivability in **K** from the analogous notion in other modal systems; if, however, it is clear from the context which system we mean, we often omit the subscript.

The syntactic part of logic consists in the first instance of making derivations. Handy expedients for keeping proofs as short as possible are the use of so called 'derived rules' and the deduction theorem, which we will presently discuss in that order:

Derived rules

Consider the following derivation of $\Box(\varphi \wedge \psi) \rightarrow \Box\varphi$:

1. $(\varphi \wedge \psi) \rightarrow \varphi$	Axiom (1)
2. $\Box((\varphi \wedge \psi) \rightarrow \varphi)$	Necessitation on 1
3. $\Box((\varphi \wedge \psi) \rightarrow \varphi) \rightarrow (\Box(\varphi \wedge \psi) \rightarrow \Box\varphi)$	Axiom (2)
4. $\Box(\varphi \wedge \psi) \rightarrow \Box\varphi$	Modus Ponens on 2, 3

It is rather obvious from this derivation that whenever $\vdash \varphi \rightarrow \psi$, also $\vdash \Box\varphi \rightarrow \Box\psi$ (cf. exercise 2). Another way to put this is to say that $\varphi \rightarrow \psi/\Box\varphi \rightarrow \Box\psi$ is a *derived rule* of **K**.

The usual rules of the propositional calculus are of course also derived rules for **K**, because **K** contains modus ponens as well as the tautologies. Some examples like the fact that

$$\frac{\varphi \rightarrow \psi, \varphi \rightarrow \chi}{\varphi \rightarrow (\psi \wedge \chi)}$$

and

$$\frac{\varphi \rightarrow \psi, \psi \rightarrow \chi}{\varphi \rightarrow \chi}$$

are derived rules of **K** are easily proved (cf. exercise 1).

As an application of this we prove:

Proposition 1 $\vdash \Box(\varphi \wedge \psi) \leftrightarrow (\Box\varphi \wedge \Box\psi)$.

Proof: $\Rightarrow$: From $\vdash \Box(\varphi \wedge \psi) \rightarrow \Box\varphi$ and $\vdash \Box(\varphi \wedge \psi) \rightarrow \Box\psi$ (which is derivable in a similar fashion) it follows, by the first derived rule from exercise 1, that $\vdash \Box(\varphi \wedge \psi) \rightarrow (\Box\varphi \wedge \Box\psi)$.
$\Leftarrow$: Given that $\vdash \varphi \rightarrow (\psi \rightarrow (\varphi \wedge \psi))$, it follows that $\vdash \Box\varphi \rightarrow \Box(\psi \rightarrow (\varphi \wedge \psi))$ (exercise 2). Now, $\Box(\psi \rightarrow (\varphi \wedge \psi)) \rightarrow (\Box\psi \rightarrow \Box(\varphi \wedge \psi))$ is an axiom (Axiom 2). So we have $\vdash \Box\varphi \rightarrow (\Box\psi \rightarrow \Box(\varphi \wedge \psi))$ by the second derived rule from exercise 1. Applying once again a derived rule (which?) we finally obtain $\vdash (\Box\varphi \wedge \Box\psi) \rightarrow \Box(\varphi \wedge \psi)$. •

Note that this proof would be much longer without the use of derived rules: each time a rule is used one would in its stead have to insert its entire derivation. In exercise 4, a syntactic variant of the substitution theorem for **K**, large numbers (twice infinitely many) of derived rules for **K** are derived at once.

Before closing our discussion of the subject of derived rules a warning is due: Derived rules φ/ψ that **K** does not share with the propositional calculus are weaker than the corresponding implication $\varphi \to \psi$. The latter can be used to conclude ψ whenever one has φ, while the first has a stronger 'antecedent': one is only allowed to conclude that ψ is a theorem if φ is a theorem (so having φ as a premise does not help us to get ψ in this case).

Why this is so, can readily be seen from our first example where we needed necessitation to get our conclusion: we are not allowed to use that rule on premises. However, since there is only a restriction on the use of the necessitation rule, this warning does not extend to the rules **K** 'inherits' from the propositional calculus.

Theorem 2 (Deduction Theorem for K)
If $\varphi_1, \ldots, \varphi_n, \psi \vdash_{\mathbf{K}} \chi$, then $\varphi_1, \ldots, \varphi_n \vdash_{\mathbf{K}} \psi \to \chi$.

Proof: Suppose $\varphi_1, \ldots, \varphi_n, \psi \vdash_{\mathbf{K}} \chi$. Then χ is derivable from $\varphi_1, \ldots, \varphi_n, \psi$ and a number of **K**-theorems $\theta_1, \ldots, \theta_m$, using only modus ponens. This means that $\theta_1, \ldots, \theta_m, \varphi_1, \ldots, \varphi_n, \psi \vdash \chi$ in the propositional calculus. (This is so, because the only rule **K** has in addition to those of the propositional calculus may not be used on premises.) By the deduction theorem we already have for that system, we get

$$\vdash \theta_1 \to (\theta_2 \to (\ldots(\theta_m \to (\varphi_1 \to (\varphi_n \to (\psi \to \chi)))) \ldots)),$$

hence by the soundness theorem for the propositional calculus this formula is a tautology and, hence an Axiom 1 of **K**. Since $\vdash_{\mathbf{K}} \theta_1, \ldots, \vdash_{\mathbf{K}} \theta_m$ (all of the θ's are theorems of **K**)) we have

$$\vdash_{\mathbf{K}} \varphi_1 \to (\ldots(\varphi_n \to (\psi \to \chi)) \ldots)$$

by applying modus ponens m times. Using modus ponens again, n times this time, we finally get $\varphi_1, \ldots, \varphi_n \vdash_{\mathbf{K}} \psi \to \chi$. •

The shortening of proofs by means of derived rules and the shortening by means of the deduction theorem work on a like principle: a regularity has been found in the way we get from theorems to other theorems, and we can prove the procedure to be correct for all cases at once. From then on we no longer have to repeat the procedure every time we need it.

Some history The somewhat dry, syntactic manner in which we started this exposition reflects a fact about the actual history of modern modal logic: in the beginning the whole explication of the logic of 'the necessary and the possible' consisted in the construction of axiom systems. This beginning can be dated to 1918, the year C.I. Lewis published his *Survey of symbolic logic*.

After a long tradition, starting with the 'father' of logic, Aristotle, himself, modal logic had been expelled from logical consideration by Frege. ('By saying that a proposition is necessary I give a hint about the grounds for my judgment. But, since this does not affect the conceptual content of the judgment, the form of the apodictic judgment has no significance for us.' Quoted from *Begriffschrift* in a translation by Stefan Bauer-Mengelberg.) The immediate cause of its revival was contained in the *Principia Mathematica*, in which Russell and Whitehead made a (now judged to be unsuccessful) attempt to reduce all mathematics to logic. What is of interest for the subject under consideration is that in this work Russell and Whitehead chose the material implication $\to$ to formalize the notion of 'entailment'. Lewis was among the people dissatisfied with this formal translation, because he thought 'entailment' to be a much stronger relation between propositions than is expressed by the arrow: 'φ entails ψ' not merely means that it is not the case that φ and not-ψ, but rather that it is impossible for this to be the case. He therefore introduced a new symbol, the so called 'fish hook' $\hookrightarrow$, which was supposed to fulfil in a much better way the task to which Russell and Whitehead had put material implication. He tried to pin down the properties of this stronger relation of *strict implication* in (a number of) axiom systems regulating the behaviour of the new symbol. This disagreement over the correct formalization of the notion of entailment may have been caused by confusion about the question at stake: perhaps the PM-translation should not be taken as an attempt to capture the overall intuitive meaning of 'φ entails ψ', but rather be looked upon as a technical device that satisfies in a minimal fashion the needs within the limited context of the PM project.

In any case, modern modal logic was born, for $\varphi \hookrightarrow \psi$ can equivalently be rendered $\Box(\varphi \to \psi)$, or $\neg\Diamond(\varphi \wedge \neg\psi)$. Soon, the study of the notions of 'possible' and 'necessary' overshadowed the original occupation with strict implication and thus modalities were brought back to the attention of logicians.

Unfortunately, during a number of decennia others followed Lewis on the road he had taken: the only products delivered by modal logic were large numbers of axiom systems, all claiming to represent *the* logic of modalities. The better known among these are:

$$\mathbf{T} = \mathbf{K} + \Box\varphi \to \varphi;$$
$$\mathbf{S4} = \mathbf{T} + \Box\varphi \to \Box\Box\varphi; \text{ and}$$
$$\mathbf{S5} = \mathbf{S4} + \Diamond\Box\varphi \to \varphi$$

In the end this course of things resulted in a veritable 'chaos', which could not be resolved within a purely syntactic framework for the following reason: Axiom systems are supposed to capture intuitions; in this case intuitions about possibility and necessity. But when the intuitions of different people do not coincide there is not much left to go on, for axioms are only the outcome of a process of reasoning, based on considerations that fall outside the scope of a syntactic theory.

This objection carries weight against syntactic methods in general, but seems to be most applicable in the case of modal logic. For among the modal systems, in contrast with, for example, the case of pure propositional calculus, there is no more or less 'generally accepted' standard system. The following considerations may explain why the choice of an axiom system is such an arduous task in the modal case:

Difficulties do not arise over relatively simple formulas: for example it must be clear to anyone who has spent some thought about the nature of necessity that $\Box\varphi \to \varphi$ can safely be counted as a truism. (This axiom is not part of **K** for reasons we will discuss further down.) But the various systems proposed differ mainly in axioms containing a stacking of modal operators (as for example the characteristic axioms of **S4** and **S5**) and it makes no sense to talk of 'intuitive plausibility' when it comes to phrases as 'It is possibly necessary that ...'(or worse), since here our intuitions seem to give out completely.

There are two ways out, both equally unsatisfactory: either one accepts some unintuitive axioms, the reason for this being that they serve the purpose of reducing the number of different (i.e. non equivalent) stackings of modal operators (but what else than intuitions can one use to decide in this matter, in the absence of any explicitation by an underlying semantics), or else one has to accept the idea that every such stacking can be used to state some fact not expressible by any other (so twenty-six times $\Box$ followed by φ means something different from twenty-seven times $\Box$ followed by φ).

Of course, in the case of modal logic proper one might even doubt the advisability of allowing stacking of modal operators at all. After all necessity is originally some kind of 'metanotion'. We will meet however many useful applications of stackings of modal operators outside the domain of modal logic proper.

The above problems could only be solved after the perspective had shifted from pure syntax to semantics; we will therefore continue these historical remarks in the next section on semantics.

We conclude this section on syntax by defining in an abstract manner what a modal logic is from a syntactical point of view:

Definition 2 A *modal logic* **S** is a set of formulas that (i) contains **K**, (ii) is closed under modus ponens, (iii) is closed under necessitation, (iv) is closed under substitution, i.e.: whenever $\varphi(p_1, \ldots, p_n) \in \mathbf{S}$ also $\varphi(\psi_1 \ldots, \psi_n) \in \mathbf{S}$, this for arbitrary $\psi_1, \ldots, \psi_n$.

Every such **S** can be obtained from **K** by adding axiom schemes. In general we drop the demand that the axiom system should be decidable, although this will be so in every special case we will treat. We also write $\varphi \in \mathbf{S}$ as $\vdash_{\mathbf{S}} \varphi$, or $\vdash \varphi$ for short, if it is clear which system we mean. Moreover, we define $\Delta \vdash_{\mathbf{S}} \psi$ as for **K**; we keep the restriction on the use of the necessitation rule for every **S**. Δ is called **S**-*consistent* if $\Delta \nvdash_{\mathbf{S}} \perp$.

That the deduction theorem is valid for every **S** can easily be seen: nowhere in the proof of this theorem for **K** did we use the fact that $\theta_1, \ldots, \theta_m$ are **K**-theorems specifically.

1.2 Semantics of modal logic

Intensional logic differs from standard logic in the semantical sphere by the use of so called 'Kripke-semantics'. This form of semantics can be characterized by two key notions: 'possible world' and 'accessibility relation'. A Kripke-model is built up from a number of possible worlds, representing as many alternative, possible states of affairs.

In such a model, truth values are no longer connected to a formula per se, but always to a formula *in a possible world*.

In determining the truth value of a purely propositional formula in a world other possible worlds play no part; only if $\Box$ occurs in a formula can it be necessary to involve other possible worlds in determining whether the formula in question is true (in the first world) or not. Even then one does not always have to look at all possible worlds in the model; which worlds may be relevant is determined by the accessibility relation. If a world is connected to a world by this relation we call the second accessible from the first, or, more graphically, visible from the first.

In the definition of a model for modal logic this idea is made precise in the following way:

Definition 3 A *frame* $\mathcal{F}$ is a tuple $\langle \mathcal{W}, \mathcal{R} \rangle$, of which
(i) W is a non-empty set (of possible worlds);
(ii) $\mathcal{R} \subseteq \mathcal{W} \times \mathcal{W}$.

Definition 4 A *model* (on $\mathcal{F}$) is a tuple $\langle \mathcal{F}, \mathcal{V} \rangle$ of which
(i) $\mathcal{F}$ is a frame $\langle \mathcal{W}, \mathcal{R} \rangle$;
(ii) $\mathcal{V}$ is a *valuation function* that assigns one of the truth values 0 or 1 to each propositional letter in every world of $\mathcal{W}$. The *valuation* $\mathcal{V}_{(\mathcal{M},w)}(\varphi)$ of a formula φ in world w given the model $\mathcal{M}$, is obtained by adding to the standard definition for the connectives of the propositional calculus the following clause:
$\mathcal{V}_{(\mathcal{M},w)}(\Box\psi) = 1$ iff for all w' such that $w\mathcal{R}w'$, $\mathcal{V}_{(\mathcal{M},w')}(\psi) = 1$.

Check that this entails that
$\mathcal{V}_{(\mathcal{M},w)}(\Diamond\psi) = 1$ iff there is a w' such that $\mathcal{V}_{(\mathcal{M},w')}(\psi) = 1$).

Instead of $\mathcal{V}_{(\mathcal{M},w)}(\varphi) = 1$ we write $w \models_{\mathcal{M}} \varphi$, or $w \models \varphi$ for short (pronounce: 'w *forces* φ' or 'φ *is true in* w'), and for $\mathcal{V}_{(\mathcal{M},w)}(\varphi) = 0$ we write $w \not\models_{\mathcal{M}} \varphi$ Therefore we also note $\mathcal{M}$ as $\langle \mathcal{F}, \models \rangle$.

Definition 5

(i) The model $\mathcal{M} = \langle\langle \mathcal{W}, \mathcal{R} \rangle, \models\rangle$ *verifies* φ (written as $\mathcal{M} \models \varphi$) is defined as: $\mathcal{M} \models \varphi \Leftrightarrow$ for all $w \in \mathcal{W}$, $w \models \varphi$.

(ii) φ *is valid on* the frame $\mathcal{F}$ (written as $\mathcal{F} \models \varphi$) is defined as: $\mathcal{F} \models \varphi \Leftrightarrow$ for all $\mathcal{M}$ on $\mathcal{F}$, $\mathcal{M} \models \varphi$.

(iii) φ *is valid within the class of frames* $\mathcal{C}$ (written as $\models_{\mathcal{C}} \varphi$) is defined as: $\models_{\mathcal{C}} \varphi \Leftrightarrow$ for all $\mathcal{F} \in \mathcal{C}$, $\mathcal{F} \models \varphi$.

(iv) φ *is valid* we write as $\models \varphi$, and we define $\models \varphi \Leftrightarrow$ for all $\mathcal{F}$, $\mathcal{F} \models \varphi$.

Given these definitions for the semantic notions we seem to have three options to characterize the notion $\Delta \models \varphi$ ('φ *follows semantically from the premise set* Δ'):

(i) $\Delta \models \varphi \Leftrightarrow$ for all $\mathcal{M}$ and $w \in \mathcal{W}$: if $w \models \psi$ for all $\psi \in \Delta$, then $w \models \varphi$;

(ii) $\Delta \models^* \varphi \Leftrightarrow$ for all $\mathcal{M}$: if $\mathcal{M} \models \psi$ for all $\psi \in \Delta$, then $\mathcal{M} \models \varphi$;

(iii) $\Delta \models^x \varphi \Leftrightarrow$ for all $\mathcal{F}$: if $\mathcal{F} \models \psi$ for all $\psi \in \Delta$, then $\mathcal{F} \models \varphi$.

These alternatives give rise to non-equivalent results: with (ii) we get $p \models^* \Box p$ (cf. exercise 8) which seems rather unattractive. To see what is wrong with (iii) we first have to prove (cf. exercise 8) that for all propositional letters $\Box p \to p$ is valid on a frame $\mathcal{F} = \langle \mathcal{W}, \mathcal{R} \rangle$ iff $\{\langle w, w \rangle | w \in \mathcal{W}\} \subseteq \mathcal{R}$. (Such frames we call *reflexive*; about this manner of characterizing properties of the accessibility relation R we come to speak in the next chapter.) It is then obvious that $\Box p \to p$ and $\Box q \to q$ are valid on exactly the same class of frames, so we have $\Box p \to p \models^x \Box q \to q$. On the other hand the following model is a counterexample against $\Box p \to p \models^* \Box q \to q$ (cf.exercize 7).

INSERT FIGURE 1

Alternative (i) is the weakest; it does not give rise to such unwelcome inferences as do the other two. We therefore adopt it as our definition from now on. Our choice is of course not a matter of technical preference only, but can also be motived by a more philosophical argument: What we want our definition to capture is the notion of 'validity' as current in logic: whenever the premises (stating this and that to be the case) are true, the conclusion has to be true too. But we are interested in this relation between premises and conclusion with respect to one world at the time only, namely the world in which we are, and not with respect to all situations we hold to be possible. So the technical advantages run parallel to the philosophically 'right' view on validity.

Some remarks on the usefullness of Kripke semantics In the section on syntax we already mentioned that the attempt to capture the logical behaviour of modalities by constructing axiom systems, guided by extra-logical intuitions on the subject only, resulted in an outburst of systems, all claiming to be the right one.

Only after the use of semantics became a current practice in modal logic, it was possible to explicate the uniform 'conceptual structure' lying behind all these different logics and to discuss their relative merits on a more solid ground. To be more definite: In the next chapter we will see that modal formulas often correspond to certain properties of the accessibility relation R: Thus $\Box\varphi \rightarrow \varphi$ corresponds to reflexiveness, $\Box\varphi \rightarrow \Box\Box\varphi$ to transitivity and $\Diamond\Box\varphi \rightarrow \varphi$ to symmetry of $\mathcal{R}$.

Instead of wondering which of the reduction axioms (all equally 'opaque' in a sense) is the right one we can now ask questions that seem to be more susceptible to reasonable argumentation, like: should the accessibility relation (for modal logics interpreted as 'is a possible alternative to') be transitive or not; that is, is it reasonable to demand that whenever w_1 has w_2 as a possible alternative, while w_3 in its turn is a possible alternative to w_2, w_3 is a possible alternative for w_1 also? If the answer is 'yes', one should add $\Box\varphi \rightarrow \Box\Box\varphi$ as an axiom and otherwise one should not.

The same goes for symmetry: is it always the case that a situation is a possible alternative to its alternatives?

The answers to these and similar questions often depend on the kind of necessity and possibility at stake: thinking of logical possibility one may be inclined to decide this question differently than when one has, say, physical possibility in mind.

There is something else that became more clear against the semantic background: even while doing syntax only, some logicians had already spotted similarities between the pair of notions 'possible' and 'necessary' and other such couples (like for example 'ought' and 'may') which are mutually interrelated in a manner similar to 'possible' and 'necessary'. All of these notions turn out to be interpretable in Kripke semantics, simply by choosing different interpretations for the worlds and the accessibility relation. Once Kripke semantics had been disjoined from the particular mode of interpretation bound to modal logic, the road was clear for 'experimenting' with other interpretations too.

In this way quite a few logics were developed, each formalizing a different set of related concepts, originating from different branches of philosophy (we mention 'deontic logic', 'tense logic', 'epistemic logic' and of course modal logic itself), computer science (which gave rise to 'dynamic logic') and foundations of mathematics (in which connection provability logic can be mentioned). Some of these logics will be discussed more extensively in chapters to come.

In different cases there seem to have been different kinds of motivations prompting the conception of these new logics: Some arose in connection with a new 'perspective' on the notion of necessity itself. This, for example, is the case in epistemic logic, which is a formalization of reasoning about knowledge. Modal expressions in a natural language often have to be evaluated with respect to a state of knowledge, rather than with respect to what is necessary in an absolute sense. (The sentence 'Mary has to be in the building somewhere, for I saw her walking in five minutes ago' may serve as an example.) Possible worlds can then be taken to represent the alternative states the world might still be in, given certain facts known to be the case; and $\mathcal{R}$ holds between two situations if the second is compatible with the knowledge contained in the first: $\Box\varphi$ then means 'φ is known to be the case'.

Also, one should hope that logical necessity is closely related to provability, or provability in a particular system to be more precise. Provability logic then is concerned with $\Box\varphi$ interpreted e.g. as 'φ is provable in Peano Arithmetic'. What the models can be taken to represent in this case is not so clear; perhaps somebody will try to puzzle this out after this logic has been treated.

Other logics seem to have originated from taking a different view on the worlds and/or the accessibility relation $\mathcal{R}$. Thus in deontic logic, the logic that does for obligations (cf. the couple of notions we already mentioned) what modal logic does for modalities, the worlds can be looked upon as representing ideal situations in stead of just possible ones and $\mathcal{R}$ consequently as fixing which situations are ideal in respect to which; in this context $\Box\varphi$ means 'φ is true in every world in which everything is morally proper'.

Again, tense logic takes possible worlds as (situations at) different moments and $\mathcal{R}$ as the relation of 'is earlier than' between these; $\Box\varphi$ then means 'It will always be the case that φ.'

Finally we should mention dynamic logic, in which the worlds represent information states (think of a machine for example) and $\mathcal{R}$ fixes which states can be reached from which by means of a "program" (so that $\mathcal{R}$ is now dependent on the particular program). Given the program π, $\boxed{\pi}\varphi$ can then be rendered as something like 'φ is the case, whenever π is successfully applied'.

Keeping this in mind we can now explain why we did not include for example $\Box\varphi \to \varphi$ (something anyone who has ever spent a thought on the nature of necessity would take to be a truism) in **K** , the first modal logic we introduced: **K** turns out to be not so much a specifically modal logic, but rather 'the logic of the underlying framework' giving everything that has to be the case, given the kind of models we have chosen. And, since we can use the framework to represent other kinds of concepts as well as the modal ones, it is clear $\Box\varphi \to \varphi$ should not be part of this logic; for we would rather not accept for example 'Everything that is true in an ideal situation, is true' as a logical truth. (As for the other logics mentioned above, you may check for yourself whether $\Box\varphi \to \varphi$ seems a reasonable axiom or not.)

1.3 Linking up syntax and semantics

First we have fixed a syntax for a modal logic and next we have specified a semantics for it.

However, as yet we cannot be sure that this semantics is the right one, in the sense that we have the desired connections between the notion of 'syntactic derivability' on the one hand and the notion of 'semantic inference' on the other hand. This turns out to be the case: in chapter 3 we prove the completeness theorems for **K** (and some other modal logics) and here we already prove the soundness theorems:

Theorem 3 (The weak soundness theorem for K)
If $\vdash_{\mathbf{K}} \varphi$, then $\models \varphi$.

Proof: The proof is simple, by induction on the length of the derivation of φ. As an example, we spell out the case of the necessitation rule:
Suppose $\mathcal{M} \not\models \Box\varphi$ for a certain formula φ. Then there must be a $w \in \mathcal{W}$ such that $w \not\models \Box\varphi$. This can only be the case if $w' \not\models \varphi$ for some $w' \in \mathcal{W}$ such that $w\mathcal{R}w'$. So, φ cannot be forced in all worlds in all models: we have $\not\models \varphi$. •

From this the soundness theorem for **K** can be deduced:

Theorem 4 (The soundness theorem for K)
If $\varphi_1, \ldots, \varphi_n \vdash_{\mathbf{K}} \psi$ then $\varphi_1, \ldots, \varphi_n \models \psi$.

Proof: Suppose $\varphi_1, \ldots, \varphi_n \vdash \psi$. Then we have by the deduction theorem:

$$\vdash \varphi_1 \to (\varphi_2 \to (\ldots(\varphi_n \to \psi)\ldots)$$

So $\models \varphi_1 \to (\varphi_2 \to ((\ldots(\varphi_n \to \psi)\ldots))$ holds by weak soundness. This means that for all $\mathcal{M}$ and $w \in \mathcal{W}$,

$$w \models \varphi_1 \to (\varphi_2 \to (\ldots(\varphi_n \to \psi)\ldots)$$

Hence, if $w \models \varphi_1, \ldots, w \models \varphi_n$ then $w \models \psi$, which means that $\varphi_1, \ldots, \varphi_n \models \psi$. •

Note that this theorem is not really stronger then the first, for to get it from the weak variant we only need the deduction theorem and we have shown already this theorem holds for all modal logics **S**. In fact we even have:

Theorem 5 (The strong soundness theorem for K)
If $\Delta \vdash_{\mathbf{K}} \varphi$, then $\Delta \models \varphi$.

Proof: To prove this we notice that derivations are always finite; so, the derivation of φ can only involve finitely many premises from Δ. •

The soundness theorems immediately allow us to use models to prove that certain inferences are not valid. If we construct a model $\mathcal{M}$ and a world w in $\mathcal{M}$ such that $w \models \psi$ for all $\psi \in \Delta$, but $w \not\models \varphi$, then $\Delta \nvdash \varphi$ in **K** (cf. exercise 6).

1.4 Exercises

Exercise 1 Prove that

$$\frac{\varphi \to \psi, \varphi \to \chi}{\varphi \to (\psi \wedge \chi)}$$

and

$$\frac{\varphi \to \psi, \psi \to \chi}{\varphi \to \chi}$$

are derived rules of **K**.

Exercise 2 Prove that replacing Axiom 2 by the axiom $\Box\top$ and the schema $(\Box\varphi \wedge \Box\psi) \to \Box(\varphi \wedge \psi)$, and the necessitation rule by the rule

$$\frac{\varphi \to \psi}{\Box\varphi \to \Box\psi}$$

yields a system equivalent to **K**

Exercise 3 Show that, if $\varphi_1, \ldots, \varphi_n \vdash \psi$, then $\Box\varphi_1, \ldots, \Box\varphi_n \vdash \Box\psi$.

Exercise 4 (A syntactic variant of the substitution theorem for **K**) Prove by induction on the complexity of χ that $\varphi \leftrightarrow \psi / \chi(\varphi) \leftrightarrow \chi(\psi)$ is a derived rule for arbitrary φ, ψ and χ.

Exercise 5 Be M a sequence of $\Box$ and $\neg$. Prove by induction on the number of occurrences of $\neg$ in M that $\varphi \to \psi / M\varphi \to M\psi$ is a derived rule of **K** if $\neg$ occurs an even number of times in M. (Hint: Formulate some hypothesis also about what happens in the case of an odd number of occurrences of $\neg$ in M, and use this in your inductive proof.)

Exercise 6 Show that in **K**,
(i) $p \nvdash \Box p$
(ii) $\Box p \nvdash p$
(iii) $\Box(p \vee q) \nvdash \Box p \vee \Box q$

Exercise 7

(i) Check that $p \models^* \Box p$.

(ii) Prove that, for all propositional letters p, $\Box p \to p$ is valid on a frame $\mathcal{F} = \langle \mathcal{W}, \mathcal{R} \rangle$ iff $\{\langle w, w \rangle | w \in \mathcal{W}\} \subseteq \mathcal{R}$, i.e. iff R is reflexive.

(iii) Verify that the model in figure 1 is a counterexample to $\Box p \to p \models^* \Box q \to q$.

(iv) Prove that $\Delta \models \varphi \Rightarrow \Delta \models^* \varphi$, and $\Delta \models^* \varphi \Rightarrow \Delta \models^x \varphi$.

Chapter 2

Properties of frames

2.1 Some introductory remarks on characterization

In the previous chapter we fixed a language and we introduced a type of semantic structure to serve as an interpretation for this language. A first question that comes to mind is: how accurately can these structures be described in the given language? Or, in other words: what information about a structure does a formula give us by being true in that structure? Since in the case of modal logic four kinds of truth have been distinguished (i.e. truth in a world, truth in a model, truth in a frame and truth in a whole class of frames; compare the truth definition we gave in the last part of 1.2.1), we could ask this question fourfold. It is immediately obvious that the first question will not be an interesting one, since fixing our gaze on one particular world at the time can bring us nothing but pure propositional calculus. Here, we intend to restrict our attention to the expressibility of those aspects of a model that are not dependent on the properties of a particular valuation, but solely depend on the structure of the underlying frame. What we need to do first therefore, is to determine which of the kinds of truth mentioned above is the appropriate one in connection with this purpose of describing structural features; clearly it will have to be one of the two definitions regarding validity on frames. Now what kind of structural aspect is exhibited by a frame? One might first think for example about the size of $\mathcal{W}$, but not much can be said about that without recourse what turns out to be the main structural aspect: the way in which the worlds in $\mathcal{W}$ are connected by $\mathcal{R}$. In fact anything modal formulas can tell us about frames goes by way of this structural aspect. We mentioned already in discussing semantics in Chapter 1 that modal formulas sometimes "characterize" (in a sense to be made precise further down) definite properties of $\mathcal{R}$, and we will see more of this in the sequel. On the other hand we will see that there are a number of simple properties which cannot be expressed in the modal language. In any way, what does get characterized this way is always a whole class of frames, sharing the property of $\mathcal{R}$ in question, and not just a single frame. Therefore the notion of truth we will focus our attention on in the remainder of this chapter is validity of formulas on a class of frames.

The first thing we have to do is to give an exact definition of the notion of characterization; after that we close this introductory section off by making some remarks in connection with that definition.

Definition 6 A set of formulas Δ *characterizes* a class of frames $\mathcal{C}$ if and only if $\mathcal{C} = \{\mathcal{F} | \mathcal{F} \models \varphi \text{ for all } \varphi \in \Delta\}$. If $\Delta = \{\varphi\}$ we also loosely say that φ, rather than $\{\varphi\}$, characterizes $\mathcal{C}$.

Note that with this definition a statement that Δ characterizes $\mathcal{C}$ comes down to stating that all the formulas in Δ are valid on each frame in $\mathcal{C}$ and on each frame not in $\mathcal{C}$ at least one formula of Δ can be falsified. One thing that is directly obvious is that, if two frames $\mathcal{F}_1$ and $\mathcal{F}_2$ are isomorphic ($\mathcal{F}_1 \cong \mathcal{F}_2$), then they validate the same formulas. This is so, because what a world is is of no importance to its truth values; the only thing that matters is which visible worlds verify which formulas. An immediate consequence is that the classes $\mathcal{C}$ do contain all isomorphic copies of their members. We state these facts in the following proposition.

Proposition 6

(i) If $\mathcal{F}_1 \cong \mathcal{F}_2$, then, for all $\varphi, \mathcal{F}_1 \models \varphi \Leftrightarrow \mathcal{F}_2 \models \varphi$;

(ii) If $\mathcal{C}$ is a characterizable class of frames, $\mathcal{F}_1 \in \mathcal{C}$ and $\mathcal{F}_1 \cong \mathcal{F}_2$, then $\mathcal{F}_2 \in \mathcal{C}$.

For the sake of clarity we point out the fact that, trivially, every set of formulas characterizes some class of frames, namely the class of frames on which it is valid (which may in some cases be empty of course). On the other hand not every class of frames is characterized by such a set. This is immediately obvious from the fact that the possible classes of frames outnumber the possible sets of formulas by far. In the sequel we will not be so much interested in just any class of frames, but in classes specified by some condition on $\mathcal{R}$. Given such a condition we will concern ourselves with the question whether a characterizing set of formulas can be found. As said, we will come across examples of conditions on $\mathcal{R}$ for which such a set can be found, and we will also meet examples for which there is no such set. We start with some instances of properties of the first kind.

2.2 Characterizable properties of $\mathcal{R}$

In discussing semantics in Chapter 1 we already mentioned the fact that

(i) $\Box\varphi \rightarrow \varphi$ characterizes the class of *reflexive* frames,

(ii) $\Box\varphi \rightarrow \Box\Box\varphi$ characterizes the class of *transitive* frames and

(iii) $\Diamond\Box\varphi \rightarrow \varphi$ characterizes the class of *symmetric* frames.

Someone who has never proved propositions of this kind, or does not remember clearly how it was done, is well advised to try these three on. (To give the general idea: when proving $\langle \mathcal{W}, \mathcal{R} \rangle \models \psi \Leftrightarrow \langle \mathcal{W}, \mathcal{R} \rangle$ satisfies P, both sides are best proved by contraposition; if $\langle \mathcal{W}, \mathcal{R} \rangle \not\models \psi$ is assumed, writing out the appropriate definitions of forcing automatically will give one the desired counterexample to '$\langle \mathcal{W}, \mathcal{R} \rangle$ satisfies P' and, if '$\langle \mathcal{W}, \mathcal{R} \rangle$ does not satisfy P' is assumed a suitable valuation on the worlds of $\mathcal{W}$ contradicting the validity of the ψ under consideration is easily produced. A somewhat more complicated example of this proof procedure is spelled out in proposition 9(iii))

The next proposition (the proof of which we leave as exercise 8) serves the purpose of showing that e.g. $\{\Box\varphi \rightarrow \varphi, \Box\varphi \rightarrow \Box\Box\varphi\}$ characterizes the class of frames with $\mathcal{R}$ reflexive as well as transitive, etc., etc.

Proposition 7 If Δ_1 characterizes $\mathcal{C}_1$ and Δ_2 characterizes $\mathcal{C}_2$, then $\Delta_1 \cup \Delta_2$ characterizes $\mathcal{C}_1 \cap \mathcal{C}_2$.

In order to be able to show that the modal logic **T** characterizes the class of the reflexive, **S4** the class of the reflexive and transitive and **S5** the class of the reflexive, transitive and symmetric frames it just remains to prove the two parts of the following proposition:

Proposition 8

(i) If Δ characterizes $\mathcal{C}$, then so does

$$\{\varphi(\psi_1, \ldots, \psi_n) | \varphi(p_1, \ldots, p_n) \in \Delta\}$$

.

(ii) If $\mathcal{C}$ is characterized by Δ, then also by $\{\psi | \vdash_{K+\Delta} \psi\}$.

Proof: Exercise 9. •

Proposition 8(i) ensures us that the formulas we used in characterizing properties of $\mathcal{R}$ can safely be replaced by the corresponding axiom schemata; proposition 8(ii) enables us to see the logic as a set of formulas, as indeed a logic is defined in 2, instead of having to rely on an axiomatization. It is perhaps worth noting that the formulas with propositional variables are easier to use in giving counterexamples, something often needed in characterizability proofs, since their truth values can be stipulated without restraint as needed; whereas the corresponding formulas with variables over formulas are often convenient in other contexts. To give some more examples of characterizable $\mathcal{R}$-properties we define some properties of relations:

Definition 7 Let $\mathcal{R}$ be a binary relation on the set $\mathcal{W}$.

(i) $\mathcal{R}$ is *serial* iff $\forall x \in \mathcal{W}\ \exists y \in \mathcal{W}\ x\mathcal{R}y$.

(ii) $\mathcal{R}$ is *intransitive* iff $\forall x, y, z \in \mathcal{W}\ ((x\mathcal{R}y \wedge y\mathcal{R}z) \to \neg x\mathcal{R}z)$.

(iii) $\mathcal{R}$ is *completely disconnected* iff $\forall x, y \in \mathcal{W}\ \neg xRy$.

(iv) $\mathcal{R}$ is *strongly connected* iff $\forall x, y \in \mathcal{W}\ (x\mathcal{R}y \vee y\mathcal{R}x)$

(v) $\mathcal{R}$ is *piecewise strongly connected* iff
$\forall x, y, z \in \mathcal{W}\ ((x\mathcal{R}y \wedge x\mathcal{R}z) \to (y\mathcal{R}z \vee z\mathcal{R}y))$.

(vi) $\mathcal{R}$ is *dense* iff $\forall x, y \in \mathcal{W}\ (x\mathcal{R}y \to \exists z \in \mathcal{W}\ (x\mathcal{R}z \wedge z\mathcal{R}y))$

(vii) $\mathcal{R}$ is *well-founded* iff there is no infinite sequence of worlds $w_0, w_1, w_2, \ldots$ in $\mathcal{W}$ such that $\ldots w_n\mathcal{R}w_{n-1}\mathcal{R}\ldots w_2\mathcal{R}w_1\mathcal{R}w_0$. (Or, alternatively, there is no infinite sequence $w_0\widetilde{\mathcal{R}}w_1\widetilde{\mathcal{R}}w_2\ldots$, where $\widetilde{\mathcal{R}}$ denotes the *converse* of $\mathcal{R}$, i.e.: $\forall x, y(x\widetilde{\mathcal{R}}y \Leftrightarrow y\mathcal{R}x)$.

An equivalent formulation of well-foundedness is: for every non-empty subset $\mathcal{W}'$ of $\mathcal{W}$ there is a $w \in \mathcal{W}'$ such that, for no $w' \in \mathcal{W}', w'\mathcal{R}w$. [1]

Proposition 9

(i) $\Box\bot$ characterizes the class of completely disconnected frames and $\Box p$ characterizes the same class.

(ii) $\Diamond\top$ characterizes the class of serial frames, as does $\Box p \to \Diamond p$.

(iii) $\Box(\Box p \to q) \vee \Box(\Box q \to p)$ characterizes the piecewise connected frames.

(iv) $\Box(\Box p \to p) \to \Box p$ characterizes the frames with $\mathcal{R}$ transitive and $\widetilde{\mathcal{R}}$ well-founded. (Keep this formula in mind; it plays an important role in provability logic, which we will treat further down.)

Proof: For an example we prove (iii); the other proofs are left as exercise 11. Actually, for (iii) we prove the obvious equivalent:
$\mathcal{F} \not\models \Box(\Box p \to q) \vee \Box(\Box q \to p)$ iff $\mathcal{F}$ is not piecewise connected.
$\Rightarrow$: Suppose $\mathcal{F} \not\models \Box(\Box p \to q) \vee \Box(\Box q \to p)$. Then there must be a $w \in W$ and a $\models$ on $\mathcal{F}$ such that $w \not\models \Box(\Box p \to q) \vee \Box(\Box q \to p)$, i.e. $w \not\models \Box(\Box p \to q)$ and $w \not\models \Box(\Box q \to p)$. This implies there must be $w', w'' \in W$ such that (i) $w\mathcal{R}w'$ and $w' \not\models \Box p \to q$, so $w' \models \Box p$ and $w' \not\models q$, and (ii) $w\mathcal{R}w''$ and $w'' \not\models \Box q \to p$, so $w'' \models \Box q$ and $w'' \not\models p$. As $w' \models \Box p$ and $w'' \not\models p$ it is impossible that $w'\mathcal{R}w''$ and, as $w'' \models \Box q$ and $w' \not\models q$ it is impossible that $w''\mathcal{R}w'$. Hence, we can conclude that $\mathcal{R}$ is not piecewise connected.
$\Leftarrow$: Suppose $\mathcal{F}$ is not piecewise connected. Then the following situation must occur somewhere in $\mathcal{F}$ (see figure 2): there are $w, w', w'' \in W$ such that both $w\mathcal{R}w'$ and $w\mathcal{R}w''$ but neither $w'\mathcal{R}w''$ nor $w''\mathcal{R}w'$. Define $\models$ on $\mathcal{F}$ by the following condition: $v \models p$ iff $w'\mathcal{R}v$, and $v \models q$ iff $w''\mathcal{R}v$, for all $v \in W$.

INSERT FIGURE 2

1. This property cannot, as can the others mentioned here, be expressed by a formula of first order predicate logic. (See exercise 10 for some connections between these properties.)

(Note that this stipulation does not define a complete forcing relation, since it only provides for the propositional letters p and q. It can easily be extended however to cover the other propositional letters in the language; an arbitrary choice for these will do. We will henceforth without further ado give such partial stipulations as if they were complete ones.) Under a forcing relation satisfying the above condition it must be so that $w' \not\models \Box p \to q$ and $w'' \not\models \Box q \to p$, so it is indeed the case that $w \not\models \Box(\Box p \to q) \vee \Box(\Box q \to p)$ and hence $\mathcal{F} \not\models \Box(\Box p \to q) \vee \Box(\Box q \to p)$. •

At the risk of labouring the obvious let us once more emphasize the fact that characterizability is by definition a notion connecting formulas and frames and not formulas and models. This really does make a difference: for example, whereas $\Box p \to p$ is clearly verified by every reflexive model. Yet a model does not have to be reflexive in order to verify $\Box p \to p$, witness the following counterexample:

INSERT FIGURE 3

2.3 The expressive power of the modal language

Forgetting for a moment the special interpretation we tacked on $\mathcal{W}$ and $\mathcal{R}$ in the context of modal logic and the possibility of speaking of the truth of formulas in elements of $\mathcal{W}$, we can look at the frame $\langle \mathcal{W}, \mathcal{R} \rangle$ as an "ordinary" model for the predicate calculus. We can then say that in the previous section we have, via the models, established a number of correspondences between modal formulas on the one hand and formulas of predicate logic on the other: some formulas of the modal propositional language turn out to be equivalent to some formulas of the standard predicate (!) calculus in the sense that they are true on exactly the same semantic structures, by first viewing these as frames for modal logic and after that as models for the predicate calculus. To be more precise: the formulas written using symbols of the predicate calculus in definition 9 were not formulas of a logical object language: we only used them as convenient abbreviations for sentences in the English metalanguage describing the *models* we had in mind. But this way of describing models is convenient in connecting the models with "real" formulas of the language of the predicate calculus, for there can be no misunderstanding which formulas those will have to be. Hence, we can conclude that some properties of $\mathcal{R}$ can be expressed in modal formulas as well as in formulas of the first order predicate calculus. Furthermore, we have already seen a counterexample against the over-hasty conclusion that both languages are equally strong in expressive power: the property of being well-founded plus transitive cannot be secured by means of a formula of first order predicate logic, while in the modal language it is exactly characterized by $\Box(\Box p \to p) \to \Box p$. The latter can do more in this case. However, it is also a mistake to jump from this one example to the general conclusion that modal propositional logic is the stronger of the two: some properties are easily expressed by means of the language of the predicate calculus, while there exists no formula in the modal language (and not even a whole set of formulas) that can do the same. Examples of such properties are in the following:

Proposition 10 There is no set Δ of formulas of the modal propositional language such that Δ characterizes:

(i) the class of frames which are not reflexive
(ii) the class of strongly connected frames
(iii) the class of frames in which $\mathcal{R}$ is universal (i.e. $\mathcal{R}$ is interpreted as $\mathcal{W} \times \mathcal{W}$)
(iv) the class of irreflexive frames
(v) the class of asymmetric frames
(vi) the class of intransitive frames.
(vii) the class of frames of which the set of worlds $\mathcal{W}$ has cardinality n.

Proofs of these claims are given in the next section, after some general remarks are made and some useful techniques have been introduced.

2.3.1 Non-characterizability proofs

What we need in order to establish propositions of the form "The class of frames of which the $\mathcal{R}$ has property P is not characterizable" is a pair of frames $\mathcal{F}$ and $\mathcal{F}'$ such that $\mathcal{F}$ has the property P in question while $\mathcal{F}'$ lacks it, but still validates all the formulas valid in $\mathcal{F}$. To see why this is sufficient for our purpose let us assume we have such a pair of frames. If we then claim that some set Δ of modal formulas characterizes the class of P-frames we are immediately faced with a contradiction. For $\mathcal{F}'$ has to falsify one of the $\varphi \in \Delta$, while $\mathcal{F}$ has to verify all of them. This contradicts our assumption that all the formulas validated by $\mathcal{F}$ are also valid on $\mathcal{F}'$. So the whole secret lies in finding such a pair of frames. In the sequel we will discuss three techniques for constructing "new" frames out of given ones, all three of which satisfy the condition assumed above: none of the formulas valid on the frame $\mathcal{F}$ we start with ceases to be valid on any frame $\mathcal{F}'$obtained from $\mathcal{F}$ by using these methods of construction. When therefore in going from $\mathcal{F}$ to $\mathcal{F}'$ a property of $\mathcal{R}$ does get lost, this will mean that that property is not expressible in the modal language. (Actually in the second method, the one of constructing disjoint unions, two frames $\mathcal{F}_1$ and $\mathcal{F}_2$ are used in the construction of $\mathcal{F}'$. This makes no essential difference, if one substitutes in the discussion above ' valid in $\mathcal{F}_1$ and $\mathcal{F}_2$' for ' valid in $\mathcal{F}'$.)

2.3.2 Generated subframes

Two ideas underly the method of generated subframes. The first one is that, if one has a frame $\mathcal{F}$, then the formulas valid on $\mathcal{F}$ are the formulas valid on all $w \in \mathcal{W}$ in all models on $\mathcal{F}$, and that hence, if one could get a frame $\mathcal{F}'$ on a subset $\mathcal{W}'$ of $\mathcal{W}$ verifying on each $w \in \mathcal{W}'$ in all models on $\mathcal{F}'$ the same formulas as in a corresponding model on $\mathcal{F}$, this would be a frame as required, because all formulas "valid on all of $\mathcal{W}$" would certainly be "valid on $\mathcal{W}'$ ", since $\mathcal{W}' \subseteq \mathcal{W}$. Now, in general, if one restricts a frame to a certain subset $\mathcal{W}'$ of $\mathcal{W}$ the relation $\mathcal{R}$ will get badly disturbed: worlds w will miss worlds w' which were visible from them in $\mathcal{W}$, but have been lost in $\mathcal{W}'$. In consequence one will not fulfill the requirement that on all models w will keep verifying the same formulas. The second idea in generated subframes is that, if in such a situation in any world w in $\mathcal{W}'$ all worlds visible in $\mathcal{W}$ from w are still present in $\mathcal{W}'$, then the first idea is successful after all. The point is that only worlds visible from w directly or via other worlds can be of any consequence for the truth values of formulas in w.

Proposition 11 following the definition of 'subframe generated by w' puts exactly into words what we have just stated.

Definition 8 Let $\mathcal{F} = \langle \mathcal{W}, \mathcal{R} \rangle$, $w \in \mathcal{W}$ and $\mathcal{M} = \langle \mathcal{F}, \models \rangle$.

(i) The *hereditary closure of* $\mathcal{R}$ (notation $\mathcal{R}^*$) is defined as follows: For all $w, w' \in \mathcal{W}$: $w\mathcal{R}^*w' \Leftrightarrow w = w'$ or there is a sequence of worlds $w = w_1\mathcal{R}w_2\mathcal{R}w_3\mathcal{R}\ldots\mathcal{R}w_n = w'$.

(ii) For $w \in \mathcal{W}$ we further define $\mathcal{W}_w = \{w' | wR^*w'\}$.

(iii) *The subframe of* $\mathcal{F}$ *generated by* w (notation: $\mathcal{F}_w$) can now be defined as $\mathcal{F}_w = \langle \mathcal{W}_w, \mathcal{R}_w \rangle$ (with $\mathcal{R}_w$ denoting $\mathcal{R}$ restricted to $\mathcal{W}_w$).

(iv) Finally, $\mathcal{M}_w$, is by definition the model that results when one restricts $\mathcal{M}$'s forcing relation to $\mathcal{W}_w$.

Proposition 11

(i) $w' \models_{\mathcal{M}_w} \varphi \Leftrightarrow w' \models_M \varphi$, for all φ and all $w' \in \mathcal{W}_w$ (So it cannot give rise to problems to just write $w' \models \varphi$, for both.)

(ii) For all φ : if $\mathcal{F} \models \varphi$, then $\mathcal{F}_w \models \varphi$.

(iii) If $\mathcal{C}$ is a characterizable class of frames and $\mathcal{F} \in \mathcal{C}$, then $\mathcal{F}_w \in \mathcal{C}$.

Proof:

(i) The proof by induction is straightforward but long-winded; therefore we omit it. That the proposition holds can easily be seen from the following facts though: the only worlds in $\mathcal{M}$ that can influence the truth of formulas in w' are precisely those in $\mathcal{W}_{w'}$. Conversely, the worlds in $\mathcal{M}_w$ that can have any impact on the truth of formulas in w' are precisely those in $(\mathcal{W}_w)_{w'}$. According to exercise 12(c) we have $(\mathcal{W}_{w'} = (\mathcal{W}_w)_{w'}$ and because also the forcing relations of the two models coincide (at least in so far as can be relevant to w'), $\models_{\mathcal{M}}$ and $\models_{\mathcal{M}_w}$ must be the same in w' for all φ.

(ii) Suppose $\mathcal{F}_w \not\models \varphi$. Then there is a model $\mathcal{M}_w$ on $\mathcal{F}_w$ such that $w' \not\models_{\mathcal{M}_w} \varphi$ for certain $w' \in \mathcal{W}_w$. If we extend $\mathcal{M}_w$ to a model $\mathcal{M}$ on $\mathcal{F}$ (any extension will do), we can apply (i) to get $w' \not\models_{\mathcal{M}} \varphi$, hence $\mathcal{F} \not\models \varphi$.

(iii) We leave the proof as exercise 13.

•

We now have at our disposal all the means needed to prove that the class of non-reflexive frames is not characterizable (claim (i) of proposition 10(i)).

Proof: Look at the frame in figure 4, which is not reflexive. The subframe generated by w_3 is (isomorphic to) the frame in figure 5 and this frame lacks the property of being non-reflexive. According to proposition 2.3.5 the latter validates all the formulas valid on the first. For the reasons explained above this implies that non-reflexiveness is not characterizable. •

INSERT FIGUREs 4 AND 5

2.3.3 Disjoint unions of frames

Before we explain this method for constructing frames, we would first like to draw attention to the fact that frames may be built up from a number of 'loose pieces': 'pieces' consisting of proper subsets of $\mathcal{W}$ of which the worlds mutually are connected by $\mathcal{R}$ in a certain fashion while there exists no connection at all between the worlds in different 'pieces'. To the truth values of the formulas in a world w the presence of such other 'pieces' that are unconnected to its own 'piece' can make no difference. This implies that formulas valid in both pieces separate will be so in the whole. This is the idea that underlies the definition (and the subsequent use made of it) of taking the disjoint unions of frames: two frames (at least in the definition as we give it below; this definition can however be extended in an obvious manner to cover a more general case) get "amalgamated" into one new frame. Thereby it is the intention to keep both component frames strictly separated, for the nice connections between the 'old' frames and the 'new' one only hold under the condition that we add *loose* pieces. This last demand, if one just takes the union of two frames by taking the union of their sets of worlds, one gets into difficulties in case the intersection of the two world-sets of the composing frames is not empty, for then we have no guarantee that they will not be connected by means of one of the 'original' accessibility relations. In order to make sure that these difficulties will not crop up the following definition contains a simple trick:

Definition 9 The *disjoint union* of $\mathcal{F}_1 = \langle \mathcal{W}, \mathcal{R} \rangle$ and $\mathcal{F}_2 = \langle \mathcal{W}', \mathcal{R}' \rangle$ (notation $\mathcal{F}_1 \sqcup \mathcal{F}_2$) is the frame with the set of worlds $\mathcal{W} \sqcup \mathcal{W}' = \{(w, 0) | w \in \mathcal{W}\} \cup \{(w, 1) | w \in \mathcal{W}'\}$ and the accessibility relation $\bar{\mathcal{R}}$ consisting of all pairs $((w_1, i), (w_2, j))$ such that, either $i = j = 0$ and $w_1 \mathcal{R} w_2$, or $i = j = 1$ and $w_1 \mathcal{R}' w_2$.

The following proposition spells out that the connection between validity on the two original frames and on the new frame is indeed as could be expected.

Proposition 12

(i) For all $\mathcal{F}$, $\mathcal{F}'$ and $w \in \mathcal{W}$, $\mathcal{F}_w$ is isomorphic to the subframe of $\mathcal{F} \sqcup \mathcal{F}'$ generated by $(w, 0)$. (The same holds of course for $\mathcal{F}'$ and $w \in \mathcal{W}'$)

(ii) $\mathcal{F} \sqcup \mathcal{F}' \models \varphi$ iff $\mathcal{F} \models \varphi$ and $\mathcal{F}' \models \varphi$.

(iii) Let $\mathcal{C}$ be a characterizable class of frames, then $\mathcal{F} \sqcup \mathcal{F}' \in \mathcal{C}$ iff $\mathcal{F} \in \mathcal{C}$ and $\mathcal{F}' \in \mathcal{C}$.

Proof:

(i) The isomorphism needed is the function f such that $f(w', 0) = w'$ for all $w' \in \mathcal{W}$.

(ii) $\Leftarrow$: Suppose $\mathcal{F} \sqcup \mathcal{F}' \not\models \varphi$. Then there exists a $\models$ on $\mathcal{F} \sqcup \mathcal{F}'$ with $(w, i) \not\models \varphi$. Suppose $i = 0$ (The argument is the same for $i = 1$.) According to the (i) part we just proved and proposition 11(i) there is a $\models$ on $\mathcal{F}$ such that $w \not\models \varphi$. Hence, $\mathcal{F} \not\models \varphi$.
$\Rightarrow$: Suppose $\mathcal{F} \not\models \varphi$. Then, for some $\models$ on $\mathcal{F}$ and some w, $w \not\models \varphi$. For a forcing relation $\models'$ on $\mathcal{F} \sqcup \mathcal{F}'$ defined by taking $(w, 0) \models' p \Leftrightarrow w \models p$ for propositional letters p and arbitrarily defining $\models'$ on the $(w', 1)$ it clearly has to be the case that $(w, 0) \not\models' \varphi$. Hence, $\mathcal{F} \sqcup \mathcal{F}' \not\models \varphi$.

(iii) $\Rightarrow$ follows already from proposition 11(iii) and part (i). The argument for $\Leftarrow$ is analogous.

•

This technique enables us to give a proof that strong connectedness of $\mathcal{R}$ cannot be expressed in terms of modal formulas, and the same holds for universality of $\mathcal{R}$.

Proof: (of 10(ii)). The frame in figure 5 is strongly connected and the frame in figure 6 is isomorphic to the disjoint union of two copies of the frame in figure 5. Hence, by propositions 10(iii) and 6(ii) the statement follows immediately. •

INSERT FIGURE 6

Proof: (of 10(iii)). We can use the same pair of frames considering them this time in a different perspective. The frame in figure 5 has a universal $\mathcal{R}$ while the one in figure 6 has not, etc. •

The proofs of the other claims mentioned in proposition 10 depend on a somewhat more complicated notion, which we presently set out to discuss:

2.3.4 P-morphisms

One of our early conclusions was that validity of formulas is of course preserved under the taking of isomorphic images. Isomorphism is a very "strong" connection between two frames however: it means that the two structures are identical from a mathematical point of view. But since (as we have seen already in the previous) not everything about the structure of a frame is susceptible to being pinned down by means of modal formulas, we need not demand such a strong kind of similarity to have the same set of valid formulas in two frames. Moreover, for the purpose of disproving characterizability we only need to construct frames in which at least (and not precisely) the "formerly" valid formulas remain to be so. Instead of producing by an isomorphism exact copies we are therefore going to produce images by functions which overlook certain dissimilarities; we are free to do so as long as we make sure that those aspects of the frame that do affect the truth of formulas are not altered in the process. The idea now is that, since the identity or distinctness of the worlds is nowhere essential in the forcing definition different worlds may get to be mapped on to the same one, but only in such a way that no structure gets disturbed: every complex of worlds visible directly or indirectly from a world w must be retraceable in the image. In the image of such a complex the rôles of two or more worlds may have been taken over by a single one. Thus it may easily happen that the new frame is smaller than the old one. In the following definition this idea is made precise:

Definition 10 A function f is a *p-morphism from* $\mathcal{F} = \langle \mathcal{W}, \mathcal{R} \rangle$ *onto* $\mathcal{F}' \langle \mathcal{W}', \mathcal{R}' \rangle$ iff f is a surjection from $\mathcal{W}$ to $\mathcal{W}'$ with the following properties:

(i) $w\mathcal{R}w' \Rightarrow f(w)\mathcal{R}'f(w')$. (In other words, f is a *homomorphism*).

(ii) $f(w)\mathcal{R}'f(w') \Rightarrow$ there is a world $w'' \in \mathcal{W}$ such that $w\mathcal{R}w''$ and $f(w') = f(w'')$.

We call $\mathcal{F}'$ a p-morphic image of $\mathcal{F}$ *by* f.

Clause (i) says that whenever a world is visible from a world in the original, the same holds for their f-images. Clause (ii) signifies that if a world v sees a world v' in the image while their counterparts in the original do not, this can only be the case because v' is the f-image of another world in $\mathcal{W}$ and this fact forces the connection on $\mathcal{R}'$ by clause (i).

Comparing the validity of formulas on a frame and its p-morphic images leads to the following results (of a by now familiar kind):

Proposition 13 Let $\mathcal{F}'$ be a p-morphic image of $\mathcal{F}$ by f, then:

(i) If for all $w \in \mathcal{W}$ and for all propositional letters p, $w \models p$ iff $f(w) \models p$, then it is also the case that $w \models \varphi$ iff $f(w) \models \varphi$ for all formulas φ.

(ii) For all φ it holds that, if $\mathcal{F} \models \varphi$, then $\mathcal{F}' \models \varphi$.

(iii) If $\mathcal{C}$ is a characterizable class of frames and $\mathcal{F} \in \mathcal{C}$, then $\mathcal{F}' \in \mathcal{C}$.

Proof:

(i) By induction on the complexity of φ. We spell out one case; the rest is completely trivial: $f(w) \not\models \Box\varphi \Leftrightarrow$ (f is a surjection) there is some w_1 such that $f(w)\mathcal{R}'f(w_1)$ and $f(w_1) \not\models \varphi \Leftrightarrow$ ($\Rightarrow$ and $\Leftarrow$ resp. properties (ii) and (i) of p-morphisms) there exists a w_2 such that $w\mathcal{R}w_2$ and $f(w_2) \not\models \varphi \Leftrightarrow$ (induction hypothesis) there is a w_2 such that $w\mathcal{R}w_2$ and $w_2 \not\models \varphi \Leftrightarrow w \not\models \Box\varphi$.

(ii) This follows immediately from (i).

(iii) This follows immediately from (ii).

•

We are now in a position where we can prove the remaining claims of proposition 10 , recapitulating, (iv) The class of irreflexive frames is not characterizable. (v) The class of intransitive frames is not characterizable. (vi) The class of asymmetric frames is not characterizable. (vii) The class of frames with exactly n worlds is not characterizable.

Proof: (iv) Compare the frame in figure 7 below with the frame in figure 5. It can easily be checked that the latter is a p-morphic image of the first. Hence we can apply Proposition 2.3.11(iii).

INSERT FIGURE 7

(v) The frame in figure 7 is intransitive and has as a p-morphic image, the frame in figure 5, which is not intransitive.

(vi) Compare the frame in figure 5 this time with the frame in figure 8 below. This frame also has the frame in figure 5 as its p-morphic image and the first is asymmetric while the latter is not, etc.

INSERT FIGURE 8

(vii) Exercise 14 •

Reconsidering the three techniques we discussed in the last three sections we can point out a common principle underlying them: We start with the idea which is at that point not much more than a conjecture that perhaps not every structural aspect of a model can be captured in the language under consideration. In order to prove this surmise we need first of all an exact specification of these aspects so that we can vary the frame to our heart's content without these changes possibly coming to light in even the most accurate description in the given language. In each of the definitions we gave (i.e. the one of generated subframes, the one of disjoint unions and the one of p-morphisms) the finger is put on such an aspect. And subsequently we were able to prove that our conjectures concerning these were right. This implies that a class of frames with a certain property cannot be characterized unless it is stable under variations with respect to these aspects; in other words, stability under these three operations is a necessary conditions for a class to be characterizable. After having spotted a number of such "elusive" (with respect to the language given of course) features it becomes a natural question to ask whether it is perhaps possible to give an exhaustive list of all features sharing this property, that is: whether there are also sufficient conditions for characterizability. This question has actually been asked and can be answered (surprisingly maybe) in the affirmative. It is however not so that the three aspects we already mentioned are sufficient. This can be seen from exercise 19. If everything were right you should able to prove that the class of frames with $\widetilde{\mathcal{R}}$ well-founded is not characterizable. This class is closed however under taking generated subframes, disjoint unions and p-morphic images, as can easily be checked. So something else must be at work here to spoil the possibility of characterization. To get an idea of what this "something else" comes down to look at what is used to disprove characterizability in this case: whenever you have a valuation that provides a counterexample on the frame that lacks converse well-foundedness you can easily transpose it (the "relevant" part of this valuation, anyway) to a frame that has the property in question. In order for a class of frames to be characterizable then, this possibility has to be precluded. The exact formulation of what has to be precluded is however rather complicated (it makes use of a construction-technique called taking the "ultrafilter extension" of a frame), so we will not give it right now. This technique added to the three we had already provides indeed sufficient conditions for characterizability of a class of frames. This yields the following elegant theorem (which we are also as yet not in a position to prove and will only state):

Proposition 14 A class of frames is characterizable if and only if it is closed under taking generated subframes, disjoint unions and p-morphic images and its complement is closed under taking ultrafilter extensions.

2.4 Exercises

Exercise 8 Prove proposition 7.

Exercise 9

(i) Show that, if $\mathcal{F} \not\models \varphi(\psi_1, \ldots, \psi_n)$, then also $\mathcal{F} \not\models \varphi(p_1, \ldots, p_n)$,. Use this to prove proposition 8(i).

(ii) Prove proposition 8(ii).

Exercise 10 Show

(i) If $\mathcal{R}$ is well-founded, then $\mathcal{R}$ is irreflexive and asymmetric,

(ii) If $\mathcal{R}$ is finite and transitive, then $\mathcal{R}$ is well-founded iff $\mathcal{R}$ is irreflexive; show further that, under the same conditions, this equivalence also holds for $\widetilde{\mathcal{R}}$,
(iii) If $\mathcal{R}$ is reflexive, then $\mathcal{R}$ is serial and dense.

Exercise 11 Proof the claims (i), (ii) and (iv) of proposition 9.

Exercise 12 To get a feeling for what generated subframes look like it is useful to prove the following:
(i) Show that $\mathcal{W}_w$ is the smallest subset $\mathcal{W}'$ of $\mathcal{W}$ such that the following holds: (i) $w \in \mathcal{W}'$; and (ii) for all $w' \in \mathcal{W}'$: if $w'\mathcal{R}w''$, then $w'' \in \mathcal{W}'$.
(ii) Show that if $\mathcal{R}$ is transitive, then $\mathcal{W}_w = \{w' \in \mathcal{W} | w = w' \text{ or } w\mathcal{R}w'\}$ and if $\mathcal{R}$ is reflexive also $\mathcal{W}_w = \{w' \in \mathcal{W} | w\mathcal{R}w'\}$.
(iii) Show that if $w' \in \mathcal{W}_w$, then $\mathcal{W}_{w'} = (\mathcal{W}_w)_{w'}$.

Exercise 13 Prove proposition 11(iii).

Exercise 14 Prove proposition 10(vii).

Exercise 15 Prove that the frames of figures 9, 10 and 11 validate the same formulas.

INSERT FIGURE 9, 10 AND 11

Exercise 16
(i) Show that the frame in figure 5 is a p-morphic image of the frames in the figures 6, 7 and 8 and of $\langle , \mathbb{N}, < \rangle$
(ii) Check whether the other frames mentioned in (i) are also p-morphic images of $\langle , \mathbb{N}, < \rangle$.

Exercise 17 Prove that if a formula φ is neither valid on the frame $\langle \{w\}, \emptyset \rangle$, nor on the frame $\langle \{w\}, \{(w, w)\} \rangle$, then φ is valid on no frame at all. Draw a conclusion.

Exercise 18
(i) Find combinations of properties of $\mathcal{R}$ whose non-characterizability can be established by means of the frames in the figures 5, 6, 7 and 8, or other simple finite frames. (For example: on the account of figure 7 it can be seen that irreflexiveness + intransitivity + symmetry is not expressible. (Note this is not trivial as it might seem, since converse well-foundedness for example cannot be expressed, whereas this property combined with transitivity can.)

(ii) If you did the (i)-part of this exercise right you have not found a finite frame to show that transitivity + irreflexiveness is not characterizable. Explain why not.

(iii) Show by means of a simple infinite frame that transitivity + irreflexiveness is not characterizable.

Exercise 19

(i) Show that well-foundedness of $\mathcal{R}$ is not characterizable.

(ii) (a little more difficult!) Show that well-foundedness of $\widetilde{\mathcal{R}}$is not characterizable. (Hint: Define a number that expresses the degree of nestedness of $\Box$ in φ. Consider the frame consisting of $\mathbb{N}$ with the successor relation S for accessibility (i.e. $nSm \Leftrightarrow n = m + 1$). Finally, investigate how a formula φ with nestedness number k behaves on $\langle, \mathbb{N}, S\rangle$).

Chapter 3

Completeness and decidability

In this chapter we will prove for several modal logics, viz. **K**, **T**, **S4** and **S5**, that they are strongly complete and decidable.

The main thrust of the proofs is the same: in each and every case we have to obtain a model satisfying a number of constraints. We will discuss two different ways of acquiring such models. Both will be introduced in connection with proofs of strong completeness in the first section of this chapter. In the second second section we will discuss a finite variant of each. By means of the finite versions 'plain' completeness can be shown to hold in cases where strong completeness fails. But these finite versions can, in addition, be used to show that the logic in question is decidable.

3.1 Strong completeness: two methods of proof

What it means for a logic to be sound and complete To prove that a given logic **S** is sound — as we did for **K** in chapter 1 — one has to show that **S** is not too strong for the semantics specified: every formula that can be derived in **S** from a (in the strong case: possibly infinite) set of premises is bound to be true in any world in which the premises are true that one can find in any **S**-frame .

The completeness theorem for a logic **S** states the converse: a formula that is true in any world of a frame that validates **S** in which every formula of a (in the strong case, possibly infinite) set of premises is true, can be derived from this set. Thus, in proving a logic to be complete we establish that the logic is not weaker than the class of frames it characterizes suggests.

Together, the soundness and completeness theorem for a logic imply that the logic is the right logic for its semantics; it accurately reflects the structural features which the frames in this class have in common: derivability and validity coincide.

(Strong) completeness defined Contrasting the case of modal logic with the case of the standard propositional calculus, the first difference we come across is that in the present case we have to deal with several logics instead of just one. This difference requires a small adjustment to be made in the concept of completeness.

The propositional calculus could be called complete without much further ado. In the case of modal logic however, we will have to set aside for each logic a subclass of the class of all possible Kripke-frames, the elements of which will be the potential counterexamples when establishing the (strong) completeness of the logic in question; henceforth, we will therefore always speak of 'completeness with respect to a certain class of frames'. As one might expect, the demarcation of a class of frames for a logic is obtained (at least in first instance) by ignoring those frames on which the logic is not valid, i.e. one concentrates on the class of frames characterized by the logic. Since in general we opt for this 'maximal' choice, we simply call a logic (strongly) complete if it is complete with respect to this class. However, once the concept of 'completeness with respect to a class of frames' is introduced, it will sometimes be interesting to consider completeness of a logic with respect to some subclass of its characteristic class; in doing so successfully for some smaller class, we gain insight into which features of the frames for the logic in question are essential to it being the logic it is.

These reflections lead to the following

Definition 11 Let **S** be some modal logic,

(a) **S** is *strongly complete with respect to a class of frames* $\mathcal{C}$ iff it holds that If $\Phi \models_{\mathcal{C}} \varphi$, then $\Phi \vdash_{\mathbf{S}} \varphi$.
If $\mathcal{C} = \{\mathcal{F}\}$ for some $\mathcal{F}$, we also say that **S** is strongly complete with respect to $\mathcal{F}$.

(b) **S** is *complete with respect to* $\mathcal{C}$ iff (a) holds for all finite Φ.

(c) The *class of frames characterized by* **S**, *Char*(**S**) is defined as $Char(\mathbf{S}) = \{\mathcal{F} \mid \text{ for all } \varphi \in \mathbf{S}, \mathcal{F} \models \varphi\}$.

(d) **S** is *(strongly) complete* iff ((a) or) (b) holds for $\mathcal{C}$=*Char*(**S**).

We conclude this section with three remarks on the above definition.

Firstly: for the sake of analogy to the definition of soundness we might have added to this definition a clause for 'weak completeness', defining this notion in the obvious manner. The reason we did not do so is that this notion coincides with "plain" completeness in the sense that every weakly complete logic is complete (and vice versa of course). It should be noted that this equivalence does not extend to strong completeness: although by definition no strongly complete logic can fail to be complete, the converse is not true. In the sequel (cf. the chapter on provability logic) we will meet a logic that is complete, but not strongly complete.

Secondly, from the fact that a logic is strongly complete it immediately follows that the logic is also compact. (A logic is *compact* if and only if for every set of formulas Φ, Φ is satisfiable iff all its finite subsets are.):

Proposition 15 Let **S** be a logic. If **S** is strongly complete it is compact.

Proof: We argue by contraposition. Assume that there is no world w in any model on any **S**-frame such that $w \models \Phi$. Then $\Phi \models_{Char(\mathbf{S})} \bot$. Strong completeness implies that $\Phi \vdash_{\mathbf{S}} \bot$. Since derivations are finite sequences there must be $\varphi_1, \ldots, \varphi_n \in \Phi$ such that $\varphi_1, \ldots, \varphi_n \vdash_{\mathbf{S}} \bot$. Hence, the finite subset $\{\varphi_1, \ldots, \varphi_n\}$ of Φ is inconsistent, whence it is not satisfiable. •

The fact that strong completeness entails compactness is often used to show that strong completeness does not hold for a logic, because disproving compactness is generally easier than disproving strong completeness directly; an example of this can be found in the chapter on provability logic.

Finally, note that once again we are concerned with truth in (classes of) frames and not with truth in models. This is natural, because we are interested in connecting what is valid, i.e. true under *every possible valuation* (in the class of frames in question) to what can be derived in the logic.

Strong completeness: two methods of proof Our object in this section is to prove for several modal logics **S**, viz. for **K**, **T**, **S4** and **S5**, that if $\Phi \models_{Char(\mathbf{S})} \psi$, then $\Phi \vdash_{\mathbf{S}} \psi$, for all Φ and ψ .

It would however be rather difficult to give a direct proof of this proposition for, since there are so many Φ, ψ, and $\mathcal{F} \in Char(\mathbf{S})$ to consider; the assumption $\Phi \models_{Char(\mathbf{S})} \psi$ is hard to handle. To get a better hold on what has to be shown, completeness theorems are nearly always proven by contraposition. So we are going to establish propositions of the form 'If $\Phi \nvdash_{\mathbf{S}} \psi$ then $\Phi \not\models_{Char(\mathbf{S})} \psi$' . What we need then is, for all Φ and ψ such that $\Phi \nvdash_{\mathbf{S}} \psi$, a model on a frame in $Char(\mathbf{S})$ with for some world w such that $w \models \varphi$ for all $\varphi \in \Phi$ while $w \not\models \psi$.

The two methods we employ to get such models are called the *method H*, in which the 'H' stands for 'Henkin', and the *method C* ('C' for 'construction'). These methods will be treated separately in some detail, in the order mentioned.

Some general remarks on the method H The principle behind the method H stems directly from the current way of proving the completeness theorem for the standard predicate calculus. The backbone of the method is, the idea to build a model from maximal consistent sets of formulas (i.e. consistent sets Δ such that no $\Delta' \supset \Delta$, Δ' is consistent). These sets link up the syntax with the semantics: a purely syntactical piece of information, the fact that a set of sentences is consistent, is used to acquire something of a purely semantic nature, a valuation.

The three cornerstones the method H shares with its analogue in standard logic, are, as we will see, the reduction of the completeness theorem to the consistency theorem, the Lindenbaum lemma and the valuation lemma .

3.1.1 The method H in modal logic

To make the method used in the classical case fit the more complicated case of modal logic, some adjustments have to be made.

A first complicating factor once stems from the fact that we have many different logics to cope with, instead of just one. Since what is to be called consistent and what is not comes down to something different for each different logic, we have to be careful in making distinctions. From now on we will therefore always speak of **S**-consistency, whenever we want to make precise statements.

Two other complicating factors are that here we need many valuations serving as possible worlds (instead of just one) and we need to define an accessibility relation between them. These difficulties are dealt with in the following way:

For every modal logic **S**, we are going to define one large model, called the 'canonical model $\mathcal{M}_{\mathbf{S}}$ for **S**', containing all at once, for every formula ψ and all sets of sentences Φ such that $\Phi \not\vdash_{\mathbf{S}} \psi$, a counterexample against the truth of ψ in some world in which all the elements of Φ are true. Since a model is composed of a set of worlds, a valuation and an accessibility relation, we have to decide for each of these what they look like. Our decisions in these matters are the following:

The set of worlds. We are going to take all maximal **S**-consistent sets as worlds. (Note that, given a countably infinite language, there are always $2^{\aleph_0}$ of these.)

The forcing relation. It is obvious we want to base the $\models$-relation on the contents of the worlds as sets, in the sense that a formula φ is true in some world iff φ is an element of the maximal **S**-consistent set representing this world.

In fact, maximal consistent sets are valuations in a sense, or rather, they can be taken to determine one unambiguously. As usual this procedure of getting a valuation via such a set of formulas starts by stipulating the truth values for the atomic propositions and letting the semantic definitions of the logical connectives do the rest. Consequently, we have to check that by this procedure membership of the set and being true under the valuation come down to the same thing. The valuation lemma is there to prove that this is indeed the case.

The accessibility relation Keeping in mind the intention to use the sets as valuations, we have already one half of the answer to the question of what to do with $\mathcal{R}$: because of the semantics of $\Box$ we have got to define $\mathcal{R}$ in such a manner that $\Gamma\mathcal{R}\Gamma'$ implies that $\varphi \in \Gamma'$ for all Γ, $\Gamma' \in \mathcal{W}$ and all φ with $\Box\varphi \in \Gamma$. Simple-mindedly reversing the arrow in addition turns out to be a right choice for the definition. The proof of this fact is in the valuation lemma.

Altogether, we get the following definition.

Definition 12

The *canonical model* $\mathcal{M}_{\mathbf{S}}$ of the logic **S** is the model $\langle\langle \mathcal{W}_{\mathbf{S}}, \mathcal{R}_{\mathbf{S}}\rangle, \models\rangle$ with

(i) $\mathcal{W}_{\mathbf{S}} = \{\Gamma \mid \Gamma \text{ is maximal } \mathbf{S}\text{-consistent}\}$

(ii) $\mathcal{R}_{\mathbf{S}} = \{\langle\Gamma, \Gamma'\rangle \mid \varphi \in \Gamma' \text{ for all } \varphi \text{ such that } \Box\varphi \in \Gamma\}$
(iii) and $\models$ defined by, $\Gamma \models p$ iff $p \in \Gamma$.
We write $\mathcal{F}_{\mathbf{S}}$ for the frame $\langle\mathcal{W}_{\mathbf{S}}, \mathcal{R}_{\mathbf{S}}\rangle$.

Under this stipulation for $\models$ it holds that

Lemma 16 (Truth Lemma) $\Gamma \models \varphi$ iff $\varphi \in \Gamma$, for all formulas φ .

To prove this we need the Lindenbaum lemma and the valuation lemma.

Lemma 17 (Lindenbaum lemma.) For all **S**-consistent sets Γ there is a maximal **S**-consistent $\Gamma' \supseteq \Gamma$

Proof: The proof is almost the same as in the classical case. Therefore, a rough outline must suffice:

Fix an enumeration of all the formulas in the language. Set $\Gamma_0 = \Gamma$. Define, for all $n \in \omega$, $\Gamma_{n+1} = \Gamma_n \cup \{\varphi_n\}$, if this set is **S**-consistent; $\Gamma_{n+1} = \Gamma_n$ otherwise. Finally, take $\Gamma' = \bigcup_{n\in\omega} \Gamma_n$. •

Lemma 18 (Valuation lemma) For all maximal **S**-consistent Γ the following holds:

(i) $\Gamma \vdash_S \varphi$ iff $\varphi \in \Gamma$.
(ii) $\varphi \in \Gamma$ iff $\neg\varphi \notin \Gamma$.
(iii) $\varphi \wedge \psi \in \Gamma$ iff $\varphi \in \Gamma$ and $\psi \in \Gamma$.
(From these two it follows that analogous propositions hold for all the other standard propositional connectives included in the language.)
(iv) $\Box\varphi \in \Gamma$ iff $\varphi \in \Gamma'$ for all Γ' with $\Gamma\mathcal{R}_{\mathbf{S}}\Gamma'$.

Proof: (i) and (ii) can be proven as in the classical case; to refresh your memory we took them in as exercise 22. As for (iii):

$\Rightarrow$: Trivial, since it follows immediately from the definition of $\mathcal{R}$.

$\Leftarrow$: Suppose $\Box\varphi \notin \Gamma$, then ($\Gamma$ is maximal **S**-consistent) $\neg\Box\varphi \in \Gamma$. So $\{\Box\psi \mid \Box\psi \in \Gamma\} \cup \{\neg\Box\varphi\}$ is **S**-consistent. We have to show that there is a $\Gamma' \in \mathcal{W}_{\mathbf{S}}$ such that $\Gamma\mathcal{R}\Gamma'$ and $\neg\varphi \in \Gamma$. We know that such a Γ' exists if $\{\psi \mid \Box\psi \in \Gamma\} \cup \{\neg\varphi\}$ is **S**-consistent, because then this set is contained in a maximal **S**-consistent set that satisfies all conditions put on the Γ' we are looking for. Now suppose that $\{\psi \mid \Box\psi \in \Gamma\} \cup \{\neg\varphi\}$ is not **S**-consistent. Then $\{\psi \mid \Box\psi \in \Gamma\} \vdash_{\mathbf{S}} \varphi$. Because every derivation in **S** is finite it follows that for some $\psi_1, \ldots, \psi_n \in \{\psi \mid \Box\psi \in \Gamma\}$, $\psi_1, \ldots, \psi_n \vdash_{\mathbf{S}} \varphi$. According to exercise 2 (check this proposition also holds for other logics than **K**), $\Box\psi_1, \ldots, \Box\psi_n \vdash_{\mathbf{S}} \Box\varphi$. This however contradicts our assumption that $\{\Box\psi \mid \Box\psi \in \Gamma\} \cup \{\neg\Box\varphi\}$ is **S**-consistent. •

From this the Truth lemma can easily be deduced, by induction on the complexity of φ.

Proof: The only new step, compared to the classical case, is $\varphi = \Box\psi$: $\Box\psi \in \Gamma$ iff (valuation lemma (iii)) for all Γ' with $\Gamma\mathcal{R}\Gamma'$, $\psi \in \Gamma'$ iff (induction hypothesis) for all these $\Gamma', \Gamma' \models \psi$ iff (semantic definition of $\Box$) $\Gamma \models \Box\psi$. •

To be able to prove strong completeness for **K** it remains to show the following.

Theorem 19 (The S-consistency theorem) If Φ is **S**-consistent there is a valuation such that for some $\mathcal{F}$, $\models$ and $w \in \mathcal{W}$, $w \models \varphi$ for all $\varphi \in \Phi$.

Proof: Assume Φ to be **S**-consistent. By the Lindenbaum lemma there is a maximal **S**-consistent $\Gamma \supseteq \Phi$. The truth lemma now gives us immediately what we want. •

We are now ready to prove that **K** is strongly complete.

Theorem 20 **K** is strongly complete.

Proof: If $\Phi \not\vdash \psi$, $\Phi \cup \{\neg\psi\}$ is **K**-consistent. Now apply the **S**-consistency theorem to $\Phi \cup \{\neg\psi\}$. •

From the proof it follows also that **K** is strongly complete with respect to $\mathcal{F}_{\mathbf{K}}$. It is good to notice that something one might be tempted to call strong completeness immediately follows for all modal logics stronger than **K**.

Corollary 21 Let **S** be any modal logic and let $\Phi \not\vdash_{\mathbf{S}} \psi$. Then there is a model $\mathcal{M}$ with a world w such that $w \models \Phi$, $w \not\models \psi$

Proof: It is sufficient to realize that **S** as a logic is a set of formulas. This means that, if $\Phi \not\vdash_{\mathbf{S}} \psi$, then $\Phi \cup \mathbf{S} \not\vdash_{\mathbf{K}} \psi$. Now apply the strong completeness theorem for **K**; for the requested model $\mathcal{M}$ one may take $\mathcal{M}_{\mathbf{K}}$. •

One can improve on this by looking at the canonical model for the modal logic **S**.

Theorem 22

(a) $\mathcal{M}_{\mathbf{S}} \models \mathbf{S}$.

(b) If $\Phi \not\vdash_{\mathbf{S}} \psi$, then, for some $w \in \mathcal{M}_{\mathbf{S}}$, $w \models \Phi$, $w \not\models \psi$.

Proof: The proof goes just as in the case of **K**. For (a) just has to note that all maximal **S**-consistent sets have to contain **S**, since otherwise they would contradict **S**. •

This is still not what we want. To really get completeness for **T**, **S4** and **S5**, it is as yet necessary to establish that $\mathcal{F}_{\mathbf{T}} \in$*Char*(**T**), $\mathcal{F}_{\mathbf{S4}} \in$*Char*(**S4**), and $\mathcal{F}_{\mathbf{S5}} \in$*Char*(**S5**); in other words that **T** is valid on the frame underlying $\mathcal{M}_{\mathbf{T}}$ etc. In the case of **K** we did not need this step, since all frames are in *Char*(**K**). The matter is not such a trivial one as it might seem at first sight. That the canonical *model* of a logic **S** itself verifies **S** by no means implies that the underlying *frame* does so for all valuations: the fact that $\mathcal{M}_{\mathbf{S}}$ verifies **S** may well depend on its particular forcing relation. And although for most of the well-known modal logics (**T**, **S4**, **S5** among them) the transition from verification of **S** by $\mathcal{M}_{\mathbf{S}}$ to validity of **S** in $\mathcal{F}_{\mathbf{S}}$ presents no problems, there do exist logics for which it actually fails; we will return to this subject in Section 3.2.

Theorem 23 **T** is strongly complete.

Proof: By the above it suffices to show that $\mathcal{R}_{\mathbf{T}}$ is reflexive, because in that case $\mathcal{F}_{\mathbf{T}} \in Char(\mathbf{T})$. In other words, we have to prove that $\Gamma\mathcal{R}_{\mathbf{T}}\Gamma$ for all maximal **T**-consistent Γ, i.e. if $\Box\varphi \in \Gamma$, then $\varphi \in \Gamma$. Now suppose $\Box\varphi \in \Gamma$. Since $\Box\varphi \vdash_{\mathbf{T}} \varphi$, also $\varphi \in \Gamma$ (see exercise 21). •

Theorem 24 **S4** is strongly complete.

Proof: We have to show that $\mathcal{R}_{\mathbf{S4}}$ is reflexive and transitive. The reflexivity part can be established by the argument used in the case of **T**. Concerning the transitivity of $\mathcal{R}_{\mathbf{S4}}$, suppose $\Gamma_1\mathcal{R}_{\mathbf{S4}}\Gamma_2$ and $\Gamma_2\mathcal{R}_{\mathbf{S4}}\Gamma_3$, for maximal **S4**-consistent Γ_1, Γ_2 and Γ_3. We have to prove that $\Gamma_1\mathcal{R}_{\mathbf{S4}}\Gamma_3$. So, suppose $\Box\varphi \in \Gamma_1$. It is the case that $\Box\varphi \vdash_{\mathbf{S4}} \Box\Box\varphi$, hence (by exercise 21) $\Box\Box\varphi \in \Gamma_1$. The definition of $\mathcal{R}$ in the canonical frame then implies that $\Box\varphi \in \Gamma_2$ and applying the definition of $\mathcal{R}$ once more gives that $\varphi \in \Gamma_3$. •

Theorem 25 **S5** is strongly complete.

Proof: We have to show that $\mathcal{R}_{\mathbf{S5}}$ is reflexive, transitive and symmetric. That the first two properties hold for $\mathcal{R}_{\mathbf{S5}}$ can be established as in the previous cases. It remains to prove that $\mathcal{R}_{\mathbf{S5}}$ is symmetric, that is, that $\Gamma_2 \mathcal{R}_{\mathbf{S5}} \Gamma_1$ whenever $\Gamma_1 \mathcal{R}_{\mathbf{S5}} \Gamma_2$. So suppose $\Gamma_1 \mathcal{R}_{\mathbf{S5}} \Gamma_2$ and $\Box\varphi \in \Gamma_2$ for some φ. Then $\neg\Box\varphi \notin \Gamma_2$ and, because $\Gamma_1 \mathcal{R}_{\mathbf{S5}} \Gamma_2$, $\Box\neg\Box\varphi \notin \Gamma_1$, which means that $\neg\Box\neg\Box\varphi \in \Gamma_1$. By definition $\neg\Box\neg\Box\varphi = \Diamond\Box\varphi$ hence, $\Diamond\Box\varphi \in \Gamma_1$. Finally, $\Diamond\Box\varphi \vdash_{\mathbf{S5}} \varphi$; so, $\varphi \in \Gamma_1$, by exercise 21. •

Again, we have shown also that **T**, **S4**, **S5** are strongly complete with respect to $\mathcal{F}_{\mathbf{T}}$, $\mathcal{F}_{\mathbf{S4}}$, $\mathcal{F}_{\mathbf{S5}}$ respectively. We could continue to prove similar theorems for many other modal logics in a similar manner, but the idea should be clear by now. Therefore, we leave the matter with these proofs and instead retrace our steps to see what we have accomplished so far.

Evaluation of the method H In a way, the method H proves something a little bit stronger than is asked for. We only have to prove that for every pair Φ and φ such that φ is non-derivable from Φ (given a logic **S** for which we are proving strong completeness to hold) there exists some counterexample against the latter entailing the first on a frame validating **S** and not (as we actually did) that there is one model that does the job for every such Φ and φ . The price paid for this strengthening is rather high. This comes to light when we make an effort to determine what the four canonical frames we have encountered so far look like.

It turns out that about the first three very little can be said: we only know that they are very large (we already mentioned how large, viz. $2^{\aleph_0}$, that $\mathcal{R}_{\mathbf{T}}$ of the canonical model for **T** can be proven to be reflexive and $\mathcal{R}_{\mathbf{S4}}$ in the canonical model for **S4** can be proven to be transitive also; about $\mathcal{R}_{\mathbf{K}}$ we know nothing at all. But it is clear that $\mathcal{R}_{\mathbf{T}}$ has a lot more structure than being a reflexive relation implies and $\mathcal{R}_{\mathbf{S4}}$ is not only reflexive and transitive: because of the way we defined the accessibility relation on the canonical world sets and on account of how the contents of the maximal consistent sets come out there are a lot of connections laid by the $\mathcal{R}$'s we cannot say much about. Both $\mathcal{R}_{\mathbf{T}}$ and $\mathcal{R}_{\mathbf{S4}}$ are rather complicated relations.

The **S5**-frame is more regular, hence we can say a little bit more about it; in order to show how to get more perspicuous results from the method H, we are going to examine this frame a little bit closer.

Examination of $\mathcal{F}_{\mathbf{S5}}$ The accessibility relation of $\mathcal{F}_{\mathbf{S5}}$ is of course an equivalence relation (i.e. it is reflexive, symmetric and transitive) and therefore it partitions the set of all maximal **S5**-consistent sets into equivalence classes (i.e. "pieces" of which all the elements are mutually connected in all possible ways, while no connection exists between the elements of two different subsets). Hence, on every equivalence class $\mathcal{R}_{\mathbf{S5}}$ is universal. Of these equivalence classes there are $2^{\aleph_0}$ many, as can be seen in the following way:

Because of the properties of $\mathcal{R}_{\mathbf{S5}}$ every piece is completely determined by the formulas of the form $\Box\varphi$ contained in the worlds that make it up, because in each piece exactly the worlds that make up that piece are visible from each world. Now consider the sets $\Delta_M = \{\Box p_m \mid m \in M\} \cup \{\neg\Box p_n \mid n \notin M\}$, for $M \subseteq \mathbb{N}$.

Clearly these sets are **S5**-consistent, so each Δ_M has a maximal **S5**-consistent extension Γ_M. Furthermore, because we know already that the worlds in one and the same piece have to contain the same $\Box$ -formulas, we know that $\Gamma_M \mathcal{R}_{\mathbf{S5}} \Gamma_{M'}$ is impossible whenever $M \neq M'$. This observation, combined with the fact there are $2^{\aleph_0}$ different subsets of $\mathbb{N}$ guarantees the existence of exactly $2^{\aleph_0}$ different equivalence classes. Actually from this examination it follows that there is a simply described frame with respect to which **S5** is strongly complete.

Proposition 26 Let $\mathcal{F} = \langle \mathcal{W}, \mathcal{R} \rangle$ be such that $|\mathcal{W}| = 2^{\aleph_0}$ and $\mathcal{R}$ is universal on $\mathcal{W}$, then **S5** is strongly complete with respect to $\mathcal{F}$.

Proof: All pieces of $\mathcal{F}_{\mathbf{S5}}$ are p-morphic images of $\mathcal{F}$ (exercise 25). Moreover, clearly $\mathcal{F}_{\mathbf{S5}}$ is the disjoint union of its pieces (we have not defined infinite disjoint unions, but there is no obstacle to doing this). This means that any counterexample on $\mathcal{F}_{\mathbf{S5}}$ can be rebuilt on $\mathcal{F}$. •

This result is the best we will get by way of the method H and only because we are lucky to have a comparably surveyable model as in the case of **S5**. We can get such nicer results from the canonical models for other logics too, but, as this entails some construction work anyway (even in the case of **S5**) we may just as well start constructing immediately. Besides this lack of transparency of the resulting models the method H simply does not always work: not every logic that is in fact (strongly) complete can be proven to be so by means of this method. Although most of the well-known modal logics do not suffer from this deficiency there are logics **S** whose canonical frame is not in *Char*(**S**).

These shortcomings justify the existence of another method of proof for establishing the strong completeness theorem which we will now describe.

3.1.2 The construction method C

The models acquired by means of the method C share with the canonical models of method H that they are built using maximal consistent sets. But the way in which the frames come into being is quite different: we may not use all maximal consistent sets of formulas as worlds, and we may use the same maximal consistent sets of formulas as more than one world. The second difference entails that we actually have to use different entities than the maximal consistent sets themselves as the worlds, because one entity cannot be two or more. It is best to think then of each of these new entities as being accompanied by an adjoined maximal consistent set. The point is now that we are relatively free in how to connect these worlds by means of an accessibility relation; the only obvious restrictions to this freedom are due to our intentions to make the construction validate the logic in question and respect the standard semantics of $\Box$

The effects of the liberties we permit ourselves in the C-case are twofold:

Diminishing the number of worlds that are included in a C-model for the logic **S** from all possible ones to some restricted number, has as a consequence that the models we obtain no longer can be taken to suffice for *every* Φ and *every* φ not derivable from Φ in **S**: we can only be sure that the model we construct falsifies φ while verifying all of Φ for the *one* pair Φ and φ we start with in each case.

This implies that in order to prove a logic to be strongly complete by means of the method C we have to use a whole set of models, whereas in the case of the method H the one canonical model by itself settled the proof. Sometimes we can make do with one frame however.

A second consequence is that we will not be able to delineate the class of models that can be obtained by applying the method C in such a clear cut fashion as we can delineate the class of canonical models, obtained by the method H; this is among other things caused by the fact that there are many different ways (even for one and the same logic) of constructing models.

The advantages of the method however outweigh these little imperfections by far, for by using it we gain much smaller models (at most countably large) and an accessibility relation that can be defined as orderly as the logical system in question allows. The best way to clarify these remarks is by inspecting an example of the method 'in action'; for the purpose of illustration we have picked out **S4**. In the following we will concentrate once again on proving this logic to be strongly complete.

Definition 13 The *infinite infinitely branching tree* $\mathcal{T}_{\omega\omega}$ is the frame $\langle \mathcal{W}, \mathcal{R} \rangle$, where $\mathcal{W}$ is the set of all finite sequences of natural numbers (including the empty sequence $\langle\rangle$) and $\sigma\mathcal{R}\tau$ iff σ is a (not necessarily proper) initial segment of τ (we write somewhat improperly $\sigma \subseteq \tau$). The elements of such a tree are called *nodes*.

New proof of proposition 24 .

Suppose $\Phi \not\vdash_{\mathbf{S4}} \varphi$. We will construct a counterexample on $\mathcal{T}_{\omega\omega}$. The set $\Phi \cup \{\neg\psi\}$ is a **S4**-consistent set, and so it has to have some maximal **S4**-consistent extension $\Gamma_{\langle\rangle}$. We will connect $\Gamma_{\langle\rangle}$ to the node $\langle\rangle$ and make that node into a world falsifying $\Phi \models_{Char(\mathbf{S4})} \psi$, by constructing a **S4**-model departing from it in the form of $\mathcal{T}_{\omega\omega}$. The first thing to note is that $\mathcal{T}_{\omega\omega} \in Char(\mathbf{S4})$, because it is reflexive and transitive. For the remainder of the construction we have to keep two things in mind:

Because we want to use maximal **S4**-consistent sets again as valuations in the same manner as before (i.e.: setting $\Gamma \models p$ iff $p \in \Gamma$, automatically has to lead to $\Gamma \models \varphi$ iff $\varphi \in \Gamma$ for all formulas φ), we first have to take care that for every formula of the form $\Box\chi$ occurring in the set Γ adjoined to a node γ, χ will occur in each set Δ adjoined to any node δ such that $\gamma \subset \delta$. Secondly we have to take care that for each formula $\neg\Box\chi$ occurring in some Γ adjoined to some γ there will be a Δ adjoined to some $\delta \supset \gamma$ such that $\neg\chi \in \Delta$; this in order to make the model-under-construction satisfy the standard semantics of $\Box$ (maximal consistency takes care of the standard semantics of the other logical connectives).

Since our further steps in the process of constructing are going to depend on the elements of $\Gamma_{\langle\rangle}$, it is convenient to fix an enumeration of its formulas: $\Gamma_{\langle\rangle} = \{\varphi_1, \varphi_2, \ldots\}$. We will now say what $\Gamma_{\langle n\rangle}$ is going to be depending on φ_n and its relationship to $\Gamma_{\langle\rangle}$. The procedure is as follows:

Check whether φ_n is the form $\neg\Box\chi$. If it is not, set $\Gamma_{\langle n\rangle} = \Gamma_{\langle\rangle}$. If it is, check whether $\neg\chi \in \Gamma_{\langle\rangle}$. If it is, again set $\Gamma_{\langle n\rangle} = \Gamma_{\langle\rangle}$. If it is not, introduce a maximal **S4**-consistent extension $\Gamma_{\langle n\rangle}$ of the set $\{\Box\psi \mid \Box\psi \in \Gamma_{\langle\rangle}\} \cup \{\neg\chi\}$ (the proof the latter set is **S4**-consistent, so that the existence of such a $\Gamma_{\langle n\rangle}$ can be assumed, we leave as exercise 24). This procedure of fixing an enumeration and checking for each formula whether it entails a need for action can be repeated for every world $\Gamma_{\langle i_1,\ldots,i_k\rangle}$ that gets introduced along the way, resulting in a countably infinite set of worlds.

It is obvious from the construction that if we define $\langle i_1, \ldots, i_k\rangle \models p$ iff $p \in \Gamma_{\langle i_1,\ldots,i_k\rangle}$, this extends to $\langle i_1, \ldots, i_k\rangle \models \varphi$ iff $\varphi \in \Gamma_{\langle i_1,\ldots,i_k\rangle}$ and hence $\langle\rangle \models \Phi$, $\langle\rangle \not\models \varphi$.

We conclude the first part of this chapter with a final remark on the adjective 'constructive' as we used it here: despite our emphasis on the fact that we are constructing the models acquired by the C-method, this 'constructiveness' should only be taken in a loose sense; it is not to be taken in the strict metamathematical one, if only because the existence of the maximal consistent sets themselves (the Lindenbaum lemma) requires non-constructive methods of proof. We called the method the construction method after the way in which the models are built up intuitively.

3.2 Exercises

Exercise 20 Deduce from the strong soundness theorem for **K** that **T**, **S4** and **S5** are sound.

Exercise 21 Show that, if Γ is maximal **S**-consistent and $\Gamma \vdash_{\mathbf{S}} \varphi$, then $\varphi \in \Gamma$.

Exercise 22 Prove the 'classical part' of the valuation lemma.

Exercise 23 Show that
(a) if $\vdash_{\mathbf{K}} \Box\varphi$, then $\vdash_{\mathbf{K}} \varphi$;
(b) if $\vdash_{\mathbf{K}} \Box\varphi \vee \Box\psi$, then $\vdash_{\mathbf{K}} \Box\varphi$ or $\vdash_{\mathbf{K}} \Box\psi$.

Exercise 24 Prove that if $\Gamma \cup \{\neg\Box\varphi\}$ is **S4**-consistent, then so is $\{\Box\varphi \mid \Box\varphi \in \Gamma\} \cup \{\neg\varphi\}$.

Exercise 25
(a) Prove that the logic $\mathbf{K}+\Diamond\top$ is strongly complete with respect to the irreflexive, intransitive, infinite and infinitely branching tree (give a definition yourself). Explain the rôle the additional axiom $\Diamond\top$ plays in the proof.
(b) Give, by means of (a), a proof of the fact that irreflexiveness + intransitivity + asymmetry is not expressible in modal formulas.

Exercise 26 Show that, if $\mathcal{R}_1$ and $\mathcal{R}_2$ are universal relations on $\mathcal{W}_1$ and $\mathcal{W}_2$ respectively, then $\langle \mathcal{W}_2, \mathcal{R}_2 \rangle$ is a p-morphic image of $\langle \mathcal{W}_1, \mathcal{R}_1 \rangle$ iff $|\mathcal{W}_1| \geq |\mathcal{W}_2|$.

3.3 Decidability

In order to understand what the decision problem for a logic is about, it is advisable to glance back at the beginning of chapter 1, where we discussed (modal) logics for the first time. In the course of our exposition there, we mentioned two different characterizations of the kind of entity a logic can be taken to be. a logic can be identified (1) with a number of axioms plus some derivation rules, or (2) with a set of formulas. Every axiom system, i.e. a specification of a logic in the first sense generates a set of theorems that can be derived, i.e. a specification of a logic in the second sense. If for a logic we can determine membership of this set effectively (i.e. by means of a mechanic procedure that can at least in principle actually be carried out) the logic in question is said to be decidable.

Definition 14 A logic **S** is *decidable* iff an effective procedure exists by means of which for every formula φ in the language it can be settled whether $\varphi \in \mathbf{S}$. A logic **S** is *effectively axiomatizable* iff there is a decidable set of axioms for **S** and a decidable set of rules (the latter meaning that there is an effective procedure by means of which, for any one of the rules, any finite sequence Γ of formulas and any formula φ , it can be settled whether φ is derivable from Γ.

In the sequel we will sketch how such procedure can be found for an effectively axiomatizable logic of any kind, provided only that it is sound with respect to some semantics and countermodels can always taken to be finite..

Definition 15 A modal logic **S** has the *finite model property* (henceforth often noted as FMP for short) iff for every φ such that $\nvdash_S \varphi$ there is some model $calM = \langle calW, \mathcal{R}, \models\rangle$ such that

(i) $\mathcal{W}$ is finite;
(ii) $\mathcal{M}$ verifies S;
(iii) $\mathcal{M}$ falsifies $|varphi$.

Proposition 27 Let **S** be effectively axiomatizable, sound, and complete. If **S** has the FMP with respect to some semantics, then **S** is decidable.

Prior to proving this proposition, we have to make two remarks on the phrase 'complete' as it occurs in our formulation of this proposition.

First, from now on in this chapter we are done with strong completeness; in this section the term is only meant to refer to plain completeness.

Second, further down we will prove that every modal logic **S** with the FMP is complete. This entailment is far from being obvious; therefore we postpone the proof to a more suitable moment, and give here the general proof which does not at all depend on the logic being a *modal* logic.

A final remark concerns the decision procedure outlined. It is very unhandy, but that derives from the fact that its existence is proved from very general assumptions. In practical applications to particular logics one can be fairly sure that a less unwieldy decision procedure will exist.

Proof: There must be a derivation for every φ in **S**. Because **S** is effectively axiomatizable, all its derivations can be enumerated in some definite order. The fact that derivations are finite objects should suffice to convince you of the effectiveness of this enumeration. The result will be a fixed sequence $\Pi_1, \Pi_2, \Pi_3, \ldots$ of all the derivations **S** admits. Each derivation has as its conclusion a formula φ_i. Furthermore, because **S** is complete, for every φ that cannot be derived in **S**, there must be some model that can serve as a counterexample to its universal validity in **S**-models. Because **S** has the FMP all these counterexamples can be taken to be among the set of all finite models for **S**. This set can also be enumerated in some determinate way; the result will be some fixed sequence $\mathcal{M}_1, \mathcal{M}_2, \mathcal{M}_3, \ldots$, this time of all finite **S**-models. (See the remark in parentheses in definition 15; that this is an effective operation is obvious in the case of modal logics, since the property characterized by the logic will be decidable for finite frames.) To decide, for arbitrary formula φ of the language whether or not $\varphi \in$**S**, we can now alternately check Π_i and $\mathcal{M}_i$ on being respectively a derivation of φ or a counterexample for it. For the effectiveness of this checking-procedure the FMP is essential, for only if the models are finite, perusing them takes only a finite amount of time. If we continue this search long enough, sooner or later we have to run into either some Π_i or some $\mathcal{M}_i$ that relates to the formula φ in question; in the first case φ is a theorem and in the second it is not. •

We focus again on the four logics that are paramount in this chapter, namely **K**, **T**, **S4** and **S5**. We now set ourselves to proving them to be decidable. By now, we have already learnt that this task can be reduced to proving them to have the FMP; the other conditions mentioned in proposition 27 are already known to hold for the logics under consideration. Remember that in the introductory remarks of the present chapter, we announced our intention to use finite variants of the methods H and C in establishing decidability. At this moment it must be clear where these come in: showing the FMP to hold for some logic **S** comes down to finding a set of finite models, one for each φ that can be falsified in some model for **S**, satisfying certain conditions concerning the truth and falsity of certain formulas. Except for the fact that the models they deliver are not of the right size, the methods H and C are perfectly suitable. The manner in which we bring about that the models constructed now make the grade with respect to the demand concerning finiteness consists in setting bounds to the number of formulas that we consider. For the purposes we have in mind, it will turn out we can in each particular case narrow our attention down to a finite number of formulas; the only formulas we then have to take care of in order to make the models do what we want them to do are the subformulas (in this circumstance also of course finite in number) of these. In this respect our proceeding shows a resemblance to the current way of defining models for ordinary propositional logic by means of truth tables. This being clarified, it is clear our needs necessitate a concept of the following kind:

Definition 16 A set of formulas Φ is *closed* iff
(i) If $\varphi \in \Phi$ and ψ is a subformula of φ , then $\psi \in \Phi$;
(ii) if $\varphi \in \Phi$ and φ itself is not a negation, then $\neg\varphi \in \Phi$.
The result of closing off a set of formulas Φ under taking subformulas and single negations is called *the closure of* Φ.

In order to given an idea of what the closure of a set looks like, we give the following example.

Example Let Φ be the set $\{(p \to \neg r), \neg(p \wedge q)\}$.
Its closure is the set $\{p, q, r, \neg p, \neg q, \neg r, p \to \neg r, \neg(p \to \neg r), p \wedge q, \neg(p \wedge q)\}$

Note that the closure of a finite set is finite; since this is what we want, we did not in the second clause simply take closure under $\neg$. (Condition (ii) of definition 16 is not directly needed, but is put in for reasons of symmetry: if $\Gamma \subseteq \Phi$, $\varphi \in \Phi$ and $\Gamma \not\models \varphi$, one likes to have a maximal consistent $\Gamma' \supseteq \Gamma$ which does not contain Γ, and the existence of such a Γ' would not be guaranteed without (ii).)

All the machinery needed to be able to expose the finite variant of the H-method, H_{fin}, are now at our disposal, so let us turn to explaining it right away.

3.3.1 The method $\mathbf{H}_{fin}$

We will employ the pure form of this method only for the modal logics **K** and **T**, for reasons to be discussed later on. First we will outline how it works.

H_{fin} only differs from its infinitary sister in that the counterparts of the canonical models it brings about stay within the limits marked out by some finite set. The issue behind this strategy is, as said before, to admit the forming of only finitely many maximal consistent sets. Taking these, in the same way as before, to be the set of worlds of the models we define, guarantees the models to be finite too, as required. We had better first of all redefine, so as to adapt them to the finite case under consideration, the notions, lemmata and propositions that adhere to the method H.

Definition 17 Let Φ and Γ be sets of formulas with $\Gamma \subseteq \Phi$. Γ is *maximal* **S**-*consistent in* Φ iff Γ is **S**-consistent and for no **S**-consistent $\Gamma' \subseteq \Phi$, it holds that $\Gamma \subset \Gamma'$.

Lemma 28 (The Lindenbaümchen lemma for S) If Φ is closed and $\Gamma \subseteq \Phi$ is **S**-consistent, then Γ has a maximal **S**-consistent extension Γ' in Φ.

Proof: Like usual. •

Lemma 29 (The finite valuation lemma for S) If a finite set of formulas Φ is closed and Γ is maximal **S**-consistent in Φ, then for all $\varphi \in \Phi$ it holds that

(i) if $\neg\varphi \in \Phi$, then $\neg\varphi \in \Gamma$ iff $\varphi \notin \Gamma$ and
(ii) if $\varphi \vee \psi \in \Phi$, then $\varphi \vee \psi \in \Gamma$ iff $\varphi \in \Gamma$ or $\psi \in \Gamma$.
(iii) If $\Box\varphi \in \Phi$, then $\Box\varphi \in \Gamma$ iff $\varphi \in \Gamma'$ for all Γ' with $\Gamma\mathcal{R}\Gamma'$.

Proof: Like usual. But note that the antecedents of (i), (ii) and (iii) cannot be left out, since finite closed sets are only closed in the downward, and not in the upward direction (with the exception of closure under single negations of course); i.e. we can never be certain whether a formula with a higher degree of complexity than some φ is among its elements and we can be sure, on account of the finiteness of Φ, that many of them certainly are not. •

Proposition 30 (Finite consistency theorem for S) If Φ is finite and **S**-consistent, there is a valuation such that for some finite $\mathcal{F}$, $\models$ and $w \in \mathcal{W}$, $w \models \varphi$ for all $\varphi \in \Phi$.

Proof: Easy. •

We are now in a position to prove the following:

Theorem 31 **K** has the finite model property.

Proof: Suppose $\not\vdash_{\mathbf{K}} \varphi$, i.e. $\{\neg\varphi\}$ is **K**-consistent. Take Φ to be the closure of $\{\neg\varphi\}$. According to the Lindenbaümchen lemma we may assume there is a $\Gamma \supseteq \{\neg\varphi\}$ that is maximal **K**-consistent in Φ. Defining the model $\mathcal{M}^{\Phi}_{\mathbf{K}}$ in much the same way as the canonical model for **K** before, by setting $\mathcal{W}^{\Phi}_{\mathbf{K}} = \{\Gamma \mid \Gamma \text{ is maximal } \mathbf{K}\text{-consistent in } \Phi\}$ and defining its $\mathcal{R}^{\Phi}_{\mathbf{K}}$ and $\models$ completely similarly, it can easily be checked that $\mathcal{M}^{\Phi}_{\mathbf{K}}$ is a model of the kind we were looking for because it is finite and $\Gamma \not\models \varphi$. •

Needless to say perhaps (we only mention this fact to contrast the present case with the one of the real canonical models) that in the proof that **K** has the FMP we obtain a whole set of models; little reflection is needed to see that different Φ's will generally get us different models.

By the way, the analogue of corollary 21 does not apply here: that **K** has the FMP does not entail that all modal logics have some property like that. This comes to light if we try to reconstruct the argument given: from $\neg\varphi$ being **S**-consistent we do get that $\mathbf{S} \cup \{\neg\varphi\}$ is **K**-consistent, but after that the reasoning gets stuck, for this set (and a fortiori its closure) does not have to be finite. We can only be certain that it is finite in case **S** is finitely axiomatizable over **K** (i.e. can be obtained from K by adding finitely many axioms). Almost all logics however lack this property; by adding only one axiom scheme already infinitely many axioms are subjoined! Exceptions are $\mathbf{K}+\ \Box\bot$ and $\mathbf{K}+\Diamond\top$, and for these logics the argument does go through completely: they do have the FMP by the above proof.

3.3.2 The method $\mathbf{C}_{fin}$

This time the defect that justifies the existence of a method C_{fin} comes from a different quarter than the reason we had in the infinitary case for the introduction of a method C, viz. perspicuity of the models acquired, and it crops up even sooner: whereas the H-method could still be applied successfully and without presenting any problems to **S4** and **S5**, presently matters stand differently. The source of the difficulty can easily be located. Consider for example the characteristic axiom scheme $\Box\varphi \to \Box\Box\varphi$ of **S4**. If we want H_{fin} to deliver a model verifying **S4**, i.e. a transitive one, we must, looking back at the proof of theorem 24, be able to rely on the fact that for every ψ, $\Box\Box\psi$ is in every maximal **S4**-consistent set that contains $\Box\psi$. But given our definition of the closure of a set $\Box\Box\psi$ might very well not be in the closed set Φ within which we are working and therefore we have no hold on its truth value. This deficiency can be resolved in a fashion similar to the one we used in connection with single negations, or also by tampering with the relation $\mathcal{R}$; we leave it to your imagination to spell out precisely how this can be done, in exercise 29 (see also chapter 4 on provability logic). But we have to arrange things properly for each logic requiring similar adjustments separately and cannot cope with this difficulty for all such logics at the same time. As this will mean going into some construction work anyway, we decided to develop the finite C-variant here.

The models we are about to construct resemble their infinitary analogues less than in the case of H. We have to make accommodations on some essential points, so it is best to spell out again most of the details from the beginning. Our example again is **S4**.

Theorem 32 **S4** has the finite model property.

Proof: Suppose there is a model verifying **S4** while falsifying some formula φ . We are going to construct a finite model with the same properties. $\{\neg\varphi\}$ clearly is **S**4-consistent and therefore it has a maximal consistent extension Γ in the closure Φ of $\{\neg\varphi\}$. We again take this Γ as our starting point out of which a **S4**-model is to be generated, although this time of course we intend to stay within the boundaries of Φ. Again, we introduce maximal **S4**-consistent sets in proportion to the elements of Γ and for the same reason: all formulas of the form $\neg\Box\chi$ in Γ necessitate the existence of a Γ' such that $\Gamma\mathcal{R}\Gamma'$ and $\neg\chi \in \Gamma'$. This time we do not use these sets merely as valuations, but really identify the worlds of the model-in-construction with them. Acting thus cannot present the same problems it did before, because now we are going to be sparse with introducing new worlds, doing so only if we cannot get away from it. This means that, whenever in the previous construction with method C we had reasons to produce identical copies of a set Γ' we already had, we do nothing at all here except relating the Γ' which is already there by $\mathcal{R}$ to the world we are operating from. (This will mean that the relation $\mathcal{R}$ constructed will not be antisymmetric and we will not get a partial order, but (exercise 27) that is unattainable in combination with finiteness for **S4** anyway). Still another disparity is that, whenever we cannot escape introducing a new world from any Δ along the way, we take it to be a maximal consistent extension in Φ, not of $\{\psi \mid \Box\psi \in \Delta\} \cup \{\neg\chi\}$ as before, but of $\{\Box\psi \mid \Box\psi \in \Delta\} \cup \{\neg\chi\}$. To show that this works we have to prove that, if $\{\Box\psi \mid \Box\psi \in \Gamma\} \cup \{\neg\Box\chi\}$ is **S4**-consistent, then so is $\{\Box\psi \mid \Box\psi \in \Gamma\} \cup \{\neg\chi\}$ (exercise 29). Thus, we manage to get round the difficulty concerning transitivity that ended H_{fin} in a deadlock. The accessibility relation on the sets obtained is simply defined to be reflexive and transitive and $\models$ is defined as in canonical models by membership. We have to check that by this indeed $\varphi \in \Gamma$ if and only if $\Gamma \models \varphi$ for all φ . A little thought will reveal that the only points to be checked are where we force the relation to be reflexive and transitive:

With regard to reflexivity this means again that we have to show that, if $\Box\varphi \in \Gamma$, then $\varphi \in \Gamma$ which goes in the same manner as with **T**. With regard to transitivity we have to check that, if $\Box\varphi \in \Gamma$, $\Gamma\mathcal{R}\Gamma'$ and $\Gamma'\mathcal{R}\Gamma$", then $\varphi \in \Gamma$". This is the reason $\mathcal{R}$ was defined differently: if $\Box\varphi \in \Gamma$, then $\Box\varphi \in \Gamma'$, and hence $\Box\varphi \in \Gamma$" and indeed $\varphi \in \Gamma$". •

A similar proof procedure goes for the analogous theorem in relation to **S5**:

Proposition 33 **S**5 has the finite model property.

Proof: This is exercise 31. •

One final remark on these methods of proof: A proof as just given of a logic having the FMP can easily be transformed in a proof of weak completeness for the same, by viewing the models obtained 'through a different pair of spectacles' so to speak. Plain completeness can be shown to hold similarly, by reasoning with the closure of $\{\psi_1, \ldots, \psi_n\} \cup \{\neg\varphi\}$ for $\psi_1, \ldots, \psi_n, \varphi$ such that $\psi_1, \ldots, \psi_n \not\vdash \varphi$.

3.4 The FMP implies completeness

In the previous section we claimed that the FMP entails completeness; a claim for which we still owe the proof. This takes quite some doing, but we will take all the trouble, if only because the preparatory work is of interest for its own sake. We start by marking off a special class of models.

Definition 18 A model $\mathcal{M} = \langle \mathcal{W}, \mathcal{R}, \models \rangle$ is *distinguishable* iff, for all $w, w' \in \mathcal{W}$ with $w \neq w'$, there is a formula φ such that $w \models \varphi$ and $w' \not\models \varphi$.

Any model can be transformed into a distinguishable equivalent.

Proposition 34 For every model $\mathcal{M}$ there is a distinguishable $\mathcal{M}_d$ such that $|\mathcal{M}_d| \leq |\mathcal{M}|$ and $\mathcal{M} \models \varphi$ iff $\mathcal{M}_d \models \varphi$, for all formulas φ.

Proof: Let $\mathcal{M}$ be a model. We will construct a distinguishable equivalent $\mathcal{M}_d = \langle \mathcal{W}_d, \mathcal{R}_d, \models_d \rangle$. Obviously, $\mathcal{M}$ can contain a number of worlds that, because they validate exactly the same formulas, function as obstacles to it being a distinghuishable model in the first place. We can get rid of these by drawing them together into equivalence classes, brought about by the relation $\cong_d$, defined on $\mathcal{W}$ in the following manner:

$w \cong_d w'$ iff $\{\varphi \mid w \models \varphi\} = \{\varphi \mid w' \models \varphi\}$

Since $\cong_d$ clearly is an equivalence relation, it is proper to define $\mathcal{W}_d = \{w_d \mid w \in \mathcal{W}\}$, where w_d is the equivalence class of w under $\cong_d$. On this set $\mathcal{W}_d$ we want to define the accessibility relation $\mathcal{R}_d$ in such a manner that, setting $w_d \models_d p$ iff $w \models p$ for all propositional letters p, the truth of all formulas in $\mathcal{M}$ is preserved. Simply stipulating $w_d \mathcal{R}_d w'_d$) iff there are w'' and w''' in $\mathcal{W}$ such that $w \cong_d w''$, $w' \cong_d w'''$ and $w'' \mathcal{R} w'''$, does the trick. That the models $\mathcal{M}$ and $\mathcal{M}_d$ are equivalent, i.e. that $w_d \models_d \varphi$ iff $w \models \varphi$ for all φ , can be shown by a simple induction on the complexity of φ . The only step interesting enough to be looked at is $\varphi = \Box\varphi$:

$\Leftarrow$: Suppose $w \not\models \Box\psi$. Then for some w' with $w \mathcal{R} w'$ it holds that $w' \not\models \psi$. The induction hypothesis gives $w'_d \not\models_d \psi$. The definition of $\mathcal{R}_d$ entails that $w_d \mathcal{R}_d w'_d$, and so $w_d \not\models_d \Box\psi$.

$\Rightarrow$: Now suppose $w_d \not\models_d \Box\psi$. Then $w'_d \not\models_d \psi$, for some w'_d such that $w_d \mathcal{R}_d w'_d$. By the definition of $\mathcal{R}_d$ there must be w'' and w''' such that $w \cong_d w'$, $w' \cong_d w'''$ and $w'' \mathcal{R} w'''$. By the induction hypothesis we have $w''' \not\models \psi$, hence $w'' \not\models \Box\psi$. Since w and w'' are in the same equivalence class, $w \not\models \Box\psi$ also holds. •

The model obtained in the proof of proposition 34 is called a *filtration* of the original one.

Distinguishable models are called so after the fact that to every ontologically distinct entity (i.e. world) in it, there corresponds a linguistically distinct entity, viz. a set of formulas that in its entirety is true in this world only. If such a model in addition is finite also, this characteristic can be strengthened to the following rather surprising result.

Proposition 35 Let $\mathcal{M} = \langle \mathcal{W}, \mathcal{R}, \models \rangle$ be a distinguishable and finite model. Then for every $w \in \mathcal{W}$ there is a single formula φ such that $w \models \varphi$, but for no other $w' \in \mathcal{W}$ $w' \models \varphi$.

Proof: Assume $\mathcal{W} = \{w_1, \ldots, w_n\}$. As $\mathcal{M}$ is distinguishable, we can for every i, j with $1 \leq i < j \leq n$ choose a formula φ_{i_j} such that $w_i \models \varphi_{i_j}$ and $w_j \not\models \varphi_{i_j}$. Additionally, we set $\varphi_{i_i} = \top$, for reasons of convenience in stating the following: it is easy to see that $w' \models \varphi_{i_1} \wedge \varphi_{i_2} \wedge \ldots \wedge \varphi_{i_n}$ iff $w' = w_i$. •

Note that the finiteness of $\mathcal{M}$ plays an essential rôle in the proof. We would need infinite conjunctions — and we don't have them in the languages we are working with — for this argument to go through for infinite models. With the help of the next proposition, we can prove the key theorem of this section.

Proposition 36 Let $\mathcal{M} = \langle \mathcal{W}, \mathcal{R}, \models \rangle$ be a finite, distinguishable model that verifies the modal logic **S**. Then **S** is valid on the underlying frame $\mathcal{F} = \langle \mathcal{W}, \mathcal{R} \rangle$.

Proof: Assume $\mathcal{F} \not\models \varphi$ for some $\varphi \in$**S**. Then there must be some $w \in \mathcal{W}$ and $\models'$ on $\mathcal{F}$ such that $w \not\models' \varphi$. We are going to specify a substitutional instance φ^* of φ such that $w \not\models \varphi^*$. We then have a contradiction, since $\varphi^* \in$**S**, because by definition **S** is closed under substitution and $\models$ is assumed to verify **S** on $\mathcal{F}$. The instance needed comes about in the following way: Consider for every atomic formula p the set $[p] = \{w \in \mathcal{W} \mid w \models' p\}$. According to proposition 35 to all $w' \in [p]$ there belongs some $\varphi_{w'}$ exclusively true in w'. We define p^* as the disjunction of all $\varphi_{w'}$ for $w' \in [p]$. We can then alternatively specify $[p]$ as the set $\{w \in \mathcal{W} \mid w \models p^*\}$; i.e. with respect to $\langle \mathcal{F}, \models \rangle p^*$ has the same truth conditions as p has with regard to $\langle \mathcal{F}, \models' \rangle$. By a trivial induction we omit (see exercise 9 though) it has to hold good that, for all p, substituting p^* for p in every formula φ containing it (to be precise: define the translation operation * on formulas φ as: p^* as specified above, $(\neg\varphi)^* = \neg\varphi^*$, $(\varphi \vee \chi)^* = (\varphi^* \vee \chi^*)$, etc.) yields formulas φ^* such that $w \models' \varphi$ iff $w \models \varphi^*$. Hence, $w \not\models' \varphi$ implies $w \not\models \varphi^*$ and this is what we had to prove. •

Theorem 37 All modal logics with the FMP are complete.

Proof: Let **S** be a logic having FMP. Connect to every φ such that $\not\vdash_{\mathbf{S}} \varphi$ a finite distinguishable model $\mathcal{M}_\varphi$ that verifies **S** and falsifies φ . We can take for granted that such a model exists because the fact that **S** has the FMP assures us of the possibility to find a finite model with the required characteristics, whereas proposition 34 then implies that we can safely take the model to be distinguishable. Let $\mathcal{C}$ be the set that contains one such model $\mathcal{M}_\varphi$ for every φ that is not a theorem of **S**. The previous proposition tells us that we are allowed to make the transition from the fact each of these models verifies **S** to the fact that their underlying frames validate **S** also. Hence, $\mathcal{C} \subseteq Char(\mathbf{S})$ and obviously **S** is complete with respect to $\mathcal{C}$. •

This theorem has a nice corollary; it enables us in some cases to prove for a whole set of logics that they are complete, an example of which is in the following:

Theorem 38 All modal logics **S** with **S5**$\subseteq$**S** are complete.(See exercise 32 to learn what such logics can look like.)

Proof: It suffices to show such logics to have the FMP. Assume **S5**$\subseteq$**S**, i.e. **S** is of the form $\mathbf{S5} \cup \Sigma$ for some set of axioms Σ. If $\not\vdash_{\mathbf{S} \cup \Sigma} \varphi$, then $\Sigma \not\vdash_{\mathbf{S}} \varphi$. Since **S5** is strongly complete with respect to the class of frames with universal $\mathcal{R}$, there must be a $\models$ on a $\mathcal{F}$ in this class such that for certain $w \in \mathcal{W}$, $w \not\models \varphi$. $\mathcal{F}$ does of course not have to be finite itself, but we can rebuild it into a model $\mathcal{M}' = \langle \mathcal{F}', \models' \rangle\rangle$ on a finite frame $\mathcal{F}' = \langle \mathcal{W}', \mathcal{R}' \rangle$, in a way resembling the one we used in connection with obtaining distinguishable models. For this purpose we proceed in the following manner: Define the equivalence relation $\cong$ on $\mathcal{W}$ such that for all $w, w' \in \mathcal{W}$: $w \cong w'$ iff for all p occurring in φ: $w \models p$ iff $w' \models p$. Write $w_\cong$ for the equivalence class of w and take $\mathcal{W}'$ to be the set $\{w_\cong \mid w \in \mathcal{W}\}$. Since φ can contain only finitely many p's, this set has to be finite. Further, take $\mathcal{R}'$ to be the universal relation on this set and chose $\models'$ such that for all p, $w_\cong \models' p$ iff $w \models p$. A trivial induction on the complexity of φ shows that that $w_\cong \models' \varphi$ iff $w \models \varphi$, for all formulas φ, whence $w_\cong \not\models' \varphi$ and so $\mathcal{M}' \not\models \varphi$. •

3.5 On completeness, canonicity and the FMP

A large portion of the present chapter has been devoted to completeness proofs: for a number of particular modal systems we established that they are strongly complete or (via the detour of the FMP plus the remark at the very end of section 3.2) plainly complete with respect to their characteristic classes. In all the cases we treated we furthermore managed to single out proper subclasses of the characteristic class, the consideration of which suffices for completeness. In relation to plain completeness, this subclass consisted of all the finite frames in *Char* (**S**); the possibility to restrict oneself to those only, we called 'having the FMP'. And in relation to strong completeness, this subclass consisted of the singleton $\{\mathcal{F}_{\mathbf{S}}\}$ for the **S** in question. Of course, it is of interest whether the latter is possible for a logic.

Definition 19 A logic **S** is *canonical* iff $\mathcal{F}_{\mathbf{S}} \in Char(\mathbf{S})$.

None of the properties mentioned apply to all modal logics. Also, few implications between the different properties under consideration exist. The situation can be pictured as in the figure below. The diagram is supposed to represent all modal logics (and, anticipating the course of our exposition, all tense logics can be included as well), the shaded areas are empty and the small letters in the picture refer to our discussion below.

INSERT FIGURE 13

Ad (a). In 1974 the first incomplete modal logics were discovered by K. Fine and S.K. Thomason. It should be recorded that incomplete modal logics are mostly artefacts in that they are the products of a diligent search, solely motivated by the wish to find examples of the incompleteness phenomenon. In chapter 4, however, we will prove the rather natural logic **GH**= **K**+ $\Box(\Box\varphi \leftrightarrow \varphi) \to \Box\varphi$ to be incomplete.

Ad (b) Canonical logics are complete since the method H can be successfully applied to canonical logics to prove them to be strongly complete.

Ad (c) We proved that all modal logics with the FMP are complete.

Ad (d) A logic that is complete but not canonical and without the FMP is the tense logic $\mathbf{K}_t$+ $H(H\varphi \to \varphi) \to H\varphi + GF\varphi \to FG\varphi$.

Ad (e) The provability logic L=**K**+$\Box(\Box\varphi \to \varphi) \to \Box\varphi$ is complete (not strongly complete) and has the FMP, but is not canonical (exercise 34).

Ad (f) **K**+$(\Box\varphi \to \varphi)$+$((\Box\varphi \wedge \Diamond\Diamond\psi) \to \Diamond(\Box\Box\varphi \wedge \Diamond\Diamond\Diamond\psi))$ — no fault of ours — is canonical, but does lack the FMP.

Ad (g) All logics treated so far are examples of logics with all three properties.

Anyone who would like to know more about the subject is referred to the article "Some kinds of modal completeness" by Johan van Benthem, to be found in *Studia Logica* 1980.

3.6 Exercises

Exercise 27

(a) Show that $\Gamma = \{\Box(p \to \Diamond\neg p), \Box(\neg p \to \Diamond p)\}$ is **S4**-consistent by constructing an infinite partially ordered frame that validates Γ.

(b) Construct also a finite reflexive and transitive frame that validates Γ.

(c) Deduce that no finite partially ordered frame exists that does the same.

Exercise 28 Prove a finite variant of exercise 21.

Exercise 29 Prove that if $\{\Box\psi \mid \Box\psi \in \Gamma\} \cup \{\neg\Box\chi\}$ is **S4**-consistent, then so is $\{\Box\psi \mid \Box\psi \in \Gamma\} \cup \{\neg\chi\}$.

Exercise 30

(a) Consider the closure Φ of the set $\{p, \Box p\}$. Construct $\mathcal{F}^{\Phi}_{\mathbf{S4}}$ and show that its accessibility relation is not transitive.

(b) (a little more difficult) Suggest an adjustment for the method H_{fin} that makes it work for **S4** also.

Exercise 31 Prove that **S5** has FMP using the method C_{fin}. Hint: set out to prove as in the case of **S4**. You will then automatically reach a point where you can get no further without a lemma similar to the one of exercise 29. If you have thought this out, you have overcome the main difficulty.

Exercise 32 For $n > 1$ let A_n stand for the following axiom-scheme:

$$\Box\varphi_1 \vee \Box(\varphi_1 \to \varphi_2) \vee \Box((\varphi_1 \wedge \varphi_2) \to \varphi_3) \vee \ldots \vee \Box((\varphi_1 \wedge \ldots \wedge \varphi_n) \to \varphi_{n+1})$$

This exercise studies the logics obtained by adding A_n to **S5**.

1. Prove that **S5+A**$_n$ is complete with respect to the class of frames $\langle \mathcal{W}, \mathcal{R}\rangle$ such that (i) $\mathcal{R}$ is universal and (ii) $|\mathcal{W}| \leq n$.
2. Prove this proposition can be strengthened to $|\mathcal{W}| = n$.
3. Argue that, if **S** is a consistent modal logic such that **S5**$\subseteq$**S**, then **S**=**S5** or **S**=**S5+A**$_n$for some n.

Exercise 33 Show that **L** is not canonical (Hint: look at exercise 35).

Chapter 4

Provability logic

In this chapter we will discuss the provability logic **L** (the L stands for M.H. Löb). This logic has interpretations in formal arithmetic, e.g. in Peano-arithmetic PA. We first go into the purely modal side of the matter.

4.1 L as a modal logic

The provability logic **L** is axiomatized by the axiom scheme $\Box(\Box\varphi \to \varphi) \to \Box\varphi$ (called Löb's axiom for reasons which will become clear later). According to proposition 9(d) $\Box(\Box\varphi \to \varphi) \to \Box\varphi$ characterizes the frames with $\mathcal{R}$ transitive, and its converse $\widetilde{\mathcal{R}}$ well-founded. According to exercise 10(b) the finite transitive, conversely well-founded frames are the same as the finite transitive, irreflexive ones. A first question which comes up is, whether $\Box\varphi \to \Box\Box\varphi$, which characterizes the transitive frames, is derivable in this logic. If it is not, **L** would be incomplete.

Proposition 39 $\vdash_{\mathbf{L}} \Box p \to \Box\Box p$

Proof: Exercise 34. •

We prove completeness and decidability of **L** with a variant of the method H_{fin}. The methods H and C are not applicable, because **L** is not strongly complete (see exercise 35).

Theorem 40 **L** is decidable and complete with respect to the finite, irreflexive (and therefore conversely well-founded), transitive frames.

Proof: Suppose $\psi_1, \ldots, \psi_m \nvdash_{\mathbf{L}} \varphi$. Take Γ as a maximally **L**-consistent set containing $\{\psi_1, \ldots, \psi_m, \neg\varphi\}$ in the closure Φ of $\{\psi_1, \ldots, \psi_m, \neg\varphi\}$. The set of worlds is as in the method H_{fin}, i.e. all maximally **L**-consistent sets in Φ. But here we set $\Delta\mathcal{R}\Delta'$ iff, for each $\Box\varphi \in \Delta$, (i) both $\Box\varphi \in \Delta'$ and $\varphi \in \Delta'$; and (ii), for at least one $\Box\varphi \in \Delta'$, $\Box\varphi \notin \Delta$ ($\Box\varphi$ is 'new'). In this way, we immediately have transitivity as in **S4**, but also converse well-foundedness, since any $\mathcal{R}$-chain will have to break off when we run out of new $\Box$ -formulas to add. The proof now runs as usual, except that because of the change in the definition of $\mathcal{R}$ the proof that if $\neg\Box\chi \in \Delta$, with Δ maximally **L**-consistent, there is a $\Delta'\widetilde{\mathcal{R}}\Delta$ with $\neg\chi \in \Delta$, changes somewhat. It is sufficient here to show that in that case $\{\psi, \Box\psi \mid \Box\psi \in \Delta\} \cup \{\Box\chi, \neg\chi\}$ is **L**-consistent. We leave this as exercise 36. Note that condition (ii) for $\mathcal{R}$ is then satisfied, because $\Box\chi$ is a new formula in Δ'. •

Here, at last, is an example of incomplete modal logic. The frames characterized by $\Box(\Box\varphi \leftrightarrow \varphi) \to \Box\varphi$ and $\Box(\Box\varphi \to \varphi) \to \Box\varphi$ are the same. Yet in **GH**, the logic axiomatized by $\Box(\Box\varphi \leftrightarrow \varphi) \to \Box\varphi$, not all formulas can be derived which are derivable in **L**, in particular, $\not\vdash_{\mathbf{GH}} \Box p \to \Box\Box p$.

Theorem 41 The frames characterized by **GH** are the transitive, conversely well-founded frames

Proof: $\Box(\Box\varphi \leftrightarrow \varphi) \to \Box\varphi$ is derivable in **L**, hence $\Box(\Box\varphi \leftrightarrow \varphi) \to \Box\varphi$ is valid on all transitive, conversely well-founded frames.

For the opposite direction, assume first that $\mathcal{F}$ is not transitive. There are worlds u, v, w such that $u\mathcal{R}v\mathcal{R}w$, but not $u\mathcal{R}w$. Define for $x \in \mathcal{W}$, $x \models p$ *unless* w is an element of the hereditary closure of $\mathcal{R}$ from x (i.e. unless there is a sequence $x = x_0\mathcal{R}x_1\mathcal{R}\ldots\mathcal{R}x_n = w$, with possibly $n = 0$, in which case $x = w$, see also Definition 8).

Firstly, it is clear that $v \not\models p$; hence $u \not\models \Box p$. To get that u does force $\Box(\Box p \leftrightarrow p)$ note that, for any $x \neq w$, w is in the hereditary closure of $\mathcal{R}$ from x iff w is in the hereditary closure of $\mathcal{R}$ from y for some y accessible from x (possibly w itself). This means that, for all $x \in \mathcal{W}$ (except possibly for $x = w$), $x \models \Box p$ iff $x \models p$, because, for $x \neq w$, $x \not\models p$ iff $x \not\models \Box p$. Because w is not accessible from u, this implies that $u \models \Box(\Box p \leftrightarrow p)$.

Next, assume that $\mathcal{R}$ is not conversely well-founded. Then there existes some infinite sequence $u_0\mathcal{R}u_1\mathcal{R}u_2\ldots$ exists. Set $x \models p$ unless there is a u_i in the hereditary closure of $\mathcal{R}$ from x. As in the first case it is clear that $u_0 \not\models \Box p$, $u_0 \models \Box(\Box p \leftrightarrow p)$. •

To show that $\not\vdash_{\mathbf{GH}} \Box p \to \Box\Box p$ we cannot simply think up a frame on which **GH** is valid and $\Box p \to \Box\Box p$ is not, because all frames on which **GH** is valid are transitive and on those $\Box p \to \Box\Box p$ is valid, too. We need a different, nonstandard notion of validity here, and introduce so-called generalized frames to that end.

Definition 20

(a) A *generalized frame* $\langle \mathcal{F}, \mathcal{U} \rangle$ with $\mathcal{F} = \langle \mathcal{W}, \mathcal{R} \rangle$ consists of a frame $\mathcal{F}$ together with a set $\mathcal{U}$ of admissible subsets of $\mathcal{W}$, for which
  (i) If $U, V \in \mathcal{U}$, then also $U \cap V \in \mathcal{U}$,
  (ii) If $U \in \mathcal{U}$, then also $\mathcal{W} \setminus U \in \mathcal{U}$,
  (iii) If $U \in \mathcal{U}$, then also $C(U) = \{v | \exists u \in U(v\mathcal{R}u)\} \in \mathcal{U}$.

(b) A model $\mathcal{M}$ on a generalized frame $\mathcal{F}$ is a model on $\mathcal{F}$ such that for all propositional letters p, $\{u \mid u \models p\} \in \mathcal{U}$.

Thus generalized frames and the models on them are just as normal frames and models, except that only certain valuations for the propositional letters are allowed. Validity of φ on a generalized frame is also defined with regard to just the admitted models. The conditions (i)–(iii) are exactly chosen in such a way that if $\langle \mathcal{F}, \mathcal{U} \rangle \models \varphi(p_1, \ldots, p_n)$, then $\langle \mathcal{F}, \mathcal{U} \rangle \models \varphi(\psi_1, \ldots, \psi_n)$ for all φ, because from (i)–(iii) it follows that also complex formulas have only admissible valuations (exercise 37). From this latter remark it already follows that we can use generalized frames to establish nonderivability, but we may just as well treat the method of using these frames a little more extensively before continuing.

From theorem 22 we can quickly derive for any logic **S** (considered as a set of formulas closed under modus ponens, substitution and necessitation):

Lemma 42 $\nvdash_S \varphi$ iff there is a model $\mathcal{M}$ such that $\mathcal{M} \models \mathbf{S}$ and $\mathcal{M} \not\models \varphi$.

Proof: $\Leftarrow$: Immediate from the soundness theorem for **K**, because $\vdash_{\mathbf{S}} \varphi$ iff $\mathbf{S} \vdash_{\mathbf{K}} \varphi$.

$\Rightarrow$: By theorem 22, $\mathcal{M}_{\mathbf{S}} \models \mathbf{S}$, and, if $\nvdash_{\mathbf{S}} \varphi$, then $\mathcal{M}_{\mathbf{S}} \not\models \varphi$. •

In practice we have however given modal logics **S** like **T**, **S4**, **S5** and **L** not as sets of formulas, but as sets of consequences of a set axiom schemes $Ax_{\mathbf{S}}$ using modus ponens and necessitation. According to the next lemma this makes no difference.

Lemma 43 Let the modal logic **S** be axiomatized by the set of axiom schemes $Ax_{\mathbf{S}}$ (i.e. **S** is exactly the set of formulas derivable from $Ax_{\mathbf{S}}$ by modus ponens and necessitation). Then $\mathcal{M} \models \mathbf{S}$ iff $\mathcal{M} \models Ax_{\mathbf{S}}$.

Proof: $\Rightarrow$: Trivial.
$\Leftarrow$: Clearly the set of formulas valid on a model is closed under modus pones. That it is closed under necessitation is obvious, too: If $w \models \varphi$ for all $w \in \mathcal{W}$, then also $w \models \Box\varphi$ for all $w \in \mathcal{W}$. •

Corollary 44 If $Ax_{\mathbf{S}}$ axiomatizes $\mathbf{S}$, then $\nvdash_{\mathbf{S}} \varphi$ iff there is a model $\mathcal{M}$ such that $\mathcal{M} \models Ax_{\mathbf{S}}$ and $\mathcal{M} \not\models \varphi$.

Proof: Immediate from lemmata 42 and 43. •

To find a counterexample to the derivability in **GH** of $\Box p \rightarrow \Box\Box p$ it suffices to find a generalized frame which verifies $\Box(\Box p \leftrightarrow p) \rightarrow \Box p$ and falsifies $\Box p \rightarrow \Box\Box p$. So far, we have only looked at axiom schemes like $\Box(\Box\varphi \leftrightarrow \varphi) \rightarrow \varphi$, but now we start considering 'single' axioms like $\Box(\Box p \leftrightarrow p) \rightarrow \Box p$.

Lemma 45 Let $Ax_{\mathbf{S}}$ axiomatize $\mathbf{S}$ and let $(Ax_{\mathbf{S}})^{Prop}$ be single axioms arising from $Ax_{\mathbf{S}}$, when one replaces the formula variables in the schemes by propositional letters. If $\mathcal{M} = \langle \mathcal{F}, \models \rangle \models Ax_{\mathbf{S}}$ and $\mathcal{M} \not\models \varphi$, then there is a generalized frame $\langle \mathcal{F},\mathcal{U} \rangle$ such that $\langle \mathcal{F},\mathcal{U} \rangle \models (Ax_{\mathbf{S}})^{Prop}$, and $\langle \mathcal{F},\mathcal{U} \rangle \not\models \varphi$.

Proof: Just take $\mathcal{U} = \{\{w \in \mathcal{W}_{\mathcal{F}} \mid w \models \varphi\} \mid \varphi$ a modal formula$\}$. Obviously, this set $\mathcal{U}$ is closed under the properties mentioned under (i)-(iii) in definition 20(a). That $\langle \mathcal{F},\mathcal{U} \rangle \models (Ax_{\mathbf{S}})^{Prop}$ follows from the fact that, if one chooses for the forcing relations on $\mathcal{F}$ for, say the propositional variables $p_1, \ldots, p_n$, the sets $U_1, \ldots, U_n$ from $\mathcal{U}$, then, each of these sets U_i is represented in $\mathcal{M}$ by a formula ψ_i forced on exactly U_i. Hence, for that $\models$, for each w, $w \models \varphi(p_1, \ldots, p_n)$ for $\varphi(p_1, \ldots, p_n) \in (Ax_{\mathbf{S}})^{Prop}$, since, for each w, $w \models_{\mathcal{M}} \varphi(p_1, \ldots, p_n)$. Finally, it is obvious that $\langle \mathcal{F},\mathcal{U} \rangle \not\models \varphi$. •

The next lemma shows that this result is good enough.

Lemma 46 $\langle \mathcal{F},\mathcal{U} \rangle \models (Ax_{\mathbf{S}})^{Prop}$ iff $\langle \mathcal{F},\mathcal{U} \rangle \models Ax_{\mathbf{S}}$.

Proof: From right to left: Obvious. From left to right: This follows from the fact that, if $\langle \mathcal{F},\mathcal{U} \rangle \models \varphi(p_1, \ldots, p_n)$, then also $\langle \mathcal{F},\mathcal{U} \rangle \models \varphi(\psi_1, \ldots, \psi_n)$ for any $\psi_1, \ldots, \psi_n$ (exercise 37). •

.

Theorem 47 $\nvdash_{\mathbf{S}} \varphi$ if and only if there is a generalized frame $\langle \mathcal{F},\mathcal{U} \rangle$ such that $\langle \mathcal{F},\mathcal{U} \rangle \models (Ax_{\mathbf{S}})^{Prop}$ and $\langle \mathcal{F},\mathcal{U} \rangle \not\models \varphi$.

Proof: Immediate from corollary 44, and lemmata 45 and 46. •

Theorem 48 ((Magari 1982)) $\nvdash_{\mathbf{GH}} \Box p \rightarrow \Box\Box p$ (and, hence **GH** is incomplete).

Proof: (Cresswell 1986): Take the following generalized frame $\langle \mathcal{W}, \mathcal{R} \rangle$ where $\mathcal{W} = \{a_0, a_1, a_2, \ldots\} \cup \{b_0, b_1, b_2, \ldots\}$.

$\mathcal{R}$ is defined as follows:

- for all $i, j, a_i \mathcal{R} b_j$,
- $b_i \mathcal{R} b_j$ iff $j < i$,
- $a_i \mathcal{R} a_j$ iff $i \leq j + 1$.

In a picture comes this down to the following:

INSERT FIGURE 14

The b-part is conversely well-founded and transitive, the a-part is neither, that is to say: it is transitive 'going to the right', but, for all $j, a_{j+2} \mathcal{R} a_{j+1} \mathcal{R} a_j$, whereas not $a_{j+2} \mathcal{R} a_j$.

The admissible sets are the finite and the co-finite (i.e. sets with a finite complement) ones. It is important that, if an admissible set U is not finite and hence co-finite, then it is the case that, for certain j and k, $\{a_j, a_{j+1}, a_{j+2}, \ldots\} \cup \{b_k, b_{k+1}, b_{k+2}, \ldots\} \subseteq U$.

To see that in fact the structure given is a generalized frame, we have to check (i)–(iii). Condition (ii) follows from the fact that the complement of a finite set is co-finite and the complement of a co-finite set is finite. Condition (i) follows from the fact that the intersection of two finite sets, as well as of a finite and a co-finite set is finite, and of the intersection two co-finite sets is co-finite. For (iii), let U be finite or co-finite; it is to be shown that $C(U)$ is finite or co-finite. First assume that $b_i \in U$ for some i. Then $C(U)$ has to be co-finite, because $C(U)$ contains everything left of b_i. Next assume that for no i, $b_i \in U$. Then by the remark above that co-finite sets have infinite tails in the a-part as well as in the b-part, U has to be finite. Let a_i be the rightmost element of U; then $C(U) = \{a_0, \ldots, a_{i+1}\}$, and hence finite.

It is simple to give a countermodel to $\Box p \rightarrow \Box\Box p$ (this is exercise 38). Next we prove that on the generalized frame $\Box(\Box p \leftrightarrow p) \rightarrow \Box p$ is valid. For, assume that there is a countermodel to $\Box(\Box p \leftrightarrow p) \rightarrow \Box p$ on it. Then there has to be a c with $c \not\models \Box(\Box p \leftrightarrow p) \rightarrow \Box p$. This c cannot lie in the b-part, since this in itself is a transitive, conversely well-founded frame (the a-part is visible nowhere from the b-part). This means that the worlds on which $\Box(\Box p \leftrightarrow p) \rightarrow \Box p$ is forced form a non-finite and hence co-finite set. This in its turn implies that there is a last a_i on which $\Box(\Box p \leftrightarrow p) \rightarrow \Box p$ is falsified, and this again that $a_i \models \Box(\Box p \leftrightarrow p)$, $a_i \not\models \Box p$. That means in the first place that for all c to the right of a_i, $c \models \Box(\Box p \leftrightarrow p)$, but also that for such a c, $c \models \Box p$, because otherwise a_i is not the last world on which $\Box(\Box p \leftrightarrow p) \rightarrow \Box p$ is falsified. The fact however that hereby $a_{i+1} \models \Box p$ implies that $a_i \models p$. But from $a_i \models \Box(\Box p \leftrightarrow p)$ it follows that $a_i \models \Box p \leftrightarrow p$, and so $a_i \models \Box p$ holds after all. Hence, the assumption that there is a countermodel has lead to a contradiction: $\Box(\Box p \leftrightarrow p) \rightarrow \Box p$ is valid on this generalized frame.
•

4.2 Arithmetization

As an example we arithmetize the *classical implicational logic* **PCI** which has as its axiom schemes:

(i) $\varphi \to (\psi \to \varphi)$
(ii) $(\varphi \to (\psi \to \chi)) \to ((\varphi \to \psi) \to (\varphi \to \chi))$
(iii) $((\varphi \to \psi) \to \varphi) \to \varphi)$

and which has *modus ponens* as its only rule.

Arithmetization of an area means the coding of the concepts of that area by means of natural numbers. This can be done with each area which is in a certain sense discrete. The basic idea is that sequences of natural numbers can be coded by natural numbers. The classical way of coding is by means of the sequence of *prime numbers* $p_0, p_1, p_2, \ldots$. Prime numbers are those natural numbers larger than 1 divisible by themselves and 1 only, i.e. $p_0 = 2, p_1 = 3, p_2 = 5, p_3 = 7, p_4 = 11$, etc. The so-called *unique factorization into prime numbers of natural numbers* is used: there is exactly one way to write a natural number n as $n = p_0^{n_0} \cdot p_1^{n_1} \ldots p_k^{n_k}$. The code number (*gödel number*) of a sequence $(a_0, \ldots, a_k)$, written as $\langle a_0, \ldots, a_k \rangle$, is $p_0^{a_0} \cdot p_1^{a_1} \ldots p_k^{a_k}$, i.e., if $n = p_0^{n_0} \cdot, p_1^{n_1} \ldots p_k^{n_k}$, then is $n = \langle n_0, \ldots, n_k \rangle$; this sequence is not determined in a completely unique way: it is also the case that $n{=}\langle n_0, \ldots, n_k, 0, \ldots, 0\rangle$, but that will not hurt us, because we will not use 0 as a code number. As an example: $108 = 4 \cdot 27 = 2^2 \cdot 3^3 = \langle 2, 3 \rangle$, but also $108 = 2^2 \cdot 3^3 \cdot 5^0 \cdot 7^0 = \langle 2, 3, 0, 0 \rangle$. We do not count the zeroes at the end in defining the *length* of the sequence coded by n. If $n = \langle n_0, \ldots, n_k \rangle$ and $n_k \neq 0$, then the length of n ($lh(n)$ for short) is $k + 1$; for example, $lh(108) = 2$. There are also inverse functions $(a)_i$ which give the i-th place in the sequence coded by a. For example: $(108)_0 = 2$, $(108)_1 = 3$, $(108)_2 = 0$, $(108)_3 = 0$, $(108)_i = 0$ for all $i \geq 4$. For $((a)_i)_j)$ we write $(a)_{i,j}$. So is e.g. $(108)_{0,0} = 1$, $(108)_{0,1} = 0$, $(108)_{1,0} = 0$, $(108)_{1,1} = 1$.

For the last element of a sequence coded by n (again not counting zeroes at the end): $(n)_{lh(n)-1}$, we also write $(n)_{last}$, for example, $(108)_{last} = 3$. We are often interested of course in the sequence that arises by putting two sequences together one behind the other, one says to *concatenate* them. The number that codes the concatenation of two sequences coded by x and y is written as $x * y$ and is found by a purely number-theoretic operation:

$$x * y = x \cdot p_{lh(x)}^{(y)_0} \cdots p_{lh(x)*lh(y)-1}^{(y)_{last}}$$

For the coding of an area we also have to give a fixed code to the basic elements of the area. In the example of classical implicational logic: $\ulcorner \to \urcorner = 2$, i.e. the gödel number of '$\to$' is 2, $\ulcorner) \urcorner = 4$, $\ulcorner (\urcorner = 6$, and for the infinite sequence of propositional letters which we will write here to avoid confusion with the prime numbers as $q_0, q_1, q_2, \ldots$: $\ulcorner q_0 \urcorner = 1, \ulcorner q_1 \urcorner = 3, \ulcorner q_2 \urcorner = 5, \ulcorner q_3 \urcorner = 7, \ulcorner q_4 \urcorner = 9$, etc. As we now see formulas as sequences of symbols, there are certain numbers that code formulas, and, whether they do so or not, can be expressed in a purely number theoretic way. For example, the formula $(q_0 \to q_1)$ is coded by $2^6 \cdot 3^1 \cdot 52 \cdot 7^3 \cdot 11^4$. To be able to find the general number theoretic expression that determines whether a number codes a formula it is important to transform the inductive definition of 'formula' into the following form: φ is a formula iff there is a sequence $\varphi_0, \varphi_1, \ldots, \varphi_n$ such that $\varphi_n = \varphi$ and, for all $i \leq n$, either φ_i is an atomic formula, i.e. one of the q_j, or φ_i is of the form $(\varphi_i \to \varphi_k)$ for a pair $j, k < i$.

Such a sequence is called a *formula forming sequence* (with φ as its last element). We say first, when a number x codes a formula forming sequence (for short: $Formulaseq(x)$:

$$\begin{array}{lll} Formulaseq(x) & \text{iff} & \forall i < lh(x)((x)_i \text{ is odd, or} \\ & & \exists j, k < i((x)_i = \langle \ulcorner (\urcorner \rangle * (x)_j * \langle \ulcorner \to \urcorner \rangle * (x)_k * \langle \ulcorner) \urcorner \rangle)) \end{array}$$

Here '$\ulcorner) \urcorner$', '$\ulcorner \to \urcorner$' and '$\ulcorner (\urcorner$' can of course be replaced by respectively '4', '2' and '6'.

Next we obtain:

$$Formula(x) \quad \text{iff} \quad \exists y(Formulaseq(y) \text{ and } (y)_{last} = x)$$

Henceforth we will write $imp(x, y)$ for $\langle \ulcorner (\urcorner \rangle * x * \langle \ulcorner \to \urcorner \rangle * y * \langle \ulcorner) \urcorner \rangle$.

A derivation is just a sequence of formulas which satisfies certain properties. The following speaks for itself.

$$\begin{array}{rll} Axiom_1(x) & \text{iff} & \exists y, z(Formula(y) \text{ and } Formula(z) \text{ and} \\ & & x = imp(y, imp(z, y))) \\ Axiom_2(x) & \text{iff} & \exists y, z, w(Formula(y) \text{ and } Formula(z) \text{ and } Formula(w) \text{ and} \\ & & x = imp(imp(y, imp(z, w)), imp(imp(y, z), imp(y, w))) \\ Axiom_3(x) & \text{iff} & \exists y, z(Formula(y) \text{ and } Formula(z) \text{ and} \\ & & x = imp(imp(imp(y, z), y), y) \\ Axiom(x) & \text{iff} & Axiom_1(x) \text{ or } Axiom_2(x) \text{ or } Axiom_3(x) \\ Derivation(x) & \text{iff} & \forall i < lh(x)(Axiom((x)_i) \text{ or} \\ & & \exists j, k < i((x)j = imp((x)_k, (x)_i))) \\ Prf_{\mathbf{PCI}}(y, x) & \text{iff} & Derivation(y) \text{ and } x = (y)_{last} \\ Prf_{\mathbf{PCI}}(x) & \text{iff} & \exists y Prf_{\mathbf{PCI}}(y, x) \end{array}$$

'$Prf_{\mathbf{PCI}}(y, x)$' stands for 'y codes a derivation of x in **PCI**', '$Prf_{\mathbf{PCI}}(x)$' for 'x codes a formula derivable in **PCI**'. Also somewhat more complicated metamathematical concepts can be formalized. As examples we give substitution and tautology. First the substitution of a formula ψ for a propositional letter q in a formula φ

$$sub(x,y,z)=w \quad \text{iff} \quad \begin{array}{l} \exists u \forall i < lh(u)((((u)_{0,0}=y \text{ and } (u)_{0,1}=x), \text{ and either} \\ \text{(i) } (u)_{i,0} \text{ is odd and } (u)_{i,0} \neq y \text{ and } (u)_{i,0}=(u)_{i,1}, \text{ or} \\ \text{(ii) } \exists j,k < i \text{ such that } (u)_{i,0}=imp((u)_{j,0},(u)_{k,0}) \text{ and} \\ (u)_{i,1}=imp((u)_{j,1},(u)_{k,1}) \text{ while } z=(u)_{last,0} \text{ and} \\ w=(u)_{last,1}) \end{array}$$

In this substitution of x for y in z one should read u as a sequence of ordered pairs, the first elements of which form a formula forming sequence for z, and the second elements a parallel construction of the resultant formula w. In a similar manner one can see a valuation of (the subformulas of) a formula φ as a sequence of pairs, the second element of which is 0 or 1 (of course satisfying certain conditions); this and the definition of a formula $Taut(x)$ expressing 'x codes a tautology' is left as an exercise.

One ought to realize that by this procedure metamathematical assertions about **PCI** simply translate into arithmetical assertions. For example the assertion that, for all φ , p, ψ, if $\vdash_{\mathbf{PCI}} \varphi(p)$, then also $\vdash_{\mathbf{PCI}} \varphi(\psi)$, translates into:

$$\forall x,y,z(\text{if } Prov_{\mathbf{PCI}}(x) \text{ and } y \text{ is odd and } Formula(z), \\ \text{then } Prov_{\mathbf{PCI}}(sub(z,y,x)))$$

And the completeness theorem translates into

$$\forall x(\text{ if } Prov_{\mathbf{PCI}}(x), \text{ then } Taut(x))$$

4.3 Formal arithmetic

The following axioms are known as the axioms of *Peano-arithmetic* **PA**:

1. $Sx \neq \mathbf{0}$
2. $Sx = Sy \rightarrow x = y$
3. $x + \mathbf{0} = x$
4. $x + Sy = S(x+y)$
5. $x \cdot \mathbf{0} = \mathbf{0}$
6. $x \cdot Sy = x \cdot y + x$
7. $(\alpha(\mathbf{0}) \wedge \forall x(\alpha(x) \rightarrow \alpha(Sx))) \rightarrow \forall x \alpha(x)$ (for each formula α)

The language used has $\mathbf{0}$ as a constant symbol, and S, $\cdot$ and $+$ as function symbols. To distinguish the formulas from modal propositional ones we will use α, β, γ, ... as metavariables over formulas for the language of **PA** instead of φ, ψ, ξ, ... The axioms come of course on top of the usual ones for standard predicate logic with identity and function symbols. If α is a theorem of **PA**, then write we also $\vdash_{\mathbf{PA}} \alpha$.

The *standard model* for the language is the set of natural numbers $\mathbb{N}$ with the successor function, sum and product. A sentence α of the language is called true, if $\mathbb{N} \models \alpha$. An important role is played by the following sequence of terms *(numerals)*:

$$\mathbf{0}, S\mathbf{0}, SS\mathbf{0}, SSS\mathbf{0}, \ldots$$

It is clear that these are the simplest terms which in the standard model $\mathbb{N}$ have as their interpretation the numbers $0, 1, 2, 3, \ldots$. We write also $\mathbf{0}$, $\mathbf{1}$, $\mathbf{2}$, $\mathbf{3}$, ... for the numerals and as metavariables over numerals we will use **m, n**, etc. (i.e. the terms with respectively m and n times S followed by a $\mathbf{0}$, etc.). In other cases also we will mostly print numerals in bold print, but sometimes, for example because printing doubly bold is not possible, we underline: the numeral corresponding to $\ulcorner\varphi\urcorner$ will for example be given as $\ulcorner\boldsymbol{\varphi}\urcorner$, but also as $\underline{\ulcorner\varphi\urcorner}$. It is very important realize that $\vdash_{\mathbf{PA}} \forall x\alpha(x)$ is a much stronger assertion than: 'for each n, $\vdash_{\mathbf{PA}} \alpha(\mathbf{n})$'; and not only in theory; the distinction plays an important role in the following.

Instead of the intuitive arithmetical language from the last section one can also use the formal language of **PA** in the arithmetization of **PCI**. This is not a trivial matter: it will have to be shown that the whole coding of sequences can be executed in the language of **PA**. That is an awful lot of work; we will not do it here, but refer to the course on incompleteness theorems. The point is that we do not only have to show that the relevant concepts can be expressed in the language of **PA**, we also will have to try to prove the relevant results in **PA** (from the axioms, of course). This is a lot of work too; and we will not do that here either. We assume that the work has been done and that for example the following formulas have been obtained: $\mathbf{Formula}(x)$, $\mathbf{Prf_{PCI}}(y, x)$, $\mathbf{Prov_{PCI}}(x)$, $\mathbf{Taut}(x)$, such that for example, the following results have been obtained

- For any n, if $Formula(n)$, then $\vdash_{\mathbf{PA}} \mathbf{Formula}(\mathbf{n})$
- For any n, if not $Formula(n)$, then $\vdash_{\mathbf{PA}} \neg\mathbf{Formula}(\mathbf{n})$.
- For any m, if m codes a derivation of φ, then $\vdash_{\mathbf{PA}} \mathbf{Prf_{PCI}}(\mathbf{m}, \ulcorner\boldsymbol{\varphi}\urcorner)$;
- For any m, if m does not code a derivation of φ, then $\vdash_{\mathbf{PA}} \neg\mathbf{Prf_{PCI}}(\mathbf{m}, \ulcorner\boldsymbol{\varphi}\urcorner)$,
- For each φ, if $\vdash_{\mathbf{PCI}} \varphi$, then $\vdash_{\mathbf{PA}} \mathbf{Prov_{PCI}}(\ulcorner\boldsymbol{\varphi}\urcorner)$,
- for each φ, $\vdash_{\mathbf{PA}} \mathbf{Taut}(\ulcorner\boldsymbol{\varphi}\urcorner) \to \mathbf{Prov_{PCI}}(\ulcorner\boldsymbol{\varphi}\urcorner)$;
- or even better: $\vdash_{\mathbf{PA}} \forall x(\mathbf{Taut}(x) \to \mathbf{Prov_{PCI}}(x))$.

The last mentioned result means that the completeness theorem for **PCI** has this after its arithmetization also been formalized. Besides assuming that we have the relevant *formulas* we also assume that we have *terms* $\mathbf{imp}(x, y)$, $\mathbf{sub}(x, y, z)$ with the desired properties. Strictly speaking this is not right: there are no such terms, but they can be added without danger (we won't go deeper into this subtle point and just assume we have them). We have then for example:

- for all formulas φ and ψ : $\vdash_{\mathbf{PA}} \ulcorner\boldsymbol{\varphi \rightarrow \psi}\urcorner = \mathbf{imp}(\ulcorner\boldsymbol{\varphi}\urcorner, \ulcorner\boldsymbol{\psi}\urcorner$, and for all formulas φ, ψ, and propositional letters q,
  $\vdash_{\mathbf{PA}} \mathbf{Prov_{PCI}}(\ulcorner\boldsymbol{\varphi}\urcorner) \rightarrow \mathbf{Prov_{PCI}}(\mathbf{sub}(\ulcorner\boldsymbol{\psi}\urcorner, \ulcorner\boldsymbol{q}\urcorner, \ulcorner\boldsymbol{\varphi}\urcorner))$.

Now **PCI** has really only been intended as an example: the real issue is that **PA** can be arithmetized and formalized in an essentially similar manner. This is somewhat, but not spectacularly, more complicated then in the case of **PCI**, because **PA** is formalized in predicate logic instead of propositional logic. We will not execute the necessary work. We just take it that we have obtained formulas in **PA** with the following properties:

(i) if n codes a derivation of α in **PA**, then $\vdash_{\mathbf{PA}} \mathbf{Prf_{PA}}(\mathbf{n}, \ulcorner\boldsymbol{\alpha}\urcorner)$,

(ii) if n does not code a derivation in **PA** of α, then $\vdash_{\mathbf{PA}} \neg\mathbf{Prf_{PA}}(\mathbf{n}, \ulcorner\boldsymbol{\alpha}\urcorner)$;

(iii) $(*)$ if $\vdash_{\mathbf{PA}} \alpha$, then $\vdash_{\mathbf{PA}} \mathbf{Prov_{PA}}(\ulcorner\boldsymbol{\alpha}\urcorner)$.

Mostly we will write $\boxplus\alpha$ for $\mathbf{Prov_{PA}}(\ulcorner\boldsymbol{\alpha}\urcorner)$. This has great advantages, because it makes the connections with modal logic clear; $(*)$ for example becomes:

$$\text{If } \vdash_{\mathbf{PA}} \alpha, \text{ then } \vdash_{\mathbf{PA}} \boxplus\alpha$$

which quite naturally shows up as a translation of the necessitation rule. A disadvantage is that it is no longer clear from the notation that the formula α does not occur as such as a subformula, but only in the form of the numeral for its gödel number.

The most important substitutions we consider in the case of **PA** are those of a numeral for a variable. We can assume here that we always substitute for the same fixed variable x. We take it that we have a **PA**-function **subst** such that

$\vdash_{\mathbf{PA}} \mathbf{subst}(\underline{\ulcorner\varphi(x)\urcorner}, \mathbf{n}) = \underline{\ulcorner\varphi(\mathbf{n})\urcorner}$.

Implicit in Gödel's work is the so-called diagonalization lemma based on this formalization of substitution:

Lemma 49 (Diagonalization lemma) For each **PA**-formula $\beta(x)$ there is a **PA**-formula α, such that $\vdash_{\mathbf{PA}} \alpha \leftrightarrow \beta(\ulcorner\boldsymbol{\alpha}\urcorner)$. (Informally expressed, the formula α says: 'I am β'.)

Proof: Let $\beta(x)$ be given, and write n for $\ulcorner\beta(\mathbf{subst}(x,x))\urcorner$.
Take for α: $\beta(\mathbf{subst}(\mathbf{n},\mathbf{n}))$.
The property of **subst** given above implies that

$\vdash_{\mathbf{PA}} \mathbf{subst}(\underline{\ulcorner\beta(\mathbf{subst}(x,x))\urcorner}, \mathbf{n}) = \underline{\ulcorner\beta(\mathbf{subst}(\mathbf{n},\mathbf{n}))\urcorner}$.

But looking carefully, one sees that $\underline{\ulcorner\beta(\mathbf{subst}(x,x))\urcorner}$ is really $\mathbf{n}$, and $\underline{\ulcorner\beta(\mathbf{subst}(\mathbf{n},\mathbf{n}))\urcorner}$ is really $\ulcorner\boldsymbol{\alpha}\urcorner$. That is to say that

$\vdash_{\mathbf{PA}} \mathbf{subst}(\mathbf{n},\mathbf{n}) = \ulcorner\boldsymbol{\alpha}\urcorner$,

and so on purely logical grounds,

$\vdash_{\mathbf{PA}} \beta(\mathbf{subst}(\mathbf{n},\mathbf{n})) \leftrightarrow \beta(\ulcorner\boldsymbol{\alpha}\urcorner)$,

but $\beta(\mathbf{subst}(\mathbf{n},\mathbf{n}))$ actually is α, so $\vdash_{\mathbf{PA}} \alpha \leftrightarrow \beta(\ulcorner\boldsymbol{\alpha}\urcorner)$ •

The formula α of the diagonalization lemma is called the *fixed point* of the formula $\beta(x)$. An immediate application of the diagonalization lemma is the first incompleteness theorem, towards with we first give a definition and a small lemma.

Definition 21 **PA** is *ω-consistent* if, for no $\alpha(x)$, $\vdash_{\mathbf{PA}} \exists x\alpha(x)$, and, for all n, $\vdash_{\mathbf{PA}} \neg\alpha(\mathrm{n})$

It is clear that, if **PA** proves only true formulas, then **PA** is ω-consistent, and that, if **PA** is ω-consistent, then **PA** is certainly consistent (otherwise anything would be derivable).

Lemma 50 If **PA** is ω-consistent and $\vdash_{\mathbf{PA}} \boxplus\alpha$, then $\not\vdash_{\mathbf{PA}} \neg\alpha$.

Proof: Assume that **PA** is ω-consistent (and hence consistent), $\vdash_{\mathbf{PA}} \boxplus\alpha$ and at the same time $\vdash_{\mathbf{PA}} \neg\alpha$. From the latter together with the consistency of **PA** it follows that $\not\vdash_{\mathbf{PA}} \alpha$, which means that, for no n, n is the gödel number of a derivation of α. According to the properties of **Prf** given above this implies in its turn that, for each n, $\vdash_{\mathbf{PA}} \neg\mathbf{Prf}(\mathbf{n}, \ulcorner\boldsymbol{\alpha}\urcorner)$. But $\vdash_{\mathbf{PA}} \boxplus\alpha$ means nothing but $\vdash_{\mathbf{PA}} \exists x\mathbf{Prf}(x, \ulcorner\boldsymbol{\alpha}\urcorner)$: an ω-inconsistency has been reached contrary to assumption. •

Theorem 51 (Gödel's First Incompleteness theorem, 1931) If **PA** is ω-consistent, then there is a sentence γ such that neither $\vdash_{\mathbf{PA}} \gamma$, nor $\vdash_{\mathbf{PA}} \neg\gamma$.

Proof: Take for γ the fixed point of $\neg\mathbf{Prov}(x)$ the existence of which is guaranteed by the diagonalization lemma; γ is a sentence which says: 'I am not provable'. The way we write it:

$\vdash_{\mathbf{PA}} \gamma \leftrightarrow \neg \boxplus \gamma.$

Assume first that $\vdash_{\mathbf{PA}} \gamma$. Then according to $(*)$, $\vdash_{\mathbf{PA}} \boxplus\gamma$, but then the fixed point equation entails also $\vdash_{\mathbf{PA}} \neg\gamma$, and this would mean, contrary to the assumption, that **PA** is inconsistent.

Next assume that $\vdash_{\mathbf{PA}} \neg\gamma$. By the fixed point equation it holds that $\vdash_{\mathbf{PA}} \neg\neg \boxplus \gamma$, and hence $\vdash_{\mathbf{PA}} \boxplus\gamma$; but according to lemma 50 this contradicts the assumption that $\vdash_{\mathbf{PA}} \neg\gamma$. •

The assumption of ω-consistency can be weakened to consistency by a trick due to Rosser. We will not show that here.

4.4 Arithmetic and modal logic

Implicitly in Gödel's work and especially in that of Hilbert/Bernays were the following facts about the derivability in **PA** of principles concerning the proof predicate. They have been explicitized by Löb, and given by him as the principles sufficient to found the existing knowledge. They are known as *Löb's principles*.

(I) If $\vdash_{\mathbf{PA}} \alpha$, then $\vdash_{\mathbf{PA}} \boxplus\alpha$
(II) $\vdash_{\mathbf{PA}} \boxplus(\alpha \to \beta) \to (\boxplus\alpha \to \boxplus\beta)$
(III) $\vdash_{\mathbf{PA}} \boxplus\alpha \to \boxplus\boxplus\alpha$

Löb's principle I we have encountered as $(*)$. Löb's principle II is not too much work to prove. The proof of principle III is very much work, mostly of a rather trivial nature: it consists in the formalization in **PA** of the facts mentioned above concerning the proof predicate $\mathbf{Prf_{PA}}$. Of course, we just assume the principles here. From these three principles Löb proved a theorem

Theorem 52 (Löb's Theorem, 1956) For each α, if $\vdash_{\mathbf{PA}} \boxplus\alpha \to \alpha$, then $\vdash_{\mathbf{PA}}$ α.

Proof: Fix α. According to the diagonalization lemma there is a β such that $\vdash_{\mathbf{PA}} \beta \leftrightarrow (\boxplus\beta \to \alpha)$, because $\boxplus\beta \to \alpha$ stands for $\mathbf{Prov}(\ulcorner\beta) \to \alpha\urcorner$. From this it follows that $\vdash_{\mathbf{PA}} \boxplus\beta \leftrightarrow \boxplus(\boxplus\beta \to \alpha)$ (Exercise 41 (a));

From this in turn it follows that $\vdash_{\mathbf{PA}} \boxplus\beta \to (\boxplus\boxplus\beta \to \boxplus\alpha)$ (Löb's principle II);

finally Löb's principle III gives: $\vdash_{\mathbf{PA}} \boxplus\beta \to \boxplus\boxplus\beta$, by which $\vdash_{\mathbf{PA}} \boxplus\beta \to \boxplus\alpha$ which together with the assumption $\vdash_{\mathbf{PA}} \boxplus\alpha \to \alpha$ gives $\vdash_{\mathbf{PA}} \boxplus\beta \to \alpha$, and hence by the fixed point equation, $\vdash_{\mathbf{PA}} \alpha$. •

One can put this theorem into a stronger form which then shows up the connection with the system **L**.

Theorem 53 $\vdash_{\mathbf{PA}} \boxplus(\boxplus\alpha \to \alpha) \to \boxplus\alpha$

Proof: Exercise 41(b). •

An immediate consequence of theorem 53 is Gödel's second incompleteness theorem. To this end we first give the formula $\neg\boxplus\bot$ the alternative name $\mathbf{Con_{PA}}$, because this formula expresses in **PA** that $\bot$ is not provable, which is equivalent to the assertion that **PA** is consistent. The second incompleteness theorem states that this is not provable in **PA** itself.

Theorem 54 (Second Incompleteness Theorem (Gödel 1931) If **PA** is consistent, then $\nvdash_{\mathbf{PA}} \mathbf{Con_{PA}}$.

Proof: Exercise 42. •

Another consequence is:

Theorem 55
If $\vdash_{\mathbf{L}} \varphi(p_1, \ldots, p_n)$, then for all **PA**-formulas $\alpha_1, \ldots, \alpha_n$, $\vdash_{\mathbf{PA}} \varphi^*(\alpha_1, \ldots, \alpha_n)$, where φ^* is formed from φ by replacing all $\Box$ by $\boxplus$.
(Or in a more official formulation: for each translation * of formulas of modal propositional logic into sentences of arithmetic which is such that:

$(\varphi \circ \psi)^* = \varphi^* \circ \psi^*$, for each connective $\circ$,
$\bot^* = \bot$,
$(\neg\varphi)^* = \neg\varphi^*$ (one says: * commutes with the standard connectives),
$(\Box\varphi)^* = \boxplus\varphi^*$ (i.e., $\mathbf{Prov_{PA}}(\underline{\varphi^*})$),

it holds that, if $\vdash_{\mathbf{L}} \varphi(p_1, \ldots, p_n)$, then $\vdash_{\mathbf{PA}} \varphi * (\alpha_1, \ldots, \alpha_n)$.)

Proof: Exercise 42. •

The area got a tremendous impulse, when the converse of this theorem was proved.

Theorem 56 (Arithmetical Completeness Theorem (Solovay, 1973))
Suppose that for any translation * defined as above, $\vdash_{\mathbf{PA}} \varphi^*(\alpha_1, \ldots, \alpha_n)$ for all **PA**-sentences $\alpha_1, \ldots, \alpha_n$. Then $\vdash_{\mathbf{L}} \varphi(p_1, \ldots, p_n)$.

Proof: Requires too much foreknowledge for this course. From a Kripke countermodel to $\varphi(p_1, \ldots, p_n)$ arithmetical formulas $\alpha_1, \ldots, \alpha_n$, are constructed for which $\nvdash_{\mathbf{PA}} \varphi^*(\alpha_1, \ldots, \alpha_n)$. The original proof relies heavily on the recursion theorem (see Introduction to Recursion Theory). Recently however proofs have been given which bypass the use of the recursion theorem. •

Another theorem in the area is the Fixed Point Theorem, a reflection into **L** of the Diagonalization Lemma. First a special case as an exercise:

Proposition 57 (writing $\boxdot\xi$ for $\xi \wedge \Box\xi$)

$\vdash_{\mathbf{L}} \boxdot(p \leftrightarrow \neg\Box p) \rightarrow (p \leftrightarrow \neg\Box\bot)$, and $\vdash_{\mathbf{L}} \neg\Box\neg\Box\bot \leftrightarrow \neg\Box\bot$.

Proof: Exercise 43. •

Theorem 58 (Fixed point theorem for L(de Jongh, Sambin 1976))
If in $\varphi(p, q_1, \ldots, q_n)$ the propositional letter p exclusively occurs under $\Box$ (i.e. in

subformulas of the form $\Box\xi$), then there is a formula $\delta(q_1,\ldots,q_n)$ such that $\vdash_{\mathbf{L}} \delta \leftrightarrow \varphi(\delta, q_1,\ldots,q_n)$.

Before giving the proof of this theorem we first give a somewhat weaker form, the implicit fixed point theorem, and two of its consequences. [1]

Theorem 59 (Implicit fixed point theorem) Let p_0 be a new propositional letter. In each model $\mathcal{M}$ on a conversely well-founded frame there is, for each formula $\varphi(p_0, q_1,\ldots,q_n)$, in which p_0 occurs only under $\Box$, an extension of the forcing relation $\models$ of $\mathcal{M}$ to p_0 such that $\mathcal{M} \models p_0 \leftrightarrow \varphi(p_0, q_1,\ldots,q_n)$, and there is a unique way to do this.

Proof: Assume $\mathcal{R}$ is well-founded in $\mathcal{M}$, and $\models$ is, for all $w'\mathcal{R}w$ so defined that $w' \models p_0 \leftrightarrow \varphi(p_0, q_1,\ldots,q_n)$. It is sufficient to prove that $\models$ can be extended in the desired manner to w. The crucial point is that by the definition of $\models$, it has been fixed already whether or not $w \models \varphi(p_0, q_1,\ldots,q_n)$, because p_0 occurs only under $\Box$, and for the determination of the value of $w \models p_0$, only the $w'\mathcal{R}w$ are of importance. But then we can as yet that $w \models p_0$ exactly when $w \models \varphi(p_0, q_1,\ldots,q_n)$, so that indeed $w \models p_0 \leftrightarrow \varphi(p_0, q_1,\ldots,q_n)$.

It is clear moreover that there is no freedom in choosing $\models$ for w: if $w \models \varphi(p_0, q_1,\ldots,q_n)$, then we have to have $w \models p_0$ as well, and if $w \not\models \varphi(p_0, q_1,\ldots,q_n)$, then we have no choice but to set $w \not\models p_0$. •

Theorem 60 ((Unicity theorem for fixed points in L) If p occurs only under $\Box$ in $\varphi(p, q_1,\ldots,q_n)$, and $\vdash_{\mathbf{L}} \delta_1 \leftrightarrow \varphi(\delta_1, q_1,\ldots,q_n)$, as well as $\vdash_{\mathbf{L}} \delta_2 \leftrightarrow \varphi(\delta_2, q_1,\ldots,q_n)$, then $\vdash_{\mathbf{L}} \delta_1 \leftrightarrow \delta_2$.

Proof: This is Exercise 43. It is simplest to prove the somewhat stronger result (writing $\boxdot\xi$ for $\xi \wedge \Box\xi$):
$\vdash_{\mathbf{L}} \boxdot(p_1 \leftrightarrow \varphi(p_1, q_1,\ldots,q_n)) \wedge \boxdot(p_2 \leftrightarrow \varphi(p_2, q_1,\ldots,q_n)) \rightarrow (p_1 \leftrightarrow p_2)$. •

Another consequence is:

1. The method for proving the implicit fixed point theorem uses *transfinite induction*. For those unacquainted with it, here is a short explanation. An equivalent of the assertion '$\mathcal{R}$ is a well-founded partial ordering on X' is: '$\mathcal{R}$ is a partial ordering on X with the property that each $Y \subseteq X$ contains at least one minimal element'. ($y \in Y$ is a *minimal* element of Y, if for no $y' \in Y$, $y'\mathcal{R}y$.)
Assume now that $\mathcal{R}$ is a well-founded relation on X. For a proof of $\forall x A(x)$, it is sufficient now to prove that $\forall x(\text{ if } \forall y \mathcal{R} x A(y), \text{ then } A(x))$. For, if $\forall x A(x)$ does not hold, then there has to be a minimal x such that $A(x)$ does not hold, and for that x it has to be the case that $\forall y \mathcal{R} x A(y)$, i.e. it is sufficient to exclude such an occurrence.

Corollary 61 The following rule is a derived rule of **L**:

$$\frac{\boxdot(p \leftrightarrow \varphi(p, q_1, \ldots, q_n)) \to \psi(q_1, \ldots, q_n)}{\psi(q_1, \ldots, q_n)}$$

All propositional variables present are indicated; in particular, p does not occur in φ. (The validity of this rule for **L** means that in deriving theorems in **L** one can use fixed points without a problem.)

Proof: Exercise 44. •

Proof of theorem 58. We will write mostly $\varphi(p)$ for $\varphi(p, q_1, \ldots, q_n)$ etc., because the variables $q_1, \ldots, q_n$ turn out to play no essential role in the proof. In the first place one ought to be clear that, if p occurs only under $\Box$ in $\varphi(p)$, $\varphi(p)$ can be written as $\xi(\Box\varphi_1(p), \ldots, \Box\varphi_k(p))$, where $\xi(r_1, \ldots, r_k)$ is a formula not containing p.

We first prove the theorem for $\varphi(p)$ of the form $\Box\varphi_1(p)$. A fixed point δ of $\Box\varphi_1(p)$ is $\Box\varphi_1(\top)$, i.e. it is the case that

$$\vdash_{\mathbf{L}} \Box\varphi_1(\top) \leftrightarrow \Box\varphi_1(\Box\varphi_1(\top))$$

To show that this is indeed the case, it is sufficient to show that the equivalence holds on all well-founded, transitive models. First assume $w \models \Box\varphi_1(\top)$. Then, because of transitivity, also $w' \models \Box\varphi_1(\top)$ for all $w'\widetilde{\mathcal{R}}w$. That means that for all $w'\widetilde{\mathcal{R}}w$ the forcing relations for $\top$ and $\Box\varphi_1(\top)$ are identical. But the $w'\widetilde{\mathcal{R}}w$ are the only worlds which are relevant to the determination of the forcing relation in w for $\Box\varphi_1(\top)$ and $\Box\varphi_1(\Box\varphi_1(\top))$, because $\top$ and $\Box\varphi_1(\top)$ occur in those formulas only under $\Box$. So it follows from $w \models \Box\varphi_1(\top)$ that also $w \models \Box\varphi_1(\Box\varphi_1(\top))$.

Next assume that $w \not\models \Box\varphi_1(\top)$. Now we use the following fact: if one runs over a $\mathcal{R}$-chain in a model from left to right, then the valuation of a $\Box$-formula changes at most once, and only from not-forced to forced, because if such a formula is forced, then it remains forced further down, because of transitivity; furthermore such a formula will end up being forced in any case, because it is surely forced in the end points. With the aid of this fact it follows now from the well-foundedness of $\mathcal{R}$ that there is a last $w'\widetilde{\mathcal{R}}w$, on which $\Box\varphi_1(\top)$ is not forced, i.e. on all $w"\mathcal{R}w'$ $\Box\varphi_1(\top)$ is forced. Now $\Box\varphi_1(\top)$ behaves for all $w"\widetilde{\mathcal{R}}w'$ exactly as $\top$. i.e., with the same reasoning as above, $w' \not\models \Box\varphi_1(\Box\varphi_1(\top))$. But then also $w \not\models \Box\varphi_1(\Box\varphi_1(\top))$, which was to be shown.

The case that $\varphi(p)$ is of the form $\xi(\Box\varphi_1(p))$ is left as Exercise 46. One has to prove (this is done syntactically, using of the first case) that the fixed point δ of this formula is $\xi(\Box\varphi_1(\xi(\top)))$.

Finally we look at the general case that $\varphi(p)$ is of the form

$$\xi(\Box\varphi_1(p), \ldots, \Box\varphi_k(p))$$

Or rather, we take $k = 2$: the really general case is not more difficult, but gives rise to notational complications which hide the issue. So, we assume that $\varphi(p)$ is of the form $\xi(\Box\varphi_1(p), \Box\varphi_2(p))$. We first look at the formula $\varphi(p_1, p_2)$, i.e. $\xi(\Box\varphi_1(p_1), \Box\varphi_2(p_2))$. For this formula we can apply the second case with respect to p_1. We then get

$$(*) \vdash_{\mathbf{L}} \delta_1(p_2) \leftrightarrow \xi(\Box\varphi_1(\delta_1(p_2)), \Box\varphi_2(p_2)),$$

where $\delta_1(p_2)$ is the formula $\xi(\Box\varphi_1(\xi(\top, \Box\varphi_2(p_2)), \Box\varphi_2(p_2))$. In this last formula $\delta_1(p_2)p_2$ occurs exclusively in the context $\Box\varphi_2(p_2)$. So also here the second case is applicable, now with respect to p_2: there is a formula δ such that $\vdash_{\mathbf{L}} \delta \leftrightarrow \delta_1(\delta)$. Substitution of δ for p_2 in $(*)$ gives $\vdash_{\mathbf{L}} \delta_1(\delta) \leftrightarrow \xi(\Box\varphi_1(\delta_1(\delta)), \Box\varphi_2(\delta))$. The fact that $\vdash_{\mathbf{L}} \delta \leftrightarrow \delta_1(\delta)$ means however that $\delta_1(\delta)$ can be replaced everywhere in provable formulas by δ while provability is preserved, and, if we do that in the last reached formula, then we get $\vdash_{\mathbf{L}} \delta \leftrightarrow \xi(\Box\varphi_1(\delta), \Box\varphi_2(\delta))$: i.e. δ is the desired fixed point.

For more information on provability logic, see C. Smorynski: *Self-Reference and Modal Logic* , Springer 1985, or G. Boolos, *The Unprovability of Consistency* , Cambridge University Press, 1979.

4.5 Exercises

Exercise 34 Prove proposition 39. (Hint: use the substitution $p \wedge \Box p$ for p in the axiom scheme.)

Exercise 35 Show that $\mathbf{L}$ is not strongly complete by showing that compactness of $\mathbf{L}$ with respect to Char($\mathbf{L}$) fails for the set:
$\{\Diamond p_0, \Box(p_0 \to \Diamond p_1), \Box(p_1 \to \Diamond p_2), \ldots, \Box(p_n \to \Diamond p_{n+1}), \ldots\}$.

Exercise 36 Prove:

(a) If $\Box\psi_1, \ldots, \Box\psi_n, \psi_1, \ldots, \psi_n, \Box\chi \vdash_{\mathbf{L}} \chi$, then $\Box\psi_1, \ldots, \Box\psi_n \vdash_{\mathbf{L}} \Box\chi$.

(b) If $\Delta \cup \{\neg\Box\chi\}$ is $\mathbf{L}$-consistent, then so is $\{\psi, \Box\psi \mid \Box\psi \in \Delta\} \cup \{\Box\chi, \neg\chi\}$.

Exercise 37 Prove that, if $\langle \mathcal{F}, \mathcal{U} \rangle \models \varphi(p_1, \ldots, p_n)$, then also $\langle \mathcal{F}, \mathcal{U} \rangle \models \varphi(\psi_1, \ldots, \psi_n)$ for any $\psi_1, \ldots, \psi_n$

Exercise 38 Give a countermodel against $\Box p \to \Box\Box p$ on the generalized frame of figure 13.

Exercise 39 Where in the proof of theorem 48 did we use the fact that we had a generalized frame and not a standard one? Show that on the corresponding standard frame $\Box(\Box p \leftrightarrow p) \to \Box p$ is not valid.

Exercise 40 Give an arithmetical representation of:

(a) $Valuation(x, y)$: y codes a valuation for x.

(b) $Taut(x)$: x is a gödel number of a tautology.

Exercise 41 (a) Prove the left out step in the proof of Löb's theorem.

(b) Prove theorem 53 in two ways: (i) by slightly adapting the proof of theorem 52; (ii) by proving 'modally' that $\Box(\Box\varphi \to \varphi) \to \Box\varphi$ follows from the scheme $\Box\varphi \to \Box\Box\varphi$ and the rule embodied by Löb's theorem: $\Box\varphi \to \varphi/\varphi$ (use a substitution!).

Exercise 42 Prove theorems 54 and 55.

Exercise 43 . Prove proposition 57. What does the proposition mean for **PA**?

Exercise 44 Prove corollary 61.

Exercise 45

(a) Find fixed points of $\Box p$, $\Box\neg p$, $\Box(p \to q)$ and show directly in **L** that the formulas found are fixed points.

(b) Find fixed points of $\neg\Box p$, $\Box p \to q$ and show directly in **L** that they are fixed points.

Exercise 46 Prove that, if $\varphi(p)$ is of the form $\xi(\Box\varphi_1(p))$ with ξ a standard connective, then a fixed point δ of that formula is $\xi(\Box\varphi_1(\xi(\top)))$.

Exercise 47 Formulate as exactly as possible what the Fixed point theorem and the Unicity theorem for fixed points imply for **PA**. In particular consider the formula $\neg\Box p$.

Exercise 48

(a) Write the fixed point δ in the proof of the fixed point theorem 58 out for the case k=2,

(b) Find a fixed point of $\Box(\Box p \to q) \wedge \neg\Box p$ (not necessarily using (a); show directly in **L** that it is a fixed point.

(c) Give the proof for k=3.

Chapter 5

Tense logic

In this chapter the focus switches from the logic of necessity and possibility to the logic of another philosophical topic, time. Although the intensional approach is not the only reasonable way to tackle problems and questions related to the logical analysis of time and temporal expressions, we will restrict ourselves (the historical exposition in the first section being an exception) to this approach. Moreover, in connection with tense logic we will concern ourselves exclusively with establishing (whenever possible, strong) completeness of a number well-known tense-logical systems, leaving all other questions, interesting or not, aside.

We start with defining the syntax and semantics of this kind of tense logic as an intensional logic. In order to get the right perspective to view the intensional approach, we follow this up with a section with a few historical remarks and some hints to other possible approaches to the logic of time. Finally, returning to Kripke-models the technical machinery obtained will be set to work on proving completeness in a number of cases.

5.1 Introduction

The syntax and semantics of tense logic Modality and time may seem quite unrelated from a philosophical point of view; from a logical point of view they resemble each other to a high degree, the syntax as well as the semantics of both logics can be defined in such a manner that they share the same features. On the syntactic side the congeniality shows in the fact that a tense-logical language, as a modal one, is obtained from a language of the standard propositional calculus by adjoining some sentence operators. In the case of tense logic two new operators have to be added: G (to be interpreted as 'It is always **g**oing to be the case that ...') and H (standing for 'It **h**as always been the case that ...'). The rôle of these operators resembles the one $\Box$ plays in modal logic and they equally have duals F (standing for 'will be the case at least once in the **f**uture', defined as

$\neg G \neg$) and P (standing for 'has been the case at least once in the past', defined as $\neg H \neg$).

On the semantic side the similarity of the two logics comes to light in the fact that their languages are interpreted in the same kind of structures. A frame for tense-logical systems also consists of a tuple $\langle \mathcal{T}, < \rangle$ with $\mathcal{T}$ a non-empty set, the elements t of which are now to be thought of as points of time, and $<$ a relation that, in accordance with this interpretation of $\mathcal{T}$, represents the 'is earlier than' relation between the instances in $\mathcal{T}$. The exact definition of the semantics can be copied from its modal analogue (cf.definition 3), apart from some notational changes and with the definition of $\mathcal{V}_{\mathcal{M},w}(\Box\psi) = 1$ in the second clause replaced by

$$\mathcal{V}_{\mathcal{M},t}(G\psi) = 1 \text{ iff for all } t' \in \mathcal{T} \text{ such that } t < t', \mathcal{V}_{\mathcal{M},t'}(\psi) = 1$$

and

$$\mathcal{V}_{\mathcal{M},t}(H\psi) = 1 \text{ iff for all } t' \in \mathcal{T} \text{ such that } t' < t, \mathcal{V}_{\mathcal{M},t'}(\psi) = 1$$

Some history Before turning to the technical side of this logic we would like to spend some words on matters falling outside the scope of strictly 'Priorean'tense logic in order to give a quick impression of what is going on in the field of the logic of time.

According to its founder, A. N. Prior, the history of modern intensional tense logic has a rather paradoxical beginning in the writings of the philosopher McTaggart who aimed to prove the unreality of time. The argument, much discussed in books and articles on tense logic, hinges on the impossibility to reduce tensed sentences (as 'It has been raining', 'It will be raining') to untensed ones, stating that some event (the raining) applies at a certain moment in the past, present or future in a non-temporal fashion.

Today it may seem obvious that temporal expressions can be dealt with as they stand, but the argument must be seen against its historical background. At that time the prevalent view on sentences was essentially Fregean: all statements were supposed to express some completely determined proposition (for sentences of a natural language the proposition expressed can in general be identified with what the sentence states to be the case, when one includes all contextual information, including the moment of utterance of the sentence), and this proposition was supposed to have a truth value for all eternity.

In 1941, J. N. Findlay, however, wrote the following much quoted sentence: 'our conventions with regard to tenses are so well worked out that we have practically the materials in them for a formal calculus' and he added: 'the calculus of tenses should have been included in the modern development of modal logics.' Through these statements, Findlay deserves the credit for having been the true prophet of the developments to come.

The first person who actually carried out such a plan (in concurrence also with the forecast affiliation to modal logic) was Prior. By his own account (in

Past, Present and Future, 1967; chapter 1) the occasion for him undertaking this project was that the above mentioned trend of 'detensing' tensed sentences was employed lavishly in the expounding of ancient texts. Thus, one reasoned that a sentence as 'Socrates is sitting' (to quote the classical example) occurring in a text of e.g. Aristotle has to be completed with a phrase like 'at point of time t' before truth or falsity can be predicated of it.

In a critical review of such an analysis Geach pointed out that this way of looking at such sentences was alien to ancient minds: sentences as the one above had been taken to be complete and fit to be called "true" or "false" as they stand; only the verdict which of two opposites applies and not the sentence itself (or its propositional content, if you like) changes in time.

This remark not only made Prior, who was up till then also under the spell of the Fregean view, acutely aware of the need for a theory as exposed in the previous section, but also directed his mind to the manner in which this theory should take shape. In Priorean tense logic then, propositional variables receive truth values relative to points of time, and the four sentence operators F, G, H and P are added to capture the differences between differently tensed variants of the same sentence.

The above mentioned Fregean view on the nature of temporal utterances can be formalized by means of a two-sorted predicate calculus. This calculus, a modern proponent of which is Needham, has besides the standard individual variables also a set of variables ranging over points of time. In this logic, propositional letters get converted into predicate variables with one argument-place interpreted as ranging over instances of time. The resulting logic is much stronger than the Priorean one; a defect according to some and an advantage according to others. In an article by van Benthem (*Tense logic and Standard logic*, to be found in *Tense Logic*, Åqvist and Guenthner (eds.), 1977), in which the relative merit of the two kinds of logic is extensively discussed, it is argued that, when to account for more phenomena one keeps increasing the complexity of Priorean tense logic, this logic converges to the predicate calculus in any case.

Another rival theory of time stems from Reichenbach. This theory claims that in principle every temporal expression needs for its interpretation three points of time from the extra-linguistic context: first there is the *moment of speech* (S) , secondly the time of occurrence of the *event* (E) the sentence describes, and thirdly there is the so called *point of reference* (R) that denotes a kind of stand-point in time from which the event is judged. The advantages gained by this approach are mainly to be found in the linguistic sphere: so, for example, it is possible in this theory to reflect the difference between e.g. 'It rained' ($E = R < S$), 'It has been raining' ($E < R = S$) and 'It had been raining' ($E < R < S$), which is not accomplished by the Priorean approach. An objection is that, in general, one point of reference does not seem to be sufficient: 'It would have been raining' (to be read in a purely temporal sense) already needs two such points. For linguistic purposes a combination of both approaches may be profitable.

A comparison of tense logic in the Priorean approach to natural language shows that the first is rather meagre in expressive power, especially when one looks at temporal adverbials; even common ones like *now*and *since* cannot be defined by intensional tense logic in its basic form. In investigations concerning the expressive power of tense logic Hans Kamp proved that Priorean tense logic can be made 'temporally complete' (a notion that qua significance is best compared with (truth-)functional completeness of the standard propositional calculus), but only by adjoining binary operators expressing *since* and *until* to the operators F, G, H and P.

Last but not least, there is a recently popular trend of loosening the language of Priorean tense logic from Kripke-semantics, replacing it with a kind of semantics in which intervals of time are primitive. Since intervals are assumed to be less abstract than instances of time, this approach is judged to make a smaller ontological commitment, and hence to be superior from a philosophical point of view. Moreover, the description of certain features of natural languages, closely related to temporal constructions, seems to rely heavily on the use of intervals. One of these features so-called *aspect*, indicates the manner in which an event takes place in time, e.g. whether an event should be considered as a unit, or as something taking a stretch of time in which other events can take place. (Compare for instance 'When Queen Anne died, the Whigs brought in George' with 'While Queen Anne was dying, the Jacobites hatched treasonable plots' to give a classical example.)

5.2 Completeness of tense logics

5.2.1 The minimal tense logic $\mathbf{K}_t$

Because a tense logic is nothing but a kind of modal logic having one 'forward-looking' $\Box$-operator (viz. G) and one 'backward-looking' (viz. H), little reflection suffices to realize that in order to get a tense-logical system for (i.e. sound and strongly complete with respect to) the class of all time frames we will at least have to duplicate the rules and axioms of the modal minimal logic **K** for both operators G and H. It only remains then to secure that the direction in which H-sentences are evaluated is the converse of the direction in which one has to look for the evaluation of G-sentences. This can be accomplished by adding as a third axiom two special schemes relating the future and past tense operators. Thus $\mathbf{K}_t$ is axiomatized by:

Axiom 1 all formulas having the form of a tautology,

Axiom 2 $G(\phi \to \psi) \to (G\phi \to G\psi)$ and its mirror-image $H(\phi \to \psi) \to (H\phi \to H\psi)$

Axiom 3 $PG\phi \to \phi$ and $FH\phi \to \phi$.

Rules : modus ponens, $\phi/G\phi$ and $\phi/H\phi$.

The verification of the claim that the axioms for $\mathbf{K}_t$ are indeed correctly chosen is merely a matter of routine by now:

Since $\mathbf{K}_t$'s soundness with respect to the class of all tense-logical frames goes almost without saying, we skip its proof. The proof of strong completeness of $\mathbf{K}_t$ with respect to the same class by means of the method H runs smoothly.

Theorem 62 $\mathbf{K}_t$ is strongly complete with respect to the class of all tense-logical frames.

Proof: All the necessary definitions, lemmas and theorems can be transferred from the analogous case of the modal logic $\mathbf{K}$. The only thing that has to be checked is, whether defining the canonical relation $\prec$ on maximal $\mathbf{K}_t$-consistent sets Γ, Δ for all formulas φ as: $\Gamma \prec \Delta$ iff for all φ, if $G\varphi \in \Gamma$, then $\varphi \in \Delta$ (just as in the modal case) also comes out right for H. (That this holds good is, of course, the raison d' être of the two schemes of axiom 3.) We give this as the next proposition. •

Proposition 63 The following four stipulations yield equivalent results:

(i) $\Gamma \prec \Delta$, i.e. for all φ, if $G\varphi \in \Gamma$, then $\varphi \in \Delta$.

(ii) For all φ: if $\varphi \in \Delta$, then $F\varphi \in \Gamma$.

(iii) For all φ: if $\varphi \in \Gamma$, then $P\varphi \in \Delta$.

(iv) For all φ: if $H\varphi \in \Delta$, then $\varphi \in \Gamma$

Proof: We prove one implication, viz. that (ii) implies (iii), leaving the others as exercise 49. Assume (ii) and suppose $P\varphi \notin \Delta$. Then $\neg P\varphi \in \Delta$, or equivalently, $H\neg\varphi \in \Delta$. Because of (ii) we can conclude to $FH\neg\varphi \in \Gamma$. Since $FH\neg\varphi \to \neg\varphi$ is an axiom 3, we then have $\neg\varphi \in \Gamma$, hence $\varphi \notin \Gamma$. •

We can now define a tense logic $\mathbf{S}$ in general to be a system containing $\mathbf{K}_t$ closed under the same operations as in the modal case: inferences by modus ponens, substitution and temporal necessitation. It will also have become clear by now that many notions, definitions, theorems and proofs for tense logic can simply be transferred from modal logic, in most cases needing no more than some slight notational adjustments. We will just refer to the chapters on modal logic, whenever it is convenient to do so.

5.2.2 The logic of linear frames: Lin

To get more interesting logical theories about time assumptions about its structure have to be joined to the bare, rudimentary, system $\mathbf{K}_t$. There are several ways in which the details of a logical theory of time can be given shape; what the best choice is may depend on the phenomena in the context of which time is considered. First of all we are going to consider a rather modest system, that is basic in the sense that it will be included in all other systems we will take into consideration further on.

Most people will find the following a more than reasonable assumption: the 'River of Time' can be represented as an axis of which every point t divides the set of all others into two complementary halves: all points to the left of ('earlier than') t, constituting t's past, and all points to the right of ('later than') t constituting t's future.

To this picture the following mathematical notion corresponds:

Definition 22 A relation $\mathcal{R}$ is a *strict linear ordering* iff (i) $\mathcal{R}$ is transitive, (ii) $\mathcal{R}$ is irreflexive and (iii) $\mathcal{R}$ is weakly connected (i.e. $\forall xy(x\mathcal{R}y \vee x = y \vee y\mathcal{R}x)$, henceforth we omit the 'weakly'). The properties (i) and (ii) define a so-called *strict partial ordering*; together they imply that $\mathcal{R}$ is asymmetric also. *Non-strict linear* and *partial orderings* are defined similarly except that they are reflexive instead of irreflexive.

Striving for a tense logic of time that forces the 'earlier than'-relation to satisfy this complex of properties presents a problem we have not encountered in the modal sphere. For one glance at this definition must, after the material offered in chapter 2, suffice to realize that this combination of properties cannot be characterized: transitivity readily translates in $G\varphi \to GG\varphi$ (or $H\varphi \to HH\varphi$), but, both irreflexivity, and weak connectedness, fall outside the scope of the kind of language we have at our disposal. So, finding a logic of the required sort becomes another kind of enterprise then it was in the modal context.

The reason why this problem crops up in connection with tense logic, whereas it did not in the modal case, can be explained as follows: In modal logic our main interest concerns acquiring a set of plausible axioms that serve as principles governing the behaviour of modalities. As argued in chapter 1, Kripke-frames can play a part when it comes to testing what the less perspicuous formulas signify, but once a reasonable set of axioms is acquired, any frame validating it will do: the picture of a set of possible worlds and the fashion in which they are connected to their alternatives is not put to the test by a similar picture of this kind already existing in our mind. The contrary holds for tense logic, where we work in the opposite direction. Here we already have a fairly clear-cut picture in mind of what a structure can, or must, look like if it is to qualify as a reasonable candidate for representing the structure of time. It is not so miraculous that such a stringent demand cannot be pinned down by the rather coarse language we use.

Although this circumstance may hamper our search for a logic fitting to the structure of the linear frames, it does not bring us to the end of our resources. For it may still be possible — and in fact it will turn out to be the case indeed — that there exists a logic that characterizes a larger class of frames $\mathcal{C}$ that includes all linear ones, but is strongly complete with respect to this subclass of $\mathcal{C}$; we know such a thing to be possible: we have for example already encountered something like this in proving by means of the method C that **S4** is strongly complete with respect to one single frame in its characteristic class, the tree $\mathcal{T}_{\omega\omega}$. This fact already suggests the technique which can be used for our enterprise, once we have spotted the logic whose existence we claimed above; the consequence that this tense logic will allow non-standard models of time is something we will simply have to live with.

In trying to get as close as possible to linearity one quickly realizes the following: with respect to irreflexiveness nothing can be accomplished, but we saw already in proposition 9(c) that we can express something very much like linearity: that is to say, we can at least force all frames to consist of (possibly more than one and possibly featuring reflexive points, but still) straight lines. The current technical term for this property is highly suggestive and is defined as follows.

Definition 23 A relation $\mathcal{R}$ is

(i) *not branching towards the future* iff
$\forall x, y, z((x\mathcal{R}y \wedge x\mathcal{R}z) \rightarrow (y\mathcal{R}z \vee y = z \vee z\mathcal{R}y))$.

(ii) *not branching towards the past* iff
$\forall x, y, z((z\mathcal{R}x \wedge y\mathcal{R}x) \rightarrow (y\mathcal{R}z \vee z = y \vee z\mathcal{R}y))$.

(iii) *not branching* iff it is neither branching towards the future nor towards the past.

These properties can be characterized by the following formulas:

Proposition 64

(i) $Fp \rightarrow G(p \vee Pp \vee Fp)$ characterizes the class of frames that are not branching towards the future.

(ii) $Pp \rightarrow H(p \vee Pp \vee Fp)$ characterizes the class of frames that are not branching towards the past.

Proof: Exercise 50. •

Adjoining these two schemes, together with the axiom for transitivity, to $\mathbf{K}_t$ yields a logic (called **Lin**) which is sound and strongly complete with respect to the class of all linear frames was already predicted by Prior, but a proof of this fact had to await the appearance of K. Segerberg. The tense logic **Lin** then, the logic axiomatized by $\mathbf{K}_t + G\varphi \rightarrow GG\varphi + F\varphi \rightarrow G(\varphi \vee P\varphi \vee F\varphi) + P\varphi \rightarrow H(\varphi \vee P\varphi \vee F\varphi)$, can be shown to be strongly complete with respect to *Char*(**Lin**) by the use of the following lemma.

Lemma 65

(i) If $F\varphi \to G(\varphi \vee P\varphi \vee F\varphi)$ is a theorem of the tense logic **S**, then its canonical relation $\prec$ (defined as $\mathcal{R}_{\mathbf{S}}$ in definition 12) is not branching towards the future.

(ii) Analogous for $P\varphi \to H(\varphi \vee P\varphi \vee F\varphi)$

Proof: Exercise 51. •

Together with the analogous proposition for $G\varphi \to GG\varphi$ proved in the modal case, this lemma is sufficient to show that the $\prec$ of the canonical model for **Lin** has all the required properties and hence **Lin** is strongly complete with respect to it. Segerberg managed to rebuild this canonical frame into a strictly linear one, thereby arriving at the desired conclusion. The problem is not in getting a linear frame (one gets one starting from a maximally **Lin**-consistent set Γ by taking $\{\Delta \mid \Delta\mathcal{R}\Gamma \text{ or } \Delta = \Gamma \text{ or } \Gamma\mathcal{R}\Delta\}$), but in changing that into a strictly linear one. We have already announced our intention however to follow another road to the same goal; the proof by means of the method C is easier in this case.

The construction procedure to be used has to deviate slightly from the one used in proving strong completeness of **S4** (cf. definition 13), because, as you will see in a moment, we do not construct the model continuously proceeding either towards the future or towards the past, but we are forced to place new points between two old ones and therefore we have to have a method to keep track of which formulas have already taken care of where.

Proposition 66 **Lin** is strongly complete with respect to the class of all linear frames.

Proof: Assume $\Delta \nvdash_{\mathbf{Lin}} \varphi$. It suffices to construct a linear frame $\langle \mathcal{T}, < \rangle$ with a maximal **Lin**-consistent set Γ_t associated to each $t \in \mathcal{T}$ in such a manner that the following conditions are met:

(a) For some $t \in \mathcal{T}$, Γ_t is a maximally **Lin**-consistent extension of $\Delta \cup \{\neg\varphi\}$.

(b) For all t, $t' \in \mathcal{T}$: if $t < t'$, then $\Gamma_t \prec \Gamma_{t'}$.

(c) For all φ and $t \in \mathcal{T}$: if $\neg G\varphi \in \Gamma_t$, then in $\mathcal{T}$ there some a $t' > t$ such that $\neg\varphi \in \Gamma_{t'}$.

(d) For all φ and $t \in \mathcal{T}$: if $\neg H\varphi \in \Gamma_t$, then in $\mathcal{T}$ there is some $t' < t$ such that $\neg\varphi \in \Gamma_{t'}$.

The frame $\langle \mathcal{T}, < \rangle$ will be constructed in stages. After each stage n we will have a (finite) linearly ordered frame $\langle \mathcal{T}_n, < \rangle$, denoted $\mathcal{T}_n$ for short, satisfying conditions (a) and (b). At stage $n+1$ a formula of the form $\neg G\varphi$ or $\neg H\varphi$ is taken care of *with respect to the result of the previous stage*: depending on the situation in this frame we add a point of time t, with associated Γ_t, at the appropriate spot, or else leave everything the way it is. Since inserting a point can again undo our efforts in relation to formulas treated at previous stages, we need an enumeration of all formulas of the form indicated in which every one of them is repeated infinitely many times. This can be done in the following way: Fix an enumeration $\varphi_0, \varphi_1, \varphi_2, \ldots$ and then take as a new enumeration $\varphi_0, \varphi_0, \varphi_1, \varphi_0, \varphi_1, \varphi_2, \varphi_0, \varphi_1, \varphi_2, \varphi_3, \varphi_0, \varphi_1, \varphi_2, \varphi_3, \varphi_4, \ldots$.

We can now formulate the construction procedure in a neat manner:

Stage 0 $\mathcal{T}_0 = \{t_0\}$ with which we associate Γ_{t_0}, a maximally **Lin**-consistent extension of $\Delta \cup \{\neg\varphi\}$.

Stage n+1 Assume the formula to be treated is of the form $\neg G\varphi$; the argument for $\neg H\varphi$ is analogous. We are going to fashion $\mathcal{T}_{n+1}$ in such a way that in relation to the elements of $\mathcal{T}_n$ condition (c) holds for $\neg G\varphi$. Three possible situations can occur:

1. $\neg G\varphi \notin \Gamma_t$ for all $t \in \mathcal{T}_n$. In that case nothing has to be done; we can take $\mathcal{T}_{n+1} = \mathcal{T}_n$.
2. $\neg G\varphi \in \Gamma_t$ for some t, but in all such cases there is already some $t' > t$ in $\mathcal{T}_n$ with $\neg\varphi \in \Gamma_{t'}$. We can again take $\mathcal{T}_{n+1} = \mathcal{T}_n$.
3. The only situation that calls for action is one in which for some t, $\neg G\varphi \in \Gamma_t$ and for all $t' > t$, $\varphi \in \Gamma_{t'}$. So, suppose this to be the case and let t be the last element (i.e. the largest in the $<$-ordering) of $\mathcal{T}_n$ with that property. As in the case of **K** there exists a maximal **Lin**-consistent Σ such that $\neg\varphi \in \Sigma$ and $\Gamma_t \prec \Sigma$. This Σ we can associate as Γ_{t^*} to a point t^* which we insert behind t as its immediate successor. If t happened to be the last element of $\mathcal{T}_n$ it is clear that t^* can be added to $\mathcal{T}_n$ at the end. In case t is not the last element of $\mathcal{T}_n$ we have to check that putting t^* into $\mathcal{T}_n$ in the way mentioned does not clash with condition (b) with respect to the other elements of $\mathcal{T}_n$. So assume that t' is the immediate successor of t in $\mathcal{T}_n$. Because $\prec$ is not branching towards the future, either $\Gamma_{t'} \prec \Gamma_{t^*}$, or $\Gamma_{t'} = \Gamma_{t^*}$ or $\Gamma_{t^*} \prec \Gamma_{t'}$ must hold. The first possibility is ruled out because $G\varphi \in \Gamma_{t'}$ (for we took t to be the last point with $\neg G\varphi$) and $\neg\varphi \in \Gamma_{t^*}$. And the second possibility can be discarded by the observation that $\neg\varphi \in \Gamma_{t^*}$. Hence we can set $\mathcal{T}_{n+1} = \mathcal{T}_n \cup \{t*\}$.

Finally, it is obvious that $\bigcup_{n\in\mathbb{N}} \mathcal{T}_n$, which clearly satisfies (a)-(d), is the kind of frame we are looking for. •

5.3 The frames $\mathbb{Q}$, $\mathbb{R}$ and $\mathbb{Z}$

To come to a full-blown picture of time, the decision to represent the succession of moments as points on a linear axis must be amplified with conditions concerning the pattern in which this succession is to be fashioned.

First of all we adopt a pair of assumptions to the effect that time has no beginning or end. From a scientific point of view this is not an uncontroversial decision, but it is rather innocent in the sense that either can easily be removed whenever this is thought to be desirable. In the meantime one can think of these assumptions as being at least on the safe side, procuring potentially all points of time we will possibly need. A more important choice has to be made between the following two options: either time consists of an infinitely divisible flow (between every two points there is a third), or the passage of time 'jumps' from one moment to the next; intuitively it is clear that whatever holds for one moment in this respect must hold for all and that consequently the alternatives must be universal statements excluding the existence of distinct moments or intervals displaying a divergent, not eternally recurring, pattern. Both assumptions can furthermore be supplemented with the demand of continuity, signifying that time uninterruptedly goes by and that no breaking points can be pointed out.

Four kinds of structures emerge from these considerations that merit closer examination. But before turning to the systems of tense logic adhering to them, we would like to settle a matter of a more philosophical nature first. Perhaps it surprises you that different incompatible pictures of time peacefully co-exist in our investigation: are not models supposed to represent things as they are, leaving no room for variation? Of course this is not a valid objection. Deciding which type of structure fits the facts best would be a problem for the empirical sciences rather than a matter of logic, whose task it is merely to explicate what choosing for a particular kind of structure out of some stock of potentially interesting ones amounts to. But apart from this, one should not forget that points of time, being highly abstract, belong to the "theoretical network" put over reality by our minds, rather than to the phenomena themselves. Therefore, we are free, at least to a large extent, to structure them as we wish, the only restriction to this freedom stemming naturally from the demand that the way in which we accomplish it must provide suitable and expedient means to represent the sort of events we are interested in. And the seeming abundance of structures borrows its right to existence precisely from the fact that different kinds of events call for different kinds of representational means (the differences consisting for example in the scale on which they take place, the purposes we have in studying them). This is also quite clearly revealed in the different ways in which different phenomena are described in natural language. To give you an impression of what is meant, each of the structures that will be discussed in the sequel will be accompanied with an example of an event for which the kind of structure in question is particularly appropriate.

5.3.1 The logic Q of dense time

We start with examining a logic that portrays the passage of time as an infinitely divisible set of points. On linear frames, this property is captured by the mathematical notion of density (see definition 7 for the definition of this property) and density can be characterized in the language of tense logic by the formula $GG\varphi \to G\varphi$ or, equivalently, by $HH\varphi \to H\varphi$. (You can check for yourself that both express this property and characterize the class of dense frames.) If we add this formula as an axiom scheme to **Lin**, in addition to the axioms $P\top$ and $F\top$, expressing that time has neither beginning, nor end, we arrive at a logic that we name **Q** (for reasons that will soon become clear). Hence **Q** is axiomatized as **Lin** + $P\top$ +$F\top$ +$GG\varphi \to G\varphi$. This logic can be shown to be (sound and) strongly complete with respect to the class of all dense and linear, serial frames. But we can do even better: since we are working with a countable language we can be sure that in establishing this proposition by means of the method C the model obtained will be countable too. Hence, we can apply Cantor's theorem that all countable dense, serial orderings are isomorphic to the rationals with their usual ordering (it is unlikely you have never seen his zigzag method proof of this fact) to get the following theorem.

Proposition 67 **Q** is strongly complete with respect to $\langle \mathbb{Q}, < \rangle$

Proof: As in the completeness proof for **Lin**, the existence of all the points necessary to comply with the semantics of G and H is secured in stages. But this time the procedure by which that is accomplished is only used at the even-numbered stages, i.e. a φ_n that needs to be treated will be dealt with at stage $2n + 2$. Since $P\top$ appears in every maximally **Q**-consistent set, and as every formula is treated infinitely many times, we can rely on it to secure that the addition of new points to the left never ceases and similarly having $F\top$ as a **Q**-axiom guarantees the ever continuing addition of points to the right, whence the resulting model will have no end-point as required. At the odd stages $2n + 1$ density can be taken care of, by simply inserting between every t_k and t_{k+1} in $\mathcal{T}_k$ a new point u, associating with it a Γ_u such that $\Gamma_{t_k} \prec \Gamma_u \prec \Gamma_{t_{k+1}}$; the proof that an appropriate Γ_u exists is exercise 52. Although none of the finite $\mathcal{T}_n$ will be dense, infinite repetition of this insertion procedure insures density in the limit $\bigcup_{n\in\mathbb{N}} \mathcal{T}_n$, which we of course take again to be the final model, satisfying (a) to (d). Finally then, Cantor's theorem declares this model to be isomorphic to $\langle \mathbb{Q}, < \rangle$. •

As promised, we provide an example of an event for which dense time is not only suitable but even necessary, for the assumption that the set of instances of time must be taken to be densely ordered gets support from one of Zeno's famous paradoxes. The argument goes like this: It is impossible for an arrow to move from one point of space to another, because movement involves changing of position but on every point of time in which this changing should take place the arrow can occupy only one fixed portion of space equal to itself and hence is in rest. This argument is cogent under the assumption that every moment has an immediate successor. If however time is assumed to be (linear and) dense, between every two points of time there is a third in which case the contradiction disappears. To account for the event of a moving arrow a dense model is a necessary condition.

5.3.2 The logic of dense and continuous time: R

Although dense linear orderings arrange the points in such a manner that they all lie very closely together, they do not necessarily completely fill out the line: in $\langle \mathbb{Q}, < \rangle$ for example there is still room left where (even uncountably many) more points can been squeezed in. Filling in all such empty spots with new points results in what is called a 'continuous dense ordering'. The property of being continuous cannot be expressed in the first order predicate calculus; this is already clear from the fact that Cantor's theorem entails that the theory LODS of dense orderings without end-points is syntactically complete (i.e. for every first order formula φ it holds that, either $LODS \vdash \varphi$ or $LODS \vdash \neg\varphi$ in the predicate calculus). So it is only possible to add conditions by using heavier artillery. The definition of continuity will be preceded by that of a couple of notions in terms of which its meaning can be fixed.

Definition 24 Let $\langle \mathcal{T}, \mathcal{R} \rangle$ be a linear order. The tuple $\langle \mathcal{T}_1, \mathcal{T}_2 \rangle$ is a *cut* iff
(i) $\mathcal{T}_1 \neq \emptyset$, $\mathcal{T}_2 \neq \emptyset$ and $\mathcal{T}_1 \cap \mathcal{T}_2 = \emptyset$,
(ii) $\mathcal{T}_1 \cup \mathcal{T}_2 = \mathcal{T}$ and
(iii) if $t \in \mathcal{T}_1$ and $t' \in \mathcal{T}_2$, then $t\mathcal{R}t'$

(So, $\mathcal{T}$ is bisected by a cut into two complementary halves in a manner that respects the $\mathcal{R}$-ordering.) The cut $\langle \mathcal{T}_1, \mathcal{T}_2 \rangle$ determines a *gap* iff $\langle \mathcal{T}_1, \mathcal{R} \restriction \mathcal{T}_1 \rangle$ has no maximum (i.e. there is no $t \in \mathcal{T}_1$ such that for all $t' \in \mathcal{T}_1$ different from t, $t'\mathcal{R}t$) and $\langle \mathcal{T}_2, \mathcal{R} \restriction \mathcal{T}_2 \rangle$ has no minimum (this notion receiving a similar definition).

Finally, $\mathcal{R}$ is *continuous* iff no cut in $\langle \mathcal{T}, \mathcal{R} \rangle$ determines a gap.

With the help of the following pair of closely related notions we can sharpen our comparison between mere density and continuity:

Definition 25 Let $\langle \mathcal{T}, \mathcal{R} \rangle$ be as before and let $\langle \mathcal{T}_1, \mathcal{T}_2 \rangle$ be a cut in $\langle \mathcal{T}, \mathcal{R} \rangle$.
$\langle \mathcal{T}_1, \mathcal{T}_2 \rangle$ determines a *jump* iff $\langle \mathcal{T}_1, \mathcal{R} \restriction \mathcal{T}_1 \rangle$ has a maximum and $\langle \mathcal{T}_2, \mathcal{R} \restriction \mathcal{T}_2 \rangle$ has a minimum.
$\langle \mathcal{T}_1, \mathcal{T}_2 \rangle$ determines a *transition* iff either $\langle \mathcal{T}_1, \mathcal{R} \restriction \mathcal{T}_1 \rangle$ has a maximum or $\langle \mathcal{T}_2, \mathcal{R} \restriction \mathcal{T}_2 \rangle$ has a minimum.

Proposition 68 Let $\langle \mathcal{T}, \mathcal{R} \rangle$ be a linear ordering. Then the following holds:
(i) $\mathcal{R}$ is dense iff no cut in $\langle \mathcal{T}, \mathcal{R} \rangle$ determines a jump.
(ii) If $\langle \mathcal{T}, \mathcal{R} \rangle$ has no end-points and is countable in addition, then not every cut in it determines a transition.
(iii) If $\mathcal{R}$ is presumed to be continuous, then every cut determines a transition

Proof: The proof of these facts is left as exercise 53 (Consult a reference book on set theory if you are not familiar with these structures.) •

The interpretation of time as a continuous and densely ordered set of points implies among other things the existence of a definite point at the beginning and ending of every event. The point at the beginning may be the initial point of the event or the last moment of time before the event starts (and similarly at the end of the event). To this circumstance adhere advantages as well as disadvantages. For an example of the former we can stick to arrows: if one wants to maintain that an arrow, first ascending and then descending after having been released from a bow, must reach a highest point somewhere in its course one should better have continuous time, for otherwise it might be the case that the event coincides with a cut in time determining a gap in such a manner that in the first half the arrow ascends, while it descends in the later half, without a turning point ever occurring (note that as in the previous example concerning arrows the existence of points in space and time is intimately connected). But in many cases the existence of precise points of beginning or ending is less fortunate: often a change of the circumstances does not seem marked by one precise moment, and, if it is, it may seem arbitrary to count the change as having just occurred or being just about to occur. Think for example extinguishing a fire, when precisely can it be said to be out? (The example is taken from Johan van Benthem's *The Logic of Time.*) The identification of the situation marked by the fire's being out with some stretch of points on a continuous ordering calls for arbitrary stipulations. ('A fire is out if such and such chemical reactions no longer take place in the larger part of the material involved.')

Let us return to logic proper. In the language of tense logic the continuous frames can be isolated within the class of all linear frames by the formula

$$\Box(G\varphi \rightarrow PG\varphi) \rightarrow (G\varphi \rightarrow H\varphi)$$

where the necessity operator is used as an abbreviation for the temporal interpretation of necessity, i.e. $\Box\varphi$ denotes $\varphi \wedge G\varphi \wedge H\varphi$. In order to gain some handiness in handling this rather long formula it is advisable to prove the above claim that it characterizes continuity; we have left this as exercise 54.

The system that results from adjoining this scheme to **Q** is known as the logic **R** and this logic can be shown to be strongly complete with respect to all continuous and dense orderings without endpoints. Analogous to the case of **Q** however, we aim at something stronger: we want to establish that all non-theorems of **R** can be refuted on one particular frame out of this class, viz., as the name already suggests, $\mathbb{R}$, the set of reals ordered as usual.

Unfortunately, demanding denseness, continuity and the lack of end-points does not suffice to pin down this structure; to fix it up to isomorphism, the (higher-order) characteristic needed in addition is that the structure must contain a countable subset lying dense in it (i.e. between every two numbers, rational or irrational, there must be some rational number). Armed with this knowledge we can set out to prove the following

Theorem 69 **R** is strongly complete with respect to $\langle\mathbb{R}, <\rangle$.

Proof: We start with the structure $\langle\mathbb{Q}, <\rangle$ acquired in the previous completeness proof, exchanging only the maximally **Q**-consistent sets associated to every point of it for maximally **R**-consistent ones; this can be done without causing problems because **Q**$\subseteq$**R**. As will be seen, the remainder of the proof relies heavily on the fact that the conditions (a) to (d) of proposition 66 are already known to hold for this $\mathbb{Q}$-part.

Next, we complete this frame to a continuous one by inserting new elements, representing the ones in $\mathbb{R} \setminus \mathbb{Q}$, in the usual places. What then remains to be done is to associate maximally **R**-consistent sets to the irrational newcomers in such a fashion that the conditions (a) to (d) hold for these too. To achieve this goal we take for all $r \in \mathbb{R} \setminus \mathbb{Q}$, Γ_r to be a maximally **R**-consistent extension of the set $\{\varphi \mid G\varphi \in \Gamma_q$ for at least one of the $q \in \mathbb{Q}$ with $q < r\} \cup \{\psi \mid H\psi \in \Gamma_{q'}$ for at least one of the $q' \in \mathbb{Q}$ with $q' > r\}$.

Note that this choice of the Γ_r's forces them to be mutually related by $\prec$ in the right order, because between every two irrational associates a rational one, relating to the other rational ones in the correct way, can be found. This means that condition (b) applies, whereas condition (a) of course holds good on account of its being valid in the revised $\mathbb{Q}$-part.

We have to show that we are allowed to proceed in the manner described, in other words, that the sets mentioned are consistent. So suppose for some r it is not. Then there must be $q_1, \ldots, q_m < r$ and $q'_1, \ldots, q'_n > r$ with $G\varphi_i \in \Gamma_{q_i}$ and $H\psi_j \in \Gamma_{q'_j}$ ($q_i, q'_j \in \mathbb{Q}$; $1 \leq i \leq m$; $1 \leq j \leq n$) such that

$$\vdash_{\mathbf{R}} \neg(\varphi_1 \wedge \ldots \wedge \varphi_m \wedge \psi_1 \wedge \ldots \wedge \psi_n)$$

Since $\langle\mathbb{Q}, <\rangle$ is dense, there must however exist some q^* in it with $q_1, \ldots, q_m < q^* < q'_1, \ldots, q'_n$ and for this q^*, $\varphi_1, \ldots, \varphi_m, \psi_1, \ldots, \psi_n \in \Gamma_{q^*}$ must obtain (again, because (b) holds surely for the $\mathbb{Q}$-part), contradicting the **R**-consistency of Γ_q .

The only thing still left is to look after the satisfaction of conditions (c) and (d). To do this for the latter is exercise 55; the first we show now:

Assume $\neg G\varphi \in \Gamma_r$ and $\varphi \in \Gamma_q$ for all $q \in \mathbb{R}$ with $q > r$, and hence for all $q \in \mathbb{Q}$ with $q > r$. Because the latter Γ_q's are known to satisfy (a)-(d) we have $G\varphi \in \Gamma_{q'}$ for all $q' \in \mathbb{Q}$ with $q' > r$. Moreover, for the same reason and the same q' 's, we have $PG\varphi \in \Gamma_{q'}$ and hence $G\varphi \to PG\varphi \in \Gamma_{q'}$. Furthermore, we have $\neg G\varphi \in \Gamma_{q''}$ for all $q'' \in \mathbb{Q}$ such that $q'' < r$ (For, if not, then $G\varphi \in \Gamma_{q''}$, whence by the transitivity-axiom $GG\varphi \in \Gamma_{q''}$, so by the definition of $\Gamma_r, G\varphi \in \Gamma_r$, contradicting our previously made assumption.). Hence, by the falsum-rule, $G\varphi \to PG\varphi \in \Gamma_{q''}$

for all $q'' < r$ also. Therefore, $\Box(G\varphi \to PG\varphi) \in \Gamma_q$ for all $q \in \mathbb{Q}$. Applying modus ponens to the characteristic axiom of **R** yields $G\varphi \to H\varphi \in \Gamma_q$ for all these q and so $H\varphi \in \Gamma_{q'}$ for all $q' > r$. But this is inconsistent with the fact that $\neg G\varphi \in \Gamma_{q''}$ for (all and on that account) for some $q'' < r$ and that the $\mathbb{Q}$-part satisfies conditions (c) and (d). •

5.3.3 The logic of discrete time: D

We exchange the idea of time as a densely ordered flow for a picture in which the passage of time is imagined to be jumping from one moment to the next. That time can be assumed to pass like that may seem highly implausible if one has the kind of events mentioned in the earlier sections in mind. But with regard to (the description of) other kinds of temporal goings on, this image can be highly appropriate. To stay close to a previous example: the history of a toxophilitic society, regarded with respect to who successively won the bow games organized on the occasion of their annual feast can only gain in tractability when modeled in this fashion. Further, one can think of representations of all kinds of events involving measurement on a fixed scale(in laboratory or elsewhere): between two reports on possible changes that have occurred between the first measurement and a second there necessarily are intervals on which no data are obtained (e.g. the record of the growth of a child).

The mathematical notion corresponding to this assumption about the behaviour of the lapse of time is *discreteness*. This notion can be defined by the first order formulas

$\forall x \exists y (x < y \wedge \neg \exists z (x < z \wedge z < y))$ and

$\forall x \exists y (x > y \wedge \neg \exists z (x > z \wedge z > y))$,

that is, if we, as we intend to, stand by the conception of time having no first or last moment; otherwise the definition must be amended by clauses making exceptions of the end-points. As can be seen fairly easily (it is instructive to spell it out for yourself) in tense logic the property of being discrete can be characterized (on the linear frames) by the pair of formulas $(\varphi \wedge G\varphi) \to PG\varphi$, signifying that there is an immediate predecessor for every point, and $(\varphi \wedge H\varphi) \to FH\varphi$, implying for every point the existence of an immediate successor. This pair of formulas, taken as schemes, added together with the of axioms concerning end-points to the system **Lin**, results in a logic named **D** (i.e. **D** is axiomatized by **Lin** $+ P\top + F\top + (\varphi \wedge G\varphi) \to PG\varphi + (\varphi \wedge H\varphi) \to FH\varphi$) which is (sound and) strongly complete with respect to the class of all the discrete and linear frames.

By now, one might expect us to propose a strengthening of this claim to strong completeness with respect to the standard-example of a structure from this class $\langle \mathbb{Z}, < \rangle$. For two reasons we delay consideration of $\mathbb{Z}$ however. The first one is that the axioms of **D** turn out be not complete for $\mathbb{Z}$, the second that no axiom system can be strongly complete for $\mathbb{Z}$; this structure is very difficult to pin down. Though this circumstance may lag behind our expectations, we are not left completely empty-handed, for in the following section we will see what can be established in connection with the integers, and now we give the proof of the claim previously made:

Theorem 70 **D** is strongly complete with respect to the class of the discrete and linear frames without endpoints.

Proof: The model we are going to construct to give a foundation to this statement comes about in much the same way as the one for **Q**: At the even stages we again follow the line of the completeness proof for **Lin**. This leaves the odd stages available for the arrangement of discreteness: the stages will be used to secure the existence of an immediate predecessor v as well as an immediate successor u for each $t \in \mathcal{T}_{2n+2}$ that still lacks one of these. We will be content with showing this can be done for the case of u in such a way that the union, finally to be taken again as the model to satisfy all our demands, still displays u as the immediate successor of t, taking the possibility to do the same for v for granted.

Take Γ_u to be a maximal consistent extension of the set $\{\varphi \mid G\varphi \in \Gamma_t\} \cup \{\neg\psi \vee \neg G\psi \mid \neg G\psi \in \Gamma_t\}$; the proof of the fact that this set is consistent is postponed till the end of the proof.

This choice of Γ_u indeed precludes the possibility of our ever getting in a position where we are forced at an even stage to insert a point between u and t. Firstly, it will never be necessary to introduce a successor of t which is not also a successor of u, as the following considerations will make clear: If $\neg G\psi \in \Gamma_t$ for some ψ, either $\neg G\psi \in \Gamma_u$, too, in which case the point constructed to verify $\neg\psi$ will be a successor of u, as well, or else $\neg\psi \in \Gamma_u$, in which there is no need to introduce a point in between t and u to verify $\neg\psi$. And secondly, it will never be necessary to introduce a predecessor of u which is not a predecessor of t also, for the following reasons. Assume $\neg H\psi \in \Gamma_u$. It suffices to prove that $\neg\psi \vee \neg H\psi \in \Gamma_t$, for knowing this to hold we can continue as before. Suppose it is not. Then, by one of the special axioms for **D**, $FH\psi \in \Gamma_t$ and this is equivalent to $\neg G\neg H\psi \in \Gamma_t$, whence by the construction of Γ_u we have $H\psi \in \Gamma_u$ or $\neg G\neg H\psi \in \Gamma_u$, i.e. $FH\psi \in \Gamma_u$. The first possibility is ruled out at once, since the contrary was assumed, and the second possibility reduces to the first, since $\vdash_{\mathbf{Lin}} FH\psi \to H\psi$ (this formula is obviously valid on linearly ordered structures and, therefore, by the completeness of **Lin** derivable in **Lin**).

It remains to establish **D**-consistency of the set $\{\varphi \mid G\varphi \in \Gamma\} \cup \{\neg\psi \vee \neg G\psi \mid \neg G\psi \in \Gamma\}$ given that Γ is consistent. Suppose it is not, then there must be $\varphi_1, \ldots, \varphi_m, \psi_1, \ldots, \psi_k \in \Gamma$ such that $\vdash_{\mathbf{D}} \varphi \to (\psi_1 \wedge G\psi_1) \vee \ldots \vee (\psi_k \wedge G\psi_k)$, where φ abbreviates the conjunction of $\varphi_1, \ldots, \varphi_m$. It then follows that:
$\vdash_{\mathbf{D}} \varphi \to PG\psi_1 \vee \ldots \vee PG\psi_k$ (by one of the **D**-axioms),
$\vdash_{\mathbf{D}} \varphi \to P(G\psi_1 \vee \ldots \vee G\psi_k)$ (by derivability in $\mathbf{K}_t$),

$\vdash_{\mathbf{D}} G\varphi \to GP(G\psi_1 \vee \ldots \vee G\psi_k)$,
$\vdash_{\mathbf{D}} G\varphi \to (G\psi_1 \vee \ldots \vee G\psi_k) \vee P(G\psi_1 \vee \ldots \vee G\psi_k)$ (by the contraposition of a **D**-axiom),
$\vdash_{\mathbf{D}} G\varphi \to (G\psi_1 \vee \ldots \vee G\psi_k)$ (by derivability in **Lin**),
and finally $\vdash_{\mathbf{D}} (G\varphi_1 \wedge \ldots \wedge G\varphi_m) \to (G\psi_1 \vee \ldots \vee G\psi_k)$ (by derivability in $\mathbf{K}_t$).

With this formula a contradiction with the consistency of Γ is obtained. •

5.3.4 The logic of integer ordered time: Z

The reason why the structure of the integers with their usual ordering is treated last is that the logic which we will prove to be complete with respect to this structure is not strongly complete. The weakness of this result, compared to the ones we established for the other number structures, is not due to a less fortunate choice of system (see exercise 58), but to facts about the structure under consideration itself: any (tense) logic coinciding with the set of (tense-logical) formulas that are valid on $\langle \mathbb{Z}, < \rangle$ by necessity fails to be strongly complete with respect to $\langle \mathbb{Z}, < \rangle$, because this set is not compact. This can be seen to hold rather easily:

1. It is obvious that $\langle \mathbb{Z}, < \rangle$ is a serial and discrete linear frame.
2. It can be verified moreover that all the frames in the class of successive and discrete linear orderings consist of linear orderings of copies of $\langle \mathbb{Z}, < \rangle$.
3. Among these, the structure $\langle \mathbb{Z}, < \rangle$ on its own is marked off by being the sole structure in this class that is continuous: It is clear that every cut in $\langle \mathbb{Z}, < \rangle$ splits the integers in two halves of which the first has a last element (or, for that matter, the second a first element also), whence there are no gaps. On the other hand, a cut that coincides with a division between two different copies of $\langle \mathbb{Z}, < \rangle$ produces two complementary sets of which the first lacks a largest element and the second a smallest one and so such cuts do determine gaps.
4. Hence, the combination of properties mentioned (i.e. seriality, discreteness and continuity) characterizes $\langle \mathbb{Z}, < \rangle$ up to isomorphism and since, as we have seen before, each of these properties is expressible on the linear frames in the language of tense logic, all instances of the formulas that express them must be among the theorems of any system **S** purporting to represent the logic of the structure in question.
5. However, there exists an infinite set of formulas (we call it Σ for future reference) implying that between some pair of points there exist infinitely many others; the specification of this set is left as exercise 56.
6. Every finite subset of the set Σ can be satisfied in conjunction with the **S**-axioms on $\langle \mathbb{Z}, < \rangle$, but to satisfy the entire set at least two copies of $\langle \mathbb{Z}, < \rangle$ are needed. Such a model is not continuous according to (3) and so it does not verify **S** after all.

Two things can be learnt from the argument above: The first is that, since by the previous proofs the logic **D** characterizes the class of serial and discrete frames and the scheme $\Box(G\varphi \to PG\varphi) \to (G\varphi \to H\varphi)$ characterizes continuity on the linear frames, (3) entails that we can take the logic of the integers to be laid down by the conjunction of the axioms just mentioned. However, for reasons of convenience a different axiomatization **Z**, with **Z**= **Lin** + $P\top$ + $F\top$ + $G(G\varphi \to \varphi) \to (FG\varphi \to G\varphi)$ + $H(H\varphi \to \varphi) \to (PH\varphi \to H\varphi)$, is current and we will use the latter in the sequel. (You can check for yourself that **Z** is equivalent to the axiom system suggested above.)

Secondly, the fact that we are dealing with a complete-but-not-strongly-complete logic carries with it that we have to switch over to a different proof technique: since C cannot work we convert to its finite variant C_{fin}. But before making the preparations necessary for its application it will be instructive to show how exactly the existence of the Σ mentioned above hampers an attempted completeness proof by means of C: Σ does not agree with the way the integers are assembled in $\langle \mathbb{Z}, < \rangle$, but because the whole infinite set is needed to obtain the contradiction, this fact cannot be derived in the logic **Z**, for derivations are finite sequences. Therefore, Σ is **Z**-consistent, so we cannot rule out the possibility of its being part of some maximal **Z**-consistent set, introduced along the way in the model we are constructing. In constructing a model, we are then forced to insert an infinite stretch of points between two points, as a consequence whereof the result will not be isomorphic to $\langle \mathbb{Z}, < \rangle$. Because with C_{fin} the construction is performed within the bounds of some finite set of relevant formulas (i.e. relevant for some specific finite set of assumptions and a formula not derivable from these assumptions) the use of this method can easily be seen to be an effective remedy against this undesirable course of events.

Although the method is basically the same as the one previously used in connection with **S4** (cf. theorem 32) some details have to be changed in order to adjust the method to the present case. In the sequel we will make clear what changes are needed and why; in the course of our exposition, a summarizing sketch of the proof procedure we intend to follow will be given.

For convenience sake, we restrict ourselves to mentioning how we are going to treat the $\neg G$-formulas only; the $\neg H$-formulas can be dealt with by a completely dual argument.

Inspecting the behaviour of truth values of a formula of the form $\neg G\varphi$ going from the left to the right on a linear, serial, discrete and continuous frame, it comes to light that they can display one of the two following patterns: Some formulas $\neg G\varphi$ cease to be true after some $\neg\varphi$ -period, possibly alternating with φ-intervals: there is one definite turning point t (continuity!) where $\neg\varphi$ is true for the last time, while its immediate successor marks the beginning of an uninterrupted and unending period in which the truth value of φ is stabilized to 'true'. On t, the structure so to speak 'flips over' from $\neg G\varphi$ to $G\varphi$; note that such a flip-over can of course only happen once. Other formulas $\neg G\psi$ never reach such a turning-point: a stretch on which ψ is true is always followed by a point verifying $\neg\psi$ and hence $\neg G\psi$ remains true throughout the entire structure.

These two kinds of formulas will receive a different kind of treatment in the model we are going to construct: in the first number of stages the former kind of formulas will be dispensed with by providing a turning-point for each $\neg G\varphi$ that can consistently have one (note that we can only do this because within a finite set there will only be finitely many formulas of this kind) and only after that we will turn to the creation of points to the right, over and over again, securing in this manner the existence of a falsifying instance for each $\neg G\psi$ of the second kind. Some reflection shows that the difference in nature between the two kinds of formulas can be read off from the corresponding FG-formula: if $FG\chi$ can be consistently assumed, then clearly $\neg G\chi$ is of the first kind whereas, conversely, if this is not the case, $\neg G\chi$ is of the second kind. Therefore, the FG-formulas are to be counted as 'relevant' (in the sense specified above) and need be included in the closure of a set. We will have to take in these more complex formulas without loosing our grip on finiteness, just as we did in the case of negations, by somehow stopping the indefinite repetition of this process of forming more complex formulas. That the method used in clause (iv) of the definition below does this without interfering with the previous requirements will only become clear in the completeness proof.

Finally, to secure seriality we have to have at our disposal $\neg G\bot$ and $\neg H\bot$. This explains the following definition of what in this case the closure of a set amounts to:

Definition 1 A set of formulas Φ is closed iff
(i) Φ is closed under the formation of subformulas,
(ii) Φ is closed under the formation of single negations,
(iii) Φ contains $G\top$ and $H\top$ and
(iv) if $G\varphi \in \Phi$ and φ is not of the form $\neg G\psi$, then $G\neg G\varphi \in \Phi$.

We also need to adapt the definition of $\prec$: First of all, the success of our previous completeness proofs depended partly on the fact that the formulation in terms of G and in terms of H of the conditions under which $\prec$ was taken to hold boiled down to the same thing. But if one takes the trouble of looking back at the proofs, one can see that formulas of a higher degree of complexity were needed to establish the equivalence; the presence of these higher complexity formulas we can not rely on this time: $FH\varphi \vdash Z\varphi$ is of no use if $FH\varphi \notin \Phi$. The definition of $\prec$ will therefore have to be given shape with clauses for G and H separately. Finally, the conditions under which $\prec$ holds will themselves have to be altered also: just like in the case of **S4**, in order to insure transitivity of the model, going to the right the G-formulas (and similarly, going to the left the H-formulas) must be taken along. These considerations motivate our redefining $\prec$ in the following way:

Definition 26 $\Gamma \prec_{\mathbf{Z}} \Delta$ iff (i) for all $\varphi \in \Phi$: if $G\varphi \in \Gamma$, then $\varphi \in \Delta$ and $G\varphi \in \Delta$ and (ii) if $H\varphi \in \Delta$, then $\varphi \in \Gamma$ and $H\varphi \in \Gamma$.

(An equivalent of the second clause is: if $\neg H\varphi \in \Gamma$ or $\neg\varphi \in \Gamma$, then $\neg H\varphi \in \Delta$, and in this form it will be used in the sequel.)

We are now in a position to prove the theorem:

Theorem 71 **Z** is complete with respect to $\langle \mathbb{Z}, < \rangle$.

Proof: Assume $\Sigma \not\vdash_{\mathbf{Z}} \varphi$, for some finite set of formulas Σ. The construction of a counterexample against Σ's entailing φ on the frame $\langle \mathbb{Z}, < \rangle$ is executed in stages again. In stage 0, we start with a point t_0 and associate a Γ_{t_0} to it that is maximally consistent in the closure Φ of $\Sigma \cup \{\neg\varphi\}$. Because Φ is finite, it can only contain a finite number of G- and H-formulas and therefore there must be among the maximally **Z**-consistent sets in Φ some Γ_l and some Γ_r such that $\Gamma_l \prec_{\mathbf{Z}} \Gamma_{t_0}, \Gamma_{t_0} \prec_{\mathbf{Z}} \Gamma_r$, where Γ_r contains a maximal number of G-formulas and a minimal number of H-formulas, while Γ_l satisfies the corresponding dual demands.

In stage 1, we add a $t_1 < t_0$ and a $t_2 > t_0$ to $\{t_0\}$ and associate Γ_l and Γ_r respectively, to them. These points will become the left-hand resp. right-hand extremes of a finite stretch of points that will eventually become the middle part of the copy of $\langle \mathbb{Z}, < \rangle$ we are constructing.

In the stages 2 to (some) k, in between t_1 and t_2 all possible turning-points (i.e. points on which $\neg\varphi$ and $G\varphi$, resp. $\neg\varphi$ and $H\varphi$, are true at the same time, for some $\neg G\varphi \in \Gamma_l$, resp. $\neg H\varphi \in \Gamma_r$) of the kind mentioned in our earlier sketch of the proof will be inserted. As usual, we forget about the past, relying on the completely dual nature of the argument required, and will just show how this is done for the $\neg G\varphi$'s that flip over somewhere; it can be argued these are precisely the formulas such that $\neg G\varphi \in \Gamma_l$ and $G\varphi \in \Gamma_r$. For those we can introduce a $t' > t$ and associate a $\Gamma_{t'}$ to it with $\Gamma_t \prec_{\mathbf{Z}} \Gamma_{t'}$ and $\neg\varphi, G\varphi \in \Gamma_{t'}$ as required, because the assumption that we cannot do this leads to a contradiction:

We start with showing that for arbitrary $\neg G\varphi$ such that $\neg G\varphi \in \Gamma_l$ and $G\varphi \in \Gamma_r$, $\neg G\neg G\varphi \in \Gamma_l$ must obtain also. At first sight there seem to be two reasons why this may fail, both leading however to an inconsistency: The first possibility is that $G\neg G\varphi \in \Gamma_l$. By the definition of $\prec_{\mathbf{Z}}$, it follows that $\neg G\varphi \in \Gamma_r$ but this conflicts with the assumption that $G\varphi \in \Gamma_r$. The second possibility is that $\neg G\neg G\varphi \notin \Phi$. This formula would have been in Φ on account of clause (ii) if $G\neg G\varphi \in \Phi$, so $G\neg G\varphi \notin \Phi$ must hold in addition. Because of clause (iv), this can only be the case if φ is of the form $\neg G\psi$, whence our previously made assumptions amount to $\neg G\neg G\psi \in \Gamma_l$ and $G\neg G\psi \in \Gamma_r$. We can conclude to $\neg G\psi \in \Gamma_r$, since $\vdash_{\mathbf{Lin}} (G\neg G\psi \wedge G\psi) \rightarrow G\bot$ and $\neg G\bot$ is among **Z**'s axioms. But because Γ_r was chosen maximally with respect to the G-formulas, this implies that $\neg G\psi \in \Delta$ for all Δ such that $\Gamma_r \prec_{\mathbf{Z}} \Delta$ and this in turn implies that

$\vdash_{\mathbf{Z}} (G\theta_1 \wedge \ldots \wedge G\theta_m \wedge \theta_1 \wedge \ldots \wedge \theta_m \wedge \neg H\chi_1 \wedge \ldots \wedge \neg H\chi_k) \rightarrow \neg G\psi$ for some $G\theta_i$, and $\neg H\chi_j$ or $\neg\chi_j \in \Gamma_l$ ($1 \leq i \leq m$ and $1 \leq j \leq k$). We can now derive $G\neg G\psi$ from Γ_l, countering the previous insight that $\neg G\neg G\psi \in \Gamma_l$, in the following way:

$\vdash_{\mathbf{Z}} G(G\theta_1 \wedge \ldots \wedge G\theta_m \wedge \theta_1 \wedge \ldots \wedge \theta_m \wedge \neg H\chi_1 \wedge \ldots \wedge \neg H\chi_k) \rightarrow G\neg G\psi$, by G-necessitation,

$\vdash_{\mathbf{Z}} (GG\theta_1 \wedge \ldots \wedge GG\theta_m \wedge G\theta_1 \wedge \ldots \wedge G\theta_m \wedge G\neg H\chi_1 \wedge \ldots \wedge G\neg H\chi_k) \rightarrow G\neg G\psi$, by 'G-distribution',

$\vdash_{\mathbf{Z}} (G\theta_1 \wedge \ldots \wedge G\theta_m \wedge G\neg H\chi_1 \wedge \ldots \wedge G\neg H\chi_k) \rightarrow G\neg G\psi$, by transitivity.

The fact that $\vdash_{\mathbf{Lin}} (\neg H\chi \vee \neg\chi) \rightarrow G\neg H\chi$ suffices to see that the entire conjunct in the antecedent of the last formula is in Γ_l and we have proved that $\neg G\neg G\varphi \equiv FG\varphi \in \Gamma_l$.

Now we return to the desired $\Gamma_{t'}$: if it would not exist, this can only be because $\{G\psi, \psi \mid G\psi \in \Gamma_l\} \cup \{\neg H\chi \mid (\neg H\chi \vee \neg\chi) \in \Gamma_l\} \cup \{G\varphi, \neg\varphi\}$ is not $\mathbf{Z}$-consistent. In much the same way as before, a contradiction with $\neg G\varphi \in \Gamma_l$ can be obtained by using the fact just established and the characteristic axiom of $\mathbf{Z}$. To derive this contradiction is exercise 57.

Note that, because of the definition of $\prec_{\mathbf{Z}}$, t' will finally fall into the right place automatically and further that, because a flip-over can at most occur once, $\neg G\varphi$ can never be up for treatment again, if dealt with in the manner described. Therefore, the insertion of all the possible turning-points will only engross our attention a finite number of stages. From this circumstance, two fortunate conclusions can be drawn:

As long as we do not interfere again with the part acquired so far (and we will not need to and do not intend to) its being finite guarantees its discreteness as well as its continuity. And the fact that we can point to a specific moment after which all $\neg G$-formulas requiring possibly turning-points taken care of for once and for all, gives us the opportunity to change the plan after this moment. From stage k+1 on, we will proceed in a different manner and extend the already obtained middle part indefinitely to the right as well as to the left, in order to get a frame that is isomorphic to $\langle\mathbb{Z}, <\rangle$. This turns out to be very simple; again we will only consider what to do for the $\neg G\varphi$'s left over. By now, we can fully exploit the fact we chose Γ_r to have a maximal number of $G-$ and the minimal number of H-formulas, for this brings along the fact that all Δ with $\Gamma_r \prec_{\mathbf{Z}} \Delta$ contain exactly the same $G\varphi$'s and $H\varphi$'s and hence the same $\neg G\varphi$'s and $\neg H\varphi$'s. The H-formulas can be ignored; they are treated at, or to the left of, Γ_r. All that remains to be done is to continually create successors falsifying φ_i for each $\neg G\varphi_i \in \Gamma_r$ in turn. It is obvious that this can be done and that the process of extending to the right never stops, since at least $\neg G\bot \in \Gamma_r$. •

5.4 Exercises

Exercise 49 Prove proposition 63.

Exercise 50 Prove proposition 64(i). The proof of (ii) is highly similar.

Exercise 51 Prove lemma 65(i). The proof of (ii) is highly similar.

Exercise 52 (a little harder). Show that for all tense logics **S**, if $GG\varphi \to G\varphi \in \mathbf{S}$, then the relation $\prec$ on the set of all maximally **S**-consistent sets is dense.

Exercise 53 Compare the quality of being dense with that of being continuous by proving proposition 68

Exercise 54 Prove that, taking only linear frames into consideration, the formula $\Box(G\varphi \to PG\varphi) \to (G\varphi \to H\varphi)$ characterizes continuity.

Exercise 55 Show that the model constructed in the completeness proof for **R** satisfies condition (d).

Exercise 56 Specify an infinite **Z**-consistent set of tense-logical formulas that semantically implies that between two (future) points there must lie infinitely many others.

Exercise 57 Derive the contradiction mentioned in the proof above.

Exercise 58 Show that, if logic **S** is strongly complete with respect to the class of frames $\mathcal{C}$ and **S'** is complete with respect to $\mathcal{C}$, then **S'** is strongly complete with respect to $\mathcal{C}$.

Chapter 6

Intuitionistic propositional logic

6.1 Introduction

The logic we address in the present chapter does not qualify as an intensional logic (in the sense of our description in section 1.1) on purely syntactic grounds: its vocabulary does not contain a non-truthfunctional logical constant beside the standard connectives of the propositional calculus. Intuitionistic logic does however display intensional aspects on a semantic level: some of the standard connectives receive a non-standard interpretation that can be specified by means of the kind of structures we used in connection with intensional logics; we think this justifies its treatment in this book

As this can be helpful in catching the intentions behind the intuitionistic line of thought and facilitate the understanding of what intuitionistic logic is all about, we will give some information about the historical background in which the origin of the intuitionistic school took place and we briefly discuss some of its basic theses, before we pass on to explaining the intuitionistic semantics of the standard connectives in detail.

6.1.1 Some history

Intuitionism came into being as a reaction against the philosophical tendency that dominated the mathematical scene about the year 1900. In the ages that had passed since the Greeks proved their first lucid and transparent geometrical theorems, mathematics had taken a high flight into the more abstract and less perspicuous spheres; mathematicians had become engaged in ever more complex forms of reasoning, involving concepts, often linked up with the infinite, whose meanings tended to be rather obscure, and thereby mathematics had lost much of its character of being the 'clear and distinct' discipline par excellence.

About the turn of the century then, the time seemed ripe for some reconsideration: in wide mathematical circles, the need was felt for a philosophical account that would explain the peculiar nature and special epistemological status the 'queen of sciences' has, or (if you like) had always been claimed to have. Therefore, much mathematical investigation was at that time devoted to the search after 'foundations': it was esteemed to be desirable to be able to point out a 'hard-core' piece of mathematics that was simple and elementary itself and to which all higher mathematics could be reduced. Thus, one hoped, the clarity and distinctness mathematics is disposed to have would indirectly be established for its more complex parts also and the certitude could be reclaimed for mathematical knowledge, for thus even the slightest inducement for doubting the correctness of its more complex parts would be taken away.

Both Frege's logicistic project (the attempt to reduce all mathematics to logic, which is judged to be the most fundamental of sciences) and Hilbert's formalistic program (the attempt to prove by simple devices of a finitistic nature the consistency of higher mathematics, thereby showing the latter to be a save means (and no more than that) to get from meaningful statements about finitistically specifiable mathematical objects to other statements about these) fall under this striving; each would have provided (had the attempt been successful) a solid basis from which a justification of (the believe in) the more complex forms of mathematical reasoning could have been derived.

In discussing the foundations of mathematics, intuitionism is always mentioned in one breath with logicism and formalism, because it was the same problem (the opaqueness of the concepts and reasonings of the higher, infinitistic, mathematics) that caused its emergency in first instance. Intuitionists however approach this problem with a deflective disposition: in stead of attempting to provide an alternative solution for it, they reject the problem altogether, or rather, they reject the situation that gave rise to it. The argument goes roughly like this: if mathematics has become such an affair that a substantial part of it can only stand firm if it is underpinned by an 'external' (i.e. not belonging to the theorem in question itself) piece of argumentation serving no other purpose than its justification, then that part is not worthy to be called 'mathematics'. Since such a procedure, which is in conflict with the nature of mathematics, can only decrease its value, it is best if this part is simply expelled from the domain of mathematics, rather than attempting to rescue it. For intuitionists, no specimen of mathematical reasoning may ever or on any point loose the quality of being directly accessible to the human mind.

Although others before him had occasionally given vent to similar ideas, Brouwer has to be credited for being the first person that formulated the intuitionistic philosophy of mathematics in a coherent and well-considered manner. Moreover, Brouwer is the first who took the intuitionistic philosophy seriously (whereas his predecessors at best practised it as a leisure activity), something of which he gave evidence by reshaping a large number of classical theorems in accordance with his intuitionistic convictions; thus showing intuitionism to be a workable theory in practice.

Our special interest is in intuitionistic *logic*, which is not from the hands of Brouwer himself (who, because of his tendency toward solipsism no doubt, in general appeared to be less than attracted to questions about language) but was distilled from Brouwer's mathematical work by Heyting. Heyting formalized the patterns of reasoning underlying this work and thereby made the sometimes rather austere writings of Brouwer accessible to other-thinking people, thus managing to draw the attention of a larger public.

6.1.2 An outline of the intuitionistic picture of mathematics

It is commonly assumed that logic precedes all, and hence all mathematical, reasoning. In the intuitionistic tradition, as we touched on before, the order of precedence is however reversed: logic, considered as a discipline by itself is conceived as the condensation of patterns of inference the (intuitionistic-minded, of course) mathematician accepts in practice. This circumstance makes intuitionistic logic for a large part dependent on the intuitionistic view on mathematics. Because the structures in which the intuitionistic semantics of the propositional calculus will be given shape are made to suit mathematical theories most of all, the specific traits of these structures can best be understood in the light of the intuitionistic ideas about mathematical reality and truth. Therefore, a short digression on this subject is warranted here:

The best way to clarify the intuitionistic conception of mathematical reality and truth is perhaps by contrasting it with its classical counterpart. The (typical) classical mathematician regards the universe of mathematical discourse as something that is given; mathematical objects with certain mathematical properties exist somewhere and somehow, independent of human interference. Again, irrespective of any intervention of the human mind, statements are true if what they say corresponds to the state of affairs in this mathematical realm and false otherwise. Mathematicians aim at expanding their knowledge (i.e. gather as many propositions out of this once and for all fixed stock of truths) by means of proofs that lead them from true statements to other true statements; in some inexplicable (or at least up till now unexplained) way this procedure is supposed to give them access to what the mathematical universe looks like.

The (typical) intuitionist holds rather divergent opinions regarding these matters. The conception of mathematics as a discipline that must be "tangible" throughout places the notion of "construction" in the centre of attention: from an intuitionistic stand-point the whole of what can count as mathematics must be identified with the knowledge of mathematical objects that can be obtained in a constructive way. Therefore, mathematical objects and the properties they display must be created by means of constructive definitions before they can be discussed. Furthermore, the only reasonable ground for taking a statement about these objects for true is having established its truth by means of a constructive proof. It must be clear that, if mathematics is equated with the set of propositions constructively known to be true, it cannot be thought of as a fixed body of facts, some of which are (still) unknown to us (as in the classical case); hence, as in time new proofs and new objects to prove statements about are invented, mathematics itself (and not only our knowledge of it) can be said to grow.

A warning is due on the term 'proof' as it is used here, because it may wrongly evoke associations with what is often referred to by the same name. For intuitionists, proofs are first of all mental operations, executed in the mind of the so-called "creative subject" (the "the" is reminiscent of Brouwers solipsism). Written proofs derive some secondary importance (after all, not all intuitionists are solipsists) from the fact they report on these mental operations, on which all mathematics is taken to hinge. Least of all however should one think in this context of proofs as derivations in some formal system. It is not difficult to see why intuitionism and formalism do not go together very well: the formalistic tenet that mathematical proofs are meaningless strings of symbols is diametrically opposed to the intuitionistic trend to avert even 'obscure meanings'.

6.2 The definition of intuitionistic semantics

The picture of intuitionistic mathematics just sketched can be exploited in motivating the way in which the structures that embody the intuitionistic semantics of the propositional calculus are given shape. First we will point-wise discus the elements out of which these structures are composed and then we will give an exact definition.

As in the case of intensional logics, a model is made up of a frame, composed of a set of points and an accessibility-relation defined on these, and a forcing-relation.

The world-set In the present context the elements of this set are called 'information states'. They can best be thought of as bundles of data, that can be 'incomplete' in the sense that the data constituting some point are possibly not enough to decide the truthvalue (indicating its being intuitionistically provable or disprovable from the data) for all the statements that can be expressed in the language at stake;this contrasts with the worlds as they figure in modal logics, which are 'complete' in this sense because every sentence sentence either conforms to the facts (whence it is true) or not (whence it is false). (Note that the 'incompleteness' only occurs on the object-language level, for every sentence is either intuitionistically implied by the data or not, so on the meta-level two-valuedness is restored.)

The accessibility-relation As accessibility-relation $\mathcal{R}$, we take a reflexive partial ordering; $i\mathcal{R}j$ must be read approximately as 'information state j can still be reached once the information of state i is already acquired'. A partial ordering enables us on the one hand to registrate the different stages in the increase of knowledge obtained (the reason why we chose an ordering), while on the other hand it does not limit us to the recording of only one (say the actual) course of information gain, but leaves room to lay down all the possible ways in which knowledge can be expanded once a certain amount of data is already obtained (the reason why the ordering must be partial only).

The forcing-relation Partially ordered frames give the opportunity to represent the possible ways in which knowledge can grow, but by themselves such frames cannot exclude the possibility to retract already gained/proven statements later on, something we do not want to be possible, and so they cannot enforce that it is indeed growth of knowledge that gets represented. This undesirable course of events can be ruled out by letting the forcing-relation obey the simple demand that every information state that is accessible from some other must force at least all the formulas the latter forces. Of course we can only make a stipulation to this effect for the propositional letters, but under the intuitionistic definition of $\models$, to be given below, this property turns out to carry over to all the formulas automatically; the proof of this we leave as exercise ?? Note this procedure entails that for the first time we demarcate a class of models (and not a class of frames) as appropriate for a logic.

The most note-worthy departure from the previous, intensional, cases is however that the definition of the forcing-relation itself must be altered drastically:

Translated in classical terms, the task for which we find ourselves appointed is that we have to fashion the models in such a manner that, in stead of the (classical) truth-predicate the (classical) predicate of being constructively provable is incorporated. (Note that for intuitionists the formulation of our task makes no sense: for them the former predicate just *is* truth, while they claim not to understand the meaning of the latter predicate.) In other words: the forcing relation must exhibit all the characteristics of 'having a constructive proof' and not that of truth, or in still other words: $\models$ must be defined in such a manner that $i \models \varphi$ can be read as 'given the information available at state i, φ can be established constructively'.

If we assume we know what must be understood by an constructive proof for the atomic propositions (if we interpret for example the atoms as arithmetical equations a calculation can count as an constructive proof) we can spell out in inductive clauses the conditions under which complex formulas must be taken to be forced:

- a constructive proof for $\varphi \wedge \psi$ can be made out of one for φ and one for ψ by concatenation;
- a constructive proof for $\varphi \vee \psi$ consists of one for φ or one for ψ;
- a constructive proof for $\varphi \to \psi$ is obtained if an effective operation exists that converts every proof for φ into one for ψ;

- $\neg\varphi$ is taken to be the equivalent of $\varphi \to \bot$, so a constructive proof for $\neg\varphi$ is an effective operation that shows every proof for φ leads to a contradiction.

How this interpretation of truth is given shape in the definition of the forcing relation can be read off from the definition of intuitionistic semantics below.

Definition 27 A *model for the intuitionistic propositional calculus* consists of a a model $\langle\langle\mathcal{I},\mathcal{R}\rangle \models\rangle$ such that

(i) $\mathcal{I}$ is a non-empty set (of information states);
(ii) $\mathcal{R}$ is a reflexive partial order;
(iii) for all $i \in \mathcal{I}$ and propositional letters p it holds that, if $i \models p$, then $j \models p$ for all j with $i\mathcal{R}j$
(iv) $\models$ is a forcing-relation such that, for complex formulas,
- $i \models \varphi \wedge \psi$ iff $i \models \varphi$ and $i \models \psi$;
- $i \models \varphi \vee \psi$ iff $i \models \varphi$ or $i \models \psi$;
- $i \not\models \bot$ and
- $i \models \varphi \to \psi$ iff for all j such that $i\mathcal{R}j$, $j \not\models \varphi$ or $j \models \psi$ (from which it follows that $i \models \neg\varphi$ iff for all j with $i\mathcal{R}j$, $j \not\models \varphi$)

We said previously that information states can be 'incomplete' in the sense that sometimes it is impossible to decide on the basis of the information available at a given state wether φ or $\neg\varphi$ holds; given the conditions for obtaining of $\neg\varphi$ in some information state i it does not hold that $i \not\models \varphi$ iff $i \models \neg\varphi$. $i \not\models \varphi$ means that no verification of φ is constructed, while $i \not\models \neg\varphi$ is much stronger: $\neg\varphi$ is only 'positively' verified by some information state if the information of no other accessible state entails φ.

The intuitionistic interpretation of $\to$-sentences (and consequently the interpretation of $\neg$-sentences) reveals the intensional character of intuitionistic semantics; the connection between intensional and intuitionistic logic will be made explicit in the final section of this chapter.

Because of the deviant definition of the semantics ('deviant' from the standard definition that is) $\wedge$,$\vee$ and $\to$ are not interdefinable, as in classical logic. Therefore, in the intuitionistic case the separate clauses are not merely there as a matter of convenience, but all of them are really needed to supply a semantics for the entire set of logical constants.

6.3 The syntax of intuitionistic logic

The relation of entailment, as it applies to formulas given the semantics we defined in the previous section, can be reproduced syntactically by the following axiom-system, known as **IPC**:

An axiom system for **IPC**

1. $\varphi \to (\psi \to \varphi)$
2. $\varphi \to (\psi \to \theta) \to ((\varphi \to \psi) \to (\varphi \to \theta))$
3. $(\varphi \wedge \psi) \to \varphi; (\varphi \wedge \psi) \to \psi$
4. $\varphi \to (\psi \to (\varphi \wedge \psi))$
5. $\varphi \to (\varphi \vee \psi); \psi \to (\varphi \vee \psi)$
6. $(\varphi \to \theta) \to ((\psi \to \theta) \to ((\varphi \vee \psi) \to \theta))$
7. $\perp \to \varphi$
8. Rule: *Modus ponens.*

From this system, an axiomatization of the classical propositional calculus **CPC** can easily be obtained by adding the axiom-scheme (8) $\neg\neg\varphi \to \varphi$ to the list above. That **IPC** should be a proper part of CPC is of course no coincidence and it can be argued for with an appeal to their respective interpretations (tacitly assuming soundness): Whatever is valid under all intuitionistic valuations, and hence cannot fail to be constructively provable, most certainly cannot fail to be true also and so it must be valid under all classical valuations too; whereas, conversely, one can easily imagine a formula to be classically valid, without there being a constructive proof for it. We can illuminate the latter phenomenon with the help of intuitionistic counter-examples falsifying some classical theorems, but only after we have proved:

Proposition 72 **IPC** is sound with respect to the class of intuitionistic models

Proof: Since the proof is straightforward but long-winded, you are asked only to run through part of it in exercise 60. •

The two most famous examples of classical principles that are not among the theorems of intuitionistic logic, are the failure of the law of double negation $\neg\neg\varphi \to \varphi$ and (its equivalent) failure of the law of the excluded middle $\varphi \vee \neg\varphi$. It is left to the reader to constuct a model in which they are falsified.

In intuitionistic logic conjunction cannot be defined in terms of disjunction and negation, and disjunction cannot be defined in terms of conjunction and negation. This, of course, is closely related to the invalidity of $\neg\neg\varphi \to \varphi$. Three of the The Morgan law's hold:
$\neg(\varphi \vee \psi) \to (\neg\varphi \wedge \neg\psi)$,
$(\neg\varphi \vee \neg\psi) \to \neg(\varphi \wedge \psi)$, and
$(\neg\varphi \wedge \neg\psi) \to \neg(\varphi \vee \psi)$,
but $\neg(\varphi \wedge \psi) \to (\neg\varphi \vee \neg\psi)$ is invalid.

More examples can be found in the exercises.

6.4 The completeness-proof for IPC

One half of the claim that **IPC** faithfully reflects the logic of the intuitionistic models, has been verified in proposition 72; we now turn to the other half. We will establish the weaker variant of completeness (the one whose proof entails decidability); the argument we provide can however by a few small interventions be converted into a proof for the stronger theorem, for **IPC** is indeed strongly complete with respect to the class of models indicated.

The method by which we will construct the counterexamples that we have to produceb in order to show that completeness holds for **IPC** is very similar to the one we used in connection with the intensional logics we have treated before: we take sequences of formulas to denote information states, define a relation with the required properties on these states and associate sets of formulas to them in such a way that the sets prescribe a valuation on such a sequence that respects the intuitionistic semantics of the propositional calculus. The main point of difference between the models we are going to construct here and the ones we employed for intensional logics is hidden in this last design. Up till now, we utilized maximally consistent sets to supply the valuations, but these are fashioned in concurrence with the classical truth-definition and hence they do not live up to the intuitionistic needs. Therefore, we must particularize a new kind of formula-sets, to which, for future reference, we already attach the name 'D-theory'. These D-theories have to take over the task maximally consistent sets perform in the intensional case, i.e. in the end we want to be able to set

$$\Delta \models \varphi \text{ iff } \varphi \in \Delta$$

The first obvious condition D-theories therefore will have to fulfil is: if $\Delta \vdash_{\mathbf{IPC}} \varphi$, then $\varphi \in \Delta$ (Δ is a *theory* in the technical sense of the word). What further conditions must be imposed on a D-theory in order to make it qualify as an allowable valuation on a particular information state we can find out by running through the definition of intuitionistic semantics clause by clause: the clause for conjunctions presents no problems, because $\varphi \wedge \psi \in \Delta$ iff $\varphi \in \Delta$ and $\psi \in \Delta$ holds in virtue of the one condition we already have; the same goes for the $\Rightarrow$-part of the clause for disjunctions. The converse is however not derivable and must be provided for explicitly, so the second condition that must be inflicted on a D-theory reads: if $\varphi \vee \psi \in \Delta$, then $\varphi \in \Delta$ or $\psi \in \Delta$ (Δ has the disjunction-property) To get $\perp \notin \Delta$ we simply demand, as the third condition, that Δ must be consistent. Finally, because the truth-value of implications chiefly depends on what is forced by other accessible information states, only a small part of the fourth clause is relevant here, namely if $\varphi \rightarrow \psi \in \Delta$, then $\varphi \notin \Delta$ or $\psi \in \Delta$ and this again is entailed by Δ 's theory-hood; the remainder of the clause must, and will, be dealt with in the construction of the models.

Definition 28 A set of formulas Δ is a *D-theory* iff

(i) Δ is a theory: If $\Delta \vdash_{\mathbf{IPC}} \varphi$. then $\varphi \in \Delta$;
(ii) Δ has the disjunction-property: If $\varphi \vee \psi \in \Delta$, then $\varphi \in \Delta$ or $\psi \in \Delta$ and
(iii) Δ is consistent.

There are two more points where the proof method needs to be adapted. The first change concerns the notion of closure. Since for D-theories Δ we will not get that $\varphi \notin \Delta$ iff $\neg\varphi \in \Delta$ any way, we can take as the closure of a set of formulas just the set of all the subformulas of these formulas; there is no need to include their negations. Secondly, we will have to establish, for the new kind of valuation-sets, something like the Lindenbaum-lemma, securing the existence of a D-theory extending some given consistent set whenever we need one. A strict analogue of the Lindenbaum-lemma would however fall short here. This becomes clear if we look at the applications we have in store for such a lemma: we want to use it of course in the construction of counterexamples to acquire a valuation in a world such that a particular formula φ is not forced. In the classical case we could get this by simply adding the negation of φ to Δ and then extend $\Delta \cup \{\neg\varphi\}$ to a maximally consistent one. Here however we have to reckon with a difference between a formula's not being forced and its negation being forced. It would be wrong to arrange on beforehand that the φ in question will not end up being an element of the D-theory to which the set in question gets extended by putting in $\neg\varphi$. Still, we will have to see to it that φ gets not included in the process of extending Δ.

Lemma 73 If $\psi_1, \ldots, \psi_n \nvdash_{\mathbf{IPC}} \varphi$, then there is a D-theory Δ in the closure of $\{\psi_1, \ldots, \psi_n, \varphi\}$ such that $\Delta \supseteq \{\psi_1, \ldots, \psi_n\}$ and $\Delta \nvdash_{\mathbf{IPC}} \varphi$.

Proof: Suppose $\psi_1, \ldots, \psi_n \nvdash_{\mathbf{IPC}} \varphi$ and let $\Phi = \{\varphi_1, \ldots, \varphi_m\}$ be the closure of $\{\psi_1, \ldots, \psi_n, \varphi\}$. We set $\Gamma_0 = \{\psi_1, \ldots, \psi_n\}$ and with induction define the sequence $\Gamma_1, \ldots, \Gamma_{m+1}$ as follows: If $\Gamma_{i-1} \cup \{\varphi_i\} \nvdash_{\mathbf{IPC}} \varphi$, then $\Gamma_i = \Gamma_{i-1} \cup \{\varphi_i\}$; otherwise, we take $\Gamma_i = \Gamma_{i-1}$ $(1 \leq i \leq m+1)$.

Γ_{m+1} is the D-theory required, because (1) most certainly $\Gamma_{m+1} \nvdash_{\mathbf{IPC}} \varphi$; and (2) Γ_{m+1} must be consistent for that reason also. Furthermore, (3) Γ_{m+1} is a theory (formulas that are **IPC**-derivable from some Γ_i that does not already entail φ cannot, by being added to that Γ_i, make φ derivable and hence these formulas must have been included already); and (4) for all ψ and χ it can only be the case that $\psi \notin \Gamma_{m+1}$ and $\chi \notin \Gamma_{m+1}$ if $\Gamma_{m+1} \cup \{\psi\} \vdash_{\mathbf{IPC}} \varphi$ and $\Gamma_{m+1} \cup \{\chi\} \vdash_{\mathbf{IPC}} \varphi$. By axiom (6) it immediately follows that $\Gamma_{m+1} \cup \{\psi \vee \chi\} \vdash_{\mathbf{IPC}} \varphi$, whence by the construction principle, $\psi \vee \chi \notin \Gamma_{m+1}$ and so Γ_{m+1} has the disjunction-property. •

We are now ready for the completeness proof; we give it in a somewhat stronger (viz. by restricting ourselves to the use of finite counterexamples only) form:

Theorem 74 IPC is complete with respect to the class of finite intuitionistic models.

Proof: Assume $\psi_1, \ldots, \psi_n \nvdash_{\mathbf{IPC}} \varphi$. We take a point $\langle\rangle$ and associate a D-theory $\Delta_{\langle\rangle}$ to it such that $\psi_1, \ldots, \psi_n \in \Delta_{\langle\rangle}$ and $\varphi \notin \Delta_{\langle\rangle}$;
lemma 73 guaranties we can produce such a D-theory.

We are going to build a finite partial ordering springing from this point $\langle\rangle$ that will make of $\langle\rangle$ an information state in an intuitionistic model. To accomplish that this model will indeed qualify as an intuitionistic one, we must keep the following in mind:

1. Its accessibility-relation $\mathcal{R}$ must be a reflexive partial ordering. For all the points $\langle\chi_i, \ldots, \chi_j\rangle$ and $\langle\theta_k, \ldots, \theta_m\rangle$ that will be introduced along the way we here already set $\langle\chi_i, \ldots, \chi_j\rangle\mathcal{R}\langle\theta_k, \ldots, \theta_m\rangle$ iff $\langle\theta_k, \ldots, \theta_m\rangle$ is of the form $\langle\chi_i, \ldots, \chi_j, \rho_1, \ldots, \rho_n\rangle$ (all sequences may be empty); after we have finished, it can checked that we named the points thus that the required properties of $\mathcal{R}$ in this way come out all right.
2. All D-theories that will eventually be associated to some point in the model satisfy:
   - (i) If $\langle\chi_i, \ldots, \chi_j\rangle\mathcal{R}\langle\theta_k, \ldots, \theta_m\rangle$ and $\Delta_{\langle\chi_i,\ldots,\chi_j\rangle} \models \varphi$, then $\Delta_{\langle\theta_k,\ldots,\theta_m\rangle} \models \varphi$, for all propositional letters φ and hence, according to exercise 59, for all formulas φ.
   - (ii) For all φ and ψ and points $\langle\chi_i, \ldots, \chi_j\rangle$ and $\langle\theta_k, \ldots, \theta_m\rangle$ such that $\langle\chi_i, \ldots, \chi_j\rangle\mathcal{R}\langle\theta_k, \ldots, \theta_m\rangle$, it holds that if $\varphi \to \psi \in \Delta_{\langle\chi_i,\ldots,\chi_j\rangle}$, then $\varphi \notin \Delta_{\langle\theta_k,\ldots,\theta_m\rangle}$ or $\psi \in \Delta_{\langle\theta_k,\ldots,\theta_m\rangle}$.
   - (iii) If for some φ, ψ and $\Delta_{\langle\chi_i,\ldots,\chi_j\rangle}$, $\varphi \to \psi \notin \Delta_{\langle\chi_i,\ldots,\chi_j\rangle}$, then there is a point $\langle\chi_i, \ldots, \chi_j, \theta_1, \ldots, \theta_n\rangle$ such that $\varphi \in \Delta_{\langle\chi_i,\ldots,\chi_j,\theta_1,\ldots,\theta_n\rangle}$ and $\psi \notin \Delta_{\langle\chi_i,\ldots,\chi_j,\theta_1,\ldots,\theta_n\rangle}$.

(i) and (ii) of condition 2 can be taken care of in one stroke by taking along all the formulas of a $\langle\varphi_i, \ldots, \varphi_j\rangle$ already obtained, whenever it is necessary to create an information state $\langle\varphi_i, \ldots, \varphi_j\chi\rangle$ 'above' it; i.e.

(*) if $\langle\varphi_i, \ldots, \varphi_j\rangle\mathcal{R}\langle\psi_k, \ldots, \psi_m\rangle$, then $\Delta_{\langle\varphi_i,\ldots,\varphi_j\rangle} \subseteq \Delta_{\langle\psi_k,\ldots,\psi_m\rangle}$.

Having to satisfy (iii) is the only thing that may compel us to introduce new points and it is easily shown we can always do this in keeping with (*). For suppose $\chi \to \theta \notin \Delta_{\langle\varphi_i,\ldots,\varphi_j\rangle}$ for some χ, θ and $\Delta_{\langle\varphi_i,\ldots,\varphi_j\rangle}$. Then, by the deduction-theorem (that we did not, but could easily have, proved for **IPC**), $\Delta_{\langle\varphi_i,\ldots,\varphi_j\rangle} \cup \{\chi\} \nvdash_{\mathbf{IPC}} \theta$, whence we can apply lemma 73.

The finite partial ordering needed is constructed in a step by step fashion; we show how to proceed after $\langle\varphi_1, \ldots, \varphi_j\rangle$ has been constructed. We fix an enumeration of all the implications $\varphi_i = \chi_i \to \theta_i$ in Φ such that $\varphi_i \notin \Delta_{\langle\varphi_1,\ldots,\varphi_j\rangle}$. These φ_i are treated in turn depending on which of the following two possibilities applies to it:

If $\chi_i \in \Delta_{\langle\varphi_1,\ldots,\varphi_k\rangle}$, then by the reflexiveness of $\mathcal{R}$ condition 2(iii) is already fulfilled (check that $\theta_i \notin \Delta_{\langle\varphi_i,\ldots,\varphi_k\rangle}$ must hold on account of consistency) and we leave everything as it is.

If $\chi_i \notin \Delta_{\langle\varphi_1,\ldots,\varphi_k\rangle}$, a new point $\langle\varphi_1,\ldots,\varphi_k,\varphi_i\rangle$ is created, with which we associate a $\Delta_{\langle\varphi_1,\ldots,\varphi_k,\varphi_i\rangle}$ such that $\Delta_{\langle\varphi_i,\ldots,\varphi_k\rangle} \subseteq \Delta_{\langle\varphi_1,\ldots,\varphi_k,\varphi_i\rangle}$, $\chi_i \in \Delta_{\langle\varphi_1,\ldots,\varphi_k,\varphi_i\rangle}$ and $\theta_i \notin \Delta_{\langle\varphi_1,\ldots,\varphi_k,\varphi_i\rangle}$

The crux in understanding why the resulting partial order is bound to be finite is that in either case φ_i never has to be treated again on the same branch: for all future $\Delta_{\langle\varphi_1,\ldots,\varphi_k,\xi_1,\ldots\xi_m\rangle}$ (in the second case $\xi_1 = \varphi_i$) $\chi_i \in \Delta_{\langle\varphi_1,\ldots,\varphi_k,\xi_1,\ldots\xi_m\rangle}$ and if $\theta_i \in \Delta_{\langle\varphi_1,\ldots,\varphi_k,\xi_1,\ldots\xi_m\rangle}$, the reflexiveness of $\mathcal{R}$ handles the matter again, because if $\theta_i \in \Delta_{\langle\varphi_1,\ldots,\varphi_k,\xi_1,\ldots\xi_m\rangle}$, the implication φ_i must be in it also, owing to $\Delta_{\langle\varphi_1,\ldots,\varphi_k,\xi_1,\ldots\xi_m\rangle}$'s theory-hood, and from then on there is no case left to be handled any more. Starting with a finite list of formulas-to-be-treated, the stock of problem cases must finally give out and therefore every branch of the tree will have only a finite height, while the same circumstance prevents a more than finite 'fanning out' in breadth. •

6.5 Relating intuitionistic logic to intensional logic

We already announced that the object of this final section would be to state precisely what the connection between intuitionistic and intensional logic is. The precise statement will take the form of a translation (i.e. a function *) relating each formula of the standard propositional language to a particular formula in the intensional language in such a manner that all theorems of **IPC** translate into theorems of an appropriately chosen intensional system **S**, while the non-theorems of **IPC** are correlated to non-**S**-derivable formulas; such a translation is said to *interpret* the first logic in the second. If we succeed in finding a translation that embeds **IPC** in some intensional logic, we thereby achieve a precise delineation of that part of the intensional logic in question that can be identified with intuitionistic logic.

In choosing an appropriate intensional system we can use our previous completeness results as a guide. Trying **S4** would seem a good start on that ground, for we already know this system to be complete with respect to the same frames that underlie the intuitionistic models, viz. the class of reflexive partial orderings. All that remains to be done then is finding a way to isolate, by the means of we have at our disposal in **S4**, those models that comply with the intuitionistic demands on the propositional letters and to imitate, with the help of $\Box$, the typical intuitionistic behaviour of the $\rightarrow$. The translation-operation * defined below does the job, as is checked in the proposition that follows it.

Definition 29 We define the function * from the standard propositional formulas to the modal propositional formulas by means of induction:

(i) $p^* = \Box p$
(ii) $(\varphi \wedge \psi)^* = \varphi^* \wedge \psi^*$
(iii) $(\varphi \vee \psi)^* = \varphi^* \vee \psi^*$
(iv) $\bot^* = \bot$
(v) $(\varphi \to \psi)^* = \Box(\varphi^* \to \psi^*)$

In words, φ^* is obtained from φ by placing $\Box$ in front of all the propositional letters and implications in φ ; exercise 63 shows however that, working in **S4**, we could have placed $\Box$ in front of the modal formulas in the other clauses as well.

Proposition 75 $\vdash_{\mathbf{IPC}} \varphi$ iff $\vdash_{\mathbf{S4}} \varphi^*$

Proof: We give a semantic argument, although we could have established the theorem in a syntactic manner also (see exercise 64 for one half of this claim):

$\Leftarrow$: Suppose $\nvdash_{\mathbf{IPC}} \varphi$. Then there is a model $\langle\langle \mathcal{I}, \mathcal{R}\rangle, \models\rangle$ with $\mathcal{I}$ finite and $\mathcal{R}$ a reflexive partial ordering, such that for some $i_0 \in \mathcal{I}$ $i_0 \not\models \varphi$. If we replace this $\models$ by a forcing-relation $\models^*$ for the modal language (i.e. $\models^*$ knows what to do with $\Box$ and acts on $\to$ in the standard fashion), we can show, setting $i \models^* p$ iff $i \models p$ for all the propositional letters p, that $i \models^* \psi^*$ iff $i \models \psi$ holds for all ψ, whence we get $i_0 \not\models^* \varphi$ and so $\nvdash_{\mathbf{S}} \varphi$. The claim can be established by a simple induction, the execution of which we leave as exercise 65.

$\Rightarrow$: The converse goes rather similar. The only thing worth noticing is that we must use the counterexamples arrived at by the method C for the formula φ that cannot be derived in **S4**, because neither the method H nor the method C_{fin} can be trusted to deliver partial orderings. •

Looking at things from another perspective suggests that we should try to establish a similar correspondence between the logic **L** of chapter 4 and **IPC**, for both logics are organized around the notion of provability; **L** is supposed to represent (classical) PA-derivability, whereas **IPC** is made to capture the logic of constructive provability. The completeness proof for **L** teaches that this logic is complete with respect to the *irreflexive* partial orderings, so in addition to the matters that had to be taken care of in connection with **S4**, this time we also have to obviate the possibility that the models on the irreflexive frames get 'in conflict' with **IPC**'s reflexive models. The translation ** defined below does precisely this; the verification of this fact is exercise 66.

Definition 30 The translation ** from propositional formulas to modal formulas is defined by:

(i) $p^{**} = \Box p \wedge p$
(ii) $(\varphi \wedge \psi)^{**} = \varphi^{**} \wedge \psi^{**}$
(iii) $(\varphi \vee \psi)^{**} = \varphi^{**} \vee \psi^{**}$

(iv) $\bot^{**} = \bot$
(v) $(\varphi \to \psi)^{**} = \Box(\varphi^{**} \to \psi^{**}) \wedge (\varphi^{**} \to \psi^{**})$

The two intensional logics mentioned here are not the only ones in which **IPC** can be interpreted; other systems, some even weaker than **S4** can be given which relate in the same manner to the intuitionistic system. However, none of these are well-known logics and therefore considering them would have been less interesting and less clarifying.

6.6 Exercises

Exercise 59 Show that the characteristic condition on $\models$ of the intuitionistic models carries over to formulas, i.e. prove that for all φ and i, j such that $i\mathcal{R}j$ it holds that if $i \models \varphi$, then $j \models \varphi$.

Exercise 60 Prove the part of the soundness theorem for **IPC** concerning the axioms 1, 2 and 3.

Exercise 61

(i) Check that the counterexample in figure 15 does the job it is claimed to do.
(ii) Show that $\not\vdash_{\mathbf{IPC}} (p \to q) \vee (q \to p)$ and $\not\vdash_{\mathbf{IPC}} (p \to (q \vee r)) \to ((p \to q) \vee (p \to r))$

.

Exercise 62

(i) Show that if $\vdash_{\mathbf{IPC}} \varphi \vee \psi$, then $\vdash_{\mathbf{IPC}} \varphi$ or $\vdash_{\mathbf{IPC}} \psi$.
(ii) Show that if $\vdash_{\mathbf{IPC}} \neg\theta \to (\varphi \vee \psi)$, then $\vdash_{\mathbf{IPC}} \neg\theta \to \varphi$ or $\vdash_{\mathbf{IPC}} \neg\theta \to \psi$, but $\not\vdash_{\mathbf{IPC}} (\neg\theta \to (\varphi \vee \psi)) \to ((\neg\theta \to \varphi) \vee (\neg\theta \to \psi))$. Compare these results for **IPC** with the standard propositional calculus.
(iii) Show that if φ is a theorem of the standard propositional calculus, then $\vdash_{\mathbf{IPC}} \neg\neg\varphi$. (Hint: consider the valuations in points from which no other points are accessible.)

Exercise 63 Show that for all φ, $\vdash_{\mathbf{S4}} \varphi^* \to \Box\varphi^*$.

Exercise 64 Use the previous exercise to prove that for **IPC**-axioms φ, $\vdash_{\mathbf{S4}} \varphi^*$ and that if $\vdash_{\mathbf{S4}} \varphi^*$ and $\vdash_{\mathbf{S4}} (\varphi \to \psi)^*$, then $\vdash_{\mathbf{S4}} \psi^*$. Note this proves one half of proposition 75 in a syntactic manner.

Exercise 65 Prove the induction-part of proposition 75

Exercise 66 Prove that the translation ** interprets **IPC** into **L**.