Lecture Notes on Modal Propositional Logic

Semantics, Proof Theory, and Correspondence

Pascal Ludwig
Sorbonne Université, SND (CNRS)

Contents

Part I
From Classical Logic to Modal Logic

1 Introduction: Non-Classical Logics

Classical propositional logic and first-order logic provide an extraordinarily powerful framework, but they rest on assumptions that are not always appropriate. There are broadly two strategies for departing from classical logic.

Revising classical logic. One can *reject* certain classical laws. Examples include many-valued logics, which drop bivalence; intuitionistic logic, which drops the law of excluded middle and double negation elimination; and relevant logics, which impose a relevance condition on the conditional.

Extending classical logic. Alternatively, one can *add* new expressive resources while retaining some or all classical laws. Second-order logic, for instance, adds quantification over properties and relations while conserving every first-order law. Modal logics add operators such as $\Box$ ("it is necessary that…") and $\Diamond$ ("it is possible that…"); in doing so they modify certain substitution principles, for reasons we shall now examine.

2 Extension, Intension, and Extensionality

2.1 Extension versus Intension

In first-order logic, every expression is assigned an *extension* relative to a structure:

- An individual constant is assigned an object in the domain (its *reference*).
- A predicate is assigned the set of objects to which it applies.
- A sentence is assigned a truth value.

Two expressions can share the same extension while differing in meaning. When this happens, we say they have distinct *intensions*.

Example 2.1. The predicates "has a heart" and "has a kidney" are co-extensive:

$$\{x \mid x \text{ has a heart}\} = \{x \mid x \text{ has a kidney}\}.$$

Yet the two expressions plainly differ in meaning; they contribute different information, and it is conceivable that they come apart. Their *intensions* are different.

2.2 The Extensionality of Classical Logic

In classical logic, the truth value of a statement is determined entirely by two factors:

1. the *extension* (reference) of its non-logical symbols, and
2. the *syntactic structure* of the statement.

No further information—no "meaning" beyond reference—plays any role. This yields a powerful substitution principle:

> **Extensionality Principle.** Expressions with the same extensions or references can be substituted for one another *salva veritate*, that is, without changing the truth value of any statement in which they occur.

2.3 Failures of Extensionality

When we move to natural-language contexts involving belief, knowledge, or necessity, the extensionality principle visibly breaks down.

Co-referent proper names. "Eric Rohmer" and "Maurice Schérer" refer to the same individual. Yet from

> Razib believes Eric Rohmer directed *La boulangère de Monceau*

we cannot conclude

> Razib believes Maurice Schérer directed *La boulangère de Monceau*.

Razib may simply not know that Eric Rohmer *is* Maurice Schérer.

Co-extensive predicates. From "Lila knows that if someone has a heart, they have a heart" it does not follow that "Lila knows that if someone has a heart, they have a kidney."

Co-denoting statements. "$2 + 2 = 4$" and "M. Ludwig has a stylish bike" may well have the same truth value, yet "It is necessary that $2+2 = 4$" is true, while "It is necessary that M. Ludwig has a stylish bike" is not.

These examples show that contexts created by "believes", "knows", "it is necessary that", etc., are *intensional*: the truth value of the whole depends on more than the extensions of its parts.

3 The Core Idea of Intensional Semantics

How should one model intensionality? The strategy is to *abandon* the assumption that each expression has a single, fixed denotation, and instead **relativise** denotations and truth values to *evaluation circumstances*.

> **Key idea.** A statement such as "M. Ludwig has a stylish bike" can be true in some circumstances and false in others, whereas "$2 + 2 = 4$" is true in all of them. The evaluation circumstances are called **indices**.

At this stage, one need not say much about what indices *are*. In an alethic (truth/necessity) reading, they are naturally thought of as *possible worlds*; in a temporal reading, they could be instants of time; in a deontic reading, they could be idealised situations in which all obligations are fulfilled. The formal machinery is deliberately neutral.

4 Syntax of Modal Propositional Logic

We extend the language $\mathcal{L}_P$ of propositional logic by adding a single new primitive symbol:

$$\Box \quad \text{("box"—read: "it is necessary that").}$$

A second operator is defined from $\Box$:

$$\Diamond \varphi \;=_{\text{Def}}\; \neg \Box \neg \varphi \quad \text{("diamond"—read: "it is possible that").}$$

The grammar of $\mathcal{L}_P$ is supplemented by a single new clause:

> If φ is a well-formed formula, then $\Box \varphi$ is a well-formed formula.

Exercise 4.1. Show that if φ is a well-formed formula, so is $\Diamond \varphi$.

Solution. Suppose φ is a wff. By the rule for negation, $\neg \varphi$ is a wff. By the new rule, $\Box \neg \varphi$ is a wff. By the rule for negation again, $\neg \Box \neg \varphi$ is a wff. Since $\Diamond \varphi$ is defined as $\neg \Box \neg \varphi$, it follows that $\Diamond \varphi$ is a wff. $\square$

Part II
Kripke Semantics

5 Intensional Models (Without Accessibility)

Before introducing the full Kripke semantics, it is instructive to work with a simpler notion of model in which the set of indices carries no additional structure.

Definition 5.1 (Intensional model)**.** An **intensional model** is a pair $\langle W, I\rangle$, where:

- W is a non-empty set of **indices**, and
- I is an **interpretation function** that assigns a truth value to each pair (propositional variable, index).

Formally, $I\colon \mathrm{Var} \times W \to \{\mathsf{T}, \mathsf{F}\}$.

Example 5.1. Consider a model $\mathcal{M} = \langle W, I\rangle$ with $W = \{\omega_1, \omega_2\}$, $I(p, \omega_1) = \mathsf{T}$, and $I(p, \omega_2) = \mathsf{F}$.

One could read p as "M. Ludwig's bike is stylish". In this model, the statement is true at ω_1 and false at ω_2.

We write $\mathcal{M}, \omega \models p$ (or equivalently $\mathcal{M} \models_\omega p$) to mean that the propositional variable p is true in $\mathcal{M}$ at index ω, and $\mathcal{M}, \omega \not\models p$ to mean that it is false there.

5.1 Tableaux for Intensional Semantics

The tableau method carries over straightforwardly: one simply annotates each signed formula with the index at which it is being evaluated. For convenience, we use numerals $(1, 2, 3, \ldots)$ to represent indices.

Example 5.2. Suppose that at some index in a model we have $\mathsf{T}\colon p \wedge q$ and $\mathsf{F}\colon q$. We build the tableau at index 1:

$$\begin{array}{l} \mathsf{T} : p \wedge q,\ 1\ \checkmark \\ \mathsf{F} : q,\ 1 \\ \quad\quad\big| \\ \mathsf{T} : p,\ 1 \\ \mathsf{T} : q,\ 1 \\ \quad\quad\times \end{array}$$

The branch closes because it contains both $\mathsf{T} : q,\ 1$ and $\mathsf{F} : q,\ 1$.

Remark 5.1. At this stage, nothing connects different indices to one another. The propositional rules operate within a single index, exactly as in classical logic. The novelty—and the need for an accessibility relation—arises when we ask what $\Box\varphi$ means.

6 Why an Accessibility Relation Is Needed

The simplest proposal for the truth-conditions of $\Box\varphi$ goes back to Leibniz:

> $\Box\varphi$ is true iff φ is true *in all indices*.

Under this definition, one immediately obtains the schema $\models \Box\varphi \to \varphi$: if φ is true in *all* indices, it is certainly true in whichever index we happen to be evaluating at.

Problem: the deontic interpretation. Consider the **deontic** interpretation of $\Box$, reading $\Box\varphi$ as "it is obligatory that φ". The schema $\Box\varphi \to \varphi$ would then say: whatever is obligatory is actual. But this is plainly incorrect—obligations can be violated.

More generally, different readings of $\Box$ validate different principles. A semantics that aims to be *general* should not build in assumptions that are appropriate only for some readings.

Solution: restrict the quantification. Instead of quantifying over *all* indices, we quantify only over the *relevant* ones—those that are **accessible** from the index of evaluation.

- Under the alethic reading, the indices accessible from w are the "metaphysically possible" worlds relative to w.
- Under the deontic reading, they are the worlds in which the norms operative at w are satisfied.
- Under the epistemic reading, they are the worlds compatible with what the agent knows at w.

Which indices count as accessible is encoded by a binary relation on the set of indices.

7 Kripke Models and the Truth Definition for K

7.1 Frames and Kripke Models

Definition 7.1 (Frame)**.** A **frame** is a pair $\mathcal{F} = \langle W, R\rangle$ where

- W is a non-empty set (the set of *possible worlds*, or *indices*),
- $R \subseteq W \times W$ is a binary relation on W (the *accessibility relation*).

A frame is simply a directed graph: the worlds are the vertices and the accessibility relation gives the edges.

Definition 7.2 (Kripke model)**.** A **Kripke model** is a triple $\mathcal{M} = \langle W, R, I\rangle$, where:

- $\langle W, R\rangle$ is a frame,
- $I\colon \mathrm{Var} \times W \to \{\mathsf{T}, \mathsf{F}\}$ is an interpretation function (or *valuation*) assigning a truth value to each propositional variable at each world.

The accessibility relation R is left completely unconstrained in the minimal system K. Imposing structural conditions on R—reflexivity, transitivity, symmetry, etc.—yields stronger modal logics (T, $S4$, $S5$, …), which we shall study in Parts VII and VIII.

7.2 Truth Definition for K

Let $\mathcal{M} = \langle W, R, I\rangle$ be a Kripke model and $w \in W$.

Definition 7.3 (Satisfaction in K)**.** The relation $\mathcal{M}, w \models \varphi$ is defined recursively:

$$\mathcal{M}, w \models p \quad \text{iff } I(p, w) = \mathsf{T} \quad (\text{for a propositional variable } p), \tag{1}$$

$$\mathcal{M}, w \models \neg\varphi \quad \text{iff } \mathcal{M}, w \not\models \varphi, \tag{2}$$

$$\mathcal{M}, w \models \varphi \wedge \psi \quad \text{iff } \mathcal{M}, w \models \varphi \text{ and } \mathcal{M}, w \models \psi, \tag{3}$$

$$\mathcal{M}, w \models \varphi \vee \psi \quad \text{iff } \mathcal{M}, w \models \varphi \text{ or } \mathcal{M}, w \models \psi, \tag{4}$$

$$\mathcal{M}, w \models \varphi \to \psi \quad \text{iff } \mathcal{M}, w \not\models \varphi \text{ or } \mathcal{M}, w \models \psi, \tag{5}$$

$$\mathcal{M}, w \not\models \bot, \tag{6}$$

$$\mathcal{M}, w \models \Box\varphi \quad \text{iff for all } w' \text{ such that } Rww',\ \mathcal{M}, w' \models \varphi, \tag{7}$$

$$\mathcal{M}, w \models \Diamond\varphi \quad \text{iff there exists } w' \text{ such that } Rww' \text{ and } \mathcal{M}, w' \models \varphi. \tag{8}$$

Remark 7.1. The clauses for the Boolean connectives (2)–(6) are exactly the classical truth-table conditions, relativised to an index. The only genuinely new clauses are (7) and (8).

Note that the clause for $\Diamond$ is derivable from those for $\neg$ and $\Box$: $\mathcal{M}, w \models \Diamond\varphi$ iff $\mathcal{M}, w \models \neg\Box\neg\varphi$ iff $\mathcal{M}, w \not\models \Box\neg\varphi$ iff it is *not* the case that for all w' with Rww' we have $\mathcal{M}, w' \models \neg\varphi$; i.e. there is some w' with Rww' and $\mathcal{M}, w' \models \varphi$.

Remark 7.2 (Boundary case). If no index is accessible from w (i.e. there is no w' with Rww'), then $\mathcal{M}, w \models \Box\varphi$ holds *vacuously* for every formula φ. Conversely, $\mathcal{M}, w \not\models \Diamond\varphi$ for every φ. This is a direct consequence of the standard treatment of vacuously true universal statements.

7.3 Semantic Concepts

Definition 7.4. Let φ be a formula and Γ a set of formulae.

1. φ is **K-satisfiable** iff there exists a Kripke model $\mathcal{M} = \langle W, R, I\rangle$ and an index $w \in W$ such that $\mathcal{M}, w \models \varphi$.
2. Γ is **K-satisfiable** iff there exist $\mathcal{M}$ and w such that $\mathcal{M}, w \models \gamma$ for every $\gamma \in \Gamma$.
3. An argument from premises $\varphi_1, \ldots, \varphi_n$ to conclusion ψ has a **counter-model** iff the set $\{\varphi_1, \ldots, \varphi_n, \neg\psi\}$ is K-satisfiable.
   It is **K-valid** iff it has no counter-model. We then write $\varphi_1, \ldots, \varphi_n \models_K \psi$.
4. φ is a **logical truth of** K (written $\models_K \varphi$) iff $\mathcal{M}, w \models \varphi$ for every Kripke model $\mathcal{M}$ and every index w in $\mathcal{M}$.

We also introduce a notion that will become central in Part VII:

Definition 7.5 (Frame validity)**.** A formula φ is **valid in a frame** $\mathcal{F} = \langle W, R\rangle$ (written $\mathcal{F} \models \varphi$) iff $\mathcal{M}, w \models \varphi$ for every model $\mathcal{M}$ built on $\mathcal{F}$ and every world w in $\mathcal{M}$.

Part III
Tableaux for K

8 Rules for $\Box$ (Box)

The tableau rules for $\Box$ follow directly from the truth definition. As in first-order logic, the key distinction is between rules with **universal import** and rules with **existential import**.

8.1 Rule T$\Box$ (Universal Import)

Rule T$\Box$

$$\frac{\mathsf{T} : \Box\varphi,\ i \qquad iRj}{\mathsf{T} : \varphi,\ j}$$

where j is an index that **already occurs** on the branch with iRj. This rule may be applied **multiple times** as new accessible indices appear.

If $\Box\varphi$ is true at i, then φ must be true at every index accessible from i. Therefore, whenever we have an accessibility link iRj on the branch, we may add $\mathsf{T} : \varphi, j$. This rule does *not* introduce new indices.

8.2 Rule F$\Box$ (Existential Import)

Rule F$\Box$

$$\frac{\mathsf{F} : \Box\varphi,\ i}{\begin{array}{c} iRj \\ \mathsf{F} : \varphi,\ j \end{array}}$$

where j is a **new** index that does **not** yet occur on the branch.

If $\Box\varphi$ is false at i, then there must be some accessible index at which φ fails. We introduce a fresh index j, declare iRj, and record $\mathsf{F} : \varphi, j$.

Remark 8.1. The analogy with first-order logic is exact. The T$\Box$ rule behaves like T$\forall$ (universal import: instantiate to all known indices); the F$\Box$ rule behaves like F$\forall$ (existential import: introduce a fresh witness).

9 Rules for $\Diamond$ (Diamond)

9.1 Rule T$\Diamond$ (Existential Import)

Rule T$\Diamond$

$$\frac{\mathsf{T} : \Diamond\varphi,\ i}{\begin{array}{c} iRj \\ \mathsf{T} : \varphi,\ j \end{array}}$$

where j is a **new** index that does **not** yet occur on the branch.

9.2 Rule F◊ (Universal Import)

Rule F◊

$$\frac{\mathsf{F} : \Diamond\varphi,\ i \qquad iRj}{\mathsf{F} : \varphi,\ j}$$

where j is an index that **already occurs** on the branch with iRj. This rule may be applied **multiple times** as new accessible indices appear.

Remark 9.1. The duality is clean: the T◊ rule is existential (introduces a new index), just as F□; the F◊ rule is universal (applies to existing accessible indices), just as T□. This reflects the equivalence $\Diamond\varphi \leftrightarrow \neg\Box\neg\varphi$.

10 Summary of Rules and Strategy

10.1 Summary of Modal Rules

Rule	Import	Index	Branching
T□	Universal	Existing j with iRj	No
F□	Existential	New j (adds iRj)	No
T◊	Existential	New j (adds iRj)	No
F◊	Universal	Existing j with iRj	No

Remark 10.1. None of the modal rules is branching. Branching only arises from the propositional rules (F∧, T∨, T →).

10.2 Order of Rule Application

Procedure for Constructing Modal Tableaux

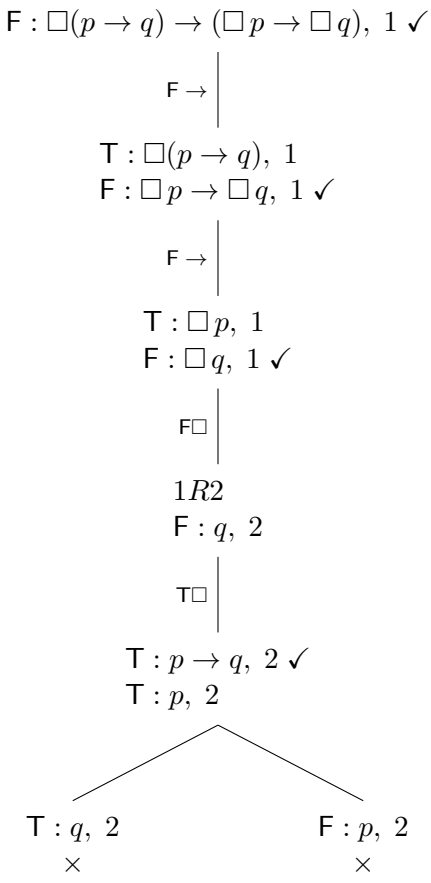
1. **Apply propositional rules** ($\neg$, $\wedge$, $\vee$, $\rightarrow$) at each index until no more can be applied.
2. **Apply rules with existential import** (T◊ and F□); each application introduces a new index. Mark these formulas: they need only be applied once.
3. **Return to step 1**: new propositional material may have appeared at the new indices.
4. **Apply rules with universal import** (T□ and F◊) to every index j such that iRj is on the branch. Do **not** mark these formulas: they may need to be re-applied as new accessible indices appear.
5. **Return to step 1**.

A tableau is **finished** when either all branches are closed, or all applicable rules have been exhausted on every open branch.

11 Worked Examples: Proving Validity

11.1 The K-Axiom: $\models_K \Box(p \rightarrow q) \rightarrow (\Box p \rightarrow \Box q)$

This is the characteristic axiom of the system K (whence the name). To show it is a logical truth, we attempt to build a counter-model by opening a tableau with its negation.



Explanation:

1. Apply $\mathsf{F} \to$ to the root: get $\mathsf{T} : \Box(p \to q)$, 1 and $\mathsf{F} : \Box p \to \Box q$, 1.
2. Apply $\mathsf{F} \to$: get $\mathsf{T} : \Box p$, 1 and $\mathsf{F} : \Box q$, 1.
3. Apply $\mathsf{F}\Box$ (existential import) to $\mathsf{F} : \Box q$, 1: introduce index 2, add $1R2$ and $\mathsf{F} : q$, 2.
4. Apply $\mathsf{T}\Box$ (universal import) to $\mathsf{T} : \Box(p \to q)$, 1 and $\mathsf{T} : \Box p$, 1 with $1R2$: get $\mathsf{T} : p \to q$, 2 and $\mathsf{T} : p$, 2.
5. Apply $\mathsf{T} \to$ to $\mathsf{T} : p \to q$, 2: branch into $\mathsf{F} : p$, 2 and $\mathsf{T} : q$, 2.

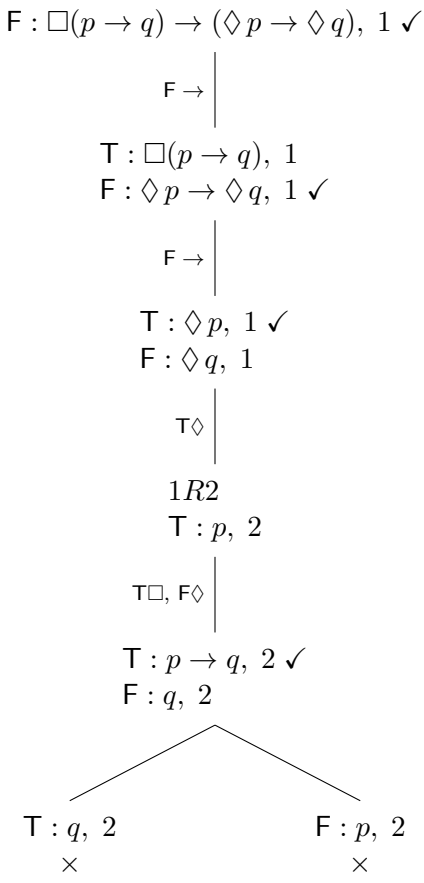
A contradiction is found on each branch:
- Left: $\mathsf{T} : p$, 2 and $\mathsf{F} : p$, 2. $\quad\times$
- Right: $\mathsf{F} : q$, 2 and $\mathsf{T} : q$, 2. $\quad\times$

All branches are closed, therefore the formula is a **logical truth of** K.

> **The Distribution Axiom.** The schema $\Box(\varphi \to \psi) \to (\Box\varphi \to \Box\psi)$ is called the *distribution axiom* (axiom K). Together with the rule of necessitation ("if φ is a theorem, so is $\Box\varphi$"), it characterises the minimal normal modal logic K.

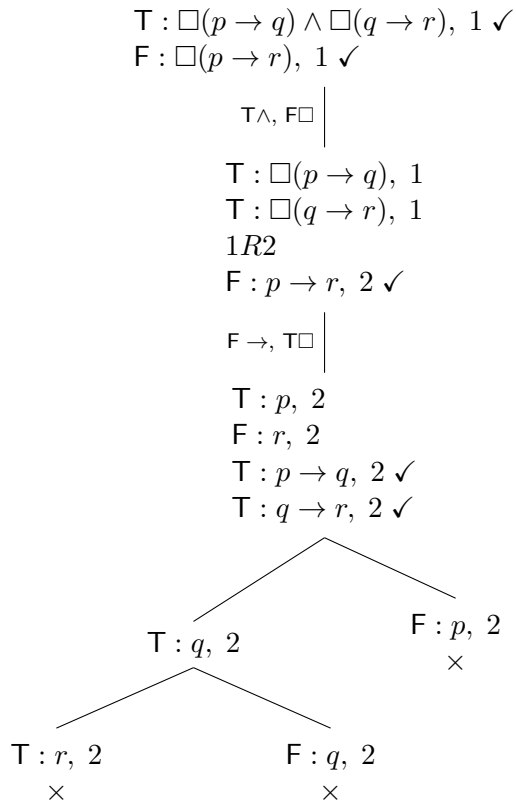
11.2 A Variant: $\models_K \Box(p \to q) \to (\Diamond p \to \Diamond q)$

We verify that necessity distributes analogously over $\Diamond$.



All branches close. The formula is a **logical truth of** K. □

11.3 Transitivity: $\Box(p \to q) \wedge \Box(q \to r) \vdash_K \Box(p \to r)$



All branches close. The argument is K-**valid**. □

11.4 The Necessitation Rule

Statement. If φ is a logical truth of K (i.e. $\models_K \varphi$), then $\Box\varphi$ is also a logical truth of K (i.e. $\models_K \Box\varphi$).

Semantic argument. Suppose $\models_K \varphi$: for every Kripke model $\mathcal{M}$ and every index w in $\mathcal{M}$, $\mathcal{M}, w \models \varphi$. Now let $\mathcal{M}$ be an arbitrary model and w an arbitrary index; we want to show $\mathcal{M}, w \models \Box\varphi$. By Definition 7.3, this requires that $\mathcal{M}, w' \models \varphi$ for every w' accessible from w. But each such w' is an index in $\mathcal{M}$, and we assumed that φ is true at *every* index in *every* model. So $\mathcal{M}, w' \models \varphi$ holds, and $\mathcal{M}, w \models \Box\varphi$ follows. Since $\mathcal{M}$ and w were arbitrary, $\models_K \Box\varphi$.

> **Necessitation Rule.** If $\models_K \varphi$, then $\models_K \Box\varphi$.
>
> Together with the distribution axiom $\Box(\varphi \to \psi) \to (\Box\varphi \to \Box\psi)$, this rule characterises the minimal normal modal logic K: every theorem of K can be derived from propositional tautologies using modus ponens, the distribution axiom, and the necessitation rule.

11.5 A Common Mistake: Necessitation Is Not a Conditional

Students sometimes confuse the necessitation rule with the schema $\varphi \to \Box\varphi$ ("if φ, then necessarily φ"). The necessitation rule is a *meta-logical* principle: it says that *if φ is a theorem* (true in all models at all indices), then $\Box\varphi$ is also a theorem. It does *not* say that whenever φ happens to be true at some index, $\Box\varphi$ is true at that index.

The schema $p \to \Box p$ is **not** a logical truth of K:

$\mathsf{F} : p \to \Box p,\ 1\ \checkmark$

$\mathsf{F}\to$

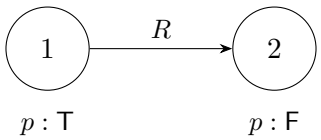
$\mathsf{T} : p,\ 1$
$\mathsf{F} : \Box p,\ 1\ \checkmark$

$\mathsf{F}\Box$

$1R2$
$\mathsf{F} : p,\ 2$

(open)

Counter-model. $W = \{1, 2\}$, $R = \{\langle 1, 2\rangle\}$, $I(p, 1) = \mathsf{T}$, $I(p, 2) = \mathsf{F}$.



At index 1, p is true, but $\Box p$ is false because p fails at the accessible index 2. So $p \to \Box p$ is false at 1.

Remark 11.1. The distinction is worth memorising: the necessitation rule is an *inference rule* ("from a theorem, derive another theorem"); the schema $\varphi \to \Box\varphi$ is a *formula* ("at any given world, if φ then $\Box\varphi$"). The former is valid in K; the latter is not.

12 Worked Examples: Building Counter-Models

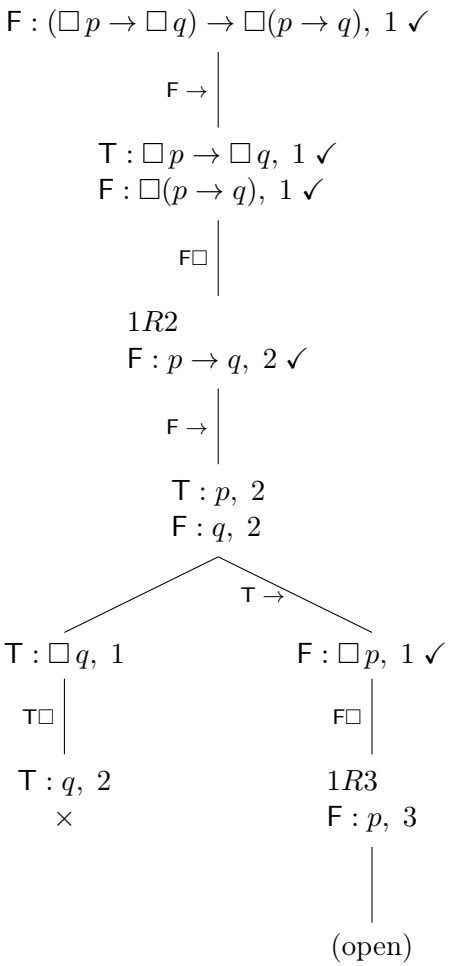
When a finished tableau has at least one *open* branch, the information on that branch can be used to construct a counter-model.

12.1 Recipe for Reading Counter-Models

1. The set of indices W consists of all indices mentioned on the open branch.
2. The accessibility relation R consists of all pairs iRj recorded on the branch.
3. For each propositional variable p and each index i on the branch: set $I(p, i) = \mathsf{T}$ if $\mathsf{T} : p$, i occurs, and $I(p, i) = \mathsf{F}$ if $\mathsf{F} : p$, i occurs. (If neither appears, the value may be chosen freely.)

12.2 Example: $\not\models_K (\Box p \to \Box q) \to \Box(p \to q)$

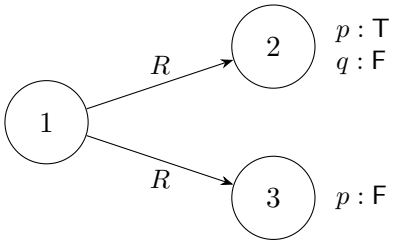
We claim that $(\Box p \to \Box q) \to \Box(p \to q)$ is *not* a logical truth of K.



Reading the Counter-Model

Following the open branch, we collect: indices $1, 2, 3$; accessibility $1R2$, $1R3$; signed atoms $\mathsf{T} : p,\ 2$; $\mathsf{F} : q,\ 2$; $\mathsf{F} : p,\ 3$.

$$W = \{1, 2, 3\}, \qquad R = \{\langle 1, 2\rangle,\ \langle 1, 3\rangle\},$$
$$I(p, 2) = \mathsf{T}, \quad I(p, 3) = \mathsf{F}, \quad I(q, 2) = \mathsf{F}.$$



Verification

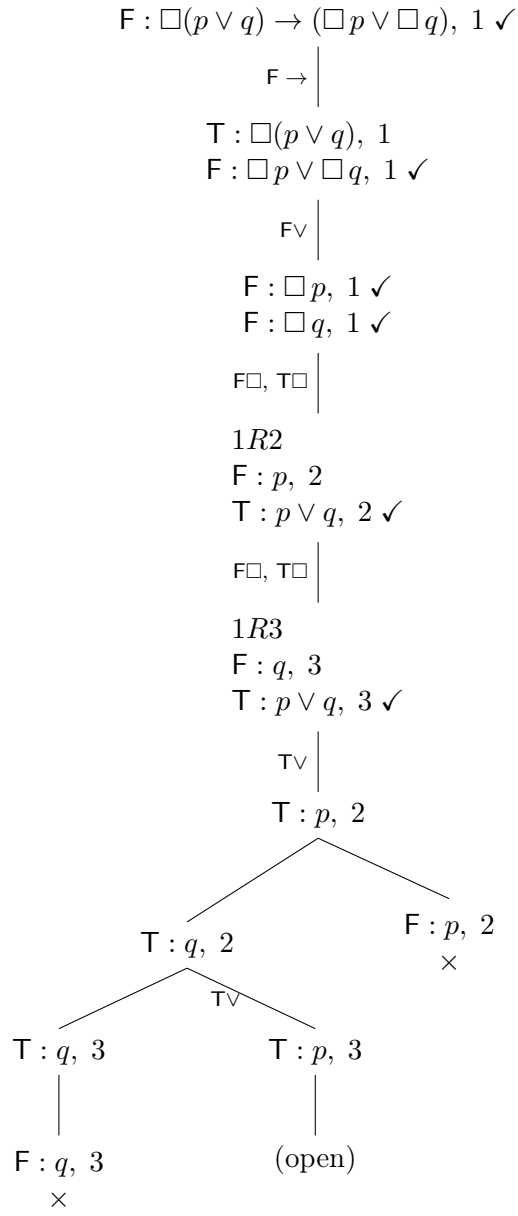
- Since $I(p, 3) = \mathsf{F}$ and $1R3$, we have $\mathcal{M}, 1 \not\models \Box p$. Therefore $\Box p \to \Box q$ is true at 1 (a conditional with a false antecedent).
- Since $I(p, 2) = \mathsf{T}$ and $I(q, 2) = \mathsf{F}$, we have $\mathcal{M}, 2 \not\models p \to q$. Combined with $1R2$, this gives $\mathcal{M}, 1 \not\models \Box(p \to q)$.

- So at index 1 the antecedent $\Box p \to \Box q$ is true and the consequent $\Box(p \to q)$ is false; the whole conditional is false.

Conclusion: $\not\models_K (\Box p \to \Box q) \to \Box(p \to q)$.

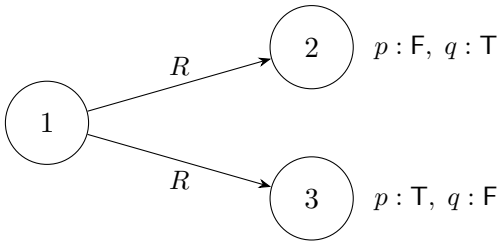
12.3 Necessity Does Not Distribute over Disjunction: $\not\models_K \Box(p \vee q) \to (\Box p \vee \Box q)$

Since $\Box$ distributes over the conditional (the K-axiom), one might expect it to distribute over disjunction as well. It does not.



Reading the Counter-Model

$W = \{1, 2, 3\}$, $R = \{\langle 1, 2\rangle, \langle 1, 3\rangle\}$, $I(p, 2) = \mathsf{F}$, $I(q, 2) = \mathsf{T}$, $I(p, 3) = \mathsf{T}$, $I(q, 3) = \mathsf{F}$.



Conclusion: $\not\models_K \Box(p \vee q) \to (\Box p \vee \Box q)$.

Remark 12.1. The counter-model illustrates the key idea: every accessible world can make $p \vee q$ true by choosing a *different* disjunct, so the disjunction is necessarily true without either disjunct being necessarily true.

13 Laws That K Does Not Validate

K is the weakest normal modal logic: it validates only those principles that hold for *every* reading of $\Box$. Principles that are true under some interpretations but not all fail in K.

13.1 The Reflexivity Axiom: $\not\models_K \Box p \to p$

This schema (known as axiom T) says that whatever is necessary is actual. In the deontic reading, it says that whatever is obligatory is the case—evidently false.

$$\mathsf{F} : \Box p \to p,\ 1\ \checkmark$$
$$\mathsf{F}\to \Big|$$
$$\mathsf{T} : \Box p,\ 1$$
$$\mathsf{F} : p,\ 1$$
$$\Big|$$
$$\text{(open)}$$

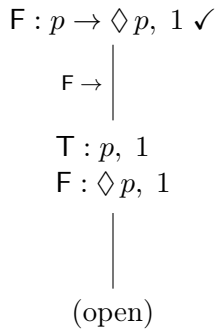
No rule forces the pair $1R1$ to appear on the branch (the accessibility relation in K is unconstrained). Without $1R1$, the universal rule $\mathsf{T}\Box$ cannot extract $\mathsf{T} : p, 1$ from $\mathsf{T} : \Box p, 1$. The branch is **open**.

Counter-model. $W = \{1\}$, $R = \varnothing$, $I(p, 1) = \mathsf{F}$. Since no world is accessible from 1, $\Box p$ is vacuously true at 1; but p is false there.

Remark 13.1. Axiom T corresponds to **reflexivity** of the accessibility relation: $\forall w.\ Rww$. Adding this condition to K yields the modal logic T (see Section 28).

13.2 The Dual: $\not\models_K p \to \Diamond p$

This schema says that whatever is actual is possible.



Counter-model. $W = \{1\}$, $R = \varnothing$, $I(p, 1) = \mathsf{T}$. The world 1 satisfies p but sees no world at all, so $\Diamond p$ is false.

13.3 The Schema $\Box p \to \Diamond p$

This schema also fails in K, by the same mechanism.

Counter-model. $W = \{1\}$, $R = \varnothing$; the value of $I(p, 1)$ does not matter. $\Box p$ is vacuously true and $\Diamond p$ is false, since no world is accessible.

The counter-model may seem artificial: a world that sees nothing at all. But this situation arises naturally in at least two well-studied interpretations of $\Box$.

Provability logic. Read $\Box\varphi$ as "φ is provable in Peano Arithmetic" and $\Diamond\varphi$ as $\neg\Box\neg\varphi$, i.e. "φ is consistent with PA" (the negation of φ is not provable). Then $\Box p \to \Diamond p$ says: if p is provable, then $\neg p$ is not provable—in other words, PA is consistent. By Gödel's second incompleteness theorem, PA cannot prove its own consistency, so this schema fails in the provability interpretation. The modal logic of provability (GL, Gödel–Löb logic) is an extension of K that validates the Löb axiom $\Box(\Box p \to p) \to \Box p$ but does not validate $\Box p \to \Diamond p$. The Kripke frames for GL are finite, transitive, and irreflexive—they contain "leaf" worlds that see nothing, which is exactly where $\Box p$ holds vacuously while $\Diamond p$ fails.

Temporal logic with a last instant. Read $\Box$ as "it will always be the case that" and $\Diamond$ as "it will at some point be the case that", with the accessibility relation being "lies in the future of". If time has an end—a final moment with no successor—then at that moment $\Box p$ is vacuously true (there is no future instant at which p could fail) and $\Diamond p$ is false (there is no future instant at all). So $\Box p \to \Diamond p$ fails at the last moment.

Remark 13.2. In both cases the failure has the same structure: a world that sees nothing, making the universal claim $\Box p$ vacuously true and the existential claim $\Diamond p$ automatically false. The system D (K + seriality: every world sees at least one world) is the weakest system that rules this out, which is why it is the usual baseline for deontic logic—the assumption that there is always at least one normatively ideal world prevents the absurdity of everything being simultaneously obligatory and nothing being permissible.

Part IV
Natural Deduction for K

14 From Tableaux to Proofs

In the preceding parts, we studied the *semantics* of K using Kripke models and the *tableau method* to test validity and build counter-models. We now turn to the *proof theory* of K: a formal system in which valid arguments can be *derived* step by step.

We present a **labelled natural deduction** system. The central idea is simple: every formula in a proof carries a **world label** (a name for a possible world), and **accessibility assertions** are handled explicitly. This approach has two key advantages:

- It connects directly to the Kripke semantics and to the tableaux we already know: the labels are the indices $1, 2, 3, \ldots$ from the tableau rules, and the accessibility pairs iRj are exactly the same.
- The rules for $\Box$ and $\Diamond$ are structurally parallel to the rules for $\forall$ and $\exists$ in first-order logic, making the analogy between quantifiers and modal operators precise.

15 Labelled Formulae

Definition 15.1 (Labelled formula)**.** A **labelled formula** is an expression of the form

$$w : \varphi$$

where w is a **world label** (typically written w, v, u, or simply $1, 2, 3$) and φ is a formula of modal propositional logic. It is read: "φ holds at world w."

Definition 15.2 (Accessibility statement)**.** An **accessibility statement** is an expression of the form

$$wRv$$

where w and v are world labels. It is read: "v is accessible from w."

A proof in our system consists of a sequence of labelled formulae and accessibility statements, each justified either as a hypothesis, as an assumption opening a sub-proof, or as following from earlier lines by a rule.

16 The Rules

16.1 Propositional Rules

All the standard propositional natural deduction rules (for $\wedge$, $\vee$, $\rightarrow$, $\neg$) carry over unchanged, **provided all the formulae involved share the same world label**. In other words, propositional reasoning operates *within a single world*.

For instance, $\rightarrow$-introduction works as follows: if, by assuming $w : \varphi$, we can derive $w : \psi$, then we may close the sub-proof and conclude $w : \varphi \rightarrow \psi$.

> **Important constraint.** Propositional rules cannot mix labels. One cannot, for instance, combine $w : \varphi$ and $v : \varphi \rightarrow \psi$ to conclude $v : \psi$: the modus ponens must happen at a single world.

16.2 Falsum Is Label-Free

There is one symbol that requires special treatment: $\bot$ (*falsum*, absurdity). In our system, $\bot$ **carries no world label**. We write simply $\bot$, never $w : \bot$. This is a deliberate design choice, and it is worth understanding why.

The semantic reason. In the Kripke semantics, every formula is evaluated *at a world*: $\mathcal{M}, w \models p$ says that p is true at world w in model $\mathcal{M}$. A formula like $p \wedge q$ can be true at one world and false at another — its truth value is *world-relative*. This is why labelled formulae carry a world label: the label tells us *where* the formula is being asserted.

But $\bot$ is different. The semantic clause (Definition 7.3, clause 6) says simply:

$$\mathcal{M}, w \not\models \bot \qquad \text{for every model } \mathcal{M} \text{ and every world } w.$$

Falsum is false *everywhere*, in *every* model. Unlike $p \wedge q$, whose truth value varies from world to world, $\bot$ has no variability at all. The world parameter in $\mathcal{M}, w \not\models \bot$ is vacuous — $\bot$ is false regardless of which world we evaluate it at. So the label on $\bot$ would carry *no information*.

The philosophical reason. Think of what a contradiction means. When we derive $\bot$ in a proof, we have shown that the *entire scenario under consideration is impossible* — not that some particular world is defective, but that the whole model cannot exist. If we assume $v : \varphi$ and eventually reach $\bot$ (perhaps via a chain of reasoning that passes through other worlds w, u, ...), the contradiction does not "belong to" any one of these worlds. It is a *global* verdict: the assumption $v : \varphi$ has led to an impossible situation, full stop.

This is why, in ordinary (non-modal) logic, a contradiction refutes any assumption in whose scope it occurs, regardless of where in the proof it arose. The same principle holds in modal logic — but only if we recognise that $\bot$ is not tied to a world. Making $\bot$ label-free is what lets this principle hold uniformly.

The proof-theoretic payoff. With label-free $\bot$, the rules for negation and *ex falso* are clean and uniform:

- $\neg$**E** (negation elimination): from $w : \varphi$ and $w : \neg\varphi$, conclude $\bot$. The two premises must share the same label w; the conclusion is label-free.
- $\neg$**I** (negation introduction): if assuming $w : \varphi$ leads to $\bot$, discharge the assumption and conclude $w : \neg\varphi$. Since $\bot$ is label-free, it does not matter *where* in the sub-proof the contradiction arose — at w, at some other world v reached via $\Box$E or $\Diamond$E, anywhere.
- $\bot$**E** (*ex falso quodlibet*): from $\bot$, conclude $w : \psi$ for any world w and any formula ψ. Since $\bot$ is label-free, this rule can produce a formula at any world — which is semantically correct, since an impossible model satisfies everything everywhere.

> With label-free $\bot$, the "same world" constraint on propositional rules holds **uniformly, with no exceptions**: $\wedge$I, $\wedge$E, $\vee$I, $\vee$E, $\rightarrow$I, $\rightarrow$E, $\neg$I, $\neg$E all require their *labelled* premises and conclusions to share the same world label. The only thing that crosses worlds is $\bot$, which has no label to conflict with anything.

Remark 16.1. This design also simplifies the interaction with $\Diamond$E. The freshness condition on $\Diamond$E requires that the conclusion χ does not mention the fresh label v. If the sub-proof ends in $\bot$, this condition is automatically satisfied: $\bot$ mentions no label at all. There is no need for a separate argument about why contradictions are allowed to "escape" the scope of v.

16.3 □-Elimination

Rule □E (□-elimination). From $w : \Box\varphi$ and wRv, conclude $v : \varphi$.

$$\frac{w : \Box\varphi \qquad wRv}{v : \varphi}\ \Box\text{E}$$

This is the simplest modal rule: if φ is necessary at w, and v is accessible from w, then φ holds at v. It directly reflects the semantic clause $\mathcal{M}, w \models \Box\varphi$ iff for all v with Rwv, $\mathcal{M}, v \models \varphi$.

The rule requires no sub-proof, no freshness condition on labels, and no side conditions. It can be applied whenever the two premises are available.

16.4 □-Introduction

Rule □I (□-introduction). Open a sub-proof by assuming a fresh accessibility statement wRv (where v is a new label that does not occur anywhere earlier in the proof). Within this sub-proof, derive $v : \varphi$. Then close the sub-proof and conclude $w : \Box\varphi$.

$$\begin{array}{l} wRv \quad (\text{fresh } v) \\ \vdots \\ v : \varphi \end{array}$$
$$\overline{\qquad w : \Box\varphi}\ \Box\text{I}$$

The idea: to show that φ is necessary at w, we show that φ holds at an *arbitrary* world accessible from w. Since v is fresh, it represents a "generic" accessible world; whatever we prove about it holds for all accessible worlds.

Remark 16.2. The freshness condition on v is exactly parallel to the freshness condition on the eigenvariable in ∀-introduction in first-order logic. In ∀-intro, one introduces an arbitrary constant a, proves $\varphi(a)$, and concludes $\forall x\,\varphi(x)$; here, one introduces an arbitrary accessible world v, proves $v : \varphi$, and concludes $w : \Box\varphi$.

What may be used inside the sub-proof. Inside the □I sub-proof, you may use:

- the assumption wRv itself,
- any formula obtained by applying □E to a boxed formula from outside the sub-proof (i.e. if $w : \Box\psi$ is available outside, you may derive $v : \psi$ inside by □E),
- anything derivable from the above by the other rules.

You may *not* import unboxed formulae from outside the sub-proof into a new label. For instance, if $w : p$ is available outside, you cannot simply write $v : p$ inside: p might be true at w without being true at v.

16.5 ◇-Introduction

Rule ◇I (◇-introduction). From wRv and $v : \varphi$, conclude $w : \Diamond\varphi$.

$$\frac{wRv \qquad v : \varphi}{w : \Diamond\varphi}\ \Diamond\text{I}$$

This is straightforward: if some accessible world satisfies φ, then φ is possible. No sub-proof is needed.

16.6 $\Diamond$-Elimination

Rule $\Diamond$E ($\Diamond$-elimination). From $w : \Diamond\varphi$, open a sub-proof assuming wRv and $v : \varphi$ (where v is a fresh label). Derive a conclusion χ that does not mention v. Then close the sub-proof and assert χ.

$$\dfrac{w : \Diamond\varphi \qquad \begin{array}{l} wRv,\ v : \varphi \quad (\text{fresh } v) \\ \vdots \\ \chi \quad (v \text{ does not occur in } \chi) \end{array}}{\chi}\ \Diamond\text{E}$$

The idea: $\Diamond\varphi$ tells us that *some* accessible world satisfies φ, but we do not know which one. We give it a fresh name v, reason about it, and derive a conclusion that does not depend on the particular identity of v.

Remark 16.3. This is exactly parallel to $\exists$-elimination in first-order logic: from $\exists x\,\varphi(x)$, assume $\varphi(a)$ for a fresh constant a, derive a conclusion not mentioning a, then discharge. The structural analogy between $\Box/\Diamond$ and $\forall/\exists$ is now complete:

	Universal	**Existential**
First-order logic	$\forall$	$\exists$
Modal logic	$\Box$	$\Diamond$
Introduction	fresh eigenvariable / fresh world	witness term / witness world
Elimination	instantiate to any term / world	assume fresh witness

16.7 Summary of the Modal Rules

Rule	**Premises**	**Conclusion**
$\Box$E	$w : \Box\varphi$, wRv	$v : \varphi$
$\Box$I	sub-proof: assume fresh wRv, derive $v : \varphi$	$w : \Box\varphi$
$\Diamond$I	wRv, $v : \varphi$	$w : \Diamond\varphi$
$\Diamond$E	$w : \Diamond\varphi$; sub-proof: assume fresh wRv, $v : \varphi$, derive χ not mentioning v	χ ($v \notin \chi$)

16.8 Why the Rules Are Sound

Each rule preserves truth in Kripke models. The propositional rules and $\Box$E/$\Diamond$I are straightforward; the two sub-proof rules ($\Box$I and $\Diamond$E) require a brief argument.

Soundness of $\Box$I. Suppose that, in every Kripke model, whenever wRv holds we can derive $v : \varphi$ using only what is available inside the sub-proof. We must show that $w : \Box\varphi$ holds, i.e. that $\mathcal{M}, v \models \varphi$ for *every* v with wRv. The **freshness condition** is what makes this work: because v does not appear anywhere else, the derivation of $v : \varphi$ cannot depend on any specific property of v. It applies to an *arbitrary* accessible world, so it applies to all of them.

What goes wrong without freshness. Consider the following incorrect "proof" of $\Diamond p \vdash_K \Box p$:

1	$w : \Diamond p$	P
2	$\quad wRv,\ \ v : p$	fresh v
3	$\quad\quad wRv$	**ERROR: v not fresh!**
4	$\quad\quad v : p$	R, 2
5	$\quad w : \Box p$	$\Box$I, 3—4
6	$w : \Box p$	$\Diamond$E, 1, 2—5

The error is at line 3: v already occurs in line 2, so it is *not* fresh. The $\Box$I sub-proof at lines 3–4 "proves" that p holds at v, but v was introduced as a specific witness for $\Diamond p$—not as a generic accessible world. The "conclusion" $\Diamond p \vdash_K \Box p$ is semantically invalid ($\Diamond p$ does not entail $\Box p$), confirming that the freshness violation produced nonsense.

Soundness of $\Diamond$E. Suppose $w : \Diamond \varphi$ holds and that, by assuming wRv and $v : \varphi$ for a fresh v, we can derive χ (not mentioning v). We must show that χ holds. Since $\Diamond \varphi$ is true at w, there exists *some* world v_0 with wRv_0 and $\mathcal{M}, v_0 \models \varphi$. The derivation of χ from wRv and $v : \varphi$ applies in particular to v_0; since χ does not mention v, the conclusion is independent of which witness was chosen.

16.9 Common Mistakes

Three errors account for the vast majority of incorrect proofs in modal natural deduction. Each corresponds to a violation of the scoping discipline.

Mistake 1: Importing unboxed formulae into a new world. If $w : p$ is available outside a $\Box$I sub-proof, one *cannot* write $v : p$ inside. Only boxed formulae can cross from w to v (via $\Box$E). In Kripke models, p may be true at w without being true at an accessible world v.

Mistake 2: Reusing a label in $\Box$I. The label v in a $\Box$I sub-proof must be *fresh*: it must not occur anywhere earlier in the proof. Reusing a label that already carries specific information (e.g. from a $\Diamond$E assumption) destroys the generality required for the universal conclusion.

Mistake 3: Applying necessitation to a non-theorem. From $\varphi \vdash_K \varphi$ one *cannot* conclude $\varphi \vdash_K \Box \varphi$. The necessitation rule only applies to formulae derived from *zero* premises. In the labelled system, this means: one may only apply $\Box$I when the sub-proof uses no undischarged assumptions other than the accessibility hypothesis wRv itself.

17 Worked Examples in Natural Deduction

17.1 $\Box p \wedge \Box q \vdash_K \Box(p \wedge q)$

1	$w : \Box p \wedge \Box q$	P
2	$w : \Box p$	$\wedge$E, 1
3	$w : \Box q$	$\wedge$E, 1
4	$\quad wRv$	fresh v
5	$\quad v : p$	$\Box$E, 2, 4
6	$\quad v : q$	$\Box$E, 3, 4
7	$\quad v : p \wedge q$	$\wedge$I, 5, 6
8	$w : \Box(p \wedge q)$	$\Box$I, 4—7

17.2 The Converse: $\Box(p \wedge q) \vdash_K \Box p \wedge \Box q$

Together with the previous proof, this establishes that $\Box p \wedge \Box q$ and $\Box(p \wedge q)$ are logically equivalent in K: necessity distributes over conjunction in both directions.

1	$w : \Box(p \wedge q)$	P
2	$\quad wRv$	fresh v
3	$\quad v : p \wedge q$	$\Box$E, 1, 2
4	$\quad v : p$	$\wedge$E, 3
5	$w : \Box p$	$\Box$I, 2—4
6	$\quad wRu$	fresh u
7	$\quad u : p \wedge q$	$\Box$E, 1, 6
8	$\quad u : q$	$\wedge$E, 7
9	$w : \Box q$	$\Box$I, 6—8
10	$w : \Box p \wedge \Box q$	$\wedge$I, 5, 9

17.3 $\Box(p \to q) \vdash_K \Box p \to \Box q$ — The Distribution Axiom

1	$w : \Box(p \to q)$	P
2	$\quad w : \Box p$	H
3	$\quad\quad wRv$	fresh v
4	$\quad\quad v : p \to q$	$\Box$E, 1, 3
5	$\quad\quad v : p$	$\Box$E, 2, 3
6	$\quad\quad v : q$	$\to$E, 4, 5
7	$\quad w : \Box q$	$\Box$I, 3—6
8	$w : \Box p \to \Box q$	$\to$I, 2—7

Remark 17.1. This proof is a derivation of the distribution axiom K: $\Box(\varphi \to \psi) \to (\Box\varphi \to \Box\psi)$. The nested sub-proofs (one for $\to$-intro, one for $\Box$I inside it) are characteristic of proofs that mix propositional and modal reasoning.

17.4 $\Box((p \wedge q) \to r),\ \Box p \vdash_K \Box(q \to r)$

1	$w : \Box((p \wedge q) \to r)$	P
2	$w : \Box p$	P
3	$\quad wRv$	fresh v
4	$\quad v : (p \wedge q) \to r$	$\Box$E, 1, 3
5	$\quad v : p$	$\Box$E, 2, 3
6	$\quad\quad v : q$	H
7	$\quad\quad v : p \wedge q$	$\wedge$I, 5, 6
8	$\quad\quad v : r$	$\to$E, 4, 7
9	$\quad v : q \to r$	$\to$I, 6—8
10	$w : \Box(q \to r)$	$\Box$I, 3—9

17.5 Working with $\Diamond$: $\Diamond(p \wedge q) \vdash_K \Diamond p \wedge \Diamond q$

1	$w : \Diamond(p \wedge q)$	P
2	$\quad wRv,\ \ v : p \wedge q$	fresh v
3	$\quad v : p$	$\wedge$E, 2
4	$\quad v : q$	$\wedge$E, 2
5	$\quad w : \Diamond p$	$\Diamond$I, 2, 3
6	$\quad w : \Diamond q$	$\Diamond$I, 2, 4
7	$\quad w : \Diamond p \wedge \Diamond q$	$\wedge$I, 5, 6
8	$w : \Diamond p \wedge \Diamond q$	$\Diamond$E, 1, 2—7

17.6 Combining $\Box$ and $\Diamond$: $\Box(p \to q) \vdash_K \Diamond p \to \Diamond q$

1	$w : \Box(p \to q)$	P
2	$\quad w : \Diamond p$	H
3	$\quad\quad wRv,\ \ v : p$	fresh v
4	$\quad\quad v : p \to q$	$\Box$E, 1, 3
5	$\quad\quad v : q$	$\to$E, 4, 3
6	$\quad\quad w : \Diamond q$	$\Diamond$I, 3, 5
7	$\quad w : \Diamond q$	$\Diamond$E, 2, 3—6
8	$w : \Diamond p \to \Diamond q$	$\to$I, 2—7

17.7 ◇E with Reductio: $\Diamond p \wedge \Box \neg p \vdash_K \bot$

1	$w : \Diamond p \wedge \Box \neg p$	P
2	$w : \Diamond p$	$\wedge$E, 1
3	$w : \Box \neg p$	$\wedge$E, 1
4	$\quad wRv, \;\; v : p$	fresh v
5	$\quad v : \neg p$	$\Box$E, 3, 4
6	$\quad \bot$	$\neg$E, 4, 5
7	$\bot$	$\Diamond$E, 2, 4—6

Remark 17.2. An immediate corollary: $\Diamond p \vdash_K \neg\Box\neg p$. Assume $\Diamond p$ and $\Box\neg p$; the proof above yields $\bot$; by $\neg$-introduction, discharge $\Box\neg p$ and conclude $\neg\Box\neg p$. This is one half of the equivalence $\Diamond\varphi \leftrightarrow \neg\Box\neg\varphi$.

17.8 The Definitional Equivalence: $\vdash_K \Diamond\varphi \leftrightarrow \neg\Box\neg\varphi$

Our labelled system treats $\Diamond$ as a primitive operator with its own rules. But the standard definition $\Diamond\varphi =_{\text{Def}} \neg\Box\neg\varphi$ should be derivable. We show both directions.

Left-to-right: $\Diamond\varphi \vdash_K \neg\Box\neg\varphi$.

1	$w : \Diamond\varphi$	P
2	$\quad w : \Box\neg\varphi$	H
3	$\quad\quad wRv, \;\; v : \varphi$	fresh v
4	$\quad\quad v : \neg\varphi$	$\Box$E, 2, 3
5	$\quad\quad \bot$	$\neg$E, 3, 4
6	$\quad \bot$	$\Diamond$E, 1, 3—5
7	$w : \neg\Box\neg\varphi$	$\neg$I, 2—6

Right-to-left: $\neg\Box\neg\varphi \vdash_K \Diamond\varphi$. This direction requires the classical rule of double negation elimination (or, equivalently, reductio ad absurdum in its classical form).

1	$w : \neg\Box\neg\varphi$	P
2	$\quad w : \neg\Diamond\varphi$	H
3	$\quad\quad wRv$	fresh v
4	$\quad\quad\quad v : \varphi$	H
5	$\quad\quad\quad w : \Diamond\varphi$	$\Diamond$I, 3, 4
6	$\quad\quad\quad \bot$	$\neg$E, 2, 5
7	$\quad\quad v : \neg\varphi$	$\neg$I, 4—6
8	$\quad w : \Box\neg\varphi$	$\Box$I, 3—7
9	$\quad \bot$	$\neg$E, 1, 8
10	$w : \neg\neg\Diamond\varphi$	$\neg$I, 2—9
11	$w : \Diamond\varphi$	$\neg\neg$E, 10

Part V
The Axiomatic Perspective

18 The Axiomatic Presentation of K

The natural deduction system we have presented is equivalent to the following axiomatic characterisation.

> The modal logic K is the smallest logic containing:
> 1. All propositional tautologies (i.e. all theorems of propositional logic LP).
> 2. The **distribution axiom**:
>
> $$\vdash_K \Box(\varphi \to \psi) \to (\Box\varphi \to \Box\psi).$$
>
> 3. The **necessitation rule**: if $\vdash_K \varphi$ (that is, φ can be derived from zero premises), then $\vdash_K \Box\varphi$.
>
> and closed under modus ponens.

We have already derived the distribution axiom in our natural deduction system (Section 17). The necessitation rule is captured by the $\Box$I rule applied with no premises other than the accessibility assumption.

Caution about necessitation. The necessitation rule may only be applied to a formula that has been derived *without any premise or assumption*. In particular, from $p \vdash_K p$ (trivially valid) one *cannot* conclude $p \vdash_K \Box p$: the derivation of p depends on the premise p itself. One may only conclude $\vdash_K \Box\varphi$ if $\vdash_K \varphi$, i.e. if φ is a *theorem*.

Philosophical remark. Since all logical theorems are necessary, all tautologies receive the highest modal status. This is a consequence of the semantic setup: a tautology is true at every world in every model, so a fortiori it is true at every accessible world.

19 Necessitation from a Model-Theoretic Perspective

Theorem 19.1. *If $\models_K \varphi$, then $\models_K \Box\varphi$.*

Proof. Suppose for contradiction that $\models_K \varphi$ but $\not\models_K \Box\varphi$. Then there exist a Kripke model $\mathcal{M}$ and an index w such that $\mathcal{M}, w \not\models \Box\varphi$. By the truth-definition for $\Box$, there exists w' with Rww' and $\mathcal{M}, w' \not\models \varphi$. But we assumed $\models_K \varphi$: φ is true at *every* index in *every* model; in particular $\mathcal{M}, w' \models \varphi$. Contradiction. □

Remark 19.1. This is a *meta-logical* result: it says that the set of logical truths of K is closed under the operation $\varphi \mapsto \Box\varphi$. It does *not* say that $\varphi \to \Box\varphi$ is valid (it is not, as we saw in Section 11.5).

Part VI
Metatheorems: Soundness, Completeness, and Decidability

The preceding parts have introduced three formal systems for modal logic: the Kripke semantics (Part II), the tableau method (Part III), and labelled natural deduction (Part IV). We now prove the fundamental metatheorems that relate them: **soundness** (the proof methods never prove something false), **completeness** (they can prove everything that is true), and **decidability** (the tableau procedure always terminates, yielding a decision algorithm).

20 Soundness of the Tableau Method

Theorem 20.1 (Soundness of tableaux for K)**.** *If $\Gamma \vdash \varphi$ (i.e. every tableau for $\Gamma \cup \{\neg\varphi\}$ closes), then $\Gamma \models \varphi$.*

In other words: *the tableau method never proves something false.*

20.1 The proof strategy: preserving satisfiability

We work with the **contrapositive**: if $\Gamma \not\models \varphi$, then $\Gamma \nvdash \varphi$. Since $\Gamma \not\models \varphi$ means $\Gamma \cup \{\neg\varphi\}$ is satisfiable, we must show that if the root of a tableau is satisfiable, the tableau stays open.

The core insight is the **Preservation Principle**: if a branch is satisfiable before a rule application, then at least one resulting branch is still satisfiable afterwards. By induction on the number of rule applications, some branch remains satisfiable at every stage. A satisfiable branch cannot close (it cannot contain both $\mathsf{T} : p, i$ and $\mathsf{F} : p, i$). Therefore the tableau stays open.

For **branching rules** (which split a branch into two), we need: at least one sub-branch is satisfiable. For **non-branching rules** (which extend a branch), we need the stronger condition: the extended branch itself is satisfiable. Since all four modal rules are non-branching, we need only the stronger condition for them.

We first make "satisfiable" precise.

Definition 20.1 (Satisfiability of a branch)**.** A branch B is **satisfiable** if there exist a Kripke model $\mathcal{M} = \langle W, R, I\rangle$ and an assignment σ mapping each world label i on B to a world $\sigma(i) \in W$, such that:

1. for every $\mathsf{T} : \psi, i$ on B: $\mathcal{M}, \sigma(i) \models \psi$;
2. for every $\mathsf{F} : \psi, i$ on B: $\mathcal{M}, \sigma(i) \not\models \psi$;
3. for every accessibility statement iRj on B: $\sigma(i)\, R\, \sigma(j)$.

20.2 The four soundness lemmas

The propositional rules were treated in the propositional logic course. We prove the four modal cases.

Universal rules. These apply to an existing accessibility link iRj already on the branch. The model and the assignment σ do not need to be modified.

Lemma 20.2 (Soundness of $\mathsf{T}\Box$)**.** *If a branch B is satisfiable and contains both $\mathsf{T} : \Box\varphi$, i and iRj, then $B' = B \cup \{\mathsf{T} : \varphi, j\}$ is satisfiable.*

Proof. Since B is satisfiable, there exist $\mathcal{M}$ and σ satisfying B. From $\mathsf{T} : \Box\varphi$, i: $\mathcal{M}, \sigma(i) \models \Box\varphi$. From iRj: $\sigma(i)\, R\, \sigma(j)$. By the semantics of $\Box$: $\mathcal{M}, \sigma(j) \models \varphi$. The same $\mathcal{M}$ and σ satisfy B'. $\square$

Lemma 20.3 (Soundness of $\mathsf{F}\Diamond$)**.** *If a branch B is satisfiable and contains both $\mathsf{F} : \Diamond\varphi$, i and iRj, then $B' = B \cup \{\mathsf{F} : \varphi,\ j\}$ is satisfiable.*

Proof. From $\mathsf{F} : \Diamond\varphi$, i: $\mathcal{M}, \sigma(i) \not\models \Diamond\varphi$. By the semantics of $\Diamond$: for every w with $\sigma(i)\, R\, w$, $\mathcal{M}, w \not\models \varphi$. In particular, $\mathcal{M}, \sigma(j) \not\models \varphi$. The same $\mathcal{M}$ and σ satisfy B'. □

Existential rules. These create a new world label j with a new accessibility link iRj. The model $\mathcal{M}$ stays the same, but σ must be extended. Freshness of j is what makes this possible.

Lemma 20.4 (Soundness of $\mathsf{F}\Box$)**.** *If a branch B is satisfiable and contains $\mathsf{F} : \Box\varphi$, i, then $B' = B \cup \{iRj,\ \mathsf{F} : \varphi,\ j\}$ (with j fresh) is satisfiable.*

Proof. From $\mathsf{F} : \Box\varphi$, i: $\mathcal{M}, \sigma(i) \not\models \Box\varphi$. By the semantics of $\Box$: there exists $w_0 \in W$ with $\sigma(i)\, R\, w_0$ and $\mathcal{M}, w_0 \not\models \varphi$. Since j is fresh, set $\sigma(j) = w_0$. This does not affect the satisfaction of anything already on B. The new items are satisfied: $\sigma(i)\, R\, \sigma(j) = \sigma(i)\, R\, w_0$ and $\mathcal{M}, w_0 \not\models \varphi$. □

Lemma 20.5 (Soundness of $\mathsf{T}\Diamond$)**.** *If a branch B is satisfiable and contains $\mathsf{T} : \Diamond\varphi$, i, then $B' = B \cup \{iRj,\ \mathsf{T} : \varphi,\ j\}$ (with j fresh) is satisfiable.*

Proof. From $\mathsf{T} : \Diamond\varphi$, i: there exists w_0 with $\sigma(i)\, R\, w_0$ and $\mathcal{M}, w_0 \models \varphi$. Set $\sigma(j) = w_0$. The same model and extended σ satisfy B'. □

Remark 20.1. The four lemmas fall into two symmetric pairs. The universal rules ($\mathsf{T}\Box$, $\mathsf{F}\Diamond$) instantiate a universal quantifier at an existing world; σ is unchanged. The existential rules ($\mathsf{F}\Box$, $\mathsf{T}\Diamond$) use a semantic witness as the interpretation of a fresh label; σ is extended. In both cases, the model itself is never modified.

20.3 The soundness theorem: full proof

Proof of Theorem 20.1. Contrapositive: suppose $\Gamma \not\models_K \varphi$. Then $\Gamma \cup \{\neg\varphi\}$ is satisfiable. The initial branch $\mathsf{T} : \gamma_1,\ 1;\ \ldots;\ \mathsf{T} : \gamma_n,\ 1;\ \mathsf{F} : \varphi,\ 1$ is satisfiable (set $\sigma(1) = w$ where w is the witnessing world). By the Preservation Principle (propositional rules: proved in the propositional course; modal rules: Lemmas above), at every stage of the tableau at least one branch remains satisfiable. A satisfiable branch cannot close. Therefore the tableau stays open: $\Gamma \not\vdash_K \varphi$. □

21 Completeness of the Tableau Method

Theorem 21.1 (Completeness of tableaux for K)**.** *If $\Gamma \models_K \varphi$, then $\Gamma \vdash_K \varphi$.*

Soundness was *atomistic*: we examined each rule individually. Completeness must be *holistic*: we show the system as a whole is powerful enough. The contrapositive formulation makes the task concrete: if $\Gamma \not\vdash_K \varphi$, then the tableau has a finished open branch B, and we must build a counter-model from B.

21.1 Building a Kripke model from an open branch

Definition 21.1 (Branch model)**.** Let B be a finished open branch. The **branch model** $\mathcal{M}_B = \langle W, R, I\rangle$ is:

- W = set of all world labels appearing on B.
- $R = \{(i, j) : iRj \text{ appears on } B\}$.
- $I(p, i) = \mathsf{T}$ if $\mathsf{T} : p$, i appears on B; $\quad I(p, i) = \mathsf{F}$ otherwise.

Remark 21.1. The valuation is well-defined because B is open: it does not contain both $\mathsf{T} : p$, i and $\mathsf{F} : p$, i for the same atom p at the same world i.

21.2 The inductive proof

Theorem 21.2 (Branch satisfiability)**.** *Let B be a finished open branch and $\mathcal{M}_B$ its branch model. For every formula φ on B:*
(a) if $\mathsf{T} : \varphi$, i is on B, then $\mathcal{M}_B, i \models \varphi$;
(b) if $\mathsf{F} : \varphi$, i is on B, then $\mathcal{M}_B, i \not\models \varphi$.

Proof. By induction on the complexity of φ.

Base case: $\varphi = p$. If $\mathsf{T} : p$, i is on B: by definition of I, $I(p, i) = \mathsf{T}$, so $\mathcal{M}_B, i \models p$. If $\mathsf{F} : p$, i is on B: since B is open, $\mathsf{T} : p$, i is not on B, so $I(p, i) = \mathsf{F}$, and $\mathcal{M}_B, i \not\models p$.

Propositional cases. These are exactly as in the propositional completeness proof, using the fact that the branch is finished (all applicable propositional rules have been applied). We omit the details.

Case $\varphi = \Box\psi$ (true-signed). We need: for every j with iRj in $\mathcal{M}_B$, $\mathcal{M}_B, j \models \psi$. Let j be such that iRj is on B. Since B is finished, $\mathsf{T}\Box$ has been applied: $\mathsf{T} : \psi$, j is on B. By the inductive hypothesis, $\mathcal{M}_B, j \models \psi$.

Case $\varphi = \Box\psi$ (false-signed). We need: there exists j with iRj and $\mathcal{M}_B, j \not\models \psi$. Since B is finished, $\mathsf{F}\Box$ has been applied: there is a label j with iRj and $\mathsf{F} : \psi$, j on B. By the inductive hypothesis, $\mathcal{M}_B, j \not\models \psi$.

Case $\varphi = \Diamond\psi$ (true-signed). Since B is finished, $\mathsf{T}\Diamond$ has been applied: there is a label j with iRj and $\mathsf{T} : \psi$, j on B. By the inductive hypothesis, $\mathcal{M}_B, j \models \psi$. So $\mathcal{M}_B, i \models \Diamond\psi$.

Case $\varphi = \Diamond\psi$ (false-signed). For every j with iRj on B: since B is finished, $\mathsf{F}\Diamond$ has been applied, so $\mathsf{F} : \psi$, j is on B. By the inductive hypothesis, $\mathcal{M}_B, j \not\models \psi$. So $\mathcal{M}_B, i \not\models \Diamond\psi$. □

Remark 21.2. The modal cases fall into two pairs. The universal rules ($\mathsf{T}\Box$, $\mathsf{F}\Diamond$) work because the branch is *finished*: the rule has been applied to every accessible world on the branch, and the worlds of $\mathcal{M}_B$ are precisely the labels on the branch. The existential rules ($\mathsf{F}\Box$, $\mathsf{T}\Diamond$) work because the rule *created* the witness.

21.3 The completeness theorem: full proof

Proof of Theorem 21.1. Contrapositive: suppose $\Gamma \nvdash_K \varphi$. The tableau for $\Gamma \cup \{\neg\varphi\}$ has a finished open branch B. Construct $\mathcal{M}_B$ (Definition 21.1). By Theorem 21.2, every T-signed formula on B is satisfied and every F-signed formula is falsified. The root formulas $\mathsf{T} : \gamma_k$, 1 and $\mathsf{F} : \varphi$, 1 are on B, so $\mathcal{M}_B, 1 \models \gamma_k$ for each $\gamma_k \in \Gamma$ and $\mathcal{M}_B, 1 \not\models \varphi$. Therefore $\Gamma \not\models_K \varphi$. □

22 Decidability

Soundness and completeness together say that $\Gamma \models_K \varphi$ iff $\Gamma \vdash_K \varphi$. But can we always *determine* which case holds? This requires that the tableau procedure **terminates**.

22.1 The subformula property

Definition 22.1 (Subformulas)**.** The set $\mathrm{Sub}(\varphi)$ of subformulas of φ is defined recursively in the obvious way. For a finite set Σ, $\mathrm{Sub}(\Sigma) = \bigcup_{\psi\in\Sigma} \mathrm{Sub}(\psi)$.

Lemma 22.1 (Subformula property)**.** *On every branch of the tableau for Σ, every formula appearing in a signed entry is a subformula of some formula in Σ.*

Proof. By inspection of each rule. Every propositional rule decomposes a formula into its immediate subformulas. Every modal rule strips away one layer of $\Box$ or $\Diamond$: from $\Box\psi$ (or $\Diamond\psi$), it produces ψ, which is a proper subformula. No rule ever creates a formula more complex than the input. □

Remark 22.1. This is where propositional modal logic differs sharply from first-order logic. In FOL, the rule for $\forall x\,\varphi(x)$ produces *instances* $\varphi(a)$, which may contain further quantifiers generating yet more constants. In modal logic, the rule for $\Box\psi$ produces the genuine *subformula* ψ — strictly simpler. This prevents the infinite regress.

22.2 Finiteness of the tableau

Definition 22.2 (Modal depth)**.** The **modal depth** $\mathrm{md}(\varphi)$ is the maximum nesting depth of $\Box$ and $\Diamond$ in φ: $\mathrm{md}(p) = 0$; $\mathrm{md}(\neg\psi) = \mathrm{md}(\psi)$; $\mathrm{md}(\psi_1 \star \psi_2) = \max(\mathrm{md}(\psi_1), \mathrm{md}(\psi_2))$; $\mathrm{md}(\Box\psi) = \mathrm{md}(\Diamond\psi) = 1 + \mathrm{md}(\psi)$.

Lemma 22.2 (Depth bound)**.** *On any branch, if a world label j is at distance d from the root (i.e. reached through a chain of d accessibility links), then every formula at j has modal depth at most $\mathrm{md}(\Sigma) - d$.*

Proof. By induction on d. At the root ($d = 0$), formulas have modal depth at most $\mathrm{md}(\Sigma)$. A new world at distance $d+1$ is created by an existential rule applied to a modal formula $\Box\psi$ or $\Diamond\psi$ at distance d; the rule produces ψ, whose modal depth is at least one less. The universal rules similarly decrease modal depth by one when propagating from an ancestor. □

Remark 22.2. When $d = \mathrm{md}(\Sigma)$, all formulas at that world have modal depth 0: they are purely propositional. No existential rule can fire, so no new world is created beyond depth $\mathrm{md}(\Sigma)$.

Lemma 22.3 (Width bound)**.** *At any world i on a branch, there are at most $2 \times |\mathrm{Sub}(\Sigma)|$ signed formulas (each subformula with sign T or F).*

Theorem 22.4 (Termination)**.** *For any finite Σ, every tableau for Σ is finite.*

Proof. The depth bound limits the chain of worlds to at most $\mathrm{md}(\Sigma)$. The width bound limits the formulas at each world. Each rule application either adds a formula at an existing world or creates a new world. Since both quantities are bounded, every branch is finite. A tableau is a finitely branching tree with finite branches, hence finite. □

22.3 The decision procedure

Theorem 22.5 (Decidability of K)**.** *There exists an algorithm that, given a finite Γ and a formula φ, decides whether $\Gamma \models_K \varphi$.*

Proof. Build the tableau for $\Gamma \cup \{\neg\varphi\}$. By Theorem 22.4, it terminates. If the tableau closes, soundness gives $\Gamma \models_K \varphi$. If it has an open branch, completeness gives $\Gamma \not\models_K \varphi$, and the branch model is an explicit finite counter-model. □

Remark 22.3 (The finite model property). A by-product: if φ is satisfiable, it is satisfiable in a *finite* Kripke model (the branch model $\mathcal{M}_B$ is always finite). This is the **finite model property** for K — a strong property that not all logics enjoy.

23 Completeness of Labelled Natural Deduction

The tableau completeness proof was easy: an open branch *is* a model. For the labelled ND system, no such by-product exists. We must construct a model from scratch, using only the abstract fact that a set of formulas is consistent. The technique, due to Henkin (1949) and adapted to modal logic by Bayart, Scott, and Kaplan, treats *maximal consistent sets* as possible worlds.

Throughout this section, $\Gamma \vdash_K \varphi$ means: there exists a labelled ND derivation of $w : \varphi$ from $\{w : \gamma \mid \gamma \in \Gamma\}$.

23.1 Maximal consistent sets

Definition 23.1 (Consistency)**.** A set Γ is **consistent** (in K) if $\Gamma \nvdash_K \bot$.

Definition 23.2 (Maximal consistent set)**.** A set ω is **maximal consistent** if it is consistent and for every formula φ, either $\varphi \in \omega$ or $\neg\varphi \in \omega$.

Think of a maximal consistent set as a complete description of a possible world: it answers every yes/no question (maximality), and its answers are never self-contradictory (consistency).

Lemma 23.1 (Properties of MCS)**.** *Let ω be a maximal consistent set. Then:*

(i) $\varphi \in \omega$ *iff* $\omega \vdash_K \varphi$.
(ii) $\varphi \notin \omega$ *iff* $\neg\varphi \in \omega$.
(iii) $\varphi \wedge \psi \in \omega$ *iff* $\varphi \in \omega$ *and* $\psi \in \omega$.
(iv) $\varphi \vee \psi \in \omega$ *iff* $\varphi \in \omega$ *or* $\psi \in \omega$.
(v) $\varphi \to \psi \in \omega$ *iff* $\varphi \notin \omega$ *or* $\psi \in \omega$.

The proofs follow from maximality and consistency using standard propositional reasoning.

23.2 Lindenbaum's Lemma

Lemma 23.2 (Lindenbaum)**.** *If Γ is consistent, there exists a maximal consistent set $\omega \supseteq \Gamma$.*

Proof. Enumerate all formulas: $\varphi_1, \varphi_2, \ldots$ Build an ascending chain $\Gamma = \Gamma_0 \subseteq \Gamma_1 \subseteq \cdots$ by setting

$$\Gamma_{n+1} = \begin{cases} \Gamma_n \cup \{\varphi_{n+1}\} & \text{if this is consistent,} \\ \Gamma_n \cup \{\neg\varphi_{n+1}\} & \text{otherwise.} \end{cases}$$

Set $\omega = \bigcup_n \Gamma_n$. Each Γ_n is consistent (by induction: if both extensions were inconsistent, Γ_n itself would be inconsistent). Therefore ω is consistent (any derivation of $\bot$ uses finitely many premises, all belonging to some Γ_N). And ω is maximal: every formula was decided at some step. □

23.3 The canonical model

Definition 23.3 (Canonical model for K)**.** The **canonical model** $\mathcal{M}^c = \langle W^c, R^c, I^c \rangle$ is:

- $W^c = \{\omega \mid \omega \text{ is a maximal consistent set}\}$.
- $\omega \, R^c \, \omega'$ iff for every φ, $\Box\varphi \in \omega \Longrightarrow \varphi \in \omega'$.
- $I^c(p, \omega) = \mathsf{T}$ iff $p \in \omega$.

Remark 23.1. The accessibility relation is defined to make the ($\Leftarrow$) direction of the box case trivial: if $\Box\psi \in \omega$ and $\omega \, R^c \, \omega'$, then $\psi \in \omega'$ by definition. The hard part is the ($\Rightarrow$) direction: showing that if ψ holds at every accessible world, then $\Box\psi \in \omega$. This requires the Existence Lemma.

23.4 Auxiliary facts

Before the Existence Lemma, we need four derivable facts in K (all proved in Part IV):

1. $\Box\varphi \wedge \Box\psi \vdash_K \Box(\varphi \wedge \psi)$.
2. Generalisation: $\Box\alpha_1 \wedge \cdots \wedge \Box\alpha_n \vdash_K \Box(\alpha_1 \wedge \cdots \wedge \alpha_n)$.
3. $\Box\varphi \wedge \Diamond\psi \vdash_K \Diamond(\varphi \wedge \psi)$.
4. If $\Diamond\varphi$ is consistent, then φ is consistent. (Contrapositive: if $\vdash_K \neg\varphi$, then by necessitation $\vdash_K \Box\neg\varphi$, and $\Diamond\varphi \vdash_K \bot$.)

23.5 The Existence Lemma

Lemma 23.3 (Existence Lemma)**.** *Let ω be a maximal consistent set. If $\Diamond\varphi \in \omega$, then there exists a maximal consistent set ω' such that $\omega\, R^c\, \omega'$ and $\varphi \in \omega'$.*

Proof. Define $\Delta_\omega = \{\psi : \Box\psi \in \omega\}$.

Step 1: $\Delta_\omega \cup \{\varphi\}$ is consistent. Suppose not. Then $\exists\, \alpha_1, \ldots, \alpha_n \in \Delta_\omega$ with $\{\alpha_1, \ldots, \alpha_n, \varphi\} \vdash_K \bot$. By Fact 4, $\Diamond(\alpha_1 \wedge \cdots \wedge \alpha_n \wedge \varphi)$ is inconsistent. But by Fact 2, $\omega \vdash_K \Box(\alpha_1 \wedge \cdots \wedge \alpha_n)$, and since $\Diamond\varphi \in \omega$, Fact 3 gives $\omega \vdash_K \Diamond((\alpha_1 \wedge \cdots \wedge \alpha_n) \wedge \varphi)$. So ω derives both $\Diamond(\alpha_1 \wedge \cdots \wedge \alpha_n \wedge \varphi)$ and its negation, contradicting consistency.

Step 2. By Lindenbaum's Lemma, extend $\Delta_\omega \cup \{\varphi\}$ to a maximal consistent set ω'. Then $\varphi \in \omega'$ and $\omega\, R^c\, \omega'$ (since $\Delta_\omega \subseteq \omega'$). □

23.6 The Truth Theorem

Theorem 23.4 (Truth Theorem)**.** *For every $\omega \in W^c$ and every formula φ:* $\mathcal{M}^c, \omega \models \varphi \iff \varphi \in \omega$.

Proof. By induction on the complexity of φ.

Base case. $\mathcal{M}^c, \omega \models p$ iff $I^c(p, \omega) = \mathsf{T}$ iff $p \in \omega$.

Propositional cases. Follow from Lemma 23.1. For example: $\mathcal{M}^c, \omega \models \psi_1 \wedge \psi_2$ iff (by induction) $\psi_1 \in \omega$ and $\psi_2 \in \omega$ iff (by Lemma 23.1(iii)) $\psi_1 \wedge \psi_2 \in \omega$.

Case $\varphi = \Box\psi$, direction ($\Leftarrow$). Suppose $\Box\psi \in \omega$. Let ω' with $\omega\, R^c\, \omega'$. By definition of R^c: $\psi \in \omega'$. By induction: $\mathcal{M}^c, \omega' \models \psi$. Since ω' was arbitrary: $\mathcal{M}^c, \omega \models \Box\psi$.

Case $\varphi = \Box\psi$, direction ($\Rightarrow$). Suppose $\mathcal{M}^c, \omega \models \Box\psi$ but $\Box\psi \notin \omega$. By maximality, $\neg\Box\psi \in \omega$, so $\Diamond\neg\psi \in \omega$. By the Existence Lemma: there exists ω' with $\omega\, R^c\, \omega'$ and $\neg\psi \in \omega'$. By induction: $\mathcal{M}^c, \omega' \not\models \psi$. But $\omega\, R^c\, \omega'$ and $\mathcal{M}^c, \omega \models \Box\psi$ imply $\mathcal{M}^c, \omega' \models \psi$. Contradiction.

Case $\varphi = \Diamond\psi$. Both directions follow from the $\Box$ case via the equivalence $\Diamond\psi \leftrightarrow \neg\Box\neg\psi$ and the inductive hypotheses for $\neg$ and $\Box$. □

23.7 The completeness theorem for labelled ND

Theorem 23.5 (Completeness of labelled ND for K)**.** *If $\Gamma \models_K \varphi$, then $\Gamma \vdash_K \varphi$.*

Proof. Contrapositive: suppose $\Gamma \nvdash_K \varphi$. Then $\Gamma \cup \{\neg\varphi\}$ is consistent. Extend it to a maximal consistent set ω (Lindenbaum). In the canonical model $\mathcal{M}^c$: by the Truth Theorem, $\mathcal{M}^c, \omega \models \gamma$ for each $\gamma \in \Gamma$ and $\mathcal{M}^c, \omega \not\models \varphi$. So $\Gamma \not\models_K \varphi$. □

23.8 Comparison: Tableau vs. canonical model completeness

	Tableau	**Canonical model**
Worlds	Labels on the open branch.	Maximal consistent sets.
Accessibility	iRj statements on the branch.	$\omega R^c \omega'$ iff $\Box\psi \in \omega \Rightarrow \psi \in \omega'$.
Valuation	T-signed atoms on the branch.	Membership: $p \in \omega$.
Key tool	Branch already finished — read off the model.	Lindenbaum + Existence Lemma.
Difficulty	Low.	High.
Advantage	Algorithmic: produces a finite counter-model.	General: works for any proof system.

Part VII
Extensions of K and Correspondence Theory

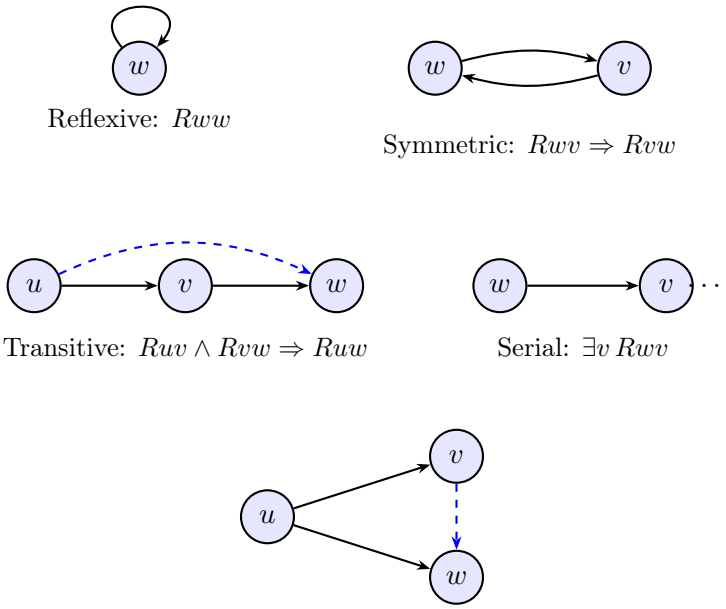
24 Properties of Frames

Since a frame (Definition 7.1) is a directed graph, we can classify frames by the structural properties of R.

Definition 24.1 (Frame properties)**.** Let $\mathcal{F} = \langle W, R \rangle$ be a frame.

Property	**FOL condition**	**Intuition**
Reflexive	$\forall w\, Rww$	every world sees itself
Transitive	$\forall u \forall v \forall w\, ((Ruv \land Rvw) \to Ruw)$	reachability chains compose
Symmetric	$\forall v \forall w\, (Rvw \to Rwv)$	access is bidirectional
Euclidean	$\forall u \forall v \forall w\, ((Ruv \land Ruw) \to Rvw)$	siblings see each other
Serial	$\forall v \exists w\, Rvw$	no dead ends

The following diagrams illustrate each property.



Reflexive: Rww

Symmetric: $Rwv \Rightarrow Rvw$

Transitive: $Ruv \land Rvw \Rightarrow Ruw$

Serial: $\exists v\, Rwv$

Euclidean: $Ruv \land Ruw \Rightarrow Rvw$

Exercise 24.1. Show that a reflexive and euclidean relation is also symmetric and transitive. Conversely, show that a reflexive, symmetric, and transitive relation is euclidean.

Hint: for symmetry from reflexivity + euclideanness, take $u = v$ in the euclidean condition; for transitivity, use symmetry first to swap an edge, then apply euclideanness.

25 Modal Correspondence: The Central Idea

Definition 25.1 (Modal correspondence)**.** A modal formula (schema) φ **corresponds to** (or **defines**) a frame property P if and only if:

$$\text{for every frame } \mathcal{F}, \quad \mathcal{F} \models \varphi \iff \mathcal{F} \text{ has property } P.$$

Here $\mathcal{F} \models \varphi$ means that φ is valid in $\mathcal{F}$ (Definition 7.5).

In other words, *the class of frames validating the formula is exactly the class of frames having the property.*

Notice the two quantifiers hidden inside "$\mathcal{F} \models \varphi$":

$$\mathcal{F} \models \varphi \quad \overset{\text{def}}{\iff} \quad \text{for } \textbf{every} \text{ model } \mathcal{M} \text{ on } \mathcal{F}, \text{ for } \textbf{every} \text{ world } w \text{ in } \mathcal{M},\ \mathcal{M}, w \models \varphi.$$

The formula must hold at every world, under every possible interpretation of the propositional variables. This is a very strong requirement, and confusing it with a weaker one is the most common mistake students make with correspondence theory.

25.1 Caution! Model validity is not frame validity

What the definition does NOT say. If a *particular model* $\mathcal{M}$ on a frame $\mathcal{F}$ satisfies

$$\mathcal{M}, w \models \Box\varphi \to \varphi \quad \text{at every world } w,$$

then $\mathcal{F}$ is reflexive. **This is false!**

What the definition DOES say. If *every possible model* $\mathcal{M}$ on $\mathcal{F}$ satisfies the formula at every world, then $\mathcal{F}$ is reflexive.

A counter-example: one world, no accessibility. Consider the frame $W = \{w\}$, $R = \emptyset$. There is a single world w, and the accessibility relation is *empty*: w does not see any world, not even itself. In particular, $\neg Rww$, so **the frame is not reflexive**.

Build a model on this frame by setting $I(w, p) = \mathsf{T}$ for every propositional variable p. Then $\mathcal{M}, w \models \Box p$ (vacuously true, no accessible worlds) and $\mathcal{M}, w \models p$ (by the valuation). So $\mathcal{M}, w \models \Box p \to p$. In fact, $\mathcal{M}, w \models \Box\varphi \to \varphi$ for every formula φ. And yet the frame is not reflexive.

Now build a different model: $I'(w, p) = \mathsf{F}$. Then $\mathcal{M}', w \models \Box p$ (still vacuously true), but $\mathcal{M}', w \not\models p$. Therefore $\mathcal{M}', w \not\models \Box p \to p$. The frame does not validate T.

The moral. The first model was a *lucky* interpretation: it happened to make p true at w, which made $\Box p \to p$ true for trivial reasons unrelated to the structure of R. The second model exposed the real problem. This is why the correspondence definition quantifies over *all* models on the frame: it ensures that the validity of the formula is due to the *structural* properties of R, not to an accidental choice of interpretation.

Rule of thumb. When working with correspondence theory, always ask:

- Does the formula hold because of the *structure of* R? $\longrightarrow$ frame property.
- Or does it hold because of a *particular choice of* I? $\longrightarrow$ model accident, says nothing about the frame.

The ($\Rightarrow$) direction of each correspondence theorem works by *contrapositive*: if R lacks the structural property, we explicitly *construct* a model (a bad choice of I) that falsifies the formula. That is why the counter-model constructions always define a specific valuation.

26 The Five Correspondences at a Glance

Axiom	Modal formula	Frame property
D	$\Box\varphi \to \Diamond\varphi$	Serial: $\forall w\, \exists v\, Rwv$
M (or T)	$\Box\varphi \to \varphi$	Reflexive: $\forall w\, Rww$
4	$\Box\varphi \to \Box\Box\varphi$	Transitive: $\forall u, v, w\ (Ruv \wedge Rvw \to Ruw)$
B	$\varphi \to \Box\Diamond\varphi$	Symmetric: $\forall v, w\ (Rvw \to Rwv)$
5	$\Diamond\varphi \to \Box\Diamond\varphi$	Euclidean: $\forall u, v, w\ (Ruv \wedge Ruw \to Rvw)$

For each correspondence, we prove two directions:

($\Leftarrow$) If the frame has the property, then the formula is valid in the frame. (Typically by *reductio*: assume a world falsifies the formula and derive a contradiction using the frame property.)

($\Rightarrow$) If the formula is valid in the frame, then the frame has the property. (Typically by *contrapositive*: assume the frame lacks the property, and construct a counter-model.)

After each model-theoretic proof, we give a proof-theoretic derivation of the axiom in the appropriate extension of K, using the labelled natural deduction system.

Remark 26.1 (Asymmetry between the two directions). The natural-deduction derivation establishes the ($\Leftarrow$) direction: *if the frame has the property, the axiom is derivable* (and therefore valid, by soundness). But the ($\Rightarrow$) direction **cannot be obtained proof-theoretically**. Its proof works by *contrapositive*: assume R lacks the property, then *construct a specific valuation I* that falsifies the formula. That construction is an inherently *semantic* operation. In short, the ND derivation shows the structural rule is **sufficient**; the model-theoretic argument shows it is **necessary**. This asymmetry reflects a deep fact about the relationship between syntax and semantics.

27 D Defines Serial Frames

Theorem 27.1. *For every frame $\mathcal{F}$,*

$$\mathcal{F} \models \Box\varphi \to \Diamond\varphi \quad \Longleftrightarrow \quad \mathcal{F} \text{ is serial.}$$

27.1 Model-theoretic proof

Direction ($\Leftarrow$): Seriality implies validity of D. Suppose $\mathcal{F}$ is serial and, seeking a contradiction, suppose there is a model $\mathcal{M}$ on $\mathcal{F}$ and a world w such that $\mathcal{M}, w \not\models \Box\varphi \to \Diamond\varphi$. Then $\mathcal{M}, w \models \Box\varphi$ and $\mathcal{M}, w \not\models \Diamond\varphi$. Since R is serial, there exists w' with Rww'. By the first conjunct, $\mathcal{M}, w' \models \varphi$. But the second conjunct says that for every v with Rwv, $\mathcal{M}, v \models \neg\varphi$; in particular $\mathcal{M}, w' \models \neg\varphi$. Contradiction. □

Direction ($\Rightarrow$): Validity of D implies seriality. Contrapositive: if $\mathcal{F}$ is not serial, there exists a "dead-end" world $v \in W$ with no successor. Since no world is accessible from v, $\mathcal{M}, v \models \Box\varphi$ vacuously and $\mathcal{M}, v \not\models \Diamond\varphi$. Therefore $\mathcal{M}, v \not\models \Box\varphi \to \Diamond\varphi$, so $\mathcal{F} \not\models D$. □

Remark 27.1. A dead-end world vacuously satisfies $\Box\varphi$ ("everything is necessary") and falsifies $\Diamond\varphi$ ("nothing is possible"). Axiom D rules out such worlds by requiring seriality.

27.2 Proof-theoretic derivation

In system D, one adds to K a **seriality rule**: from nothing, assert wRv for a fresh v.

1	$w : \Box\varphi$	H
2	wRv (fresh v, by seriality)	
3	$v : \varphi$	$\Box$E, 1, 2
4	$w : \Diamond\varphi$	$\Diamond$I, 2, 3
5	$w : \Box\varphi \to \Diamond\varphi$	$\to$I, 1—4

28 M (T) Defines Reflexive Frames

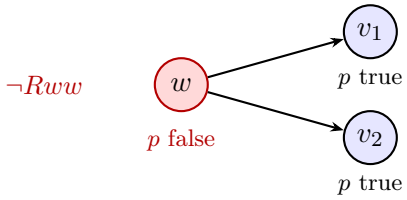
Theorem 28.1. *For every frame $\mathcal{F}$,*

$$\mathcal{F} \models \Box\varphi \to \varphi \quad \iff \quad \mathcal{F} \text{ is reflexive.}$$

28.1 Model-theoretic proof

Direction ($\Leftarrow$). Suppose $\mathcal{F}$ is reflexive. Let $\mathcal{M}$ be any model on $\mathcal{F}$ and $w \in W$. Suppose $\mathcal{M}, w \models \Box\varphi$. Since R is reflexive, Rww. By the semantics of $\Box$, $\mathcal{M}, w \models \varphi$. Therefore $\mathcal{M}, w \models \Box\varphi \to \varphi$. □

Direction ($\Rightarrow$). Contrapositive: assume $\mathcal{F}$ is not reflexive. Then $\exists w \in W$, $\neg Rww$. Choose I such that $I(p, w) = \mathsf{F}$ and for every $i \neq w$ with Rwi, $I(p, i) = \mathsf{T}$. Then $\mathcal{M}, w \models \Box p$ (since p is true at every accessible world—w itself is not accessible from w). But $\mathcal{M}, w \not\models p$. So $\mathcal{F} \not\models M$. □



28.2 Proof-theoretic derivation

In system T ($= K +$ reflexivity), we add a **reflexivity rule**: at any world w, assert Rww.

1	$w : \Box\varphi$	H
2	wRw	Refl.
3	$w : \varphi$	$\Box$E, 1, 2
4	$w : \Box\varphi \to \varphi$	$\to$I, 1—3

Boxes can be eliminated. A striking consequence: in T, any number of nested boxes can be stripped away.

1	$w : \Box\Box p$	P
2	$w : \Box p$	$\Box$E-M, 1
3	$w : p$	$\Box$E-M, 2

A theorem of T: $\varphi \to \Diamond \varphi$. In T, whatever is actual is possible:

1	$w : \varphi$	H
2	$\quad w : \neg\Diamond \varphi$	H
3	$\quad w : \Box \neg\varphi$	Def. $\Diamond$, 2
4	$\quad wRw$	Refl.
5	$\quad w : \neg\varphi$	$\Box$E, 3, 4
6	$\quad \bot$	$\neg$E, 1, 5
7	$w : \neg\neg\Diamond \varphi$	$\neg$I, 2—6
8	$w : \Diamond \varphi$	$\neg\neg$E, 7
9	$w : \varphi \to \Diamond \varphi$	$\to$I, 1—8

The key step is line 4: reflexivity gives Rww, which lets $\Box$E extract $\neg\varphi$ from $\Box\neg\varphi$ at the same world. This step is *not* available in K.

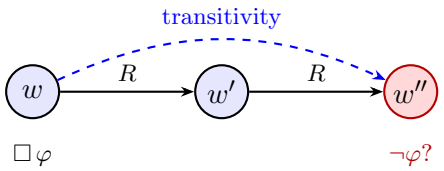
29 Axiom 4 Defines Transitive Frames ($S4$)

Theorem 29.1. *For every frame $\mathcal{F}$,*

$$\mathcal{F} \models \Box\varphi \to \Box\Box\varphi \quad \Longleftrightarrow \quad \mathcal{F} \text{ is transitive.}$$

29.1 Model-theoretic proof

Direction ($\Leftarrow$). Suppose $\mathcal{F}$ is transitive. Let $\mathcal{M}, w \models \Box\varphi$. By *reductio*: suppose $\mathcal{M}, w \not\models \Box\Box\varphi$, i.e. $\mathcal{M}, w \models \Diamond\Diamond\neg\varphi$. Then there exist w', w'' with Rww', $Rw'w''$, and $\mathcal{M}, w'' \models \neg\varphi$. By transitivity, Rww''. But $\mathcal{M}, w \models \Box\varphi$ together with Rww'' gives $\mathcal{M}, w'' \models \varphi$. Contradiction. $\square$



Direction ($\Rightarrow$). Contrapositive: suppose R is not transitive. Then $\exists u, v, w$ with Ruv, Rvw, but $\neg Ruw$. Define I so that $I(p, w) = \mathsf{F}$ and for every $x \neq w$ with Rux, $I(p, x) = \mathsf{T}$. Then $\mathcal{M}, u \models \Box p$ but $\mathcal{M}, u \models \Diamond\Diamond\neg p$ (via v then w). So $\mathcal{F} \not\models$ axiom 4. $\square$

29.2 Proof-theoretic derivation

$S4 = T +$ transitivity rule: from wRv and vRu, conclude wRu.

1	$w : \Box p$	P
2	$\quad wRv$	fresh v
3	$\quad\quad vRu$	fresh u
4	$\quad\quad wRu$	Trans., 2, 3
5	$\quad\quad u : p$	$\Box$E, 1, 4
6	$\quad v : \Box p$	$\Box$I, 3—5
7	$w : \Box\Box p$	$\Box$I, 2—6

Explanation: The transitivity rule (line 4) gives wRu, which lets us apply $\Box$E to the premise $\Box p$ at w and obtain p at u.

Remark 29.1. Without transitivity, line 4 is unavailable: we have wRv and vRu but not wRu, so $\Box$E cannot extract p at u from $\Box p$ at w. This is precisely why $\Box p \to \Box\Box p$ fails in K.

Converse in $S4$: $\Box\Box p \vdash_{S4} \Box p$. This direction uses reflexivity (from T, which is included in $S4$):

1	$w : \Box\Box p$	P
2	$\quad wRv$	fresh v
3	$\quad v : \Box p$	$\Box$E, 1, 2
4	$\quad vRv$	Refl.
5	$\quad v : p$	$\Box$E, 3, 4
6	$w : \Box p$	$\Box$I, 2—5

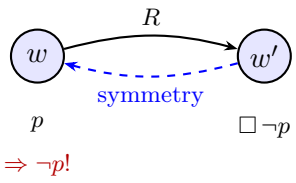
30 Axiom B Defines Symmetric Frames

Theorem 30.1. *For every frame $\mathcal{F}$,*

$$\mathcal{F} \models \varphi \to \Box\Diamond\varphi \quad \iff \quad \mathcal{F} \text{ is symmetric.}$$

30.1 Model-theoretic proof

Direction ($\Leftarrow$). Suppose $\mathcal{F}$ is symmetric. Let $\mathcal{M}, w \models p$. By *reductio*: suppose $\mathcal{M}, w \not\models \Box\Diamond p$, i.e. $\mathcal{M}, w \models \Diamond\Box\neg p$. Then $\exists w'$ with Rww' and $\mathcal{M}, w' \models \Box\neg p$. By symmetry, $Rw'w$. Therefore $\mathcal{M}, w \models \neg p$. Contradiction. □



Direction ($\Rightarrow$). Contrapositive: suppose R is not symmetric. Then $\exists v, w\ (Rvw \wedge \neg Rwv)$. Define I such that $I(p, v) = \mathsf{T}$ and for every index x accessible from w, $I(p, x) = \mathsf{F}$ (possible since v is not accessible from w). Then $\mathcal{M}, v \models p$, but $\mathcal{M}, w \models \Box\neg p$, so $\mathcal{M}, v \models \Diamond\Box\neg p$, i.e. $\mathcal{M}, v \not\models \Box\Diamond p$. Therefore $\mathcal{F} \not\models B$. □

30.2 Proof-theoretic derivation

$B = T +$ symmetry rule: from wRv, conclude vRw.

1	$w : p$	P
2	$\quad wRv$	fresh v
3	$\quad vRw$	Sym., 2
4	$\quad v : \Diamond p$	$\Diamond$I, 3, 1
5	$w : \Box\Diamond p$	$\Box$I, 2—4

Remark 30.1 (On the use of line 1 inside the sub-proof). At line 4, we use $w : p$ *inside* the $\Box$I sub-proof. This is not a violation of the rule against importing unboxed formulae: $w : p$ keeps its own label w; we use it in combination with vRw to derive $v : \Diamond p$ via $\Diamond$I.

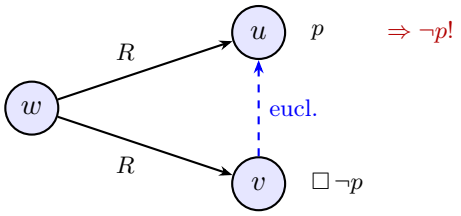
31 Axiom 5 Defines Euclidean Frames ($S5$)

Theorem 31.1. *For every frame $\mathcal{F}$,*

$$\mathcal{F} \models \Diamond\varphi \to \Box\Diamond\varphi \quad \Longleftrightarrow \quad \mathcal{F} \text{ is euclidean.}$$

31.1 Model-theoretic proof

Direction ($\Leftarrow$). Suppose $\mathcal{F}$ is euclidean. Let $\mathcal{M}, w \models \Diamond p$. By *reductio*: suppose $\mathcal{M}, w \not\models \Box\Diamond p$, i.e. $\mathcal{M}, w \models \Diamond\Box\neg p$. Then there exist u with Rwu and $\mathcal{M}, u \models p$, and v with Rwv and $\mathcal{M}, v \models \Box\neg p$. By euclideanness, $Rwu \wedge Rwv \to Rvu$. Therefore $\mathcal{M}, u \models \neg p$. Contradiction. $\square$



Direction ($\Rightarrow$). Contrapositive: suppose R is not euclidean. Then $\exists u, v, w\ (Ruv \wedge Ruw \wedge \neg Rvw)$. Define I so that $I(p, w) = \mathsf{T}$ and for every $x \neq w$ with Rvx, $I(p, x) = \mathsf{F}$ (consistent since $\neg Rvw$). Then $\mathcal{M}, u \models \Diamond p$ (via w), $\mathcal{M}, v \models \Box\neg p$ (every successor of v makes p false), hence $\mathcal{M}, u \models \Diamond\Box\neg p$, i.e. $\mathcal{M}, u \not\models \Box\Diamond p$. So $\mathcal{F} \not\models$ axiom 5. $\square$

31.2 Proof-theoretic derivation

$S5 = T +$ euclideanness rule: from wRv and wRu, conclude vRu.

1	$w : \Diamond p$	P
2	$\quad wRv, \;\; v : p$	fresh v
3	$\quad\quad wRu$	fresh u
4	$\quad\quad uRv$	Eucl., 3, 2
5	$\quad\quad u : \Diamond p$	$\Diamond$I, 4, 2
6	$\quad w : \Box\Diamond p$	$\Box$I, 3—5
7	$w : \Box\Diamond p$	$\Diamond$E, 1, 2—6

Remark 31.1 ($S5$ and equivalence relations). In $S5 = T +$ axiom 5, the accessibility relation is reflexive and euclidean. A reflexive euclidean relation is also symmetric and transitive (cf. Exercise 24.1), hence it is an *equivalence relation*. This means:

- Iterated modalities collapse: $\Box\Box\varphi \leftrightarrow \Box\varphi$ and $\Diamond\Diamond\varphi \leftrightarrow \Diamond\varphi$.
- Mixed modalities simplify: $\Box\Diamond\varphi \leftrightarrow \Diamond\varphi$ and $\Diamond\Box\varphi \leftrightarrow \Box\varphi$.
- Every string of $\Box$'s and $\Diamond$'s reduces to a single $\Box$ or $\Diamond$.

32 Summary of Modal Systems and Frame Conditions

System	**Axiom**	**Frame condition**	**ND rule added to K**
K	$\Box(\varphi \to \psi) \to (\Box\varphi \to \Box\psi)$	none	(base system)
D	$\Box\varphi \to \Diamond\varphi$	seriality	seriality rule
T (M)	$\Box\varphi \to \varphi$	reflexivity	Refl. rule
$S4$	$\Box\varphi \to \Box\Box\varphi$	refl. + transitivity	Refl. + Trans. rule
B	$\varphi \to \Box\Diamond\varphi$	refl. + symmetry	Refl. + Sym. rule
$S5$	$\Diamond\varphi \to \Box\Diamond\varphi$	equivalence relation	Refl. + Eucl. rule

Part VIII
Applications

33 Deontic Logic: The System D

New notation. We introduce a deontic operator O ("it is obligatory that") which satisfies the same rules as $\Box$ in K, plus one additional axiom. The dual operators are:

- $P\varphi =_{\text{Def}} \neg O\neg\varphi$: "it is *permitted* that φ".
- $F\varphi =_{\text{Def}} O\neg\varphi$: "it is *forbidden* that φ".

The accessibility relation is read as: Rwv means "v is a world in which all the norms holding at w are satisfied." The worlds accessible from w are the *deontically ideal* alternatives to w.

Axiom D.

$$O\varphi \to P\varphi \qquad \text{equivalently:} \qquad O\varphi \to \neg O\neg\varphi$$

What is obligatory is permitted. (Remember Kant: "ought implies can.") This corresponds to **seriality**: every world sees at least one deontically ideal world.

Why not axiom T? The schema $O\varphi \to \varphi$ ("what is obligatory is actual") would say that obligations are always fulfilled. This is plainly false. Deontic logic uses D, not T.

Necessitation in deontic logic. Because of the necessitation rule, all logical theorems are obligatory: $\vdash_D O(p \vee \neg p)$, for instance. Is this correct? One might argue: logical truths are trivially fulfilled and cost nothing to obey, so their obligatory status is harmless. Others find it philosophically suspect that logical necessities should count as moral obligations.

34 Tense Logic: A Case Study in Expressive Power and Its Limits

Tense logic provides the richest illustration of modal correspondence theory at work. It is also where the *limits* of modal definability become most visible and most philosophically significant.

> **Question.** The standard temporal ordering ("occurring before") is a *dense, strict linear order without endpoints*. Can tense logic axiomatize all of these properties?

34.1 The operators

In tense logic, we introduce two pairs of modal operators:

- **Future**: $G\varphi$ ("it is always going to be the case that φ") and $F\varphi$ ("it will be the case that φ").
- **Past**: $H\varphi$ ("it has always been the case that φ") and $P\varphi$ ("it has been the case that φ").

As with $\Box$ and $\Diamond$:

$$F\varphi =_{\text{Def}} \neg G\neg\varphi \qquad\qquad P\varphi =_{\text{Def}} \neg H\neg\varphi$$

Here, G and H each have their own accessibility relation: R_F ("strictly later than") and R_P ("strictly earlier than"), where R_P is the converse of R_F. Both operators satisfy the rules of K.

The past and future operators are linked by two **interaction axioms** expressing that the past is the converse of the future:

$$(GP) \qquad \varphi \to GP\varphi \quad \text{"If } \varphi \text{ is true now, it will always have been true."} \tag{9}$$

$$(HF) \qquad \varphi \to HF\varphi \quad \text{"If } \varphi \text{ is true now, it always was going to be true."} \tag{10}$$

Why axiom T fails. $G\varphi \to \varphi$ says: if φ will always be the case, then φ is the case now. But the future tense operator quantifies over *future* instants, not the present one. If R is the "strictly later than" relation, the current instant is not in its own future, so reflexivity fails and axiom T does not hold.

34.2 The target: properties of temporal order

What should a temporal ordering look like? The classical answer is that the "occurring before" relation $<$ should satisfy the following properties:

Property	**FOL condition**	**Intuition**
Transitivity	$a < b \wedge b < c \Rightarrow a < c$	temporal order chains compose
Irreflexivity	$\neg(a < a)$	no instant is before itself
Asymmetry	$a < b \Rightarrow \neg(b < a)$	time doesn't run in circles
Seriality	$\forall a\, \exists b\, (a < b)$ and $\forall a\, \exists b\, (b < a)$	no first, no last instant
Connectivity	$a \neq b \Rightarrow (a < b \vee b < a)$	time is a line, not branching
Density	$a < b \Rightarrow \exists c\, (a < c \wedge c < b)$	between any two instants, a third

Together, these properties axiomatize a *dense strict linear order without endpoints*. The standard model is $(\mathbb{Q}, <)$.

34.3 What tense logic can capture

Transitivity. $G\varphi \to GG\varphi$ and $H\varphi \to HH\varphi$ (instances of axiom 4).

Seriality: no last instant, no first instant. $G\varphi \to F\varphi$ and $H\varphi \to P\varphi$ (instances of axiom D).

Connectivity (linearity). $(PF\varphi \vee FP\varphi) \to (P\varphi \vee \varphi \vee F\varphi)$. This axiom forces the ordering to be linear: time is a single thread, not a branching tree.

Density. $Fp \to FFp$ and $Pp \to PPp$. If something will be the case, then it will be that it will be the case—there must be an intermediate instant.

Taking stock. At this point, tense logic has captured a remarkable amount of structure. The axiom system consisting of K axioms for G and H, the interaction axioms GP and HF, transitivity, seriality, connectivity, and density characterizes exactly the class of *dense linear orders without endpoints*. For a purely *propositional* language with just two pairs of modal operators, this is a striking degree of expressive power. It captures the topology of $(\mathbb{Q}, <)$.

34.4 What tense logic cannot capture

Two properties remain: **irreflexivity** and **asymmetry**. These are precisely the properties that make the temporal order *strict*.

> **Theorem** (Goldblatt–Thomason). Irreflexivity ($\forall w\, \neg Rww$) **cannot be defined by any modal formula**.

This means: there is no tense-logical formula φ such that $\mathcal{F} \models \varphi$ iff $\mathcal{F}$ is irreflexive. Similarly, asymmetry is not modally definable.

Why? The deep reason is that modal formulas are *bisimulation-invariant*: if two models are bisimilar, they satisfy the same modal formulas. But irreflexivity is not preserved by bisimulations. More precisely, the Goldblatt–Thomason theorem states that a first-order property of frames is definable by a set of modal formulas if and only if the corresponding class of frames is closed under generated subframes, disjoint unions, and *ultrafilter extensions* (and its complement is closed under ultrapowers). The class of irreflexive frames fails the ultrafilter-extension condition.

34.5 The philosophical moral

The situation is both impressive and frustrating:

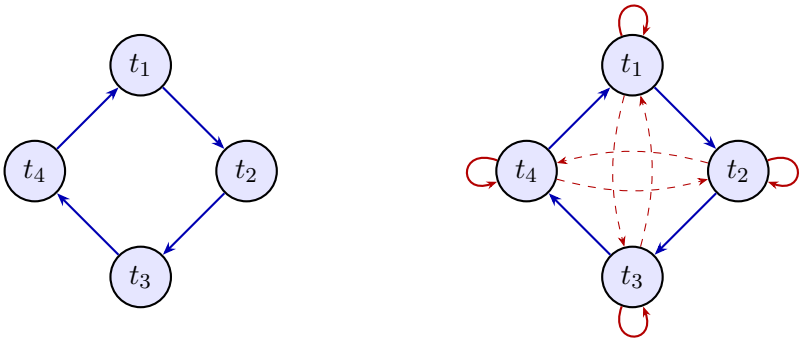
- Tense logic can capture the **topology** of time: dense, linear, no endpoints, transitive.
- But it **cannot capture the strictness** of temporal precedence: irreflexivity and asymmetry escape modal definability.

This means that tense logic *cannot distinguish a strict order ($<$) from a non-strict order ($\leq$).* A dense linear preorder without endpoints validates exactly the same tense-logical formulas as a dense strict linear order without endpoints. From the perspective of tense logic, the two are indistinguishable.

The very property that makes temporal ordering *temporal*—that no instant precedes itself—is precisely what escapes the expressive power of modal logic. We now make this concrete with a striking example.

34.6 Circular time: a concrete consequence

Since asymmetry is not modally definable, nothing prevents the temporal ordering from looping back on itself. Consider four instants arranged in a cycle: $t_1 R\, t_2$, $t_2 R\, t_3$, $t_3 R\, t_4$, $t_4 R\, t_1$.



(a) The successor relation *(b) The transitive closure*

Now recall that transitivity ($G\varphi \to GG\varphi$) is a tense-logical axiom. Taking the transitive closure of the cycle produces new edges: from $t_1 R\, t_2$ and $t_2 R\, t_3$ we get $t_1 R\, t_3$; continuing, $t_1 R\, t_4$ and finally $t_1 R\, t_1$. The self-loops (in red in diagram (b)) are the critical point: transitivity has forced reflexivity. Repeating for every starting point, the transitive closure is the *complete relation*—every instant sees every instant, including itself. The cycle has collapsed into a single $S5$ equivalence class.

Why modal logic cannot exclude this. The argument is immediate: to exclude cycles, we would need asymmetry ($aRb \to \neg(bRa)$), which is not modally definable. But there is also an intuitive explanation. Modal logic is fundamentally *local*: a formula of modal depth k can inspect the accessibility structure at most k steps outward from a given instant. On a cycle

of size $n > k$, every instant has the same local view as a point in $\mathbb{Z}$—there is always a past, always a future, always density. No finite formula can detect that the chain eventually loops. The difference between Aristotelian time (a strict linear order flowing irreversibly forward) and Nietzschean time (an eternal recurrence in which every moment eventually returns) is real, but *modally invisible*.

What would be needed. Richer formalisms can express irreflexivity. In *hybrid logic*, which adds nominals naming individual instants and a satisfaction operator $@_i$, one can write $@_i \neg F\, i$ ("instant i does not lie in its own future"). In first-order logic, one simply writes $\forall x\, \neg(x < x)$. In both cases, the key is the ability to *name* individual points in the frame and compare a world with itself—a resource that propositional modal logic, which treats worlds as anonymous evaluation points, does not have.

35 Logics for Belief and Knowledge

Belief operators. We introduce a family of modal operators B_i for each agent i: $B_i\varphi$ is read "agent i believes that φ." Each B_i satisfies the rules of K; in particular, the necessitation rule applies.

Logical omniscience. Since necessitation holds, if φ is a logical truth, then $B_i\varphi$ for every agent i: every agent believes every logical truth. This is the **principle of logical omniscience**. It is descriptively unrealistic (human agents are not logically omniscient), but may be defensible as a normative idealisation (a rational agent *should* believe logical truths).

35.1 Axiom D for Beliefs

$$B_i\varphi \to \neg B_i \neg \varphi$$

If an agent accepts a proposition, the agent does not also accept its negation. This is a *consistency* requirement on beliefs. It corresponds to seriality: for every doxastic state, there is at least one world compatible with the agent's beliefs.

35.2 Axiom T/M and Epistemic Logic

$$B_i\varphi \to \varphi \qquad ???$$

This principle is *not* appropriate for belief: as Plato famously remarked in the *Gorgias*, beliefs can be false. But it does hold for **knowledge**, because knowledge is a *factive* concept: if an agent knows that φ, then φ is true. Writing K_i for "agent i knows that":

$$K_i\varphi \to \varphi$$

35.3 Positive Introspection

$$K_i\varphi \to K_iK_i\varphi \qquad\qquad B_i\varphi \to B_iB_i\varphi$$

If an agent knows (or believes) something, the agent knows (or believes) that they know (or believe) it. This corresponds to axiom 4 (transitivity).

35.4 Negative Introspection

$$\neg K_i\varphi \to K_i\neg K_i\varphi \qquad\qquad \neg B_i\varphi \to B_i\neg B_i\varphi$$

If an agent does not know (or believe) something, the agent knows (or believes) that they do not. This corresponds to axiom 5 (euclideanness). Adding both positive and negative introspection to epistemic logic yields the system $S5$ for knowledge—the standard framework for epistemic logic in game theory, computer science, and formal epistemology.

Bibliography

Textbooks and introductions

Blackburn, de Rijke & Venema (2001). P. Blackburn, M. de Rijke, and Y. Venema, *Modal Logic*. Cambridge Tracts in Theoretical Computer Science 53. Cambridge University Press. — The standard graduate-level reference. Comprehensive treatment of correspondence theory, bisimulations, and the van Benthem characterisation theorem.

Chellas (1980). B. F. Chellas, *Modal Logic: An Introduction*. Cambridge University Press. — A classic introduction with careful attention to the axiomatic approach and frame conditions.

Fitting & Mendelsohn (1998). M. Fitting and R. L. Mendelsohn, *First-Order Modal Logic*. Synthese Library 277. Kluwer. — Extends the tableau and model-theoretic methods to first-order modal logic. Includes a thorough treatment of propositional modal tableaux as a foundation.

Garson (2013). J. W. Garson, *Modal Logic for Philosophers*. 2nd edition. Cambridge University Press. — An accessible introduction designed for philosophy students, with careful discussion of the philosophical motivations for different modal systems.

Priest (2008). G. Priest, *An Introduction to Non-Classical Logic: From If to Is*. 2nd edition. Cambridge University Press. — Broad coverage of non-classical logics including modal, intuitionistic, relevant, and many-valued logics. Uses the tableau method throughout and includes labelled tableaux for modal logics.

van Benthem (2010). J. van Benthem, *Modal Logic for Open Minds*. CSLI Lecture Notes 199. CSLI Publications, Stanford. — An informal and philosophically rich introduction by one of the founders of modern modal logic. Emphasises the "modal logic as a fragment of FOL" perspective and bisimulation invariance.

Zach (2019). R. Zach, *Boxes and Diamonds: An Open Introduction to Modal Logic*. Open Logic Project. Available at https://openlogicproject.org. — A freely available, open-source textbook covering Kripke semantics, tableaux, axiomatic systems, and correspondence theory at the introductory level.

On labelled deduction systems

Fitting (1983). M. Fitting, *Proof Methods for Modal and Intuitionistic Logics*. Synthese Library 169. D. Reidel. — The pioneering monograph on prefixed (labelled) tableau and proof methods for modal logics. The labelled approach used in these notes descends from Fitting's framework.

Negri (2005). S. Negri, "Proof Analysis in Modal Logic." *Journal of Philosophical Logic*, 34(5–6):507–544. — Develops labelled sequent calculi for modal logics with explicit accessibility statements, closely related to the labelled natural deduction system presented here. Proves cut elimination and establishes the connection between structural rules and frame conditions.

Viganò (2000). L. Viganò, *Labelled Non-Classical Logics*. Kluwer. — A systematic study of labelled deductive systems for modal, relevance, and substructural logics, with detailed proof-theoretic analysis.

On specific topics

Goldblatt (1992). R. Goldblatt, *Logics of Time and Computation*. 2nd edition. CSLI Lecture Notes 7. CSLI Publications, Stanford. — Covers temporal logic, dynamic logic, and the limits of modal definability, including the Goldblatt–Thomason theorem discussed in these notes.

Hintikka (1962). J. Hintikka, *Knowledge and Belief: An Introduction to the Logic of the Two Notions*. Cornell University Press. — The foundational work on epistemic logic, introducing the possible-worlds analysis of knowledge and belief that motivates the systems discussed in Section 35.

Kripke (1963). S. A. Kripke, "Semantical Considerations on Modal Logic." *Acta Philosophica Fennica*, 16:83–94. — The paper that introduced the relational semantics for modal logic (now called Kripke semantics), which forms the foundation of these lecture notes.