4 Models and Proofs

4.1 Soundness and completeness

You may find that this chapter is harder and more abstract than the previous chapters. Feel free to skip or skim it if you're mostly interested in philosophical applications.

We have introduced several kinds of validity: S5-validity, K-validity, T-validity, and so on. All of these are defined in terms of models. K-validity means truth at all worlds in all Kripke models. T-validity means truth at all worlds in all reflexive Kripke models. S5-validity means truth at all worlds in all universal Kripke models (equivalently, at all worlds in all "basic" models). And so on.

If you want to show that a sentence is, say, K-valid, you could directly work through the clauses of definition 3.2, showing that there is no world in any Kripke model in which the sentence is false. The tree method regiments and simplifies this process. If you construct a tree for your sentence in accordance with the K-rules and all branches close, then the sentence is K-valid. If some branch remains open, the sentence isn't K-valid.

Or so I claimed. But these claims aren't obvious. The tree rule for the diamond, for example, appears to assume that if $\Diamond A$ is true at a world then A is true at some accessible world *that does not yet occur on the branch*. Couldn't $\Diamond A$ be true because A is true at an accessible "old" world instead? Also, why do we expand $\Diamond A$ nodes only once? Couldn't A be true at multiple accessible worlds?

In the next two sections, we are going to lay any such worries to rest. We are going to prove that (1) if all branches on a K-tree close then the target sentence is K-valid; conversely, (2) if some branch on a fully developed K-tree remains open, then the target sentence is not K-valid. (1) establishes the *soundness* of the tree rules for K, (2) establishes their *completeness*.

When you use the tree method, you don't have to think of what you are doing as exploring Kripke models. I could have introduced the method as a purely syntactic game. You start the game by writing down the negation of the target sentence, followed by '(w)' (and possibly '1.' to the left and '(Ass.)' to the right, although in this chapter we will mostly ignore these book-keeping annotations.) Then you repeatedly apply the tree rules until either all branches are closed or no rule can be applied any more. At no point in the game do you need to think about what any of the symbols you are writing might mean.

Soundness and completeness link this syntactic game with the "model-theoretic" concept of validity. Soundness says that if the game leads to a closed tree (a tree in which all branches are closed) then the target sentence is true at all worlds in all models. Completeness says that if the game doesn't lead to a closed tree then the target sentence is not true at all worlds in all models. This is called completeness because it implies that every valid sentence can be shown to be valid with the tree method.

In general, a proof method is called **sound** if everything that is provable with the method is valid. A method is **complete** if everything that is valid is provable. Strictly speaking, we should say that a method is sound or complete *for a given concept of validity*. The tree rules for K are sound and complete for *K-validity*, but not for T-validity or S5-validity.

The tree method is not the only method for showing that a sentence is K-valid (or T-valid, or S5-valid). Instead of constructing a K-tree, you could construct an axiomatic proof, trying to derive the target sentence from some instances of (Dual) and (K) by (Nec) and (CPL). This, too, can be done as a purely syntactic exercise, without attending to the meaning of the relevant sentences. In section 4.4, we will show that the axiomatic calculus for K is indeed sound and complete for K-validity: all and only the K-valid sentences can be derived from (Dual) and (K) by (Nec) and (CPL). The 'all' part is completeness, the 'only' part soundness. Having shown soundness and completeness for both the tree method and the axiomatic method, we will have shown that the two methods are equivalent. Anything that can be shown with one method can also be shown with the other.

There are other styles of proof besides the axiomatic and the tree format. Two famous styles that we won't cover are "natural deduction" methods and "sequence calculi". Logicians are liberal about what qualifies as a proof method. The only non-negotiable condition is that there must be a mechanical way of checking whether something (usually, some configuration of symbols) is or is not a proof of a given target sentence.

Exercise 4.1

What do you think of the following proposals for new proof methods?

(a) In *method A*, every $\mathfrak{L}_M$-sentence is a proof of itself: To prove an $\mathfrak{L}_M$-sentence with this method, you simply write down the sentence.

(b) In *method B*, every $\mathfrak{L}_M$-sentence that is an instance of $\Box(A \vee \neg A)$ is a proof of itself. Nothing else is a proof in method B.

(c) In *method C*, a proof of a sentence A is a list of $\mathfrak{L}_M$-sentences terminating with A and in which every sentence occurs in some logic textbook.

Which of these qualify as genuine proof methods by the criterion I have described?

Exercise 4.2

Which, if any, of the methods from the previous exercise are sound for K-validity? Which, if any, are complete?

4.2 Soundness for trees

We are now going to show that the tree method for K is sound – that every sentence that can be proved with the method is K-valid. A proof in the tree method is a tree in which all branches are closed. So this is what we have to show:

> Whenever all branches on a K-tree close then the target sentence is K-valid.

By a *K-tree* I mean a tree that conforms to the K-rules from the previous chapter.

I'll first explain the proof idea, then I'll fill in the details. We will assume that there is a K-tree for some target sentence A on which all branches close. We need to show that A is K-valid. To this end, we suppose for reductio that A is *not* K-valid. By definition 3.3, a sentence is K-valid iff it is true at all worlds in all Kripke models. Our supposition that A is not K-valid therefore means that A is false at some world in some Kripke model. Let's call that world 'w' and the model 'M'. Note that the closed tree begins with

$$1.\quad \neg A \quad (w)$$

If we take the world variable 'w' on the tree to pick out world w in M, then node 1 is a correct statement about M, insofar as $\neg A$ is indeed true at w in M. Now we can show the following:

> *If all nodes on some branch of a tree are correct statements about M, and the branch is extended by the K-rules, then all nodes on at least one of the resulting branches are still correct statements about M.*

Since our closed tree is constructed from node 1 by applying the K-rules, it follows that all nodes on some branch of the tree are correct statements about M. But every branch of a closed tree contains a pair of contradictory statements, which can't both be correct statements about M. This completes the reductio.

Let's fill in the details. We first define precisely what it means for the nodes on a tree branch to be correct statements about a model.

Definition 4.1

A tree node is an **correct statement about** a Kripke model $M = \langle M, R, V \rangle$ **under** a function f that maps world variables to members of W iff either the node has the form $\omega R \upsilon$ and $f(\omega) R f(\upsilon)$, or the node has the form $A\ (\omega)$ and A is true at $f(\omega)$ in M.

A tree branch **correctly describes** a model M iff there is a function f under which all nodes on the branch are correct statements about M.

We now prove the italicised statement above:

Soundness Lemma

If some branch β on a tree correctly describes a Kripke model M, and the branch is extended by applying a K-rule, then at least one of the resulting branches correctly describes M.

Proof: We have to go through all the K-rules. In each case we assume that the rule is applied to some node(s) on a branch β that correctly describes M, so that there is a function f under which all nodes on the branch are correct statements about M. We show that once the rule has been applied, at least one of the resulting branches still correctly describes M.

- Suppose β contains a node of the form $A \wedge B\ (\omega)$ and the branch is extended by two new nodes $A\ (\omega)$ and $B\ (\omega)$. Since $A \wedge B\ (\omega)$ is a correct statement about M under f, we have $M, f(\omega) \models A \wedge B$. By clause (c) of definition 3.2, it follows that $M, f(\omega) \models A$ and $M, f(\omega) \models B$. So the extended branch still correctly describes M.
- Suppose β contains a node of the form $A \vee B\ (\omega)$ and the branch is split into two, with $A\ (\omega)$ appended to one end and $B\ (\omega)$ to the other. Since the expanded node is a correct statement about M under f, we have $M, f(\omega) \models A \vee B$. By clause (d) of definition 3.2, it follows that either $M, f(\omega) \models A$ or $M, f(\omega) \models B$. So at least one of the resulting branches also correctly describes M.

The proof for the other non-modal rules is similar. Let's look at the rules for the modal operators.

- Suppose β contains nodes of the form $\Box A\ (\omega)$ and $\omega R \upsilon$, and the branch is extended by adding $A\ (\upsilon)$. Since $\Box A\ (\omega)$ and $\omega R \upsilon$ are correct statement about M under f, we have $M, f(\omega) \models \Box A$ and $f(\omega) R f(\upsilon)$. By clause (g) of definition 3.2, it follows that $M, f(\upsilon) \models A$. So the extended branch correctly describes M.
- Suppose β contains a node of the form $\Diamond A\ (\omega)$ and the branch is extended by adding nodes $\omega R \upsilon$ and $A\ (\upsilon)$, where υ is new on the branch. Since $\Diamond A\ (\omega)$ is a correct statement about M under f, we have $M, f(\omega) \models \Diamond A$. By clause (h) of definition 3.2, it follows that $M, v \models A$ for some v in W such that $f(\omega) R v$. Let f' be the same as f except that $f'(\upsilon) = v$. The newly added nodes are correct statements about M under f'. Since υ is new on the branch, all earlier nodes on the branch are also correct statements about M under f'. So the expanded branch correctly describes M.

The cases for $\neg\Box$ and $\neg\Diamond$ are similar to the previous two cases. □

With the help of this lemma, we can prove that the method of K-trees is sound.

Theorem: Soundness of K-trees

If a K-tree for a target sentence closes, then the target sentence is K-valid.

Proof: Suppose for reductio that some K-tree for some target sentence A closes even though A is not K-valid. Then $\neg A$ is true at some world w in some Kripke model M. The first node on the tree, $\neg A \ (w)$, is a correct statement about M under the function that maps the world variable 'w' to w. Since the tree is created from the first node by applying the K-rules, the Soundness Lemma implies that some branch β on the tree correctly describes M: all nodes on the tree are correct statements about M under some function f. But the tree is closed. This means that β contains contradictory nodes of the form

n.	B	(v)
m.	$\neg B$	(v)

If both of these are correct statements about M under f, then $M, f(v) \models B$ and also $M, f(v) \models \neg B$. This is impossible by definition 3.2. □

Exercise 4.3

Spell out the cases for $A \to B$ and $\neg\Diamond A$ in the proof of the Soundness Lemma.

Exercise 4.4

Draw the K-tree for target sentence $\Box p$. The tree has a single open branch. Does this branch correctly describe the Kripke model in which there is just one world w, w has access to itself, and all sentence letters are false at w?

The soundness proof for K-trees is easily adapted to other types of trees. The tree rules for system T, for example, are all the K-rules plus the Reflexivity rule, which allows adding $\omega R \omega$ for every world ω on the branch. Suppose we want to show that everything that is provable with the T-rules is T-valid – true at every world in every reflexive Kripke model. All the clauses in the Soundness Lemma still hold if we assume that the model M is reflexive. We only need to add a further clause for the Reflexivity rule, to confirm that if a branch correctly describes a reflexive model M, and the branch is extended by adding $\omega R\omega$, then the resulting branch also correctly describes M. This is evidently the case.

Exercise 4.5

How would we need to adjust the soundness proof to show that the tree rules for K4 are sound with respect to K4-validity?

4.3 Completeness for trees

Let's now show that the tree rules for K are complete – that whenever a sentence is K-valid then there is a closed K-tree for that sentence. In fact, we will show something stronger:

> If a sentence is K-valid, then every fully developed K-tree for the sentence is closed.

By a *fully developed* tree, I mean a tree on which every node on any open branch that can be expanded (in any way) has been expanded (in this way). A fully developed tree may be infinite.

We will prove the displayed sentence by proving its contraposition:

> If a fully developed K-tree for a sentence does not close, then the sentence is not K-valid.

Assume, then, that some fully developed K-tree for some target sentence has at least one open branch. We want to show that the target sentence is false at some world in some Kripke model.

We already know how to read off a countermodel from an open branch. All we need to do is show that this method for generating countermodels really works. Let's first define the method more precisely.

Definition 4.2

The model **induced by** a tree branch is the Kripke model (W, R, V) where

(a) W is the set of world variables on the branch,
(b) ωRv holds in the model iff a node ωRv occurs on the branch,

(c) for any sentence letter P, $V(P)$ is the set of world variables ω for which a node $P\ (\omega)$ occurs on the branch.

Next we show that all nodes on any open branch on a fully developed tree are correct statements about the Kripke model induced by the branch.

Completeness Lemma

Let β be an open branch on a fully developed K-tree, and let $M = \langle W, R, V \rangle$ be the model induced by β. Then $M, \omega \models A$ for all sentences A and world variables ω for which $A\ (\omega)$ is on β.

We have to show that whenever $A\ (\omega)$ occurs on β then $M, \omega \models A$. The proof is by induction on the length of A. We first show that the claim holds for sentence letters and negated sentence letters. Then we show that *if* the claim holds for all sentences shorter than A (this is our induction hypothesis), *then* it also holds for A itself.

- If A is a sentence letter then the claim is true by clause (c) of definition 4.2 and clause (a) of definition 3.2.
- If A is the negation of a sentence letter B, then $B\ (\omega)$ does not occur on β, otherwise β would be closed. By clause (c) of definition 4.2, it follows that ω is not in $V(B)$, and so $M, \omega \models A$ by clauses (a) and (b) of definition 3.2.
- If A is a doubly negated sentence $\neg\neg B$, then β contains a node $B\ (\omega)$, because the tree is fully developed. By induction hypothesis, $M, \omega \models B$. By clause (b) of definition 3.2, it follows that $M, \omega \models A$.
- If A is a conjunction $B \wedge C$, then β contains nodes $B\ (\omega)$ and $C\ (\omega)$. By induction hypothesis, $M, \omega \models B$ and $M, \omega \models C$. By clause (c) of definition 3.2, it follows that $M, \omega \models A$.
- If A is a negated conjunction $\neg(B \wedge C)$, then β contains either $\neg B\ (\omega)$ or $\neg C\ (\omega)$. By induction hypothesis, $M, \omega \models \neg B$ or $M, \omega \models \neg C$. Either way, clauses (b) and (c) of definition 3.2 imply that $M, \omega \models A$.

I will skip the cases where A is a disjunction, a conditional, a biconditional, or a negated disjunction, conditional, or biconditional. The proofs are similar to one (or both) of the previous two cases.

- If A is a box sentence $\Box B$, then β contains a node B (v) for each world variable v for which $\omega R v$ is on β (because the tree is fully developed). By induction hypothesis, $M, v \models B$, for each such v. By definition 4.2, it follows that $M, v \models B$ for all worlds v such that $\omega R v$. By clause (g) of definition 3.2, it follows that $M, \omega \models \Box B$.
- If A is a diamond sentence $\Diamond B$, then there is a world variable v for which $\omega R v$ and B (v) are on β. By induction hypothesis, $M, v \models B$. And by definition 4.2, $\omega R v$. By clause (h) of definition 3.2, it follows that $M, \omega \models \Diamond B$.

For the case where A has the form $\neg\Box B$ or $\neg\Diamond B$, the proof is similar to one of the previous two cases. □

To establish completeness, we need to verify one more point: that one can always construct a fully developed tree for any invalid target sentence. Let's call a K-tree *regular* if it is constructed by (i) first applying all rules for the truth-functional connectives until no more of them can be applied (without adding only nodes to a branch that are already on the branch), then (ii) applying the rules for $\Diamond$ and $\neg\Box$ until no more of them can be applied, then (iii) applying the rules for $\Box$ and $\neg\Diamond$ until no more of them can be applied, then starting over with (i), and so on.

Observation 4.1: Every regular open K-tree is fully developed.

Proof: When constructing a regular tree, every iteration of (i), (ii), and (iii) only allows expanding finitely many nodes. So every node on every open branch that can be expanded in any way is eventually expanded in this way by some iteration of (i), (ii), and (iii). □

Now we have all the ingredients to prove completeness.

Theorem: Completeness of K-trees

If a sentence is K-valid, then there is a closed K-tree for that sentence.

Proof: Let A be any K-valid sentence, and suppose for reductio that there is no closed K-tree for A. In particular, then, every regular K-tree for A remains open. Take any such tree. By observation 4.1, the tree is fully expanded. Choose any open branch on the tree. By the Completeness Lemma, A is false at w in the model induced by that branch. So A is not true at all worlds in all Kripke models. Contradiction. □

Exercise 4.6

Fill in the cases for $B \to C$ and $\neg\Diamond B$ in the proof of the Completeness Lemma.

Like the soundness proof, the completeness proof for K is easily adapted to other logics. To show that the T-rules are complete with respect to T-validity, for example, we merely need check that the model induced by any open branch on a fully developed T-tree is reflexive. It must be, because an open branch on a fully developed T-tree contains $\omega R \omega$ for each world variable ω on the branch.

Exercise 4.7

What do we need to check to show that the K4-rules are complete with respect to K4-validity?

Exercise 4.8

A Kripke model is *acyclical* if you can never return to the same world by following the accessibility relation. Show that if a sentence is true at some world in some Kripke model, then it is also true at some world in some acyclical Kripke model.
(Hint: If A is true at some world in some Kripke model then $\neg A$ is K-invalid. By the soundness theorem, there is a fully developed K-tree for $\neg A$ with an open branch. Now consider the model induced by this branch.)

Exercise 4.9

The S5 tree rules from chapter 2 are sound and complete for S5-validity: all and only the S5-valid sentences can be proven. Are the rules sound for K-validity? Are they complete for K-validity?

4.4 Soundness and completeness for axiomatic calculi

Next, we are going to show that the axiomatic calculus for system K is sound and complete for K-validity. In the axiomatic calculus, a proof is a list of sentences each of which is either an instance of (Dual) or (K) or can be derived from earlier sentences on the list by application of (CPL) or (Nec). Expressed as a construction rule, (Nec) says that whenever a list contains a sentence A then one may append $\Box A$. (CPL) says that one may append any truth-functional consequence of sentences that are already on the list. (This is an acceptable rule because there is a simple mechanical test – the truth-table method – for checking whether a sentence is a truth-functional consequence of finitely many other sentences.)

Soundness is easy. We want to show that everything that is derivable from some instances of (Dual) and (K) by applications of (CPL) and (Nec) is K-valid. We show this by showing that (1) every instance of (Dual) and (K) is K-valid, and (2) every sentence that is derived from K-valid sentences by (CPL) or (Nec) is itself K-valid.

Theorem: Soundness of the axiomatic calculus for K

Any sentence that is provable in the axiomatic calculus for K is K-valid.

Proof: We first show that every instance of (Dual) and (K) is K-valid.

1. (Dual) is the schema $\neg\Diamond A \leftrightarrow \Box\neg A$. By clauses (b), (g), and (h) of definition 3.2, a sentence $\neg\Diamond A$ is true at a world w in a Kripke model M iff $\Box\neg A$ is true at w in M. It follows by clauses (f) and (e) that all instances of (Dual) are true at all worlds in all Kripke models.
2. (K) is the schema $\Box(A \to B) \to (\Box A \to \Box B)$. By clause (e) of definition 3.2, a sentence $\Box(A \to B) \to (\Box A \to \Box B)$ is false at a world w in a Kripke model M only if $\Box(A \to B)$ and $\Box A$ are both true at w while B is false. By clause (g) of

definition 3.2, $\Box B$ is false at w only if B is false at some world v accessible from w. But if $\Box(A \to B)$ and $\Box A$ are both true at w, then $A \to B$ and A are true at every world accessible from w, again by clause (g). And there can be no world at which $A \to B$ and A are true while B is false, by clause (e) of definition 3.2.

Next we show that (CPL) and (Nec) preserve K-validity.

1. By definition 3.2, the truth-functional operators have their standard truth-table meaning at every world in every Kripke model. It follows that all truth-functional consequences of sentences that are true at a world are themselves true at that world. In particular, if some sentences are true at every world in every Kripke model, then any truth-functional consequence of these sentences is also true at every world every Kripke model.

2. Let w be an arbitrary world in an arbitrary Kripke model. If A is true at every world in every Kripke model, then A is true at every world accessible from w, in which case $\Box A$ is true at w by clause (g) of definition 3.2. So if A is K-valid, then $\Box A$ is also K-valid. □

The soundness proof for K is easily extended to other modal systems. Since all instances of (Dual) and (K) are true at all worlds in all Kripke models, they are also true at all worlds in any more restricted class of Kripke models. The arguments for (CPL) and (Nec) also go through if we replace 'every Kripke model' by 'every Kripke model of such-and-such type'. So if we want to show that, say, the axiomatic calculus for T is sound with respect to the concept of T-validity – that is, if we want to show that anything that is derivable from (Dual), (K), and (T) by (CPL) and (Nec) is true at all worlds in all reflexive Kripke models – all that is left to do is to show that every instance of the (T)-schema is true at all worlds in all reflexive Kripke model. (We've already shown this: see observation 3.2.)

Exercise 4.10

Outline the soundness proof for the axiomatic calculus for S4, whose axiom schemas are (Dual), (K), (T), and (4).

Let's turn to completeness. We are going to show that every K-valid sentence is derivable from some instances of (Dual) and (K) by (CPL) and (Nec). As in section 4.3, we argue by contraposition. We will show that any sentence that cannot be derived from (Dual) and (K) by (CPL) and (Nec) is not K-valid. To show that a sentence is not K-valid, we will give a countermodel – a Kripke model in which the sentence is false at some world. In fact, we will give the *same* countermodel for every sentence that isn't derivable in the calculus. You might think we need different countermodels for different sentences, but it turns out that there is a particular model in which every K-invalid sentence is false at some world. This model is called the *canonical model* for K.

In order to define the canonical model, let's introduce some shorthand terminology. We'll say that an $\mathfrak{L}_M$-sentence is *K-provable* if it can be proved in the axiomatic calculus for K. A set of $\mathfrak{L}_M$-sentences is *K-inconsistent* if it contains a finite number of sentences $A_1, \ldots, A_n$ such that $\neg(A_1 \wedge \ldots \wedge A_n)$ is K-provable. A set is *K-consistent* if it is not K-inconsistent.

(For example, the set $\{\Box(p \wedge q), q \to p, \neg\Box q\}$ is K-inconsistent, because it contains two sentences, $\Box(p \wedge q)$ and $\neg\Box q$ whose conjunction is refutable in K, in the sense that the negation $\neg(\Box(p \wedge q) \wedge \neg\Box q)$ of their conjunction is derivable from some instances of (Dual) and (K) by (CPL) and (Nec).)

A set of $\mathfrak{L}_M$-sentences is called *maximal* if it contains either A or $\neg A$ for every $\mathfrak{L}_M$-sentence A. A set is *maximal K-consistent* if it is both maximal and K-consistent.

Exercise 4.11

Which, if any, of these sets are K-consistent? (a) $\{p\}$, (b) $\{\neg p\}$, (c) the set of all sentence letters, (d) the set of all $\mathfrak{L}_M$-sentences.

Now here's the canonical model for K.

Definition 4.3

The **canonical model** M_K for K is the Kripke model $\langle W, R, V \rangle$, where

- W is the set of all maximal K-consistent sets of $\mathfrak{L}_M$-sentences,
- wRv iff v contains every sentence A for which w contains $\Box A$,
- for every sentence letter P, $V(P)$ is the set of all members of W that contain P.

The "worlds" in the canonical model are sets of $\mathfrak{L}_M$-sentences. The interpretation function makes a sentence letter true at a world iff the letter is a member of the world. As we are going to see, this generalizes to arbitrary sentences:

(1) A world w in M_K contains all and only the sentences that are true at w in M_K.

We will also prove the following:

(2) If some sentence cannot be proved in the axiomatic calculus for K, then its negation is a member of some world in M_K.

Together, these two lemmas will establish completeness for the axiomatic calculus. Fact (2) tells us that if a sentence A isn't K-provable, then $\neg A$ is a member of some world w in the canonical model M_K. By fact (1), we can infer that $\neg A$ is true at w in M_K, which means that A is false at w in M_K. So any sentence that isn't K-provable isn't K-valid.

We are going to prove (2) first. We'll need the following observation.

Observation 4.2: If a set Γ is K-consistent, then for any sentence A, either $\Gamma \cup \{A\}$ or $\Gamma \cup \{\neg A\}$ is K-consistent.

($\Gamma \cup \{A\}$, called the *union* of Γ and $\{A\}$, is the smallest set that contains all members of Γ as well as A.)

Proof: Let Γ be any K-consistent set and A any sentence. Suppose for reductio that $\Gamma \cup \{A\}$ and $\Gamma \cup \{\neg A\}$ are both K-inconsistent.

That $\Gamma \cup \{A\}$ is K-inconsistent means there are sentences $A_1, \ldots, A_n$ in $\Gamma \cup \{A\}$ such that $\neg(A_1 \land \ldots \land A_n)$ is K-provable. Since Γ itself is K-consistent, one of the sentences $A_1, \ldots, A_n$ must be A. Let B be the conjunction of the other sentences in $A_1, \ldots, A_n$, all of which are in Γ. So $\neg(B \land A)$ is K-provable.

That $\Gamma \cup \{\neg A\}$ is K-inconsistent means that there are sentences $A_1, \ldots, A_n$ in $\Gamma \cup \{\neg A\}$ such that $\neg(A_1 \land \ldots \land A_n)$ is K-provable. As before, one of these sentences must be $\neg A$. Let C be the conjunction of the others, all of which are in Γ. So $\neg(C \land \neg A)$ is K-provable.

If $\neg(B \land A)$ and $\neg(C \land \neg A)$ are both K-provable, then so is $\neg(B \land C)$, because it is a truth-functional consequence of $\neg(B \land A)$ and $\neg(C \land \neg A)$. But $B \land C$ is a conjunction of sentences from Γ. So Γ itself is K-inconsistent, contradicting our assumption. □

Now we can prove fact (2).

Lindenbaum's Lemma

Every K-consistent set is a subset of some maximal K-consistent set.

Proof: Let S_0 be some K-consistent set of sentences. Let $A_1, A_2, \ldots$ be a list of all $\mathfrak{L}_M$-sentences in some arbitrary order. For every number $i \geq 0$, define

$$S_{i+1} = \begin{cases} S_i \cup \{A_i\} & \text{if } S_i \cup \{A_i\} \text{ is K-consistent} \\ S_i \cup \{\neg A_i\} & \text{otherwise.} \end{cases}$$

This gives us an infinite list of sets $S_0, S_1, S_2, \ldots$. Each set in the list is K-consistent: S_0 is K-consistent by assumption. And if some set S_i in the list is K-consistent, then either $S_i \cup \{A_i\}$ is K-consistent, in which case $S_{i+1} = S_i \cup \{A_i\}$ is K-consistent, or $S_i \cup \{A_i\}$ is not K-consistent, in which case S_{i+1} is $S_i \cup \{\neg A_i\}$, which is K-consistent by observation 4.2. So if any set in the list is consistent, then the next set in the list is also consistent. It follows that $S_0, S_1, S_2, \ldots$ are all K-consistent.

Now let S be the set of sentences that occur in at least one of the sets $S_0, S_1, S_2, S_3 \ldots$. (That is, let S be the union of $S_0, S_1, S_2, S_3, \ldots$.) Evidently, S_0 a subset of S. And S is maximal. Moreover, S is K-consistent. For if S were not K-consistent, then it would contain some sentences $B_1, \ldots, B_n$ such that $\neg(B_1 \wedge \ldots \wedge B_n)$ is K-provable. All of these sentences would have to occur somewhere on the list $A_1, A_2, \ldots$. Let A_j be a sentence from $A_1, A_2, \ldots$ that occurs after all the $B_1, \ldots, B_n$. If $B_1, \ldots, B_n$ are in S, they would have to be in S_j already, so S_j would be K-inconsistent. But we've seen that all of $S_0, S_1, S_2, \ldots$ are K-consistent. □

Notice that the proof of Lindenbaum's Lemma does not turn on any assumptions about the axiomatic calculus for K except that (CPL) is one of its rules. The lemma holds for every calculus with (CPL) as a (possibly derived) rule.

To prove fact (1), we need another observation, which relies on the presence of (K) and (Nec), besides (CPL).

Observation 4.3: If Γ is a maximal K-consistent set of sentences that does not contain $\Box A$, and Γ^- is the set of all sentences B for which $\Box B$ is in Γ, then $\Gamma^- \cup \{\neg A\}$ is K-consistent.

Proof: We show that if $\Gamma^- \cup \{\neg A\}$ is not K-consistent, then neither is Γ. If $\Gamma^- \cup \{\neg A\}$ is not K-consistent, then there are sentences $B_1, \ldots, B_n$ in Γ^- such that $\neg(B_1 \wedge \ldots \wedge B_n \wedge \neg A)$ is K-provable. And then $(B_1 \wedge \ldots \wedge B_n) \to A$ is K-provable, because it is a truth-functional consequence of $\neg(B_1 \wedge \ldots \wedge B_n \wedge \neg A)$. By repeated application of (Nec), (K), and (CPL), one can derive $(\Box B_1 \wedge \ldots \wedge \Box B_n) \to \Box A$ from $(B_1 \wedge \ldots \wedge B_n) \to A$. Another application of (CPL) yields $\neg(\Box B_1 \wedge \ldots \wedge \Box B_n \wedge \neg\Box A)$. So $\{\Box B_1, \ldots, \Box B_n, \neg\Box A\}$ is K-inconsistent. But $\Box B_1, \ldots, \Box B_n$ are in Γ^-. And since $\Box A$ is not in Γ and Γ is maximal, $\neg\Box A$ is in Γ. So $\{\Box B_1, \ldots, \Box B_n, \neg\Box A\}$ is a subset of Γ. And so Γ is K-inconsistent. □

Here, then, is fact (1):

Canonical Model Lemma

For any world w in M_K and any sentence A, A is in w iff $M_K, w \models A$.

Proof: The proof is by induction on complexity of A. We first show that the claim (that A is in w iff $M_K, w \models A$) holds for sentence letters. Then we show that if the claim holds for the immediate parts of a complex sentence (this is our induction hypothesis), then the claim also holds for the sentence itself.

- Suppose A is a sentence letter. By definition 4.3, $w \in V(A)$ iff $A \in w$. So by clause (a) of definition 3.2, $M_K, w \models A$ iff $A \in w$. ('$\in$' means 'is a member of the set'.)
- Suppose A is a negation $\neg B$. By clause (b) of definition 3.2, $M_K, w \models \neg B$ iff $M_K, w \not\models B$. By induction hypothesis, $M_K, w \not\models B$ iff $B \notin w$. Since w is maximal K-consistent, $B \notin w$ iff $\neg B \in w$. So $M_K, w \models \neg B$ iff $\neg B \in w$.
- Suppose A is a conjunction $B \wedge C$. By clause (c) of definition 3.2, $M_K, w \models B \wedge C$ iff $M_K, w \models B$ and $M_K, w \models C$. By induction hypothesis, $M_K, w \models B$ iff $B \in w$, and $M_K, w \models C$ iff $C \in w$. Since w is maximal K-consistent, B and C are in w iff $B \wedge C$ is in w. So $M_K, w \models B \wedge C$ iff $B \wedge C \in w$.

The cases for the other truth-functional connectives are similar.

- Suppose A is a box sentence $\Box B$, and that $\Box B \in w$. By definition 4.3, it follows that $B \in v$ for all v with wRv. By induction hypothesis, this means that $M_K, v \models B$ for all v with wRv. And then $M_K, w \models \Box B$, by clause (g) of definition 3.2.

  For the converse direction, suppose $\Box B \notin w$. Let Γ^- be the set of all sentences C for which $\Box C \in w$. By observation 4.3, $\Gamma^- \cup \{\neg B\}$ is K-consistent. By definition 4.3 and Lindenbaum's Lemma, it follows that there is some $v \in W$ such that wRv and $\neg B \in v$. Since v is K-consistent, $B \notin v$. By induction hypothesis, it follows that $M_K, v \not\models B$. And so $M_K, w \not\models \Box B$, by clause (g) of definition 3.2.

- Suppose A is a diamond sentence $\Diamond B$, and that $\Diamond B \in w$. By (Dual) and (CPL), any set that contains both $\Diamond B$ and $\Box\neg B$ is K-inconsistent. So $\Box\neg B \notin w$. By observation 4.3 and Lindenbaum's Lemma (as in the previous case), it follows that there is some $v \in W$ such that wRv and $B \in v$. By induction hypothesis, $M, v \models B$. So $M_K, w \models \Diamond B$, by clause (h) of definition 3.2.

  For the converse direction, suppose $\Diamond B \notin w$. Then $\Box\neg B \in w$, by (Dual), (CPL), and the fact that w is maximal K-consistent. By definition 4.3, it follows that $\neg B \in v$ for all v with wRv. Since all such v are maximal K-consistent, none of them contain B. By induction hypothesis, B is not true at any of them. By clause (h) of definition 3.2, it follows that $M_K, w \not\models \Diamond B$. □

The completeness of the axiomatic calculus for K follows immediately from the previous two lemmas, as foreshadowed above:

Theorem: Completeness of the axiomatic calculus for K

If A is K-valid, then A is provable in the axiomatic calculus for K.

Proof: We show that if a sentence is not K-provable then it is not K-valid. Suppose A is not K-provable. Then $\{\neg A\}$ is K-consistent. It follows by Lindenbaum's Lemma that $\{\neg A\}$ is included in some maximal K-consistent set S. By definition 4.3, that set is a world in M_K. Since $\neg A$ is in S, it follows from the Canonical Model Lemma that $M_K, S \models \neg A$. So $M_K, S \not\models A$. So A is not true at all worlds in all Kripke models. □

Done!

Once again, the proof is easily adjusted to many axiomatic calculi for logics stronger than K. All we have assumed about the K-calculus is that it contains (Dual), (K), (Nec), and (CPL). So if we're interested in, say, whether the axiomatic calculus for T is complete, we can simply replace 'K-consistent' by 'T-consistent' throughout the proof, and almost everything goes through as before. We only have to add a small step at the end.

By adapting the argument for K, we can show that if a sentence A is not T-provable then A is false at some world in the canonical model for T. This shows that A is not K-valid. But we want to show that A is not T-valid – meaning that A is not true at all worlds in all reflexive Kripke models. To complete the proof, we need to show that the canonical model M_T for T is reflexive.

This isn't hard. Given how accessibility in canonical models is defined, a world w in a canonical model is accessible from itself iff whenever $\Box A \in w$ then $A \in w$. Since the worlds in M_T are maximal T-consistent sets of sentences, and every such set contains every instance of the (T) schema $\Box A \to A$, there is no world in M_T that contains $\Box A$ but not A. So every world in M_T has access to itself.

In general, to show that a calculus that extends the K-calculus by further axiom schemas is complete, we only need to show that the canonical model for the calculus satisfies the frame conditions that correspond to the added axiom schemas. This is usually the case. But not always. Sometimes, an axiomatic calculus is sound and complete with respect to some class of Kripke models, but the canonical model of the calculus is not a member of that class. (An example is the calculus for the system GL, which I will describe at the very end of this chapter.) Completeness must then be established by some other means.

Exercise 4.12

Outline the completeness proof for the axiomatic calculus for S5.

Exercise 4.13

The set of all $\mathfrak{L}_M$-sentences is a system of modal logic. Let's call this system X (for "explosion"). (a) Describe a sound and complete proof method for X. (b) Explain why X does not have a canonical model.

4.5 Loose ends

You will remember from observation 1.1 in chapter 1 that claims about entailment can be converted into claims about validity. A entails B iff $A \to B$ is valid; A_1 and A_2 together entail B iff $A_1 \to (A_2 \to B)$ – equivalently, $(A_1 \land A_2) \to B$ – is valid; and so on. But what if there are infinitely many premises $A_1, A_2, A_3, \ldots$? Sentences of $\mathfrak{L}_M$ are always finite, so we can't convert the claim that $A_1, A_2, A_3, \ldots$ entail B into a claim that some $\mathfrak{L}_M$-sentence is valid.

We also can't use the tree method or the axiomatic method to directly show that a conclusion follows from infinitely many premises. A proof in either method is a finite object that can only invoke finitely many sentences.

As it turns out, this is not a serious limitation. In many logics – including classical propositional and predicate logic and all the modal logics we have so far encountered – a sentence is entailed by infinitely many premises only if it is entailed by a finite subset of these premises. Logics with this property are called **compact**.

Let's show that K is compact. To this end, I'll say that a sentence B is *K-derivable* from a (possibly infinite) set of sentences Γ if there are finitely many members $A_1, \ldots, A_n$ of Γ for which $(A_1 \land \ldots \land A_n) \to B$ is provable in the axiomatic calculus for K. Now we first show that whenever $\Gamma \models_K B$ then B is K-derivable from Γ. This is called *strong completeness* because it is stronger than the ("weak") kind of completeness that we have established in the previous section.

Theorem: Strong completeness of the axiomatic calculus for K

Whenever $\Gamma \models_K B$ then B is K-derivable from Γ.

Proof: Suppose B is not K-derivable from Γ. Then there are no $A_1, \ldots, A_n$ in Γ such that $(A_1 \land \ldots \land A_n) \to B$ is K-provable. This means that $\Gamma \cup \{\neg B\}$ is K-consistent. By Lindenbaum's Lemma, it follows that $\Gamma \cup \{\neg B\}$ is included in some maximal K-consistent set and thereby in some world in the canonical model M_K for K. (Lindenbaum's lemma says that every K-consistent set of $\mathfrak{L}_M$-sentences, even if it is infinite, is included in a maximal K-consistent set.) By the Canonical Model Lemma, $M_K, w \models_K A$ for all A in Γ, and $M_K, w \not\models_K B$. Thus $\Gamma \not\models_K B$. □

Theorem: Compactness of K

If a sentence B is K-entailed by some sentences Γ, then B is K-entailed by a finite subset of Γ.

Proof: Suppose $\Gamma \models_K B$. By strong completeness, it follows that there are finitely many sentences $A_1, \ldots, A_n$ in Γ for which $(A_1 \land \ldots \land A_n) \to B$ is K-provable. By the soundness of the K-calculus, $(A_1 \land \ldots \land A_n) \to B$ is valid. So $A_1, \ldots, A_n \models_K B$, by observation 1.1. □

Compactness is surprising. It is easy to think of cases in which a conclusion is entailed by infinitely many premises, but not by any finite subset of these premises. For example, suppose I like the number 0, I like the number 1, I like the number 2, and so on, for all natural numbers 0,1,2,3,…. Together, these assumptions entail that I like every natural number. But no finite subset of the assumptions has this consequence.

Exercise 4.14

A set of sentences Γ is called *K-satisfiable* if there is a world in some Kripke model at which all members of Γ are true. Show that an infinite set of sentences Γ is K-satisfiable iff every finite subset of Γ is K-satisfiable.

Our proofs of soundness, completeness, compactness, etc. were informal. We have not translated the relevant claims into a formal language, nor have we used a formal method of proof. In principle, however, this can be done. All our proofs could be formalized in an axiomatic calculus for predicate logic with a few additional axioms about sets. A well-known calculus of that kind is ZFC (named after Ernst Zermelo, Abraham Fraenkel, and the Axiom of Choice). ZFC is strong enough to prove not just soundness and completeness in modal logic, but practically everything that can be proved in any branch of maths.

An interesting feature of ZFC is that it can not only prove facts about what's provable in simpler axiomatic calculi; it can also prove facts about what's provable in ZFC itself. For example, one can prove in ZFC that one can prove in ZFC that 2+2=4.

This gives us an interesting application of modal logic. Let's read the box as 'it is mathematically provable that', which we understand as provability in ZFC. One can easily show (in ZFC) that this operator has all the properties of the box in the basic logic K. For example, all instances of the (K)-schema are provable in ZFC. (The language of ZFC doesn't have a box symbol. But one can encode the (K)-schema into a schema of ZFC, given the present reading of the box, and all instances of that schema are ZFC-provable.)

So the logic of mathematical provability is at least as strong as K. In fact, it is stronger. One can prove in ZFC that whenever a sentence is ZFC-provable then it is ZFC-provable that the sentence is ZFC-provable. This gives us the (4)-schema $\Box A \to \Box\Box A$.

You might expect that we also have the (T)-schema $\Box A \to A$ or the (D)-schema $\Box A \to \Diamond A$. The latter says that if something is provable then its negation isn't provable (since $\Diamond A$ means $\neg\Box\neg A$). And surely ZFC can't prove both a sentence and its negation – which would make ZFC inconsistent. I say 'surely', but can we prove (in ZFC) that ZFC is consistent? The answer is no. More precisely, one can prove that if one can prove that ZFC is consistent then ZFC is *in*consistent. This bizarre fact is a consequence of *Gödel's second incompleteness theorem*, established by Kurt Gödel in 1931. It is reflected by the following schema (named after Gödel and Martin Löb), all whose instances are provable in ZFC:

(GL) $\quad \Box(\Box A \to A) \to \Box A$

The system GL, which is axiomatized by (K), (GL), (Nec), and (CPL), completely captures what ZFC can prove about provability in ZFC. (Schema (4) isn't needed as a separate axiom schema because it can be derived.)

Exercise 4.15

Suppose ZFC can prove its own consistency, so that there is a proof of $\neg\Box(p \land \neg p)$. Explain how this proof could be extended to a proof of $\Box(p \land \neg p)$, showing that ZFC is inconsistent. You need each of (GL), (Nec), and (CPL).