6.3 Mathematical Induction

To complete our proofs of the soundness and completeness of propositional modal logic, it is worth briefly discussing the principles of mathematical induction, which we assume in our metalanguage.

A natural number is a nonnegative whole number, or a number in the series 0, 1, 2, 3, 4, …

The principle of mathematical induction: If
(φ is true of 0) and (for all natural numbers n, if φ is true of n, then φ is true of n + 1), then φ is true of all natural numbers.

To use the principle mathematical induction to arrive at the conclusion that something is true of all natural numbers, one needs to prove the two conjuncts of the antecedent:

1. (Base step) φ is true of 0
2. (Induction step) for all natural numbers n, if φ is true of n, then φ is true of n + 1.

Typically, the induction step is proven by means of a conditional proof in which it is assumed that φ is true of n, and from this assumption it is shown that φ must be true of n + 1. In the context of this conditional proof, the assumption that φ is true of n is called the inductive hypothesis.

From the principle of mathematical induction, one can derive a related principle:

The principle of complete induction: If (for all natural numbers n, whenever φ is true of all numbers less than n, φ is also true of n) then φ is true of all natural numbers.

Complete induction is also called strong induction. (Arguments that use these principles are actually deductive arguments.)

In this class, we rarely use these principles as stated. Instead, we use some corollaries that come in handy in the study of logical systems. Garson does not give these principles special names, but I will.

The principle of wff induction: For a given logical language, if φ holds of the simplest sentences or well-formed formulas (wffs) of that language, and φ holds of any complex wff provided that φ holds of those simpler wffs out of which it is constructed, then φ holds of all wffs.

This principle is often used in logical metatheory. It is a corollary of mathematical induction. Actually, it is a version of it. Let φ’ be the property a number has if and only if all wffs of the logical language having that number of logical operators and atomic statements are φ. If φ is true of the simplest well-formed formulas, i.e., those that contain zero operators, then 0 has φ’. Similarly, if φ holds of any wffs that are constructed out of simpler wffs provided that those simpler wffs are φ, then whenever a given natural number n has φ’ then n + 1 also has φ’. Hence, by mathematical induction, all natural numbers have φ’, i.e., no matter how many operators a wff contains, it has φ. In this way wff induction simply reduces to mathematical induction.

Similarly, this principle is usually utilized by proving:

1. (Base step) φ is true of the simplest well-formed formulas (wffs) of that language; and
2. (Induction step) φ holds of any wffs that are constructed out of simpler wffs provided that φ holds of those simpler wffs.

Again, the assumption made when establishing (2) that φ holds of the simpler wffs is called the inductive hypothesis.

For technical reasons, in the logical languages studied in this course, we define “simpler” to mean “containing fewer signs, not counting occurrences of ‘⊥’ towards the total”.

Well also be using:

The principle of proof induction: In a system containing derivations or proofs, if φ holds of a given step of the proof whenever φ holds of all previous steps of the proof, then φ holds of all steps of the proof.

The principle of proof induction is an obvious corollary of the principle of complete induction. The steps in a proof can be numbered; were just applying complete induction to those numbers.

6.4 Soundness

We now turn to showing that for the various logical systems for propositional modal logic we’ve been considering, the following holds:

If \( A_1, A_2, \ldots, A_n \vdash B \) then \( A_1, A_2, \ldots, A_n \models B \).

In other words, if a proof exists for a given argument, then that argument is valid (in the relevant sense).
1. Assume for conditional proof (in the metalanguage) that $A_1, A_2, \ldots, A_n \vdash B$. This means there is a proof with $A_1, A_2, \ldots, A_n$ as hypotheses and $B$ as conclusion. We need to show that there is no modal model (of the appropriate kind) $(W, R, a)$ and world $w$ in $W$ such that $a_w(A_1) = T$ and $a_w(A_2) = T \ldots$ and $a_w(A_n) = T$ but $a_w(B) = F$, or in other words, that for every modal model $(W, R, a)$ and world $w$ in $W$, if $a_w(A_1) = T$ and $a_w(A_2) = T \ldots$ and $a_w(A_n) = T$ then $a_w(B) = T$.

2. Assume that $a_w(A_1) = T$ and $a_w(A_2) = T \ldots$ and $a_w(A_n) = T$ for arbitrarily chosen world $w$ in an arbitrary modal model $(W, R, a)$ of the appropriate kind.

3. There exists a proof of $B$ from $A_1, A_2, \ldots, A_n$. We require that this proof be set out in full, including all (Reit) steps, without using any defined notation, or citing derived rules or theorems. (Proofs that use these shortcuts can be transformed into those that do not.)

4. We shall prove by proof induction that every step of this proof has the following property $\phi$: for any world $v$ in $W$, if all the relevant assumptions of the subproof in which this step appears are true at $v$, then so is the statement at this step. The “relevant assumptions” of a subproof can be defined as follows:

   - The relevant assumptions of the main subproof are $A_1, A_2, \ldots, A_n$.
   - The relevant assumptions of a given conditional subproof are those of the prior subproof in which it appears along with the statement that heads the subproof.
   - The relevant assumptions of a given boxed subproof are those statements in it that are introduced into it by means of the (□Out) rule.

5. We assume as inductive hypothesis that $\phi$ holds for every step in the proof prior to a given one. We need to show that $\phi$ holds of it as well. Let us call the statement written at the current line $C$, and the subproof in which it appears $s$. Let us call the subproof in which $s$ is embedded (if there is one) $s'$. There are eight cases to consider:

   Case (a): $C$ is a hypothesis or assumption for conditional proof. In that case, $C$ is one of the relevant assumptions of $s$, and so it holds trivially that it is true in all worlds $v$ that make all relevant assumptions true.

   Case (b): $C$ is the result of applying (MP) to previous steps within subproof $s$. These steps must take the form $D$ and $D \rightarrow C$ for some statement $D$. By the inductive hypothesis, for every world $v$ making the relevant assumptions of $s$ true, $a_v(D \rightarrow C) = T$ and $a_v(D) = T$. By the semantics for $\rightarrow$, $a_v(D \rightarrow C) = T$ iff either $a_v(D) = F$ or $a_v(C) = T$; since it cannot be the former, it must be the latter, and so $a_v(C) = T$ for all such worlds $v$.

   Case (c): $C$ is the result of applying (DN) to a previous step within subproof $s$. By the inductive hypothesis, for all worlds $v$ in which the relevant hypotheses of $s$ are true, $a_v(\sim \sim C) = T$, and so $a_v(\sim C) = F$ and finally, $a_v(C) = T$.

   Case (d): $C$ is the result of closing a (CP). In that case, $C$ takes the form $D \rightarrow E$ and there is a conditional subproof immediately prior to the current step headed by $D$ that ends in $E$. By the inductive hypothesis, every world $v$ in $W$ that makes all of the relevant assumptions of $s$ true and makes $D$ true also makes $E$ true. Hence, for every world $v$ making all the relevant assumptions of $s$ true, it must be that either $a_v(D) = F$ or $a_v(E) = T$; either way, $a_v(D \rightarrow E) = T$, i.e., $a_v(C) = T$.

   Case (e): $C$ is the result of a (Reit) step from subproof $s'$, and $s$ must be a conditional subproof. By the inductive hypothesis, $C$ is true in every world $v$ that makes all of the relevant assumptions of $s'$ true. Since the relevant assumptions of $s$ include those of $s'$, every world that makes all of its relevant assumptions true is also a world that makes those of $s'$ true, and hence for each such world, $a_v(C) = T$.

   Case (f): $C$ is the result of a (□Out) step. Hence $s$ must be a world subproof, and $C$ is one of its relevant hypotheses, so it holds trivially that every world $v$ that makes all its relevant hypotheses true makes it true as well.

   Case (g): $C$ is the result of (□In), i.e., it takes the form $\Box D$, and is the result of closing a world subproof whose last line is $D$. By the inductive hypothesis, $D$ is true in every world that makes the relevant assumptions of this boxed subproof true. These relevant assumptions are results of applying (□Out) from steps
within $s$ itself. It also holds by the inductive hypothesis that for every world $v$ that makes the relevant assumptions of $s$ true must make statements of the form $\Box E_1, \Box E_2, \ldots, \Box E_n$, true where $E_1, E_2, \ldots, E_n$ are the relevant assumptions of the boxed subproof. By the semantics of $\Box$-statements, this means that in every world $u$ such that $vRu$, $E_1, E_2, \ldots, E_n$ are true. Hence, for every world $u$ such that $vRu$, $a_u(D) = T$ as well, and hence, it follows that $a_v(\Box D) = T$, i.e., $a_v(C) = T$.

**Case (h):** $C$ is an axiom of $T$, $K4$, $S4$, $KB$, $B$, $K5$ or $S5$ (and our proof is in one of these systems). In that case, $C$ is true in all worlds in all models of the relevant sort, and hence certainly must be true on all worlds making the relevant assumptions true. This can be seen as follows:

- Any instance of axiom schema (M) must be true in all worlds in all M-models. Suppose for *reductio ad absurdum* that this was not the case. Then there would be an M-model $(W, R, a)$ and world $w$ in $W$, where $a_w(\Box A) = F$. It follows by the semantics of ‘$\Box$’ that $a_w(\Box A) = T$ and $a_w(A) = F$. By the semantics of ‘$\Box$’, it follows that $a_v(A) = T$ for all worlds $v$ such that $wRv$. Since this is an M-model, $wRw$, and so $a_w(A) = T$, which yields a contradiction, and our supposition is impossible.

- Since $B$-, $S4$- and $S5$-models are also M-models, any instance of (M) must be true in every world in all these sorts of models as well.

- Any instance of axiom schema (B) must be true in all KB-models. Suppose for *reductio* that this were not so. Then there would be a KB-model $(W, R, a)$ and world $w$ such that $a_w(A \rightarrow \Box \Box A) = F$. Hence, $a_w(A) = T$ and $a_w(\Box \Box A) = F$. By the semantics of ‘$\Box$’, this means there is a world $v$ such that $wRv$ and $a_v(\Box A) = F$. Hence, for every world $u$ such that $vRu$, $a_u(A) = F$. However, since it is a KB-model, $R$ is symmetric, and so $vRw$, and thus $a_w(A) = F$, which is impossible.

**Homework:** Show that something similar holds for axiom schemata (4) and (5) relative to 4-models and 5-models.

6. It follows by proof induction that $\phi$ holds for all steps of the proof, and thus the last step, which is $B$ in the main subproof. Hence, $B$ is true in every world $v$ in the model that makes the relevant assumptions of the main subproof true. Since $w$ is such a world, $a_w(B) = T$.

7. This completes the conditional proof begun at 2.

### 6.5 Tree Modeling

To complete our triangle, we need only prove:

- If $A_1, A_2, \ldots, A_n \models B$ then the tree for $A_1, A_2, \ldots, A_n, \neg B$ closes.

Rather than proving this outright, we shall instead prove something equivalent—its contrapositive:

- If the tree for $A_1, A_2, \ldots, A_n, \neg B$ has an open branch, then there is a counterexample to the argument $A_1, A_2, \ldots, A_n, \vdash B$.

The essence of this result has already been explained. If a tree has an open branch, that branch represents a model making the various statements on that branch true in the worlds in whose world box they appear. If the tree is headed by the premises and negation of the conclusion of an argument, it provides a counterexample to the argument.

While this is intuitively correct, we haven’t actually proven it in a rigorous fashion. We now do so.

1. Let us assume that the tree for $A_1, A_2, \ldots, A_n, \neg B$ has an open branch. This means that even with all possible tree rules applied to the statements on that branch, there is no occurrence of ‘$\bot$’ anywhere on the branch. We require that the tree be written out in full, with no defined notation.

2. Let us define the *tree model* for this branch as the appropriate sort of modal model (depending on the relevant notion of validity) $(W, R, a)$ such that:

- $W$ is the set of worlds represented by world boxes on the branch.

- $wRv$ holds for any worlds $w$ and $v$ in $W$ just in case there is an accessibility arrow from $w$ to $v$ on the tree. (We assume all the appropriate arrows have been inserted for the appropriate kind of validity being checked. If the tree is a tree for checking M-validity, for example, $wRw$ will hold for all worlds $w$ in $W$.)
• For every atomic statement $P$ and world $w$ in $W$, $a_w(P) = T$ if $P$ occurs on the branch inside the world box for $w$, and $a_w(P) = F$, otherwise. (For complex statements, their value for $a$ at $w$ depends upon the rules governing all modal models; see page 18 of the handouts.)

Our goal is to show that the tree model is a counterexample.

3. Let $\phi$ be the property a statement $A$ has just in case for all worlds $w$ in $W$ of the tree model, if $A$ appears on the open branch inside the world box for $w$, then $a_w(A) = T$. We shall use **wff induction** to show that $\phi$ holds of all statements of propositional modal logic.

4. **(Base step)** $\phi$ holds of all atomic statements $P$ by the definition of the tree model. $\phi$ holds of ‘⊥’ trivially, since the antecedent is always false (since the branch is open). These are the simplest statements of modal propositional logic.

5. **(Induction step)** Suppose as inductive hypothesis, $a_w(Q) = T$ for all these worlds, By the truth-conditions for statements of the form $\Box Q$, it follows that $a_w(\Box Q) = T$, i.e., $a_w(S) = T$.

Case (a): $S$ takes the form $\Box Q$. Since all tree rules have been applied, $Q$ must occur on the branch in all world boxes for all worlds $v$ such that $wRv$. By the inductive hypothesis, $a_w(Q) = T$ for all these worlds, By the truth-conditions for statements of the form $\Box Q$, it follows that $a_w(\Box Q) = T$, i.e., $a_w(S) = T$.

Case (b): $S$ takes the form $C \to \bot$ (i.e., ‘$\neg C$’). There are three cases to consider:

- **(Subcase i)** $C$ takes the form $\Box D$. Since $\neg C$ occurs on the tree in $w$, and all tree rules have been applied, there is a world box $v$ where $wRv$ containing $\neg D$. By the inductive hypothesis, $a_v(\neg D) = T$, and so $a_v(D) = F$. Hence it is not the case that $a_w(D) = T$ for all worlds $v$ such that $wRv$, and so $a_w(\Box D) = F$, i.e., $a_w(C) = F$. Hence, $a_w(\neg C) = T$, i.e., $a_w(S) = T$.

- **(Subcase ii)** $C$ takes the form $D \to \bot$, and hence $S$ takes the form $\neg \neg D$. Since all tree rules have been applied, $D$ occurs on the tree in world $w$. By the inductive hypotheses, $a_w(D) = T$, and hence $a_v(\neg D) = F$, and finally $a_w(\neg \neg D) = T$, i.e., $a_w(S) = T$.

- **(Subcase iii)** $C$ takes the form $D \to E$ where $E$ is not $\bot$. Since all tree rules have been applied, and world $w$ contains $S$, i.e., $\neg (D \to E)$, it must contain both $D$ and $\neg E$. By the inductive hypothesis, $a_w(D) = T$ and $a_w(\neg E) = T$, and so $a_w(E) = F$. By the semantics for ‘$\to$’ , it follows that $a_w(D \to E) = F$, i.e., $a_w(C) = F$. It follows that $a_w(\neg C) = T$, i.e., $a_w(S) = T$.

- **(Subcase iv)** $C$ is $\bot$. Therefore, $a_w(C) = F$ and $a_w(S) = T$.

- **(Subcase v)** $C$ is an atomic statement. Since $S$, i.e., $\neg C$, appears in the world box for $w$ on the branch, and the branch is open, $C$ must not appear. By the definition of the tree model, it follows that $a_w(C) = F$, and so $a_w(S) = T$.

Case (c): $S$ takes the form $C \to D$, where $D$ is not $\bot$. Since all tree rules have been applied, the open branch must either have $\neg C$ or $D$ on the branch. Because $D$ is not $\bot$, $\neg C$ is simpler than $S$. Hence, by the inductive hypothesis, either $a_w(\neg C) = T$ or $a_w(D) = T$. Hence, either $a_w(C) = F$ or $a_w(D) = T$. By the semantics of ‘$\to$’, it follows that $a_w(C \to D) = T$, i.e., $a_w(S) = T$.

These cases cover every possible complex statement, since ‘$\to$’ or ‘$\Box$’ is the main operator of any complex statement in propositional modal logic (in undefined notation). Hence $\phi$ holds of any complex statement $S$ assuming it holds of all simpler statements.

6. It follows by wff induction that $\phi$ holds of every statement of modal propositional logic. Hence it holds of $A_1, A_2, \ldots, A_n$, and $\neg B$. Since these statements appear in the starting world box, $w$, of the open branch, $a_w(A_1) = T$ and $a_w(A_1) = T$ and $a_w(A_n) = T$ and $a_w(\neg B) = T$, and therefore, $a_w(B) = F$.

7. The tree model $(W, R, a)$ therefore provides a counterexample to the argument $A_1, A_2, \ldots, A_n / \therefore B$, since world $w$ makes the premises true and conclusion false.

This completes our proof of the contrapositive of our desired result (which also therefore follows). The results holds for all the conceptions of validity we’ve considered (K-validity, M-validity, 4-validity,
S4-validity, etc.)—these just put constraints on what accessibility arrows must be drawn in when creating the tree. Hence:

- If $A_1, A_2, \ldots, A_n \models K B$ then the K-tree for $A_1, A_2, \ldots, A_n, \neg B$ closes.
- If $A_1, A_2, \ldots, A_n \models S_5 B$ then the S5-tree for $A_1, A_2, \ldots, A_n, \neg B$ closes.
- And so on for the others (M, KB, B, K4, S4, K5).

6.6 Summary

We now have completed showing our “triangle”:

$$A_1, A_2, \ldots, A_n \vdash B \quad \text{soundness} \quad A_1, A_2, \ldots, A_n \models B$$

The two lower parts of the triangle, put together, yield completeness: every valid argument has a proof in the relevant system.

It follows that in the logical systems we have been discussing semantic validity coincides with derivability.

As a limiting case of the above:

$$\vdash A \quad \models A$$

The tree for $\neg A$ closes

Every theorem is a logically valid statement, or a logical truth (on the relevant conception of validity). Every logically valid statement has a closed tree for its negation, and every statement that has a closed tree for its negation is a theorem of the relevant modal system.

A corollary of these results is that $\vdash A$ iff $\models A$. A statement is a theorem if and only if it is logically valid.

Another corollary is the consistency of these systems.

**Consistency:** A deductive system $S$ is consistent if and only if $\nvdash S \bot$.

We can prove the consistent of our systems in the following way:

Assume for *reductio ad absurdum* that $\vdash \bot$. By soundness, $\vdash \bot$. By the definition of validity, for every model modal $(W, R, a)$ of the relevant sort and world $w$ in $W$, $a_w(\bot) = T$. But, by definition, for every model modal $(W, R, a)$ and world $w$, $a_w(\bot) = F$, so it is not possible that $a_w(\bot) = T$. This is a contradiction, and we must conclude that $\nvdash \bot$ instead.

Together these results can be seen as establishing the “adequacy” of the logical systems we’ve discussed in so far as they allow for the derivability of the things we want to be able to derive (assuming the operant definition of validity is appropriate), and nothing that we don’t.

The results, however, don’t settle all doubts or questions we might have about propositional modal logic:

- They don’t establish that our formal definition of truth in a model and logical validity in terms of modal models capture our pre-theoretic conceptions of valid reasoning involving necessity and possibility (or the best conceptions of these notions).
- They don’t provide any insight into the nature of necessity itself, and whether or not it is reducible or explicable in terms of anything else.
- They don’t tell us what the true or ultimate nature of “possible worlds” are.
- They don’t settle which of the various rival systems best captures any given conception of modality, or which is most appropriately applied in what contexts.
- They don’t rule out the possibility of other conceptions of “necessity” or “possibility” that don’t cohere with any of the systems we’ve examined. (For example, none have the result that $\vdash \Box p & \Box \neg p$ for every atomic statement ‘$p$’, or that $\vdash \Box(p & q)$ for every two atomic statements ‘$p$’ and ‘$q$’; as you might want on certain conceptions of possibility inspired by logical atomism.)
- They don’t justify the use of ‘$\Box$’ as an object-language operator rather than a metalinguistic predicate.

These are issues that can only be addressed philosophically, not with mathematical or quasi-mathematical “proofs”. Questions such as these are even more poignant in the case of Quantified Modal Logic, which we turn to next.