Chapter 9

The Deductive System PLM

In this chapter, we introduce the deductive system PLM by combining the axioms of the previous chapter with a primitive rule of inference and defining the notions of derivation, proof, and theorem. We then develop a series of theorems, the proofs of which are often facilitated by the introduction of metarules, i.e., special metatheorems that allow us to infer the existence of derivations and proofs. The actual derivations, proofs of the theorems, and the justifications of the metarules all appear in the main Appendix, while the proofs of other metatheorems continue to appear in chapter appendices. The Appendices are gathered together in Part IV.

9.1 Primitive Rule of PLM: Modus Ponens

(42) **Primitive Rule**: Modus Ponens. PLM employs just a single primitive rule of inference:

> **Modus Ponens** (**Rule MP**)
> $\varphi, \varphi \rightarrow \psi \ / \ \psi$

i.e., ψ follows from the formulas φ and $\varphi \rightarrow \psi$.

9.2 (Modally Strict) Proofs and Derivations

(43) **Metadefinitions:** Derivations, Proofs, and Theorems of PLM. In what follows, we say that φ is an *axiom* of PLM whenever φ is one of the axioms asserted in Chapter 8. The set of axioms of PLM is recursive and we introduce the following symbol to refer to it:

$$\boldsymbol{\Lambda} = \{\varphi | \varphi \textit{ is an axiom}\}$$

Then we define:

(.1) A *derivation* in PLM *of* φ *from* a set of formulas Γ is any sequence of formulas $\varphi_1, \ldots, \varphi_n$ such that φ is φ_n and for each i $(1 \leq i \leq n)$, φ_i is either an element of $\mathbf{\Lambda} \cup \Gamma$ or follows from two of the preceding members of the sequence by Rule MP. A formula φ *is derivable from*, or *follows from*, the set Γ in PLM, written $\Gamma \vdash \varphi$, just in case there exists a derivation of φ from Γ.

We now adopt the following conventions. We call the members of Γ the *premises* or *assumptions* of the derivation. We often write $\varphi_1, \ldots, \varphi_n \vdash \psi$ when $\Gamma \vdash \psi$ and Γ is the set $\{\varphi_1, \ldots, \varphi_n\}$. We often write $\Gamma, \psi \vdash \varphi$ when $\Gamma \cup \{\psi\} \vdash \varphi$. We often write $\Gamma_1, \Gamma_2 \vdash \varphi$ when $\Gamma_1 \cup \Gamma_2 \vdash \varphi$. Whenever a sequence $\varphi_1, \ldots, \varphi_n$ is a derivation of φ from Γ, we say the sequence is a *witness to the claim* that $\Gamma \vdash \varphi$ (henceforth, *a witness to* $\Gamma \vdash \varphi$), and we call the members of the sequence the *lines* of the derivation.

In virtue of the above definition, $\vdash$ refers to *derivability*; this is a multigrade, metatheoretic relation between zero or more formulas and a formula designated as the *conclusion*.

Now using the definition of *derivation*, we may define the notions of *proof* in PLM and *theorem* of PLM:

(.2) A *proof* of φ in PLM is any derivation of φ from Γ in PLM in which Γ is the empty set $\varnothing$. A formula φ is a *theorem* of PLM, written $\vdash \varphi$, if and only if there exists a proof of φ in PLM.

Two simple consequences of our definitions are:

(.3) $\vdash \varphi$ if and only if there is a sequence of formulas $\varphi_1, \ldots, \varphi_n$ such that φ is φ_n and for each i $(1 \leq i \leq n)$, φ_i is either an element of $\mathbf{\Lambda}$ or follows from two of the preceding members of the sequence by Rule MP.

(.4) If $\Gamma = \varnothing$, then $\Gamma \vdash \varphi$ if and only if $\vdash \varphi$.

(44) **Metadefinitions:** Modally Strict Proofs, Theorems, and Derivations. To precisely identify those derivations and proofs that make no appeal to the necessitation-averse axiom (30)★, we first say that φ is a *necessary* axiom whenever φ is any axiom for which we've taken *all* the closures and not just the $\Box$-free closures. So far, (30)★ is the only axiom for which we've taken just the $\Box$-free closures, but one could extend the system with other such necessitation-averse axioms. We introduce the following symbol to refer to the set of necessary axioms:

$$\mathbf{\Lambda}_\Box = \{\varphi \mid \varphi \textit{ is a necessary axiom}\}$$

Then the following definitions of modally-strict derivations and proofs mirror (43.1) and (43.2), with the exception that the definientia refer to $\mathbf{\Lambda}_\Box$ instead of $\mathbf{\Lambda}$ and the ensuing definitions are all correspondingly modified:

(.1) A *modally-strict derivation* (or $\Box$-*derivation*) *of* φ *from* a set of formulas Γ in PLM is any sequence of formulas $\varphi_1,\ldots,\varphi_n$ such that $\varphi = \varphi_n$ and for each i $(1 \leq i \leq n)$, φ_i is either an element of $\mathbf{\Lambda}_\Box \cup \Gamma$ or follows from two of the preceding members of the sequence by Rule MP. A formula φ *is strictly derivable* (or $\Box$-derivable) *from* the set Γ in PLM, written $\Gamma \vdash_\Box \varphi$, just in case there exists a modally-strict derivation of φ from Γ.

(.2) A *modally-strict proof* (or $\Box$-*proof*) of φ in PLM is any modally-strict derivation of φ from Γ when Γ is the empty set. A formula φ is a *modally-strict theorem* (or $\Box$-*theorem*) of PLM, written $\vdash_\Box \varphi$, if and only if there exists a modally-strict proof of φ in PLM.

These two definitions have simple consequences analogous to (43.3) and (43.4). We shall suppose that all of the conventions introduced in (43) concerning $\vdash$ also apply to $\vdash_\Box$.

(**45**) **Remark:** Metarules of Inference. In what follows, we often introduce and prove certain metatheorems about derivations. These facts all have the following form:

If conditions … hold, then there exists a derivation of φ from Γ.

We call facts having this form *metarules of inference* (as opposed to *rules of inference*), since instead of allowing us to infer φ from zero or more formulas, they allow us to infer the existence of a derivation of φ given certain conditions. Metarules often shorten the reasoning we use in the Appendix to establish that $\Gamma \vdash \varphi$ since, frequently, in the process of deriving φ from Γ, we reach a point in the reasoning where what we have established thus far meets the conditions of a metarule whose consequent asserts that there is a sequence of formulas constituting a derivation of φ from Γ.

Consequently, when we reason with metarules to establish the claim that $\Gamma \vdash \varphi$, we don't actually produce a witness to the claim. However, the proof of the metarule in the Appendix (they are, after all, metatheorems) shows how to construct such a witness. We call such proofs *justifications* of the metarule. The justification shows that an appeal to a metarule during the course of reasoning can always be converted into a *bona fide* derivation.

(**46**) **Metarules:** Modally Strict Derivations are Derivations. It immediately follows from our definitions that: (.1) if there is a modally-strict derivation of φ from Γ, then there is a derivation of φ from Γ, and (.2) if there is a modally-strict proof of φ, then there is a proof of φ:

(.1) If $\Gamma \vdash_{\Box} \varphi$, then $\Gamma \vdash \varphi$

(.2) If $\vdash_{\Box} \varphi$, then $\vdash \varphi$

Clearly, however, the converses are not true in general, since derivations and proofs in which (necessitation-averse) axiom (30)$\star$ is used are not modally-strict. Consequently, modally-strict derivations and proofs constitute a proper subset, respectively, of all derivations and proofs.

(47) **Metarules:** Fundamental Properties of $\vdash$ and $\vdash_{\Box}$. The following facts are particularly useful as we prove new theorems and justify new metarules of PLM. Note that these facts come in pairs, with one member of the pair governing $\vdash$ and the other member governing $\vdash_{\Box}$:

(.1) If $\varphi \in \mathbf{\Lambda}$, then $\vdash \varphi$. ("Axioms are theorems")
If $\varphi \in \mathbf{\Lambda}_{\Box}$, then $\vdash_{\Box} \varphi$. ("Necessary axioms are modally-strict theorems")

(.2) If $\varphi \in \Gamma$, then $\Gamma \vdash \varphi$. (Note the special case: $\varphi \vdash \varphi$)
If $\varphi \in \Gamma$, then $\Gamma \vdash_{\Box} \varphi$. (Note the special case: $\varphi \vdash_{\Box} \varphi$)

(.3) If $\vdash \varphi$, then $\Gamma \vdash \varphi$.
If $\vdash_{\Box} \varphi$, then $\Gamma \vdash_{\Box} \varphi$.

(.4) If $\varphi \in \mathbf{\Lambda} \cup \Gamma$, then $\Gamma \vdash \varphi$.
If $\varphi \in \mathbf{\Lambda}_{\Box} \cup \Gamma$, then $\Gamma \vdash_{\Box} \varphi$.

(.5) If $\Gamma \vdash \varphi$ and $\Gamma \vdash (\varphi \to \psi)$, then $\Gamma \vdash \psi$.
If $\Gamma \vdash_{\Box} \varphi$ and $\Gamma \vdash_{\Box} (\varphi \to \psi)$, then $\Gamma \vdash_{\Box} \psi$.

(.6) If $\vdash \varphi$ and $\vdash (\varphi \to \psi)$, then $\vdash \psi$.
If $\vdash_{\Box} \varphi$ and $\vdash_{\Box} (\varphi \to \psi)$, then $\vdash_{\Box} \psi$.

(.7) If $\Gamma \vdash \varphi$ and $\Gamma \subseteq \Delta$, then $\Delta \vdash \varphi$.
If $\Gamma \vdash_{\Box} \varphi$ and $\Gamma \subseteq \Delta$, then $\Delta \vdash_{\Box} \varphi$.

(.8) If $\Gamma \vdash \varphi$ and $\varphi \vdash \psi$, then $\Gamma \vdash \psi$.
If $\Gamma \vdash_{\Box} \varphi$ and $\varphi \vdash_{\Box} \psi$, then $\Gamma \vdash_{\Box} \psi$.

(.9) If $\Gamma \vdash \varphi$, then $\Gamma \vdash (\psi \to \varphi)$, for any ψ.
If $\Gamma \vdash_{\Box} \varphi$, then $\Gamma \vdash_{\Box} (\psi \to \varphi)$, for any ψ.

(.10) If $\Gamma \vdash (\varphi \to \psi)$, then $\Gamma, \varphi \vdash \psi$.
If $\Gamma \vdash_{\Box} (\varphi \to \psi)$, then $\Gamma, \varphi \vdash_{\Box} \psi$.

Notice here that in the special case of (47.2), $\varphi \vdash_{\Box} \varphi$ holds even if φ isn't a necessary truth. In general, there can be modally-strict derivations in which neither the premises nor conclusion are necessary truths. The 3-element sequence $Pa, Pa \to Qb, Qb$ is a modally-strict derivation of Qb from the assumptions Pa

and $Pa \to Qb$ whether or not the premises and the conclusion are necessary truths.

(**48**) **Remark:** $\star$-Proofs and $\star$-Theorems. For the most part, we shall be interested in proofs generally, not just modally-strict ones, since our primary interest is what claims we can prove (*simpliciter*). But since most of the theorems we prove in what follows are modally-strict $\square$-theorems, it is useful to mark the ones that are not. So the reader may assume that all of the items marked **Theorem** in what follows have modally-strict proofs *unless explicitly annotated otherwise*! We use the symbol $\star$ to make it explicit that a proof of a theorem is *not* modally-strict:

(a) we sometimes refer to such sequences as $\star$-*proofs*,

(b) we use the annotation $\star$**Theorem** when the theorem is first introduced,

(c) we refer to such theorems as $\star$-*theorems*, and

(d) we put a $\star$ after its item number when referencing a $\star$-theorem.[94]

Similarly, in the case where the derivation of φ from Γ is not modally-strict, we may speak of $\star$-derivations and say that φ is $\star$-derivable from Γ.

(**49**) **Metadefinition:** Dependence. It is sometimes useful to indicate the difference between $\square$-derivations and $\star$-derivations by saying that in the latter, the conclusion *depends* upon a necessitation-averse axiom, or depends upon a $\star$-theorem that in turn depends upon a necessitation-averse axiom, etc. To make this talk of dependence precise, we define the conditions under which one formula depends upon another within the context of a derivation:

> Let the sequence $\varphi_1, \dots, \varphi_n$ be a derivation in PLM of φ $(= \varphi_n)$ from the set of premises Γ and let ψ be a member of this sequence. Then we say that φ_i $(1 \leq i \leq n)$ *depends upon* the formula ψ *in* this derivation iff either (a) $\varphi_i = \psi$, or (b) φ_i follows by Rule MP from two previous members of the sequence at least one of which depends upon ψ.

When $\Gamma = \varnothing$, this reduces to a definition of: φ depends on a formula ψ in a given proof of φ.

It follows from our definition that if a sequence S is a witness to $\Gamma \vdash \varphi$, then S is a witness to $\Gamma \vdash_\square \varphi$ if and only if φ doesn't depend upon any instance of necessitation-averse axiom (30)$\star$ in S. This holds even if, in S, φ depends upon a premise in Γ that isn't necessary. The sequence $S = Pa, Pa \to Qb, Qb$ is a witness to $Pa, Pa \to Qb \vdash Qb$ and even if Pa fails to be a necessary truth, it follows that $Pa, Pa \to Qb \vdash_\square Qb$, since Qb doesn't depend on (30)$\star$ in S. Thus, the $\star$-theorems below ultimately depend upon axiom (30)$\star$, either directly or because they depend on other $\star$-theorems that depend upon (30)$\star$, and so on.

[94]The first such theorems are (90.1)$\star$ – (90.2)$\star$ and (92)$\star$ – (97.3)$\star$ below.

9.3 Two Fundamental Metarules: GEN and RN

(**50**) **Metarule**: The Rule of Universal Generalization. The Rule of Universal Generalization (GEN) asserts that whenever there is a derivation of a claim (involving the variable α) of the form $\ldots\alpha\ldots$ from a set of premises Γ, and none of the premises in Γ is a special assumption about α, then there is a derivation from Γ of the claim $\forall\alpha(\ldots\alpha\ldots)$:

> **Rule of Universal Generalization** (**GEN**)
> If $\Gamma \vdash \varphi$ and α doesn't occur free in any formula in Γ, then $\Gamma \vdash \forall\alpha\varphi$.

When $\Gamma = \varnothing$, then GEN asserts that if a formula φ is a theorem, then so is $\forall\alpha\varphi$:

> If $\vdash \varphi$, then $\vdash \forall\alpha\varphi$

Note the proviso to GEN when φ is derived from Γ, namely, that the variable α doesn't occur free in any premise in Γ. This prohibits, for example, the meta-inference from $Rx \vdash Rx$ to $Rx \vdash \forall xRx$. (Henceforth we shall not distinguish between inferences and meta-inferences, since it is clear which ones are under discussion.) We know $Rx \vdash Rx$ by the special case of (47.2), but intuitively, we don't want $Rx \vdash \forall xRx$: from the premise that Rx is true (i.e., that some unspecified value of x exemplifies the property R), it doesn't follow that every individual exemplifies R. The proviso to GEN, of course, is unnecessary when φ is a theorem since φ is then derivable from the empty set of premises. Whenever any formula φ with free variable α is a theorem, we may invoke GEN to conclude that $\forall\alpha\varphi$ is also a theorem. For example, we shall soon prove that $\varphi \to \varphi$ is a theorem (53), so that the instance $Px \to Px$ is a theorem. From this latter it follows by GEN that $\forall x(Px \to Px)$.

Here is an example of GEN in action. The following reasoning sequence establishes that $\forall x(Qx \to Px)$ is derivable from the premise $\forall xPx$, even though strictly speaking, the sequence is not a witness to this derivability claim:

1.	$\forall xPx$	Premise
2.	$(\forall xPx) \to Px$	Instance, Axiom (29.1)
3.	Px	MP, 1,2
4.	$Px \to (Qx \to Px)$	Instance, Axiom (21.1)
5.	$Qx \to Px$	MP, 3,4
6.	$\forall xPx \vdash \forall x(Qx \to Px)$	GEN, 1–5

Lines 1–5 constitute a *bona fide* derivation of $Qx \to Px$ from $\forall xPx$: each of these lines is either an axiom, a member of Γ, or follows by MP from two previous lines. GEN, at this point, then tells us that given lines 1–5 and the fact that x doesn't appear free in the premise, there is a derivation of $\forall x(Qx \to Px)$ from the premise $\forall xPx$. This is what is asserted on line 6.

Though lines 1–6 above do not constitute a derivation of $\forall x(Qx \to Px)$ from $\forall xPx$, the justification (i.e., metatheoretic proof) of GEN given in the Appendix shows us how to convert the reasoning into a sequence of formulas that is a *bona fide* witness to the derivability claim. By studying the metatheoretic proof, it becomes clear that the above reasoning with GEN can be converted to the following derivation, in which no such appeal is made:[95]

Witness to $\forall xPx \vdash \forall x(Qx \to Px)$

1.	$\forall xPx$	Premise
2.	$\forall x(Px \to (Qx \to Px))$	Inst. Ax. (21.1)
3.	$\forall x(Px \to (Qx \to Px)) \to (\forall xPx \to \forall x(Qx \to Px))$	Inst. Ax. (29.3)
4.	$\forall xPx \to \forall x(Qx \to Px)$	MP, 2,3
5.	$\forall x(Qx \to Px)$	MP, 1,4

This sequence is a *bona fide* derivation of $\forall x(Qx \to Px)$ from $\forall xPx$, in the style of Frege and Hilbert. (In this particular example, the derivation is actually shorter by one step than the meta-derivation above that cited GEN. Most of the time, however, the meta-derivations that invoke GEN are shorter than *bona fide* derivations that don't. Of course, the reasoning with GEN already looks a bit more straightforward than the reasoning without it.) In any case, this example shows how we can use a metarule with our deductive calculus to show that universal claims are derivable.

In light of the above facts, we shall adopt the following convention for reasoning with GEN. In the Appendix, we henceforth establish a claim such as $\forall xPx \vdash \forall x(Qx \to Px)$ as follows:

> From the premise $\forall xPx$ and the instance $\forall xPx \to Px$ of axiom (29.1), it follows that Px, by MP. From this last conclusion and the instance $Px \to (Qx \to Px)$ of axiom (21.1), it follows by MP that $Qx \to Px$. Since x isn't free in our premise, it follows that $\forall x(Qx \to Px)$, by GEN. $\bowtie$

Note that this takes the liberty of treating GEN as if it were a rule of inference instead of a metarule. However, the above discussion should have made it clear just what has and has not been accomplished in this piece of reasoning. We sometimes deploy other metarules in just this way.

(**51**) **Remark:** Conventions Regarding Metarules. Although GEN was formulated to apply to $\vdash$, it also applies to $\vdash_\Box$. The following version can be proved by a trivial reworking of the justification for (50):

- If $\Gamma \vdash_\Box \varphi$ and α doesn't occur free in any formula in Γ, then $\Gamma \vdash_\Box \forall\alpha\varphi$.
- If $\vdash_\Box \varphi$, then $\vdash_\Box \forall\alpha\varphi$

[95] Lines 2 and 3 in the following derivation are indeed instances of the axiom schemata cited since we've take the closures of the instances of the schema as axioms.

In light of these results, the question arises whether we should make it explicit that GEN governs both $\vdash$ and $\vdash_\Box$, say by formulating the rule with a parenthetical box '$(\Box)$' subscript on the $\vdash$ symbol. For example, we could have formulated GEN as follows:

- If $\Gamma \vdash_{(\Box)} \varphi$ and α doesn't occur free in any formula in Γ, then $\Gamma \vdash_{(\Box)} \forall\alpha\varphi$.
- If $\vdash_{(\Box)} \varphi$, then $\vdash_{(\Box)} \forall\alpha\varphi$

Such a statement of the rule would perhaps better signpost the fact that the rule applies both to derivations and proofs generally and to modally-strict ones as well. However, this dual-purpose formulation of the rule is more difficult to read and somewhat confusing.

Consequently, we henceforth adopt the conventions:

(.1) Whenever a metarule of inference is formulated generally, so as to apply to $\vdash$, we omit the statement of the rule for the case of $\vdash_\Box$.

(.2) No metarule is to be adopted if the justification of the rule makes an appeal to the necessitation-averse axiom (30)$\star$.

Though these conventions will be discussed more fully later, the following brief remarks may be sufficient for now. As noted previously, the justifications of metarules provided in the Appendix show how to convert reasoning with the metarules into *bona fide* derivations that don't use them. A justification of a rule stated for $\vdash$ can be repurposed, with just a few obvious and trivial changes, to a justification of the analogous rule for $\vdash_\Box$, i.e., to a proof that any modally-strict reasoning using the metarule can be converted into a modally-strict derivation or proof that doesn't use the rule. As long as the justification doesn't appeal to the necessitation-averse axiom (30)$\star$ or any theorem derived from it, then any metarule of inference that applies to derivations and proofs *generally* will be a metarule of inference that also applies to modally-strict derivations and proofs.

The next rule we consider, Rule RN, contrasts with GEN because it is not a metarule that applies generally to all derivations and proofs. The antecedent of the Rule of Necessitation requires the existence of a modally-strict derivation or proof for the metarule to be applied.

(**52**) **Metarule:** Rule of Necessitation. The Rule of Necessitation (RN) is formulated in a way that prevents its application, in a derivation or proof, to any formula that depends on an a claim that isn't necessary. RN is based on a simple idea: if the derivation of a formula φ from the premise ψ is modally-strict (i.e., if φ doesn't depend on the necessitation-averse axiom (30)$\star$ in this derivation), then there is a derivation of $\Box\varphi$ from $\Box\psi$.

To formulate RN generally, however, let us introduce a definition:

- $\Box\Gamma =_{df} \{\Box\psi \mid \psi \in \Gamma\}$ (Γ any set of formulas)

So $\Box\Gamma$ is the result of prefixing a $\Box$ to every formula in Γ. Then the RN may be stated as follows:

Rule of Necessitation (**RN**)
If $\Gamma \vdash_{\Box} \varphi$, then $\Box\Gamma \vdash \Box\varphi$.

In other words, if there is a modally-strict derivation of φ from Γ, then there is a derivation of $\Box\varphi$ from $\Box\Gamma$. In the case where $\Gamma = \varnothing$, then the RN reduces to:

If $\vdash_{\Box} \varphi$, then $\vdash \Box\varphi$.

That is, if there is a modally-strict proof of φ, then there is a proof of $\Box\varphi$.

As with GEN, the justification of RN in the Appendix shows us how to turn reasoning that appeals to RN into reasoning that does not. Here is an example application of the metarule in which $\Gamma = \{Pa, Pa \to Qb\}$ and $\varphi = Qb$:

Example 1

1.	Pa	Premise
2.	$Pa \to Qb$	Premise
3.	Qb	MP, 1,2
4.	$\Box Pa, \Box(Pa \to Qb) \vdash \Box Qb$	RN, 1–3

Lines 1–3 in this example constitute a witness to $Pa, Pa \to Qb \vdash_{\Box} Qb$ since (a) Qb follows by MP from two previous members of the sequence, both of which are in Γ, and (b) the derivation of Qb from the premises doesn't involve necessitation-averse axiom (30)★. RN then asserts that from lines 1–3 it follows that $\Box Pa, \Box(Pa \to Qb) \vdash \Box Qb$, i.e., there is a derivation of $\Box Qb$ from $\Box Pa$ and $\Box(Pa \to Qb)$.

Although the reasoning in the above example doesn't qualify as a witness to the derivability claim on line 4, the justification of RN in the Appendix shows us how to convert line 4 into the following 5-element annotated sequence that does so qualify:

Witness to $\Box Pa, \Box(Pa \to Qb) \vdash \Box Qb$

1.	$\Box Pa$	Premise in $\Box\Gamma$
2.	$\Box(Pa \to Qb)$	Premise in $\Box\Gamma$
3.	$\Box(Pa \to Qb) \to (\Box Pa \to \Box Qb)$	Instance, Axiom (32.1)
4.	$\Box Pa \to \Box Qb$	MP, 2,3
5.	$\Box Qb$	MP, 1,4

This conversion works generally for any formulas φ and ψ: since there is a modally strict derivation of ψ from φ and $\varphi \to \psi$, there is a derivation of $\Box\psi$ from $\Box\varphi$ and $\Box(\varphi \to \psi)$.

Given the above discussion, it should be straightforward to see why we shall adopt the following, less formal reasoning pattern in the Appendix when presented with a case like Example 1:

> Assume $\Box Pa$ and $\Box(Pa \rightarrow Qb)$ as 'global' premises. Then note that by taking Pa and $Pa \rightarrow Qb$ as 'local' premises, it follows by MP that Qb. Since this is a modally-strict derivation of Qb from Pa and $Pa \rightarrow Qb$, it follows from our global premises that $\Box Qb$, by RN. $\bowtie$

In effect, we have reasoned by producing a 'sub-derivation' showing $Pa, Pa \rightarrow Qb \vdash Qb$, within the larger derivation of $\Box Qb$ from $\Box Pa$ and $\Box(Pa \rightarrow Qb)$.

Here is an example that uses RN to conclude that $\Box\forall xPx \vdash \Box\forall x(Qx \rightarrow Px)$ given that $\forall xPx \vdash_{\Box} \forall x(Qx \rightarrow Px)$; it involves a slight variant of the example we used to illustrate GEN.

Example 2

1.	$\forall xPx$	Premise
2.	$(\forall xPx) \rightarrow Px$	Instance, Axiom (29.1)
3.	Px	MP, 1,2
4.	$Px \rightarrow (Qx \rightarrow Px)$	Instance, Axiom (21.1)
5.	$Qx \rightarrow Px$	MP, 3,4
6.	$\forall xPx \vdash_{\Box} \forall x(Qx \rightarrow Px)$	GEN, 1–5
7.	$\Box\forall xPx \vdash \Box\forall x(Qx \rightarrow Px)$	RN, 6

Note that since lines 1–5 constitute a modally-strict derivation of $Qx \rightarrow Px$ from $\forall xPx$, we apply, on line 6, the version of GEN that governs $\vdash_{\Box}$, which was discussed in (51). So line 6 satisfies the condition for the application of RN, which then implies the conclusion on line 7. The justification of RN itself shows us how to convert line 7 into a witness for $\Box\forall xPx \vdash \Box\forall x(Qx \rightarrow Px)$:

Witness to $\Box\forall xPx \vdash \Box\forall x(Qx \rightarrow Px)$

1.	$\Box\forall xPx$	Premise
2.	$\Box[\forall x(Px \rightarrow (Qx \rightarrow Px)) \rightarrow (\forall xPx \rightarrow \forall x(Qx \rightarrow Px))]$	Inst. Ax. (29.3)
3.	$\Box[\forall x(Px \rightarrow (Qx \rightarrow Px)) \rightarrow (\forall xPx \rightarrow \forall x(Qx \rightarrow Px))] \rightarrow$ $(\Box\forall x(Px \rightarrow (Qx \rightarrow Px)) \rightarrow \Box(\forall xPx \rightarrow \forall x(Qx \rightarrow Px)))$	Inst. Ax. (32.1)
4.	$\Box\forall x(Px \rightarrow (Qx \rightarrow Px)) \rightarrow \Box(\forall xPx \rightarrow \forall x(Qx \rightarrow Px))$	MP, 2,3
5.	$\Box\forall x(Px \rightarrow (Qx \rightarrow Px))$	Inst. Ax. (21.1)
6.	$\Box(\forall xPx \rightarrow \forall x(Qx \rightarrow Px))$	MP, 4,5
7.	$\Box(\forall xPx \rightarrow \forall x(Qx \rightarrow Px)) \rightarrow (\Box\forall xPx \rightarrow \Box\forall x(Qx \rightarrow Px))$	Inst. Ax. (32.1)
8.	$\Box\forall xPx \rightarrow \Box\forall x(Qx \rightarrow Px)$	MP, 6,7
9.	$\Box\forall x(Qx \rightarrow Px)$	MP, 1,8

This is a *bona fide* derivation of the conclusion from the premise since every line is either an axiom, a premise, or follows from previous lines by MP. It should now be clear how the reasoning using GEN and RN in Example 2 is far easier to develop, or even grasp, when compared to the above reasoning. Indeed, we may compress the reasoning in Example 2 even further. The reasoning used in the Appendix for examples like this goes as follows:

> Let $\Box\forall xPx$ be our global premise. Now from the local premise $\forall xPx$ and the instance $\forall xPx \to Px$ of axiom (29.1), it follows that Px, by MP. From this last conclusion and the instance $Px \to (Qx \to Px)$ of axiom (21.1), it follows by MP that $Qx \to Px$. Since x isn't free in our local premise, it follows that $\forall x(Qx \to Px)$, by GEN. Since this constitutes a modally-strict derivation of $\forall x(Qx \to Px)$ from the local premise $\forall xPx$, it follows from our global premise, by RN, that $\Box\forall x(Qx \to Px)$. ⋈

Given the preceding discussion, this reasoning should be transparent; though it is not an actual derivation, it shows us how to construct one. Consequently, we have a way to show derivability claims without producing actual derivations.

Though RN allows us to derive and prove necessary truths, it is important to note that for some φ, we can prove $\Box\varphi$ in a way that isn't modally strict. To see this, recall that the necessitation-averse axiom for actuality (30)★ asserts $\mathcal{A}\psi \equiv \psi$. So its $\Box$-free closures are axioms, and in particular, $\mathcal{A}(\mathcal{A}\psi \equiv \psi)$ is an axiom. But $\mathcal{A}(\mathcal{A}\psi \equiv \psi) \to \Box\mathcal{A}(\mathcal{A}\psi \equiv \psi)$ is an instance of axiom (33.1). So it follows from these two axioms by MP that $\Box\mathcal{A}(\mathcal{A}\psi \equiv \psi)$. Hence, where $\varphi = \mathcal{A}(\mathcal{A}\psi \equiv \psi)$, we have established $\vdash \Box\varphi$. But our proof isn't a witness to $\vdash_\Box \Box\varphi$. This is an example that nicely demonstates that necessary truths can be established via proofs that aren't modally-strict.

9.4 The Theory of Negations and Conditionals

(**53**) **Theorems:** A Useful Fact. The following fact is derivable and is crucial to the proof of the Deduction Theorem:

$$\varphi \to \varphi$$

Although the notion of a *tautology* is a semantic notion and isn't officially defined in our formal system, we saw in Section 6.2 that the notion can be precisely defined if one takes on board the required semantic notions. It won't hurt, therefore, if we use the notion unofficially and label the above claim a tautology. Other tautologies will be derived below. As we will see, all tautologies are derivable, but it will be some time before we have assembled all the facts needed for the prove of this metatheoretic fact.

(**54**) **Metarule:** Deduction Theorem and Conditional Proof (CP). If there is a derivation of ψ from a set of premises Γ together with an additional premise φ, then there is a derivation of $\varphi \to \psi$ from Γ:

Rule CP
If $\Gamma, \varphi \vdash \psi$, then $\Gamma \vdash (\varphi \to \psi)$.

This rule is most-often used when $\Gamma = \varnothing$:

If $\varphi \vdash \psi$, then $\vdash \varphi \to \psi$.

When we cite this metarule in the proof of other metarules, we reference it as the *Deduction Theorem*. However, we shall adopt the following convention: during the course of reasoning, once we have produced a derivation of ψ from φ, we shall *infer* $\varphi \to \psi$ and cite *Conditional Proof* (CP), as opposed to concluding $\vdash \varphi \to \psi$ and citing the Deduction Theorem. The proof of the Deduction Theorem in the Appendix guarantees that we can indeed construct a proof of the conditional $\varphi \to \psi$ once we have derived ψ from φ.[96]

(**55**) **Metarules:** Corollaries to the Deduction Theorem. The following metarules are immediate consequences of the Deduction Theorem. They help us to prove the tautologies in (58) and (63). Recall that 'Γ_1, Γ_2' indicates '$\Gamma_1 \cup \Gamma_2$':

(.1) If $\Gamma_1 \vdash \varphi \to \psi$ and $\Gamma_2 \vdash \psi \to \chi$, then $\Gamma_1, \Gamma_2 \vdash \varphi \to \chi$

(.2) If $\Gamma_1 \vdash \varphi \to (\psi \to \chi)$ and $\Gamma_2 \vdash \psi$, then $\Gamma_1, \Gamma_2 \vdash \varphi \to \chi$

It is interesting that the above metarules have the following Variant forms, respectively:

(.3) $\varphi \to \psi,\ \psi \to \chi \vdash \varphi \to \chi$

(.4) $\varphi \to (\psi \to \chi),\ \psi \vdash \varphi \to \chi$

(55.3) is a Variant of (55.1) because we can derive each one from the other.[97] Similarly, (55.4) is a Variant of (55.2).

[96]Metatheorem ⟨6.7⟩, which is proved in the Appendix to Chapter 6, establishes that $\Gamma, \varphi \models \psi$ if and only if $\Gamma \models (\varphi \to \psi)$. Furthermore, Metatheorem ⟨6.8⟩, which is also proved in the Appendix to Chapter 6, establishes that $\varphi \models \psi$ if and only if $\models \varphi \to \psi$.

[97] Here is a proof. ($\hookrightarrow$) Assume (55.1), i.e., if $\Gamma_1 \vdash \varphi \to \psi$ and $\Gamma_2 \vdash \psi \to \chi$, then $\Gamma_1, \Gamma_2 \vdash \varphi \to \chi$. We want to show $\varphi \to \psi,\ \psi \to \chi \vdash \varphi \to \chi$. If we let Γ_1 be $\{\varphi \to \psi\}$, then since by (47.2) we know $\varphi \to \psi \vdash \varphi \to \psi$, we have $\Gamma_1 \vdash \varphi \to \psi$. By similar reasoning, if we let Γ_2 be $\{\psi \to \chi\}$, then we have $\Gamma_2 \vdash \psi \to \chi$. Hence, by (55.1), it follows that $\Gamma_1, \Gamma_2 \vdash \varphi \to \chi$. But, this is just $\varphi \to \psi, \psi \to \chi \vdash \varphi \to \chi$. ($\hookleftarrow$) Assume (55.3), i.e., $\varphi \to \psi,\ \psi \to \chi \vdash \varphi \to \chi$. Then by (47.7), it follows that:

$\Gamma_1, \Gamma_2, \varphi \to \psi,\ \psi \to \chi \vdash \varphi \to \chi$

From this, by two applications of the Deduction Theorem, we have:

(ϑ) $\Gamma_1, \Gamma_2 \vdash (\varphi \to \psi) \to ((\psi \to \chi) \to (\varphi \to \chi))$

Now to show (55.1), assume $\Gamma_1 \vdash \varphi \to \psi$ and $\Gamma_2 \vdash \psi \to \chi$. So by (47.7), it follows, respectively, that:

(a) $\Gamma_1, \Gamma_2 \vdash \varphi \to \psi$

(b) $\Gamma_1, \Gamma_2 \vdash \psi \to \chi$

By (a) and (ϑ), it follows by (47.5) that:

(ξ) $\Gamma_1, \Gamma_2 \vdash (\psi \to \chi) \to (\varphi \to \chi)$

From (ξ) and (b), it follows by (47.5) that $\Gamma_1, \Gamma_2 \vdash \varphi \to \chi$. ⋈

(**56**) **Derived Rules:** Hypothetical Syllogism. Note that the Variants (55.3) and (55.4) are somewhat different from the stated metarules (55.1) and (55.2): the Variants don't have the form of a conditional but instead simply assert the existence of a derivation. Of course they can be put into the traditional metarule form by conditionalizing them upon the triviality "If any condition holds" or "Under all conditions". But, given that these metarules hold without preconditions, we may transform them into *derived* rules of inference, i.e., rules of inference, like Modus Ponens, that can be used to infer formulas (as opposed to metarules, which only let us infer the existence of derivations):

- $\varphi \to \psi, \psi \to \chi \,/\, \varphi \to \chi$ [**Hypothetical Syllogism**]
- $\varphi \to (\psi \to \chi), \psi \,/\, \varphi \to \chi$

Thus, (55.3), for example, can be reconceived as a derived rule and not just a metarule. We may justifiably use this rule within derivations. The justification of (55.3) in the Appendix establishes that any derivation that yields a conclusion by an application of the above derived rule of Hypothetical Syllogism can be converted to a derivation in which this rule isn't used.

(**57**) **Remark:** Metarules vs. Derived Rules. In (55) we observed that some metarules have equivalent, variant versions that assert the existence of derivations without antecedent conditions. In (56), we observed that the variant versions could then be reconceived as *derived* rules instead of as metarules, and that one can count the justification of the metarule as a proof of the derived rule. This pattern will be repeated in this section; many of the metarules for reasoning with negation and conditionals have unconditional variants that will be regarded as derived rules. We shall formulate the metarule and its variants, and then leave the formulation of the derived rule as an obvious transformation of the variant. In the Appendix, however, we always reason with the *derived* form of the rule whenever it is available.

(**58**) **Theorems:** More Useful Tautologies. The tautologies listed below (and their proofs) follow the presentation in Mendelson 1997 (Lemma 1.11, pp. 38–40). We present them as a group because they are needed in the Appendix to this chapter to establish Lemma ⟨9.1⟩ and Metatheorem ⟨9.2⟩, i.e., that every tautology is derivable.

(.1) $\neg\neg\varphi \to \varphi$

(.2) $\varphi \to \neg\neg\varphi$

(.3) $\neg\varphi \to (\varphi \to \psi)$

(.4) $(\neg\psi \to \neg\varphi) \to (\varphi \to \psi)$

(.5) $(\varphi \to \psi) \to (\neg\psi \to \neg\varphi)$

(.6) $(\varphi \to \neg\psi) \to (\psi \to \neg\varphi)$

(.7) $(\neg\varphi \to \psi) \to (\neg\psi \to \varphi)$

(.8) $\varphi \to (\neg\psi \to \neg(\varphi \to \psi))$

(.9) $(\varphi \to \psi) \to ((\neg\varphi \to \psi) \to \psi)$

(.10) $(\varphi \to \neg\psi) \to ((\varphi \to \psi) \to \neg\varphi)$

Note that (58.5) is used to prove Modus Tollens (59).

(**59**) **Metarules/Derived Rules:** Modus Tollens. We formulate Modus Tollens (MT) as two metarules:

Rules of Modus Tollen (**MT**)
(.1) If $\Gamma_1 \vdash (\varphi \to \psi)$ and $\Gamma_2 \vdash \neg\psi$, then $\Gamma_1, \Gamma_2 \vdash \neg\varphi$
(.2) If $\Gamma_1 \vdash (\varphi \to \neg\psi)$ and $\Gamma_2 \vdash \psi$, then $\Gamma_1, \Gamma_2 \vdash \neg\varphi$

By reasoning analogous to that in footnote 97, the following are equivalent, variant versions, respectively:

$$\varphi \to \psi,\ \neg\psi \vdash \neg\varphi$$
$$\varphi \to \neg\psi,\ \psi \vdash \neg\varphi$$

In light of Remark (57), these may be transformed into well-known derived rules.

(**60**) **Metarules/Derived Rules:** Contraposition. These metarules also come in two forms:

Rules of Contraposition
(.1) $\Gamma \vdash \varphi \to \psi$ if and only if $\Gamma \vdash \neg\psi \to \neg\varphi$
(.2) $\Gamma \vdash \varphi \to \neg\psi$ if and only if $\Gamma \vdash \neg\psi \to \varphi$

If we define $\chi \dashv\vdash \theta$ ('χ is interderivable with θ') to mean $\chi \vdash \theta$ and $\theta \vdash \chi$, then the equivalent, variant versions of (.1) and (.2) are, respectively:[98]

$$\varphi \to \psi \dashv\vdash \neg\psi \to \neg\varphi$$
$$\varphi \to \neg\psi \dashv\vdash \neg\psi \to \varphi$$

[98] Although the reasoning is again analogous to that in footnote 97, we show here the left-to-right direction of (.1) is equivalent to the variant $\varphi \to \psi \vdash \neg\psi \to \neg\varphi$. ($\hookrightarrow$) Assume metarule (.1): if $\Gamma \vdash \varphi \to \psi$, then $\Gamma \vdash \neg\psi \to \neg\varphi$. Now let Γ be $\{\varphi \to \psi\}$. Then we have $\Gamma \vdash \varphi \to \psi$, by the special case of (47.2). But then it follows from our assumption that $\Gamma \vdash \neg\psi \to \neg\varphi$, i.e., $\varphi \to \psi \vdash \neg\psi \to \neg\varphi$. ($\hookleftarrow$) Assume $\varphi \to \psi \vdash \neg\psi \to \neg\varphi$. Then by (47.7), it follows that $\Gamma, \varphi \to \psi \vdash \neg\psi \to \neg\varphi$. By the Deduction Theorem, it follows that $\Gamma \vdash (\varphi \to \psi) \to (\neg\psi \to \neg\varphi)$. But from this fact, we can derive $\Gamma \vdash \neg\psi \to \neg\varphi$ from the assumption that $\Gamma \vdash \varphi \to \psi$, by (47.5). ⋈ We leave the other direction, and the proof of the equivalence of (.2) and its variant, as exercises.

By Remark (57), we henceforth use the derived rules based on these variants.

(**61**) **Metarules/Derived Rules:** Reductio Ad Absurdum. Two classic forms of Reductio Ad Absurdum (RAA) are formulated as follows:

Rules of Reductio Ad Absurdum (**RAA**)
(.1) If $\Gamma_1, \neg\varphi \vdash \neg\psi$ and $\Gamma_2, \neg\varphi \vdash \psi$, then $\Gamma_1, \Gamma_2 \vdash \varphi$
(.2) If $\Gamma_1, \varphi \vdash \neg\psi$ and $\Gamma_2, \varphi \vdash \psi$, then $\Gamma_1, \Gamma_2 \vdash \neg\varphi$

The equivalent, variant versions are, respectively:[99]

$$\neg\varphi \to \neg\psi,\ \neg\varphi \to \psi \ \vdash\ \varphi$$
$$\varphi \to \neg\psi,\ \varphi \to \psi \ \vdash\ \neg\varphi$$

We may therefore use the derived rules based on these variants, if needed.

(**62**) **Metarules/Derived Rules**. Alternative Forms of RAA. It is also useful to formulate Reductio Ad Absurdum in the following forms:

(.1) If $\Gamma,\ \varphi,\ \neg\psi \vdash \neg\varphi$, then $\Gamma, \varphi \vdash \psi$ [Variant: $\varphi,\ \neg\psi \to \neg\varphi \vdash \psi$]

(.2) If $\Gamma,\ \neg\varphi,\ \neg\psi \vdash \varphi$, then $\Gamma, \neg\varphi \vdash \psi$ [Variant: $\neg\varphi,\ \neg\psi \to \varphi \vdash \psi$]

(**63**) **Theorems:** Other Useful Tautologies. Since we not only have a standard axiomatization of negations and conditionals but also employ the standard definitions of the connectives &, $\vee$, and $\equiv$, many classical and other useful tautologies governing these connectives are derivable:

(.1) Principles of Noncontradiction:

---

[99] We can show that (61.1) is equivalent to the variant as follows. ($\hookrightarrow$) Assume (61.1), i.e., if $\Gamma_1, \neg\varphi \vdash \neg\psi$ and $\Gamma_2, \neg\varphi \vdash \psi$, then $\Gamma_1, \Gamma_2 \vdash \varphi$. Now to derive the variant, note that if we let $\Gamma_1 = \{\neg\varphi \to \neg\psi\}$ and $\Gamma_2 = \{\neg\varphi \to \psi\}$, then we know by MP both that $\Gamma_1, \neg\varphi \vdash \neg\psi$ and $\Gamma_2, \neg\varphi \vdash \psi$. Hence by our assumption, $\Gamma_1, \Gamma_2 \vdash \varphi$. ($\hookleftarrow$) Assume the variant version, i.e., $\neg\varphi \to \neg\psi,\ \neg\varphi \to \psi \ \vdash\ \varphi$. By (47.7), it follows that:

$$\Gamma_1, \Gamma_2, \neg\varphi \to \neg\psi,\ \neg\varphi \to \psi \ \vdash\ \varphi$$

So by two applications of the Deduction Theorem, we know:

(ϑ) $\Gamma_1, \Gamma_2 \vdash (\neg\varphi \to \neg\psi) \to ((\neg\varphi \to \psi) \to \varphi)$

Now to show (61.1), assume $\Gamma_1, \neg\varphi \vdash \neg\psi$ and $\Gamma_2, \neg\varphi \vdash \psi$. Then by (47.7), it follows, respectively, that $\Gamma_1, \Gamma_2, \neg\varphi \vdash \neg\psi$ and $\Gamma_1, \Gamma_2, \neg\varphi \vdash \psi$. By applying the Deduction Theorem to each of these, we obtain, respectively:

(ξ) $\Gamma_1, \Gamma_2 \vdash \neg\varphi \to \neg\psi$

(ζ) $\Gamma_1, \Gamma_2 \vdash \neg\varphi \to \psi$

But from (ϑ) and (ξ), it follows by (47.5) that:

$$\Gamma_1, \Gamma_2 \vdash (\neg\varphi \to \psi) \to \varphi$$

And from this last conclusion and (ζ) it follows again by (47.5) that $\Gamma_1, \Gamma_2 \vdash \varphi$. $\bowtie$ We leave the proof of the equivalence of (.2) and its variant as an exercise.

(.a) $\neg(\varphi \,\&\, \neg\varphi)$

(.b) $\neg(\varphi \equiv \neg\varphi)$

(.2) Principle of Excluded Middle: $\varphi \vee \neg\varphi$

(.3) Idempotent, Commutative, and Associative Laws of &, $\vee$, and $\equiv$:

(.a) $(\varphi \,\&\, \varphi) \equiv \varphi$ (Idempotency of &)

(.b) $(\varphi \,\&\, \psi) \equiv (\psi \,\&\, \varphi)$ (Commutativity of &)

(.c) $(\varphi \,\&\, (\psi \,\&\, \chi)) \equiv ((\varphi \,\&\, \psi) \,\&\, \chi)$ (Associativity of &)

(.d) $(\varphi \vee \varphi) \equiv \varphi$ (Idempotency of $\vee$)

(.e) $(\varphi \vee \psi) \equiv (\psi \vee \varphi)$ (Commutativity of $\vee$)

(.f) $(\varphi \vee (\psi \vee \chi)) \equiv ((\varphi \vee \psi) \vee \chi)$ (Associativity of $\vee$)

(.g) $(\varphi \equiv \varphi) \equiv \varphi$ (Idempotency of $\equiv$)

(.h) $(\varphi \equiv \psi) \equiv (\psi \equiv \varphi)$ (Commutativity of $\equiv$)

(.i) $(\varphi \equiv (\psi \equiv \chi)) \equiv ((\varphi \equiv \psi) \equiv \chi)$ (Associativity of $\equiv$)

(.4) Simple Biconditionals:

(.a) $\varphi \equiv \varphi$

(.b) $\varphi \equiv \neg\neg\varphi$

(.5) Conditionals and Biconditionals:

(.a) $(\varphi \to \psi) \equiv \neg(\varphi \,\&\, \neg\psi)$

(.b) $\neg(\varphi \to \psi) \equiv (\varphi \,\&\, \neg\psi)$

(.c) $(\varphi \to \psi) \to ((\psi \to \chi) \to (\varphi \to \chi))$

(.d) $(\varphi \equiv \psi) \equiv (\neg\varphi \equiv \neg\psi)$

(.e) $(\varphi \equiv \psi) \to ((\varphi \to \chi) \equiv (\psi \to \chi))$

(.f) $(\varphi \equiv \psi) \to ((\chi \to \varphi) \equiv (\chi \to \psi))$

(.g) $(\varphi \equiv \psi) \to ((\varphi \equiv \chi) \equiv (\psi \equiv \chi))$

(.h) $(\varphi \equiv \psi) \to ((\chi \equiv \varphi) \equiv (\chi \equiv \psi))$

(.i) $(\varphi \equiv \psi) \equiv ((\varphi \,\&\, \psi) \vee (\neg\varphi \,\&\, \neg\psi))$

(.j) $\neg(\varphi \equiv \psi) \equiv ((\varphi \,\&\, \neg\psi) \vee (\neg\varphi \,\&\, \psi))$

(.k) $(\varphi \to \psi) \equiv (\neg\varphi \vee \psi)$

(.6) De Morgan's Laws:

(.a) $(\varphi \,\&\, \psi) \equiv \neg(\neg\varphi \vee \neg\psi)$

(.b) $(\varphi \vee \psi) \equiv \neg(\neg\varphi \,\&\, \neg\psi)$

(.c) $\neg(\varphi \,\&\, \psi) \equiv (\neg\varphi \vee \neg\psi)$

(.d) $\neg(\varphi \vee \psi) \equiv (\neg\varphi \,\&\, \neg\psi)$

(.7) Distribution Laws:

(.a) $(\varphi \,\&\, (\psi \vee \chi)) \equiv ((\varphi \,\&\, \psi) \vee (\varphi \,\&\, \chi))$

(.b) $(\varphi \vee (\psi \,\&\, \chi)) \equiv ((\varphi \vee \psi) \,\&\, (\varphi \vee \chi))$

(.8) Exportation and Importation:

(.a) $((\varphi \,\&\, \psi) \to \chi) \to (\varphi \to (\psi \to \chi))$ (Exportation)

(.b) $(\varphi \to (\psi \to \chi)) \to ((\varphi \,\&\, \psi) \to \chi)$ (Importation)

(.9) Conjunction Simplification:

(.a) $(\varphi \,\&\, \psi) \to \varphi$

(.b) $(\varphi \,\&\, \psi) \to \psi$

(.10) Other Miscellaneous Tautologies:

(.a) $\varphi \to (\psi \to (\varphi \,\&\, \psi))$ (Adjunction)

(.b) $(\varphi \to (\psi \to \chi)) \equiv (\psi \to (\varphi \to \chi))$ (Permutation)

(.c) $(\varphi \to \psi) \to ((\varphi \to \chi) \to (\varphi \to (\psi \,\&\, \chi)))$ (Composition)

(.d) $(\varphi \to \chi) \to ((\psi \to \chi) \to ((\varphi \vee \psi) \to \chi))$

(.e) $(\varphi \to \psi) \to ((\chi \to \theta) \to ((\varphi \,\&\, \chi) \to (\psi \,\&\, \theta)))$ (Double Composition)

We leave the proof of these tautologies as exercises.

(**64**) **Metarules/Derived Rules:** The Classical Introduction and Elimination Rules. Our standard axiomatization of negation and conditionalization and standard definitions of the connectives $\&$, $\vee$, and $\equiv$ allow us to reason using all the *classical* introduction and elimination rules. However, we formulate them, in the first instance, as metarules.

Note that the metarules for the introduction and elimination of $\to$ and $\neg$ have already been presented. (47.5) is the metarule for $\to$Elimination ($\to$E) and the Deduction Theorem is the metarule for $\to$Introduction ($\to$I). Reductio Ad Absurdum, when formulated as in (61.1), is a metarule for $\neg$Elimination ($\neg$E), and when formulated as in (61.2), is a metarule for $\neg$Introduction ($\neg$I). So we formulate below the introduction and elimination metarules for $\&$, $\vee$, $\equiv$, and double negation. We also state the variant versions in each case, though we leave the proof that they are equivalent to the stated metarules for the reader. We also assume that the variant metarules can be transformed into derived rules (in which / has been substituted for $\vdash$), in the manner described in Remark (57).

(.1) &Introduction (&I):

If $\Gamma_1 \vdash \varphi$ and $\Gamma_2 \vdash \psi$, then $\Gamma_1, \Gamma_2 \vdash \varphi \,\&\, \psi$ [Variant: $\varphi, \psi \vdash \varphi \,\&\, \psi$]

(.2) &Elimination (&E):

(.a) If $\Gamma \vdash \varphi \,\&\psi$, then $\Gamma \vdash \varphi$ [Variant: $\varphi \,\&\, \psi \vdash \varphi$]

(.b) If $\Gamma \vdash \varphi \,\&\psi$, then $\Gamma \vdash \psi$ [Variant: $\varphi \,\&\, \psi \vdash \psi$]

(.3) $\vee$Introduction ($\vee$I):

(.a) If $\Gamma \vdash \varphi$, then $\Gamma \vdash \varphi \vee \psi$ [Variant: $\varphi \vdash \varphi \vee \psi$]

(.b) If $\Gamma \vdash \psi$, then $\Gamma \vdash \varphi \vee \psi$ [Variant: $\psi \vdash \varphi \vee \psi$]

(.4) $\vee$Elimination ($\vee$E):

(.a) Reasoning by Cases:
If $\Gamma_1 \vdash \varphi \vee \psi$, $\Gamma_2 \vdash \varphi \to \chi$, and $\Gamma_3 \vdash \psi \to \chi$, then $\Gamma_1, \Gamma_2, \Gamma_3 \vdash \chi$
[Variant: $\varphi \vee \psi,\ \varphi \to \chi,\ \psi \to \chi \vdash \chi$]

(.b) Disjunctive Syllogism:
If $\Gamma_1 \vdash \varphi \vee \psi$ and $\Gamma_2 \vdash \neg\varphi$, then $\Gamma_1, \Gamma_2 \vdash \psi$ [Variant: $\varphi \vee \psi,\ \neg\varphi \vdash \psi$]

(.c) Disjunctive Syllogism:
If $\Gamma_1 \vdash \varphi \vee \psi$ and $\Gamma_2 \vdash \neg\psi$, then $\Gamma_1, \Gamma_2 \vdash \varphi$ [Variant: $\varphi \vee \psi,\ \neg\psi \vdash \varphi$]

(.d) Disjunctive Syllogism:
If $\Gamma_1 \vdash \varphi \vee \psi$, $\Gamma_2 \vdash \varphi \to \chi$, and $\Gamma_3 \vdash \psi \to \theta$, then $\Gamma_1, \Gamma_2, \Gamma_3 \vdash \chi \vee \theta$
[Variant: $\varphi \vee \psi,\ \varphi \to \chi,\ \psi \to \theta \vdash \chi \vee \theta$]

(.e) Disjunctive Syllogism:
If $\Gamma_1 \vdash \varphi \vee \psi$, $\Gamma_2 \vdash \varphi \equiv \chi$, and $\Gamma_3 \vdash \psi \equiv \theta$, then $\Gamma_1, \Gamma_2, \Gamma_3 \vdash \chi \vee \theta$
[Variant: $\varphi \vee \psi,\ \varphi \equiv \chi,\ \psi \equiv \theta \vdash \chi \vee \theta$]

(.5) $\equiv$Introduction ($\equiv$I):

If $\Gamma_1 \vdash \varphi \to \psi$ and $\Gamma_2 \vdash \psi \to \varphi$, then $\Gamma_1, \Gamma_2 \vdash \varphi \equiv \psi$
[Variant: $\varphi \to \psi,\ \psi \to \varphi \vdash \varphi \equiv \psi$]

(.6) $\equiv$Elimination ($\equiv$E) (Biconditional Syllogisms):

(.a) If $\Gamma_1 \vdash \varphi \equiv \psi$ and $\Gamma_2 \vdash \varphi$, then $\Gamma_1, \Gamma_2 \vdash \psi$ [Variant: $\varphi \equiv \psi,\ \varphi \vdash \psi$]

(.b) If $\Gamma_1 \vdash \varphi \equiv \psi$ and $\Gamma_2 \vdash \psi$, then $\Gamma_1, \Gamma_2 \vdash \varphi$ [Variant: $\varphi \equiv \psi,\ \psi \vdash \varphi$]

(.c) If $\Gamma_1 \vdash \varphi \equiv \psi$ and $\Gamma_2 \vdash \neg\varphi$, then $\Gamma_1, \Gamma_2 \vdash \neg\psi$
[Variant: $\varphi \equiv \psi,\ \neg\varphi \vdash \neg\psi$]

(.d) If $\Gamma_1 \vdash \varphi \equiv \psi$ and $\Gamma_2 \vdash \neg\psi$, then $\Gamma_1, \Gamma_2 \vdash \neg\varphi$
[Variant: $\varphi \equiv \psi,\ \neg\psi \vdash \neg\varphi$]

(.e) If $\Gamma_1 \vdash \varphi \equiv \psi$ and $\Gamma_2 \vdash \psi \equiv \chi$, then $\Gamma_1, \Gamma_2 \vdash \varphi \equiv \chi$
[Variant: $\varphi \equiv \psi, \psi \equiv \chi \vdash \varphi \equiv \chi$]

(.f) If $\Gamma_1 \vdash \varphi \equiv \psi$ and $\Gamma_2 \vdash \varphi \equiv \chi$, then $\Gamma_1, \Gamma_2 \vdash \chi \equiv \psi$
[Variant: $\varphi \equiv \psi, \varphi \equiv \chi \vdash \chi \equiv \psi$]

(.7) Double Negation Introduction ($\neg\neg$I):

If $\Gamma \vdash \varphi$, then $\Gamma \vdash \neg\neg\varphi$ [Variant: $\varphi \vdash \neg\neg\varphi$]

(.8) Double Negation Elimination ($\neg\neg$E):

If $\Gamma \vdash \neg\neg\varphi$, then $\Gamma \vdash \varphi$ [Variant: $\neg\neg\varphi \vdash \varphi$]

We leave the justification of these metarules and their variants as exercises and henceforth use the corresponding derived rules within proofs and derivations.

(**65**) **Remark:** Not All Tautologies Are Yet Derivable. Rule MP and our axioms (21.1) – (21.3) for negations and conditionals are not yet sufficient for deriving all of the formulas that qualify as tautologies, as the latter notion was defined in Section 6.2. We discovered in that section that our system contains a new class of tautologies that arise in connection with 0-place relation terms of the form $[\lambda\, \varphi^*]$. Instances of the following schemata are members of this new class of tautologies: $[\lambda\, \varphi^*] \to \varphi^*$, $[\lambda\, \varphi^*] \equiv \varphi^*$, $[\lambda\, \varphi^*] \to \neg\neg\varphi^*$, etc. To derive these tautologies, we must first prove that $[\lambda\, \varphi^*] \equiv \varphi^*$ is a theorem (126.2), and to do that, we will need to show that $[\lambda\, \varphi^*] = \varphi^*$ is a theorem (126.1). The derivations of these latter theorems appeal to η-Conversion, GEN, Rule $\forall$E (a rule of quantification theory derived in item (77) below), and Rule SubId (i.e., the rule of substitution of identicals, derived in item (74.2) below). Once we've derived all of these key principles, and (126.2) in particular, we will be in a position to prove Metatheorem ⟨9.2⟩, i.e., that all tautologies are derivable. This Metatheorem is proved in the Appendix to this chapter. With such a result, we can derive Rule T, which is formulable using the semantic notions defined in Section 6.2, as a rule for our system:

Rule T
If $\Gamma \vdash \varphi_1$ and ... and $\Gamma \vdash \varphi_n$, then if $\{\varphi_1, \ldots, \varphi_n\}$ tautologically implies ψ, then $\Gamma \vdash \psi$.

Rule T asserts that ψ is derivable from Γ whenever the formulas of which it is a tautological consequence are all derivable from Γ. We won't use this rule in proving theorems, since it requires semantic notions. But it is a valid shortcut. Rule T is proved as Metatheorem ⟨9.4⟩ in the Appendix to this chapter.

9.5 The Theory of Identity

(**66**) **Theorems:** Necessarily, Every Individual and Relation Exists, and Necessarily Exists. Where α and β are both variables of the same type, it is a consequence of our axioms and rules that:

(.1) $\forall\alpha\exists\beta(\beta=\alpha)$

(.2) $\Box\exists\beta(\beta=\alpha)$

(.3) $\Box\forall\alpha\exists\beta(\beta=\alpha)$

(.4) $\forall\alpha\Box\exists\beta(\beta=\alpha)$

(.5) $\Box\forall\alpha\Box\exists\beta(\beta=\alpha)$

Note that when α and β are the individual variables x and y, respectively, then given one standard reading of the quantifiers, the above assert that: (.1) every individual exists; (.2) necessarily x exists; (.3) necessarily, every individual exists; (.4) every individual necessarily exists; and (.5) necessarily, every individual necessarily exists. We get corresponding readings when α and β are relation terms of the same arity. Clearly, on these readings, the symbol $\exists$ is being used to assert the logical existence, and not the physical existence, of the entities in question.

(**67**) **Theorems:** Identity for Properties, Relations, and Propositions is Classical. Our axioms and definitions imply that identity for properties, relations and propositions is reflexive, symmetric, and transitive.

(.1) $F^1=F^1$

(.2) $F^1=G^1 \rightarrow G^1=F^1$

(.3) $F^1=G^1 \;\&\; G^1=H^1 \rightarrow F^1=H^1$

(.4) $F^n=F^n$ $\qquad (n \geq 2)$

(.5) $F^n=G^n \rightarrow G^n=F^n$ $\qquad (n \geq 2)$

(.6) $F^n=G^n \;\&\; G^n=H^n \rightarrow F^n=H^n$ $\qquad (n \geq 2)$

(.7) $p=p$

(.8) $p=q \rightarrow q=p$

(.9) $p=q \;\&\; q=r \rightarrow p=r$

(**68**) **Metarule/Derived Rule:** Substitution of Alphabetically-Variant Relation Terms. Our principles of α-Conversion (36.1) and substitution of identicals (25) allow us to formulate a new metarule of inference: if we can derive a fact about a complex relation term τ from some premises Γ, then for any alphabetic variant τ' of τ, a corresponding fact about τ' can be derived from Γ:

> **Substitution of Alphabetically-Variant Relation Terms**[100]
> Where (a) τ is any complex n-place relation term ($n \geq 0$), (b) τ' is an alphabetic variant of τ, (c) τ and τ' are both substitutable for the n-place relation variable α in φ, and (d) φ' is the result of substituting τ' for zero or more occurrences of τ in φ^{τ}_{α}, then if $\Gamma \vdash \varphi^{\tau}_{\alpha}$, then $\Gamma \vdash \varphi'$.
> [Variant: $\varphi^{\tau}_{\alpha} \vdash \varphi'$]

As a simple example of the Variant version, consider:

$$\Box[\lambda x\, \Diamond E!x]a \vdash \Box[\lambda y\, \Diamond E!y]a$$

In this example:

$$\begin{array}{lcl}
\varphi & = & \Box Fa \\
\alpha & = & F \\
\tau & = & [\lambda x\, \Diamond E!x] \\
\tau' & = & [\lambda y\, \Diamond E!y] \\
\varphi^{\tau}_{\alpha} & = & \Box[\lambda x\, \Diamond E!x]a \\
\varphi' & = & \Box[\lambda y\, \Diamond E!y]a
\end{array}$$

In the manner of (57), we can transform the variant version into a derived rule and appeal to the latter in the course of proving theorems.

(**69**) **Theorems:** Useful Theorems About Identity$_E$ and Identity. The following are simple, but useful theorems: (.1) $x =_E y$ if and only if x exemplifies being ordinary, y exemplifies being ordinary, and x and y necessarily exemplify the same properties; (.2) whenever objects are identical$_E$, they are identical; and (.3) whenever objects are identical, then either they are both ordinary objects that necessarily exemplify the same properties or they are both abstract objects that encode the same properties:

(.1) $x =_E y \equiv (O!x \,\&\, O!y \,\&\, \Box\forall F(Fx \equiv Fy))$

(.2) $x =_E y \rightarrow x = y$

[100] ($\hookrightarrow$) To see that the stated version implies the Variant version, assume the stated version, i.e., that if $\Gamma \vdash \varphi^{\tau}_{\alpha}$, then $\Gamma \vdash \varphi'$. Then to see that the Variant version holds, note that we have as a special case of (47.4) that $\varphi^{\tau}_{\alpha} \vdash \varphi^{\tau}_{\alpha}$. But if we let $\Gamma = \{\varphi^{\tau}_{\alpha}\}$, then our stated version has the instance: if $\varphi^{\tau}_{\alpha} \vdash \varphi^{\tau}_{\alpha}$, then $\varphi^{\tau}_{\alpha} \vdash \varphi'$. Hence $\varphi^{\tau}_{\alpha} \vdash \varphi'$. ($\hookleftarrow$) Conversely, to see that the Variant version implies the stated version, assume the Variant version, i.e., $\varphi^{\tau}_{\alpha} \vdash \varphi'$. Now assume the antecedent of the stated version, i.e., $\Gamma \vdash \varphi^{\tau}_{\alpha}$. Then it follows by (47.8) that $\Gamma \vdash \varphi'$.

(.3) $x = y \equiv [(O!x \,\&\, O!y \,\&\, \Box\forall F(Fx \equiv Fy)) \vee (A!x \,\&\, A!y \,\&\, \Box\forall F(xF \equiv yF))]$

The proof of (69.1) can be shortened once we establish the strengthened principle of β-Conversion (123). But at present, we haven't established the lemmas needed for its proof. So we have developed a proof of (69.1) in strict form, since this fact is needed for the theorem that follows (70.1).

(**70**) **Theorems:** Identity for Objects is an Equivalence Condition. Our axioms and definitions imply that identity for objects is reflexive, symmetric, and transitive:

(.1) $x = x$

(.2) $x = y \rightarrow y = x$

(.3) $x = y \,\&\, y = z \rightarrow x = z$

Together with the subsitutiton of identicals, these guarantee that identity is classical.

(**71**) **Theorems:** General Identity is an Equivalence Condition. We have now established that identity for relations (67) and for objects (70) is an equivalence condition. We may represent each group of facts in terms of single schemata, where α, β, γ are any three distinct variables of the same type:

(.1) $\alpha = \alpha$

(.2) $\alpha = \beta \rightarrow \beta = \alpha$

(.3) $\alpha = \beta \,\&\, \beta = \gamma \rightarrow \alpha = \gamma$

Though these claims would remain true even if the variables aren't all distinct, they wouldn't express the idea that identity is reflexive, symmetric and transitive, respectively.

(**72**) **Theorems:** Self-Identity and Necessity. It is a consequence of the foregoing that (.1) necessarily everything is self-identical, and that (.2) everything is necessarily self-identical. Where α is any variable:

(.1) $\Box\forall\alpha(\alpha = \alpha)$

(.2) $\forall\alpha\Box(\alpha = \alpha)$

These well-known principles of self-identity and necessity are thus provable.

(**73**) **Theorems:** Term Identities Imply Logical Propriety. For any terms τ and τ', if $\tau = \tau'$, then both τ and τ' are logically proper:

(.1) $\tau = \tau' \rightarrow \exists\beta(\beta = \tau)$, provided β isn't free in τ

(.2) $\tau = \tau' \to \exists\beta(\beta = \tau')$, provided β isn't free in τ'

Note that this theorem holds even when τ or τ' is a definite description. If they appear in a true identity statement, then they are logically proper.

(**74**) **Metarules/Derived Rules:** Rules for Reasoning with Identity. We now formulate two metarules: one for the reflexivity of identity and one for the substitution of identicals:

(.1) **Rule for the Reflexivity of Identity** (**Rule ReflId**)
$\vdash \tau = \tau$, where τ is any term other than a description.

(.2) **Rule of Substitution for Identicals** (**Rule SubId**)
If $\Gamma_1 \vdash \varphi^\tau_\alpha$ and $\Gamma_2 \vdash \tau = \tau'$, then $\Gamma_1, \Gamma_2 \vdash \varphi'$, whenever τ and τ' are any terms substitutable for α in φ, and φ' is the result of replacing zero or more occurrences of τ in φ^τ_α with occurrences of τ'. [Variant: φ^τ_α, $\tau = \tau' \vdash \varphi'$]

Since Rule ReflId requires no special conditions of application, we may regard it as a derived rule that asserts: $\tau = \tau$, for any term τ other than a descriptionm, is a theorem. The Variant of Rule SubId can also be converted to a derived rule: φ' follows from φ^τ_α and $\tau = \tau'$, where φ' is the result of replacing zero or more occurrences of τ in φ^τ_α with occurrences of τ'.

Note that in Rule SubId, τ and τ' might both be definite descriptions: as long as $\tau = \tau'$ is an assumption, one may substitute τ' for τ in φ even when one or both of τ, τ' is a description. Under the assumption that $\tau = \tau'$, both terms are logically proper, as theorems (73.1) and (73.2) establish, and so one may freely substitute τ and τ' for one another in that context.

Note also that when φ' is the result of replacing *all* of the occurrences of τ in φ^τ_α by τ', then φ' just is $\varphi^{\tau'}_\alpha$ and we have the following special case of Rule SubId:

(.3) **Rule SubId Special Case**
If $\Gamma_1 \vdash \varphi^\tau_\alpha$ and $\Gamma_2 \vdash \tau = \tau'$, then $\Gamma_1, \Gamma_2 \vdash \varphi^{\tau'}_\alpha$ [Variant: $\varphi^\tau_\alpha, \tau = \tau' \vdash \varphi^{\tau'}_\alpha$]

(**75**) **Theorems:** Identity and Necessity. Where α, β are both variables of the same type, we can establish that α and β are identical if and only if it is necessary that they are identical:

$$\alpha = \beta \equiv \Box\alpha = \beta$$

The left-to-right direction of this theorem is the famous *necessity of identity* principle (Kripke 1971); it applies not only to individuals but also to relations. We've already seen that the *definitions* of object identity (15) and relation identity (16) ground the reflexivity of identity (71.1). The reflexivity of identity is one of the key facts used in the proof of the necessity of identity principle.

Thus, our proof not only validates and extends Kripke (1971), but the result is derived within a theory in which identity is not a primitive.

The right-to-left direction of (75) follows by the T schema (32.2). Thus, we have modally strict proofs of both directions of (75).

(**76**) **Theorems:** Identity, Necessity, and Descriptions. It is an interesting fact that the necessity of identity holds even when objects are described:

$$\imath x\varphi = \imath y\psi \equiv \Box \imath x\varphi = \imath y\psi$$

Notice that the theorem is not restricted to logically proper descriptions. To prove each direction of the biconditional, we need only consider the case where the antecedent of that direction is true. However, in the case where the descriptions fail to be logically proper, both sides of the biconditional are false. Thus, the necessity of identity principle and its converse, which are combined in (75), apply to every pair of terms of the same type.

9.6 The Theory of Quantification

(**77**) **Metarules/Derived Rules:** $\forall$Elimination ($\forall$E). The elimination rule for the universal quantifier has two forms (with the first being the primary form): (.1) legitimizes the instantiation of any term τ with the same type as the quantified variable provided τ is logically proper, while (.2) states that every term of the same type as the quantified variable, other than a description, is instantiable:

Rule $\forall$E
(.1) If $\Gamma_1 \vdash \forall\alpha\varphi$ and $\Gamma_2 \vdash \exists\beta(\beta = \tau)$, then $\Gamma_1, \Gamma_2 \vdash \varphi^{\tau}_{\alpha}$, provided τ is substitutable for α in φ [Variant: $\forall\alpha\varphi, \exists\beta(\beta = \tau) \vdash \varphi^{\tau}_{\alpha}$]

(.2) If $\Gamma \vdash \forall\alpha\varphi$, then $\Gamma \vdash \varphi^{\tau}_{\alpha}$, provided τ is substitutable for α in φ and τ is not a description [Variant: $\forall\alpha\varphi \vdash \varphi^{\tau}_{\alpha}$]

In the usual manner, we may convert the variants into the derived rules of $\forall$E and use them to produce genuine derivations.

Rule (77.2) and its Variant have special cases when τ is the variable α. Since α is not a description and is substitutable for itself in any formula φ with the result that $\varphi^{\alpha}_{\alpha} = \varphi$, the following special cases obtain:

- If $\Gamma \vdash \forall\alpha\varphi$, then $\Gamma \vdash \varphi$
- $\forall\alpha\varphi \vdash \varphi$

(**78**) **Remark:** A Misuse of Rule $\forall$E. Note that the following attempt to derive a contradiction involves a misuse of Rule $\forall$E. Suppose we let φ be the formula $\neg yF$ and formulate the following instance of Object Comprehension:

$$\exists x(A!x \,\&\, \forall F(xF \equiv \neg yF))$$

So by GEN, we may derive as a theorem:

(ϑ) $\forall y \exists x(A!x \,\&\, \forall F(xF \equiv \neg yF))$

Since the variable x is not a description, one may now be tempted to instantiate the quantifier $\forall y$ in (ϑ) to x, by applying Rule $\forall$E Variant (77.2), to obtain:

(ξ) $\exists x(A!x \,\&\, \forall F(xF \equiv \neg xF))$

This is easily shown to be a contradiction, for assume a is an arbitrary such object, so that we know $A!a \,\&\, \forall F(aF \equiv \neg aF)$. By &E, it follows that $\forall F(aF \equiv \neg aF)$. Now for any property you pick, say P, it follows that $aP \equiv \neg aP$, which by (63.1.b) is a contradiction.

What prevents such reasoning within our system is the fact that Rule $\forall$E Variant (77.2) has been incorrectly applied in the move from (ϑ) to (ξ). The rule states: $\forall \alpha \psi \vdash \psi^{\tau}_{\alpha}$, provided τ is substitutable for α in ψ and τ is not a description. In the present case, α is y, τ is x, (ϑ) has the form $\forall y \psi$, where $\psi = \exists x(A!x \,\&\, \forall F(xF \equiv \neg yF))$, and ($\xi$) has the form ψ^{x}_{y}. The move from (ϑ) to (ξ) obeys the condition that x not be a description, but it doesn't obey the condition that x be substitutable for y in ψ. The definition of *substitutable for* in (24) requires that for x to be substitutable for y in ψ, every variable that occurs free in the term x must remain free after we substitute x for y in ψ. But x is itself a variable that is free in the term x, yet it doesn't remain free when x is substituted for y in ψ (i.e., to produce ψ^{x}_{y}). Instead, x is captured by (i.e., falls within the scope of) the existential quantifier $\exists x$ in ψ^{x}_{y}. So the above reasoning fails to correctly apply (77.2) in the move from (ϑ) to (ξ).

(**79**) **Theorems:** Classical Quantifier Axioms as Theorems. Our principles yield two classical quantifier axioms as theorems:

(.1) $\forall \alpha \varphi \to \varphi^{\tau}_{\alpha}$, provided τ is substitutable for α in φ and is not a description

(.2) $\forall \alpha(\varphi \to \psi) \to (\varphi \to \forall \alpha \psi)$, provided α is not free in φ

These are used in Mendelson 1997, for example, as the two principal axioms of classical quantification theory (his formulation of predicate logic with function terms assumes that all terms have a denotation). Also, given what was said above about $\varphi^{\alpha}_{\alpha}$, formulas of the form:

- $\forall \alpha \varphi \to \varphi$

are special cases of theorem (79.1).

(**80**) **Metarule:** Generalization or Universal Introduction ($\forall$I) on Constants. We introduce and explain the $\forall$Introduction Rule for Constants by way of analogy.

One simple way to establish $\forall F \forall x(Fx \equiv Fx)$ is to argue that it follows by two applications of GEN from $Fx \equiv Fx$, where the latter is a theorem by being an instance of the tautology $\varphi \equiv \varphi$ (63.4.a). However, we often argue as follows: assume P is an arbitrary property and a an arbitrary object; then $Pa \equiv Pa$ is an instance of our tautology; since P and a are arbitrary, the claim holds for all properties and objects; hence, $\forall F \forall x(Fx \equiv Fx)$. In this reasoning, universal generalization is permissible in the last step because in the course of the reasoning, we haven't invoked any special assumptions about the arbitrarily chosen entities; no special facts about the property P or the individual a played a part in our conclusion that $Pa \equiv Pa$.

To formulate Rule $\forall$I generally, we first introduce some notation. Where τ is any constant and α any variable of the same type as τ:

- φ^{α}_{τ} is the result of replacing every occurrence of the constant τ in φ by an occurrence of α

We then have:

> **Rule $\forall$I**
> If $\Gamma \vdash \varphi$ and τ is a constant that does not occur in Γ or $\mathbf{\Lambda}$, then $\Gamma \vdash \forall \alpha \varphi^{\alpha}_{\tau}$, provided α is a variable that does not occur in φ.

Note that our axioms in $\mathbf{\Lambda}$ count as special facts about the named individuals or relations (such as $E!$) that might appear in them, and so the names of those individuals and relations are not acceptable instances of τ.

Here is an example of how we will use this rule. Consider the following reasoning that shows $\forall x(Px \rightarrow Qx), \forall y Py \vdash \forall x Qx$:

1.	$\forall x(Px \rightarrow Qx)$	Premise
2.	$\forall y Py$	Premise
3.	$Pa \rightarrow Qa$	$\forall$E, 1
4.	Pa	$\forall$E, 2
5.	Qa	MP, 3, 4
6.	$\forall x(Px \rightarrow Qx), \forall y Py \vdash \forall x Qx$	$\forall$I, 1–5

In this example, we set $\Gamma = \{\forall x(Px \rightarrow Qx), \forall y Py\}$, $\varphi = Qa$, and $\tau = a$. Given that $\forall$E is a derived rule and not just a metarule, lines 1–5 constitute a genuine derivation that is a witness to $\Gamma \vdash \varphi$. Since a doesn't occur in Γ and x doesn't occur in φ, we have an instance of the Rule $\forall$I in which α is the variable x, which we can then apply to lines 1–5 to infer the derivability claim on line 6.

Given that $\forall$I is a metarule, we could have reached the conclusion on line 6 without it, by adopting the following reasoning, which doesn't use the constant a:

1.	$\forall x(Px \to Qx)$	Premise
2.	$\forall yPy$	Premise
3.	$Px \to Qx$	$\forall$E, 1
4.	Px	$\forall$E, 2
5.	Qx	MP, 3, 4
6.	$\forall x(Px \to Qx), \forall yPy \vdash \forall xQx$	GEN, 1–5

The application of GEN on line 6 is legitimate since we have legitimately derived Qx on line 5 from the premises $\forall x(Px \to Qx)$ and $\forall yPy$ and the variable x doesn't occur free in the premises. Of course, GEN itself is a metarule, but we already know how to eliminate it.

(**81**) **Lemmas:** Re-replacement Lemmas. In the following re-replacement lemmas, we assume that α, β, and τ are all of the same type:

(.1) If β is substitutable for α in φ and β doesn't occur free in φ, then α is substitutable for β in φ^{β}_{α} and $(\varphi^{\beta}_{\alpha})^{\alpha}_{\beta} = \varphi$.

(.2) If τ is a constant symbol that doesn't occur in φ, then $(\varphi^{\tau}_{\alpha})^{\beta}_{\tau} = \varphi^{\beta}_{\alpha}$.

(.3) If β is substitutable for α in φ and doesn't occur free in φ, and τ is any term substitutable for α in φ, then $(\varphi^{\beta}_{\alpha})^{\tau}_{\beta} = \varphi^{\tau}_{\alpha}$.

It may help to read the following Remark before attempting to prove the above.

(**82**) **Remark:** Explanation of the Re-replacement Lemmas. By discussing (81.1) in some detail, (81.2) and (81.3) become more transparent and less in need of commentary. It is relatively easy to show that, in general, in the absence of any preconditions, $(\varphi^{\beta}_{\alpha})^{\alpha}_{\beta} \neq \varphi$. The variable α may occur in $(\varphi^{\beta}_{\alpha})^{\alpha}_{\beta}$ at a place where it does not occur in φ, and α may occur in φ at a place where it does not occur in $(\varphi^{\beta}_{\alpha})^{\alpha}_{\beta}$. Here is an example of each case:

- $\varphi = Ryx$. Then $\varphi^{y}_{x} = Ryy$ and though x is substitutable for y in φ^{y}_{x} (y doesn't fall under the scope of any variable-binding operator that binds x), $(\varphi^{y}_{x})^{x}_{y} = Rxx$. Hence $(\varphi^{y}_{x})^{x}_{y} \neq \varphi$. In this example, x occurs at a place in $(\varphi^{y}_{x})^{x}_{y}$ where it does not occur in φ.

- $\varphi = \forall yRxy$. Then $\varphi^{y}_{x} = \forall yRyy$. Since x is trivially substitutable for y in φ^{y}_{x} (there are no free occurrences of y in φ^{y}_{x}), $(\varphi^{y}_{x})^{x}_{y} = \varphi^{y}_{x} = \forall yRyy$. By inspection, then, $(\varphi^{y}_{x})^{x}_{y} \neq \varphi$. In this example, x occurs at a place in φ where it does not occur in $(\varphi^{y}_{x})^{x}_{y}$.

These two examples nicely demonstrate why the two antecedents of (81.1) are crucial. The first example fails the proviso that y not occur free in Ryx; the second example fails the proviso that y be substitutable for x in $\forall yRxy$. But here is an example of (81.1) in which the antecedents obtain:

- $\varphi = \forall yPy \to Qx$. In this example, the free occurrence of x is not within the scope of the quantifier $\forall y$. So y is substitutable for x in φ and y does not occur free in φ. Thus, $\varphi^y_x = \forall yPy \to Qy$, and since y has a free occurrence in φ^y_x not under the scope of a variable-binding operator binding x, x is substitutable for y in φ^y_x. Hence $(\varphi^y_x)^x_y = \forall yPy \to Qx$, and so $(\varphi^y_x)^x_y = \varphi$.

These remarks and the proof of (81.1) should suffice to clarify the remaining two replacement lemmas. (81.1) is used to prove (83.12), (86.10), and the Rule of Alphabetic Variants (111). Lemma (81.3) is used in the proof of (86.8).

(**83**) **Theorems:** Basic Theorems of Quantification Theory. The following are all basic consequences of our quantifier axioms and (derived) rules:

(.1) $\forall\alpha\forall\beta\varphi \equiv \forall\beta\forall\alpha\varphi$

(.2) $\forall\alpha(\varphi \equiv \psi) \equiv (\forall\alpha(\varphi \to \psi) \,\&\, \forall\alpha(\psi \to \varphi))$

(.3) $\forall\alpha(\varphi \equiv \psi) \to (\forall\alpha\varphi \equiv \forall\alpha\psi)$

(.4) $\forall\alpha(\varphi \,\&\, \psi) \equiv (\forall\alpha\varphi \,\&\, \forall\alpha\psi)$

(.5) $\forall\alpha_1 \ldots \forall\alpha_n\varphi \to \varphi$

(.6) $\forall\alpha\forall\alpha\varphi \equiv \forall\alpha\varphi$

(.7) $(\varphi \to \forall\alpha\psi) \equiv \forall\alpha(\varphi \to \psi)$, provided α is not free in φ

(.8) $(\forall\alpha\varphi \vee \forall\alpha\psi) \to \forall\alpha(\varphi \vee \psi)$

(.9) $(\forall\alpha(\varphi \to \psi) \,\&\, \forall\alpha(\psi \to \chi)) \to \forall\alpha(\varphi \to \chi)$

(.10) $(\forall\alpha(\varphi \equiv \psi) \,\&\, \forall\alpha(\psi \equiv \chi)) \to \forall\alpha(\varphi \equiv \chi)$

(.11) $\forall\alpha(\varphi \equiv \psi) \equiv \forall\alpha(\psi \equiv \varphi)$

(.12) $\forall\alpha\varphi \equiv \forall\beta\varphi^\beta_\alpha$, provided β is substitutable for α in φ and doesn't occur free in φ

The two provisos on (83.12) can be explained by referencing and adapting the examples used in Remark (82) that helped us to understand the antecedent of the Re-replacement Lemma (81.1):

- In the formula $\varphi = Rxy$, y is substitutable for x but also occurs free. Without the second proviso in (83.12), we could set α to x and β to y and obtain the instance: $\forall xRxy \equiv \forall yRyy$. Clearly, this is not valid: the left side asserts that everything bears R to y while the right asserts that everything bears R to itself.

- In the formula $\varphi = \forall yRxy$, y is not substitutable for x even though it does not occur free. Without the first proviso in (83.12), we could set α to x and β to y and obtain the instance: $\forall x\forall yRxy \equiv \forall y\forall yRyy$. Again, clearly, this is not valid: the left side is true when everything bears R to everything while the right side, which by (83.6) is equivalent to $\forall yRyy$, is true only when everything bears R to itself.

(83.12) is a special case of the interderivability of alphabetic variants; indeed, it is a special case of a special case. The interderivability of alphabetically-variant universal generalizations is a special case of the interderivability of alphabetically-variant formulas of arbitrary complexity. But within that special case, there are two basic ways in which a universal generalization of the form $\forall\alpha\varphi$ can have an alphabetic variant. (83.12) concerns one of those ways, namely, alphabetic variants of the form $\forall\beta\varphi^{\beta}_{\alpha}$. But it follows from Metatheorem $\langle 8.3\rangle$(e) (in Chapter 8) that $\forall\alpha\varphi$ can also have alphabetic variants of the form $\forall\alpha\varphi'$, where φ' is an alphabetic variant of φ. We aren't yet in a position to prove the interderivability of the latter, much less prove the interderivability of alphabetically-variant formulas of *arbitrary* complexity. The case proved above tells us only that whenever we have established a theorem of the form $\forall\alpha\varphi$, we may infer any formula with the same exact form but which differs only by the choice of the variable bound by the leftmost universal quantifier, provided the choice of new variable β is a safe one, i.e., one that will preserve the meaning of the original formula when the substitution is carried out.

(**84**) **Metarules/Derived Rules:** $\exists$Introduction ($\exists$I). The metarules of $\exists$Introduction allow us to infer the existence of derivations of existential generalizations, though their variant forms yield derived rules that let us existentially generalize, within a derivation, on any term τ that is logically proper. In that regard, it is perfectly standard. But $\exists$I rules have to be formulated carefully. There is a form that applies to any term whatsoever and a restricted form that applies to any term other than a description:

Rule $\exists$I

(.1) If $\Gamma_1 \vdash \varphi$ and $\Gamma_2 \vdash \exists\beta(\beta = \tau)$, then $\Gamma_1, \Gamma_2 \vdash \exists\alpha\varphi'$, whenever α is a variable of the same type as τ and φ' is obtained from φ by substituting α for zero or more occurrences of τ, provided both (1) when τ is a variable, all of the replaced occurrences of τ in φ are free occurrences, and (2) all of the substituted occurrences of α are free in φ'.
[Variant: $\varphi, \exists\beta(\beta = \tau) \vdash \exists\alpha\varphi'$]

(.2) If $\Gamma \vdash \varphi$, then $\Gamma \vdash \exists\alpha\varphi'$, whenever φ' is obtained from φ by substituting the variable α for zero or more occurrences of some term τ of the same type as α, provided (1) τ is not a description, (2) when τ is a variable, all of the replaced occurrences of τ are free in φ, and (3) all of the substituted occurrences of α are free in φ'.

[Variant: $\varphi \vdash \exists\alpha\varphi'$]

The simplest two examples of the variant version of (.2) are $Gy \vdash \exists xGx$ and $Gy \vdash \exists F(Fy)$. In the first case, φ is Gy, φ' is Gx, τ is y, and α is x. Condition (1) in the rule is met because y is not a description; condition (2) is met because all of the replaced occurrences of y are free in Gy; and condition (3) is met because all of the substituted occurrences of x are free in Gx.

Note also that the inference from $P\imath xQx$ and $\exists y(y = \imath xQx)$ to $\exists xPx$ is justified by the variant version of (84.1), whereas the inference from $Pa \,\&\, Pb$ to $\exists F(Fa \,\&\, Pb)$ is justified by (84.2). Since this rule is covered in detail in basic courses on predicate logic, we omit both the formulation of more complex examples and further explanation of the conditions that must be satisfied for the rule to be applied.

(**85**) **Metarule:** $\exists$Elimination ($\exists$E) on Constants. If we have asserted $\exists\alpha\varphi$ as a theorem or premise, we often continue reasoning by saying "Assume τ is an arbitrary such φ, so that we know φ^{τ}_{α}," where τ is a 'fresh' constant that hasn't previously appeared in the context of reasoning or even in our axioms. If we then validly reason our way to ψ from (some premises and) φ^{τ}_{α} without making any special assumptions about τ other than φ^{τ}_{α}, then $\exists$E allows us to discharge our assumption about τ and validly conclude that we can derive ψ from (the premises we used and) $\exists\alpha\varphi$:

Rule $\exists$E
If $\Gamma, \varphi^{\tau}_{\alpha} \vdash \psi$, then $\Gamma, \exists\alpha\varphi \vdash \psi$, provided τ is a constant that does not occur in φ, ψ, Γ, or $\boldsymbol{\Lambda}$.

In other words, if there is a derivation of ψ from $\Gamma \cup \{\varphi^{\tau}_{\alpha}\}$, where Γ, φ and ψ make no special assumptions about τ (i.e., τ is arbitrary with respect to Γ, $\boldsymbol{\Lambda}$, φ and ψ), then there is a derivation of ψ from $\Gamma \cup \{\exists\alpha\varphi\}$.

(**86**) **Theorems:** Further Theorems of Quantification Theory. The foregoing series of rules for quantification theory facilitate the derivation of many of the following theorems:

(.1) $\forall\alpha\varphi \rightarrow \exists\alpha\varphi$

(.2) $\neg\forall\alpha\varphi \equiv \exists\alpha\neg\varphi$

(.3) $\forall\alpha\varphi \equiv \neg\exists\alpha\neg\varphi$

(.4) $\neg\exists\alpha\varphi \equiv \forall\alpha\neg\varphi$

(.5) $\exists\alpha(\varphi \,\&\, \psi) \rightarrow (\exists\alpha\varphi \,\&\, \exists\alpha\psi)$

(.6) $\exists\alpha(\varphi \vee \psi) \equiv (\exists\alpha\varphi \vee \exists\alpha\psi)$

(.7) $\exists\alpha\varphi \equiv \exists\beta\varphi^{\beta}_{\alpha}$, provided β is substitutable for α in φ and doesn't occur free in φ

(.8) $\varphi \equiv \exists\beta(\beta = \alpha \;\&\; \varphi^{\beta}_{\alpha})$, provided β is substitutable for α in φ and doesn't occur free in φ.

(.9) $\varphi^{\tau}_{\alpha} \equiv \exists\alpha(\alpha = \tau \;\&\; \varphi)$, provided τ is any term other than a description and is substitutable for α in φ.

(.10) $(\varphi \,\&\, \forall\beta(\varphi^{\beta}_{\alpha} \rightarrow \beta = \alpha)) \equiv \forall\beta(\varphi^{\beta}_{\alpha} \equiv \beta = \alpha)$,
provided α, β are distinct variables of the same type, and β is substitutable for α in φ and doesn't occur free in φ.

(.11) $(\forall\alpha\varphi \,\&\, \forall\alpha\psi) \rightarrow \forall\alpha(\varphi \equiv \psi)$

(.12) $(\neg\exists\alpha\varphi \,\&\, \neg\exists\alpha\psi) \rightarrow \forall\alpha(\varphi \equiv \psi)$

(.13) $(\exists\alpha\varphi \,\&\, \neg\exists\alpha\psi) \rightarrow \neg\forall\alpha(\varphi \equiv \psi)$

A simple example of (86.8) is $Px \equiv \exists y(y = x \;\&\; Py)$, and a simple example (86.9) is $Qa \equiv \exists x(x = a \;\&\; Qx)$. But these theorems also apply to relation terms; as simple examples we have $Fa \equiv \exists G(G = F \,\&\, Ga)$ and $Pa \equiv \exists F(F = P \,\&\, Fa)$, respectively. The reader should produce examples in which φ has greater complexity. Note that (86.9) is restricted to terms τ other than descriptions. An example that illustrates why this is necessary comes readily to hand. Let φ be $Px \rightarrow Px$, and let τ be $\imath yQy$, so that $\varphi^{\imath yQy}_{x}$ is $P\imath yQy \rightarrow P\imath yQy$. Then $\varphi^{\imath yQy}_{x}$ is a tautology and so true no matter whether $\imath yQy$ is logically proper or not. But in the case where $\imath yQy$ is not logically proper, the following would be an invalid instance of (86.9): $(P\imath yQy \rightarrow P\imath yQy) \equiv \exists x(x = \imath yQy \,\&\, (Px \rightarrow Px))$. When nothing is the unique Q object, then it is not the case that something x is both the unique Q object and such that if Px then Px.

Theorem (86.10) is noteworthy because the two sides of the main biconditional are equivalent ways of asserting an important claim. When α is an individual variable, both sides of the biconditional assert the claim that *α is a unique individual such that φ*, and when α is a relation variable, they both represent the claim that *α is a unique relation such that φ*. (In the formal mode, we would say, in both cases, that α uniquely satisfies φ.) The left condition comes to us from Russell's (1905) classic analysis of uniqueness, whereas the right condition is a slightly more efficient way of expressing the uniqueness claim.

(**87**) **Definitions:** Unique Existence Quantifier. As a consequence of the preceding observation, we introduce, in the usual way, a special quantifier $\exists!$ to conveniently assert that there exists a unique α such that φ. Where β is substitutable for α in φ and doesn't occur free in φ, we stipulate that:

(.1) $\exists!\alpha\varphi =_{df} \exists\alpha(\varphi \,\&\, \forall\beta(\varphi^{\beta}_{\alpha} \rightarrow \beta = \alpha))$

It is important not to confuse the defined unique-existence quantifier '$\exists!$' with the simple predicate '$E!$' in what follows. Moreover, in light of (86.10), the above definition is equivalent to:

(.2) $\exists!\alpha\varphi =_{df} \exists\alpha\forall\beta(\varphi^{\beta}_{\alpha} \equiv \beta = \alpha)$

That is, there exists a unique α such that φ if and only if there exists an α such that all and only the entities which are such that φ are identical to α.

9.7 The Theory of Actuality and Descriptions

Although the theorems in this section sometimes involve the necessity operator, no special principles for necessity other than the axioms and rule (RN) introduced thus far are required to prove the basic theorems and metarules governing actuality.

(**88**) **Metarule:** Rule of Actualization (RA). We first define:

- $\mathcal{A}\Gamma = \{\mathcal{A}\psi \mid \psi \in \Gamma\}$ (Γ any set of formulas)

Thus, $\mathcal{A}\Gamma$ is the result of adding the actuality operator to the front of every formula in Γ. We then have the following metarule:

Rule of Actualization (RA)
If $\Gamma \vdash \varphi$, then $\mathcal{A}\Gamma \vdash \mathcal{A}\varphi$

We most often use this rule in the form in which Γ is empty:

- If $\vdash \varphi$, then $\vdash \mathcal{A}\varphi$

In other words, whenever φ is a theorem, so is $\mathcal{A}\varphi$. Given this rule, whenever φ is a theorem, so is $\Box\mathcal{A}\varphi$, since RA yields that $\mathcal{A}\varphi$ is a theorem and axiom (33.1), $\mathcal{A}\varphi \rightarrow \Box\mathcal{A}\varphi$, allows us to derive $\Box\mathcal{A}\varphi$. Indeed, this holds even if φ is a $\star$-theorem. We'll see an example of this shortly: theorem (90.1)$\star$ asserts $\neg\mathcal{A}\varphi \equiv \neg\varphi$. From this $\star$-theorem, it follows by RA that $\mathcal{A}(\neg\mathcal{A}\varphi \equiv \neg\varphi)$. So by (33.1), it follows that $\Box\mathcal{A}(\neg\mathcal{A}\varphi \equiv \neg\varphi)$. But, of course, this is a $\star$-theorem. So here we have another example of a necessary truth the proof of which is not modally-strict. We saw other examples of this phenomenon earlier, in the discussion following the presentation of RN, in (52).

(**89**) **Remark:** Discussion of RA. It is important to recognize why RA is formulated as above, as opposed to the following alternative:

If $\Gamma \vdash \varphi$, then $\Gamma \vdash \mathcal{A}\varphi$

Here, the consequent of the rule doesn't require $\mathcal{A}\Gamma$ but only Γ. One can prove that this version of the rule is semantically valid.[101] However, the justification of this rule appeals to the necessitation-averse axiom of actuality.[102] Clearly, the use of this alternative rule would undermine modally-strict derivations. For example, given that $\varphi \vdash \varphi$ is an instance of derived rule (47.4), this alternative version of RA would allow us to conclude $\varphi \vdash \mathcal{A}\varphi$, which by the Deduction Theorem (54) yields $\varphi \to \mathcal{A}\varphi$ as a theorem. But we certainly don't want this to be a modally-strict theorem—we know that its necessitation, $\Box(\varphi \to \mathcal{A}\varphi)$, fails to be valid. By formulating the consequent of RA with $\mathcal{A}\Gamma$, we forestall such a derivation. All that follows from $\varphi \vdash \varphi$ via RA, as officially formulated in (88), is that $\mathcal{A}\varphi \vdash \mathcal{A}\varphi$, which by the Deduction Theorem, yields only $\vdash \mathcal{A}\varphi \to \mathcal{A}\varphi$. Moreover, this is a modally-strict proof and we may happily apply RN to derive a valid necessary truth.

Call those rules whose justifications require the necessitation-averse axiom (30)$\star$ (like the alternative to RA formulated above) *non-strict rules*. If we were to use non-strict rules, we would have to tag any theorem proved by means of such a rule a $\star$-theorem and, indeed, tag the rule itself as a $\star$-rule. This explains why we adopted the first convention described in Remark (51), namely, that we avoid metarules whose justification makes an appeal to the necessitation-averse axiom of actuality. With such a convention in place, we don't have to worry about redefining modal-strictness and *dependence* to ensure that derivations that depend on a necessitation-averse axiom or on a non-strict rule fail to be modally-strict.

Note that there are other valid but non-strict metarules that we shall eschew because they violate our convention. Consider, for example, the following rule:

If $\Gamma \vdash \mathcal{A}\varphi$, then $\Gamma \vdash \varphi$.

This rule can be justified from the basis we now have.[103] Again, however, the argument justifying the rule makes an appeal to the necessitation-averse axiom (30)$\star$. Since this argument shows us how to turn a proof using the metarule into a proof that doesn't use that rule, it becomes apparent that any proof that

---

[101] It is provable that if $\Gamma \models \varphi$, then $\Gamma \models \mathcal{A}\varphi$. Intuitively, if $\Gamma \models \varphi$, i.e., if φ is true at the distinguished world in every interpretation in which all the formulas in Γ are true at the distinguished world, then it follows that in every interpretation in which all the formulas in Γ are true at the distinguished world, $\mathcal{A}\varphi$ is true at the distinguished world, i.e., it follows that $\Gamma \models \mathcal{A}\varphi$.

[102] To see this, assume the antecedent, i.e., $\Gamma \vdash \varphi$. Now since $\varphi \to \mathcal{A}\varphi$ is a simple consequence of the necessitation-averse axiom of actuality (30)$\star$, $\mathcal{A}\varphi \equiv \varphi$, we know $\vdash \varphi \to \mathcal{A}\varphi$. So by (47.10), we have $\varphi \vdash \mathcal{A}\varphi$. But from $\Gamma \vdash \varphi$ and $\varphi \vdash \mathcal{A}\varphi$, it follows by (47.8) that $\Gamma \vdash \mathcal{A}\varphi$.

[103] To see this, assume the antecedent, i.e., $\Gamma \vdash \mathcal{A}\varphi$. Note that since the necessitation-averse axiom of actuality (30)$\star$ asserts that $\mathcal{A}\varphi \equiv \varphi$, it follows by (47.1) that $\vdash \mathcal{A}\varphi \equiv \varphi$. So by (47.3), it follows that $\Gamma \vdash \mathcal{A}\varphi \equiv \varphi$. Then from $\Gamma \vdash \mathcal{A}\varphi$ and $\Gamma \vdash \mathcal{A}\varphi \equiv \varphi$, it follows by biconditional syllogism (64.6.a) that $\Gamma \vdash \varphi$.

uses the above rule implicitly involves an appeal to the necessitation-averse axiom (30)★. Unless we take further precautions, this rule could permit us to derive an invalidity.[104] So instead of taking such precautions as tagging the rule with a ★ (to mark it as non-strict) and tagging any derivations involving the rule as non-strict, we simply avoid non-strict rules altogether.

(**90**) ★**Theorems:** Actuality and Negation. The following are simple consequences of (30)★:

(.1) $\neg\mathcal{A}\varphi \equiv \neg\varphi$

(.2) $\neg\mathcal{A}\neg\varphi \equiv \varphi$

Given that the proofs of these theorems depend on (30)★, we may not apply RN to either theorem.

(**91**) **Theorems**. Modally Strict Theorems of Actuality.

(.1) $\Box\varphi \rightarrow \mathcal{A}\varphi$

(.2) $\mathcal{A}(\varphi \,\&\, \psi) \equiv (\mathcal{A}\varphi \,\&\, \mathcal{A}\psi)$

(.3) $\mathcal{A}(\varphi \equiv \psi) \equiv (\mathcal{A}(\varphi \rightarrow \psi) \,\&\, \mathcal{A}(\psi \rightarrow \varphi))$

(.4) $(\mathcal{A}(\varphi \rightarrow \psi) \,\&\, \mathcal{A}(\psi \rightarrow \varphi)) \equiv (\mathcal{A}\varphi \equiv \mathcal{A}\psi)$

(.5) $\mathcal{A}(\varphi \equiv \psi) \equiv (\mathcal{A}\varphi \equiv \mathcal{A}\psi)$

(.6) $\Diamond\varphi \equiv \mathcal{A}\Diamond\varphi$

(.7) $\mathcal{A}\varphi \equiv \Box\mathcal{A}\varphi$

(.8) $\mathcal{A}\Box\varphi \rightarrow \Box\mathcal{A}\varphi$

(.9) $\Box\varphi \rightarrow \Box\mathcal{A}\varphi$

(.10) $\mathcal{A}(\varphi \vee \psi) \equiv (\mathcal{A}\varphi \vee \mathcal{A}\psi)$

(.11) $\mathcal{A}\exists\alpha\varphi \equiv \exists\alpha\mathcal{A}\varphi$

Note that one can develop far simpler proofs of some of the above theorems than the ones given in the Appendix by using the necessitation-averse axiom (30)★. But such uses would prevent us from applying RN to these theorems. The proofs we give in the Appendix are modally-strict.

[104]Here is a simple example. As an instance of $\varphi \vdash \varphi$ (47.4), we know: $\mathcal{A}\varphi \vdash \mathcal{A}\varphi$. So the rule of actuality elimination would allow one to infer $\mathcal{A}\varphi \vdash \varphi$. But by the Deduction Theorem (54), it would follow that $\mathcal{A}\varphi \rightarrow \varphi$ is a modally-strict theorem. We know that the necessitation of this claim is invalid, but without further constraints on RN, it would follow that $\Box(\mathcal{A}\varphi \rightarrow \varphi)$ is a theorem.

For example, we saw in Section 6.7 why we didn't want to prove axiom (33.2), i.e., $\Box\varphi \equiv \mathcal{A}\Box\varphi$, as a theorem simply by commuting the instance of (30)★ that results by substituting $\Box\varphi$ for φ. For then (33.2) would have been a ⋆-theorem and its necessitation wouldn't have been provable. Similarly, the following proof of (91.1) isn't modally-strict: assume $\Box\varphi$, infer φ by the T schema, infer $\mathcal{A}\varphi$ by our ⋆-axiom for actuality (30)★, and conclude $\Box\varphi \to \mathcal{A}\varphi$ by conditional proof. From such a proof, one may not apply RN to the conclusion to obtain $\Box(\Box\varphi \to \mathcal{A}\varphi)$.

Finally, we note that (91.1) – (91.5) are used in the derivation of the Rule of Substitution (109) below. This is an important rule that can and should be derived in a way that complies with our convention of eschewing non-strict rules.

(**92**) ⋆**Lemmas:** A Consequence of the Necessitation-Averse Equivalence of φ and $\mathcal{A}\varphi$. It is a straightforward consequence of the necessitation-averse axiom for actuality (30)★ that an object x is uniquely such that $\mathcal{A}\varphi$ if and only if x is uniquely such that φ:

> $\forall z(\mathcal{A}\varphi^z_x \equiv z = x) \equiv \forall z(\varphi^z_x \equiv z = x)$, provided z is substitutable for x in φ and doesn't occur free in φ

(**93**) ⋆**Theorems:** Fundamental Theorems Governing Descriptions. It follows from the previous lemma that x is the individual that is (in fact) such that φ just in case x is uniquely such that φ:

> $x = \imath x\varphi \equiv \forall z(\varphi^z_x \equiv z = x)$, provided z is substitutable for x in φ and doesn't occur free in φ

The proof of this theorem goes by way of the axiom for descriptions (34) and lemma (92)★, and hence depends on the necessitation-averse axiom for actuality (30)★. So we may not apply RN to derive its necessitation. But from this ⋆-theorem, we may prove other important and famous principles involving descriptions which, in the present context, are also ⋆-theorems. Examples are Hintikka's schema (94)★ and Russell's analysis of descriptions (95)★. Though the classical statement of these famous principles are derived in a way that is not modally-strict, the principles can be slightly modified so as to be derivable by modally-strict proofs. This will become apparent below.

(**94**) ⋆**Theorems:** Hintikka's Schema. We may derive the instances of Hintikka's schema for definite descriptions (1959, 83), though in contrast to his schema, the following one involves a defined rather than primitive identity sign:[105]

[105]Note that we've changed one of the variables in Hintikka's scheme so as to simplify the statement of the theorem.

$x = \imath x\varphi \equiv (\varphi \,\&\, \forall z(\varphi^z_x \rightarrow z = x))$, provided z is substitutable for x in φ and doesn't occur free in φ

This asserts: x is the individual (in fact) such that φ just in case φ and everything that is such that φ is identical to x.

The proof of Hintikka's schema appeals to the ⋆-theorem (93)⋆ and so the schema fails to be a modally-strict theorem. When φ is within the scope of a rigidifying operator like $\imath x$ or $\mathcal{A}$ on one side of a true conditional (or biconditional) but not within the scope of such an operator on the other side, the necessitation of the conditional (or biconditional) is invalid. For the discussion of an example involving a description, see Section 5.5.2. The example discussed in detail there, $Q\imath xPx \rightarrow \exists yPy$, should help one to see why Hintikka's schema, though valid, doesn't have a valid necessitation.

(**95**) ⋆**Theorems:** Russell's Analysis of Descriptions. Our derived quantifier rules also help us to more easily prove, as a *theorem*, a version of Russell's famous (1905) analysis of definite descriptions:

$\psi^{\imath x\varphi}_x \equiv \exists x(\varphi \,\&\, \forall z(\varphi^z_x \rightarrow z = x) \,\&\, \psi)$, provided (a) ψ is either an exemplification formula $\Pi^n\kappa_1 \ldots \kappa_n$ $(n \geq 1)$ or an encoding formula $\kappa_1\Pi^1$, (b) x occurs in ψ and only as one or more of the κ_i $(1 \leq i \leq n)$, and (c) z is substitutable for x in φ and doesn't appear free in φ

This asserts: the individual (in fact) such that φ is ψ if and only if something x is such that φ, everything such that φ just is x, and x is such that ψ. It is also important to note that with rigid definite descriptions, Russell's analysis is a ⋆-theorem; the proof relies on Hintikka's schema, which in turn depends on the ⋆-theorem (93)⋆.

(**96**) ⋆**Theorems:** Logically Proper Descriptions and Uniqueness. There exists something that is identical to the individual (in fact) such that φ if and only if there exists a unique x such that φ:

$\exists y(y = \imath x\varphi) \equiv \exists!x\varphi$, provided y doesn't occur free in φ

The left-to-right direction captures definition ∗14·02 in *Principia Mathematica*, since we know, from the discussion on p. 31 (Introduction, Chapter I) of that text, that the intent of Whitehead and Russell's predicate $E!$ in the definiendum $E!\imath x\varphi$ is to assert that the x such that φ exists. However, note that since the proof of this theorem appeals to Hintikka's schema, this ⋆-theorem is not subject to RN. Both directions of the biconditional fail to be necessary, and there are counterexamples to their necessitations. We can find formulas φ and interpretations for which the following is a counterexample to the necessitation of the left-to-right direction: *possibly*, something is the (actual) φ and nothing is uniquely φ. And we can find a φ and an interpretation for which the following is a counterexample to the necessitation of the right-to-left direction: *possibly*, something is uniquely φ and nothing is the (actual) φ.

(**97**) $\star$**Theorems:** Logically Proper Descriptions Apply To Themselves.

(.1) $x = \imath x\varphi \rightarrow \varphi$

(.2) $z = \imath x\varphi \rightarrow \varphi^z_x$, provided z is substitutable for x in φ and doesn't occur free in φ

(.3) $\exists y(y = \imath x\varphi) \rightarrow \varphi^{\imath x\varphi}_x$, provided y doesn't occur free in φ

These theorems license substitutions into the matrix of a description under certain conditions. The last one tells us we can substitute a description into its own matrix if we know that the description is logically proper. Note that since the proofs appeal to Hintikka's schema, they are all $\star$-theorems.

(**98**) **Lemmas:** Consequence of the Necessary Equivalence of $\mathcal{A}\varphi$ and $\mathcal{A}\mathcal{A}\varphi$. One of the necessary axioms for actuality is (31.4), namely, $\mathcal{A}\varphi \equiv \mathcal{A}\mathcal{A}\varphi$. It is a straightforward, modally-strict consequence of this axiom that an object x is uniquely such that $\mathcal{A}\varphi$ if and only if x is uniquely such that $\mathcal{A}\mathcal{A}\varphi$:

$\forall z(\mathcal{A}\varphi^z_x \equiv z = x) \equiv \forall z(\mathcal{A}\mathcal{A}\varphi^z_x \equiv z = x)$, provided z is substitutable for x in φ and doesn't occur free in φ

The necessitation-averse axiom for actuality (30)$\star$ is not needed to prove this lemma.

(**99**) **Theorems:** Fundamental Theorems About Rigid Descriptions and Actuality. It is provable that (.1) x is identical to the individual (in fact) such that φ if and only if x is identical to the individual (in fact) such that actually φ, and (.2) If there is something which is the individual (in fact) such that φ, then it is identical to the individual (in fact) such that actually φ:

(.1) $x = \imath x\varphi \equiv x = \imath x\mathcal{A}\varphi$

(.2) $\exists y(y = \imath x\varphi) \rightarrow \imath x\varphi = \imath x\mathcal{A}\varphi$, provided y doesn't occur free in φ

These are modally-strict theorems.

(**100**) **Theorems:** Sufficient Conditions for Logically Proper Descriptions with Actuality.

(.1) $\psi^{\imath x\varphi}_x \rightarrow \exists\nu(\nu = \imath x\mathcal{A}\varphi)$, provided (a) ψ is either an exemplification formula $\Pi^n\kappa_1 \ldots \kappa_n$ $(n \geq 1)$ or an encoding formula $\kappa_1\Pi^1$, (b) x occurs in ψ and only as one or more of the κ_i $(1 \leq i \leq n)$, and (c) ν is any individual variable that doesn't occur free in φ.

(.2) $\psi_x^{\imath x\varphi} \to \imath x\varphi = \imath x\mathcal{A}\varphi$, provided (a) ψ is either an exemplification formula $\Pi^n\kappa_1 \dots \kappa_n$ ($n \geq 1$) or an encoding formula $\kappa_1\Pi^1$ and (b) x occurs in ψ and only as one or more of the κ_i ($1 \leq i \leq n$).

The first is a variant of axiom (29.1). These are modally-strict theorems.

(**101**) **Theorems:** Modally Strict Versions of Hintikka's Schema. The following versions of Hintikka's schema, which deploy the actuality operator, are $\Box$-theorems:

(.1) $x = \imath x\varphi \equiv \mathcal{A}\varphi \,\&\, \forall z(\mathcal{A}\varphi_x^z \to z = x)$, provided z is substitutable for x in φ and doesn't occur free in φ.

(.2) $x = \imath x\mathcal{A}\varphi \equiv \mathcal{A}\varphi \,\&\, \forall z(\mathcal{A}\varphi_x^z \to z = x)$, provided z is substitutable for x in φ and doesn't occur free in φ.

The proof of (101.1) appeals to the necessary axiom (34) instead of the contingently-proved theorem (93)★ and so (101.1) is a modally-strict theorem! (101.2) follows from (101.1) by virtue of (99.1) and biconditional syllogism.

(**102**) **Theorems:** Identity of Equivalent Descriptions. The modally strict version of Hintikka's scheme allows us to formulate and prove a nice theorem in connection with descriptions $\imath x\varphi$ and $\imath x\psi$ whose matrices are actually equivalent. If $\mathcal{A}(\varphi \equiv \psi)$, then x is identical to the individual (in fact) such that φ if and only if x is identical to the individual (in fact) such that ψ:

$$\mathcal{A}(\varphi \equiv \psi) \to \forall x(x = \imath x\varphi \equiv x = \imath x\psi)$$

Note that we can't prove $\mathcal{A}(\varphi \equiv \psi) \to \imath x\varphi = \imath x\psi$, since the identity would imply that the descriptions are logically proper, something that is not guaranteed by the antecedent. If, for example, both $\neg\exists x\mathcal{A}\varphi$ and $\neg\exists x\mathcal{A}\varphi$, then by (86.12), it would follow that $\forall x(\mathcal{A}\varphi \equiv \mathcal{A}\varphi)$. In such a case, $\mathcal{A}\varphi \equiv \mathcal{A}\psi$ would be true, the descriptions $\imath x\varphi$ and $\imath x\psi$ would fail to be logically proper, and $\imath x\varphi = \imath x\psi$ would be false.

(**103**) **Theorems:** Modally Strict Versions of Russell's Analysis of Descriptions. By appropriately inserting the actuality operator, we have:

(.1) $\psi_x^{\imath x\varphi} \equiv \exists x(\mathcal{A}\varphi \,\&\, \forall z(\mathcal{A}\varphi_x^z \to z = x) \,\&\, \psi)$, provided (a) ψ is either an exemplification formula $\Pi^n\kappa_1 \dots \kappa_n$ ($n \geq 1$) or an encoding formula $\kappa_1\Pi^1$, (b) x occurs in ψ and only as one or more of the κ_i ($1 \leq i \leq n$), and (c) z is substitutable for x in φ and doesn't occur free in φ

(.2) $\psi_x^{\imath x\mathcal{A}\varphi} \equiv \exists x(\mathcal{A}\varphi \,\&\, \forall z(\mathcal{A}\varphi_x^z \to z = x) \,\&\, \psi)$, provided (a) ψ is either an exemplification formula $\Pi^n\kappa_1 \dots \kappa_n$ ($n \geq 1$) or an encoding formula $\kappa_1\Pi^1$, (b) x occurs in ψ and only as one or more of the κ_i ($1 \leq i \leq n$), and (c) z is substitutable for x in φ and doesn't occur free in φ

These are $\Box$-theorems that don't require an appeal to the axiom (30)★.

(**104**) **Theorems:** Modally Strict Theorems About Logically Proper Descriptions and Actuality. Where y doesn't occur free in φ, we have:

(.1) $\exists y(y = \imath x\varphi) \equiv \exists! x\mathcal{A}\varphi$

(.2) $\exists y(y = \imath x\mathcal{A}\varphi) \equiv \exists! x\mathcal{A}\varphi$

(.3) $\exists y(y = \imath x\varphi) \to \mathcal{A}\varphi_x^{\imath x\varphi}$

(.4) $\exists y(y = \imath x\mathcal{A}\varphi) \to \mathcal{A}\varphi_x^{\imath x\mathcal{A}\varphi}$

These are modally-strict theorems.

9.8 The Theory of Necessity

(**105**) **Theorems:** Tautologies Proved Thus Far Are Necessary.

- The tautologies proved in Section 9.4, in items (53), (58), and (63), are all necessary.

In each case, the necessitation follows by an application of RN. This applies to any other tautology derivable from our axioms for negations and conditionals.

But as noted in Remark (65), we haven't yet shown that *all* tautologies are derivable, for there are some tautologies involving 0-place λ-expressions that are not derivable from the axioms for negations and conditionals alone. The remaining principles needed for the proof that every tautology is derivable are established in the present section. In particular, once we have the proof of (126.2) in hand (i.e., the 0-place case of β-Conversion, namely, $[\lambda\, \varphi^*] \equiv \varphi^*$), Metatheorems $\langle 9.2 \rangle$ and $\langle 9.3 \rangle$ (in the Appendix to this chapter) provide proofs, respectively, of the facts that *every* tautology is derivable and that every tautology is a necessary truth.

(**106**) **Metarules:** Rules RM and RM$\Diamond$. The classical rule RM of the logic of necessity asserts that if $\vdash \varphi \to \psi$, then $\vdash \Box\varphi \to \Box\psi$. However, in our system, rule RM has to be adjusted slightly to accommodate necessitative-averse axioms and contingent premises.

To see why, take a simple case in which one extends our theory by asserting both ψ and $\Diamond\neg\psi$ as axioms. Consider the 3-element sequence of formulas $\psi, \psi \to (\varphi \to \psi), \varphi \to \psi$. This sequence is a proof of $\varphi \to \psi$, since the first member is an axiom by hypothesis, the second member is an instance of axiom schema (21.1) for conditionals, and the third member follows from the first two by MP. So this establishes $\vdash \varphi \to \psi$. By the classical rule RM, it would follow that $\vdash \Box\varphi \to \Box\psi$. But if φ were, say, an instance of a necessary axiom, so that $\Box\varphi$ is also an axiom, then by (47.1), it would follow that $\vdash \Box\varphi$ and, by (47.6), that $\vdash \Box\psi$. But, by hypothesis, $\Diamond\neg\psi$. So the unmodified rule RM would allow us to derive the necessitation of a formula (ψ) that is possibly false.

The problem in this case, of course, is that in the initial proof of $\varphi \to \psi$, the conclusion depends on ψ, which is known to be true but possibly false. If we formulate RM so that it applies only to conditional theorems that have modally-strict proofs or derivations, we can forestall the potential derivation of a falsehood (from a truth). But we first formulate the rule for derivations generally:

(.1) **Rule RM**:
If $\Gamma \vdash_{\Box} \varphi \to \psi$, then $\Box\Gamma \vdash \Box\varphi \to \Box\psi$.

In other words, if there is a modally-strict derivation of $\varphi \to \psi$ from Γ, then there is a derivation of $\Box\varphi \to \Box\psi$ from the necessitations of the formulas in Γ. When $\Gamma = \varnothing$, then RM reduces to the principle:

If $\vdash_{\Box} \varphi \to \psi$, then $\vdash \Box\varphi \to \Box\psi$

i.e., if $\varphi \to \psi$ is a modally-strict theorem, then $\Box\varphi \to \Box\psi$ is a theorem.

RM$\Diamond$ is a corresponding rule:

(.2) **Rule RM$\Diamond$**:
If $\Gamma \vdash_{\Box} \varphi \to \psi$, then $\Box\Gamma \vdash \Diamond\varphi \to \Diamond\psi$.

In other words, if there is a modally-strict derivation of $\varphi \to \psi$ from Γ, then there is a derivation of $\Diamond\varphi \to \Diamond\psi$ from the necessitations of the formulas in Γ. When $\Gamma = \varnothing$, then RM$\Diamond$ reduces to the principle:

If $\vdash_{\Box} \varphi \to \psi$, then $\vdash \Diamond\varphi \to \Diamond\psi$

i.e., if $\varphi \to \psi$ is a modally-strict theorem, then $\Diamond\varphi \to \Diamond\psi$ is a theorem.

(**107**) **Theorems:** Basic K Theorems. The presentation and proof of some of the following basic theorems that depend upon the K schema has been informed by Hughes and Cresswell 1968 and 1996:

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

(.2) $\Box\neg\varphi \to \Box(\varphi \to \psi)$

(.3) $\Box(\varphi \,\&\, \psi) \equiv (\Box\varphi \,\&\, \Box\psi)$

(.4) $\Box(\varphi \equiv \psi) \equiv (\Box(\varphi \to \psi) \,\&\, \Box(\psi \to \varphi))$

(.5) $(\Box(\varphi \to \psi) \,\&\, \Box(\psi \to \varphi)) \to (\Box\varphi \equiv \Box\psi)$

(.6) $\Box(\varphi \equiv \psi) \to (\Box\varphi \equiv \Box\psi)$

(.7) $(\Box\varphi \,\&\, \Box\psi) \rightarrow \Box(\varphi \equiv \psi)$

If we define φ *necessarily implies* ψ as $\Box(\varphi \rightarrow \psi)$, then theorem (107.1) asserts that if φ is necessarily true, then every claim formulable in our language necessarily implies φ. Similarly, (107.2) asserts that if φ is necessarily false, then φ necessarily implies every claim whatsoever. Where the classic notion of *strict implication* is understood in terms of the above definition of *necessary implication*, these results are among the classical "paradoxes of strict implication" (Lewis and Langford 1932 [1959], 511): a truth that is necessarily true is strictly implied by everything and a falsehood that is necessarily false strictly implies everything. However, given the meaning of the conditional, they are harmless.[106] (107.3) establishes that the necessity operator distributes over the conjuncts of a conjunction (and vice versa!), while (107.4) and (107.5) are lemmas needed for the proof of (107.6), which asserts that the necessity operator distributes over a biconditional.

Note that the converse of (107.6), $(\Box\varphi \equiv \Box\psi) \rightarrow \Box(\varphi \equiv \psi)$, is not a theorem: the material equivalence of $\Box\varphi$ and $\Box\psi$ doesn't imply that the biconditional $\varphi \equiv \psi$ is necessary. It is important to be familiar with a counterexample that demonstrates this for, as we shall see, there is a special case (121.4) in which the converse of (107.6) holds. But, in the general case, the converse of (107.6) fails in an interpretation in which there are two worlds, $\boldsymbol{w}_0$ and $\boldsymbol{w}_1$, such that (a) φ is true at $\boldsymbol{w}_0$ and false at $\boldsymbol{w}_1$, and (b) ψ is false at $\boldsymbol{w}_0$ and true at $\boldsymbol{w}_1$. Then clearly, both $\Box\varphi$ and $\Box\psi$ are false at $\boldsymbol{w}_0$, and since $\Box\varphi$ and $\Box\psi$ have the same truth value at $\boldsymbol{w}_0$, $\Box\varphi \equiv \Box\psi$ is true at $\boldsymbol{w}_0$. But the claim $\Box(\varphi \equiv \psi)$ is false at $\boldsymbol{w}_0$, since the conditional $\varphi \equiv \psi$ fails at both worlds given that φ and ψ have different truth values at each world. By constrast, (107.7) ensures that the biconditional $\varphi \equiv \psi$ is necessary if both φ and ψ are necessary.

(**108**) **Metarules:** Rules of Necessary Equivalence. Theorems (63.5.d), (63.5.e), (63.5.f), (83.3), (91.5), and (107.6) each play a crucial role in establishing one of the cases of the following rule:

(.1) If $\vdash \Box(\psi \equiv \chi)$, then:

- (.a) $\vdash \neg\psi \equiv \neg\chi$
- (.b) $\vdash (\psi \rightarrow \theta) \equiv (\chi \rightarrow \theta)$
- (.c) $\vdash (\theta \rightarrow \psi) \equiv (\theta \rightarrow \chi)$
- (.d) $\vdash \forall\alpha\psi \equiv \forall\alpha\chi$
- (.e) $\vdash \mathcal{A}\psi \equiv \mathcal{A}\chi$

[106](107.1) is provably equivalent to the claim: $\Box\varphi \rightarrow \neg\Diamond(\psi \,\&\, \neg\varphi)$. This just follows *a fortiori* from the fact that $\Box\varphi \rightarrow \neg\Diamond\neg\varphi$, which is established below as item (113.3). Similarly, (107.2) is provably equivalent to the claim: $\Box\neg\varphi \rightarrow \neg\Diamond(\varphi \,\&\, \neg\psi)$. This just follows *a fortiori* from the fact that $\Box\neg\varphi \rightarrow \neg\Diamond\varphi$, which is established below as item (113.4).

(.f) $\vdash \Box\psi \equiv \Box\chi$

The above rule can be used to give an informal proof of the following rule, though we give both the informal and strict proof in the Appendix:

(.2) If $\vdash \Box(\psi \equiv \chi)$, then if φ' is the result of substituting the formula χ for zero or more occurrences ψ where the latter is a subformula of φ, then $\vdash \varphi \equiv \varphi'$.

Here are some examples of (108.2):

Example.
If $\vdash \Box(A!x \equiv \neg\Diamond E!x)$, then $\vdash \exists x A!x \equiv \exists x \neg\Diamond E!x$

Example.
If $\vdash \Box(Rxy \equiv (Rxy \,\&\, (Qa \vee \neg Qa)))$, then
$\vdash (Pa \,\&\, \Box Rxy) \equiv (Pa \,\&\, \Box(Rxy \,\&\, (Qa \vee \neg Qa)))$

In the second example, we've conjoined a tautology $Qa \vee \neg Qa$ with the formula Rxy and the result is necessarily equivalent to Rxy. Hence we can take the formula, $Pa \,\&\, \Box Rxy$ (= φ), replace Rxy in this formula with its necessary equivalent, and the result φ' is materially equivalent to φ.

Not only does (108.2) play an important role in the proof of the next, key derived rule, but many of the other theorems and derived rules proved thus far also play a direct or indirect role in the proof.

(**109**) **Metarule:** Rule of Substitution.

Rule of Substitution
If $\vdash_\Box \psi \equiv \chi$ and φ' is the result of substituting the formula χ for zero or more occurrences of ψ where the latter is a subformula of φ, then if $\Gamma \vdash \varphi$, then $\Gamma \vdash \varphi'$.

[Variant: If $\vdash_\Box \psi \equiv \chi$, then $\varphi \vdash \varphi'$]

(109) tells us that if we can derive φ from the set Γ, then if there is a modally-strict proof of $\psi \equiv \chi$, we can derive φ' from Γ, where φ' has χ substituted (though not necessarily uniformly) for some subformula ψ in φ.

Note that when $\vdash_\Box \psi \equiv \chi$, the Rule of Substitution does *not* allow us to substitute χ for ψ in any context whatsoever, but rather only when they occur as subformulas of some given formula. The Rule does *not*, for example, permit the substitution of χ for ψ within the formula $[\lambda y\, \psi \,\&\, \theta]x$ to obtain $[\lambda y\, \chi \,\&\, \theta]x$ (since ψ is not a subformula of $[\lambda y\, \psi \,\&\, \theta]x$). In this particular case, it just so happens that the substitution is valid: we may use β-Conversion on $[\lambda y\, \psi \,\&\, \theta]x$ to obtain $(\psi \,\&\, \theta)^x_y$, i.e., $\psi^x_y \,\&\, \theta^x_y$; then use the Rule of Substitution to obtain $\chi^x_y \,\&\, \theta^x_y$, which again by β-Conversion, yields $[\lambda y\, \chi \,\&\, \theta]x$. But such substitutions are not generally valid; just consider the similar case but with encoding formulas. From $\vdash_{\Box} \psi \equiv \chi$, we may not use the Rule of Substitution to infer $x[\lambda y\, \chi \,\&\, \theta]$ from $x[\lambda y\, \psi \,\&\, \theta]$, nor could such an inference be validly derived by other means. Of course, if ψ is a formula *defined* by χ, then rules for substituting definiens for definiendum would permit us to substituted χ for ψ in both $[\lambda y\, \psi \,\&\, \theta]x$ and $x[\lambda y\, \psi \,\&\, \theta]$. See the discussion in (203) and, in particular, the discussion of Rule SubDefForm in (203.5).

(**110**) **Remark:** Legitimate and Illegitimate Uses of the Rule of Substitution. Here are some legitimate examples of (109):

Example 1.
If $\Gamma \vdash \neg A!x$ and $\vdash_{\Box} A!x \equiv \neg\Diamond E!x$, then $\Gamma \vdash \neg\neg\Diamond E!x$.

Example 2.
If $\Gamma \vdash p \rightarrow Rxy$ and $\vdash_{\Box} Rxy \equiv (Rxy \,\&\, (Qa \vee \neg Qa))$, then $\Gamma \vdash p \rightarrow (Rxy \,\&\, (Qa \vee \neg Qa))$.

Example 3.
If $\vdash \exists x A!x$ and $\vdash_{\Box} A!x \equiv \neg\Diamond E!x$, then $\vdash \exists x \neg\Diamond E!x$.

Example 4.
If $\vdash \mathcal{A}\neg\neg Px$ and $\vdash_{\Box} \neg\neg Px \equiv Px$, then $\vdash \mathcal{A}Px$.

Example 5.
If $\vdash \Box(\varphi \rightarrow \psi)$ and $\vdash_{\Box} (\varphi \rightarrow \psi) \equiv (\neg\psi \rightarrow \neg\varphi)$, then $\vdash \Box(\neg\psi \rightarrow \neg\varphi)$.

Example 6.
If $\vdash \Box(\varphi \rightarrow \psi)$ and $\vdash_{\Box} \psi \equiv \chi$, then $\vdash \Box(\varphi \rightarrow \chi)$.

Example 7.
If $\vdash \Box(\varphi \rightarrow \varphi)$ and $\vdash_{\Box} \varphi \equiv \neg\neg\varphi$, then $\vdash \Box(\neg\neg\varphi \rightarrow \varphi)$.

Though the following is also a legitimate instance of the Rule of Substitution, it can only be invoked in certain circumstances:

- If $\vdash \Box\mathcal{A}\varphi$ and $\vdash_{\Box} \mathcal{A}\varphi \equiv \varphi$, then $\vdash \Box\varphi$

For arbitrary φ, we can't establish that $\vdash_{\Box} \mathcal{A}\varphi \equiv \varphi$, and so we can't generally substitute φ for $\mathcal{A}\varphi$ in $\Box\mathcal{A}\varphi$ to obtain $\vdash \Box\varphi$ from $\vdash \Box\mathcal{A}\varphi$. One exception is the case where φ is some necessary truth, say $\Box\psi$, for then axiom (33.2) is that $\Box\psi \equiv \mathcal{A}\Box\psi$, which by contraposition is $\mathcal{A}\Box\psi \equiv \Box\psi$. Since there is, by definition, a modally strict proof of this axiom, we can substitute $\Box\psi$ for $\mathcal{A}\Box\psi$ in $\Box\mathcal{A}\Box\psi$ to obtain $\Box\Box\psi$.[107]

Note also the following *illegitimate* instance of the Rule of Substitution:

[107] Another exception will become clear once we reach item (121.8), where we prove, via a modally strict proof, that $\mathcal{A}xF \equiv xF$. Thus, we can substitute xF for $\mathcal{A}xF$ whenever $\mathcal{A}xF$ appears somewhere as a subformula, and vice versa.

- If $\vdash Pa = Pa$ and $\vdash_\Box Pa \equiv (Pa \,\&\, (Qb \vee \neg Qb))$, then $\vdash Pa = (Pa \,\&\, (Qb \vee \neg Qb))$

Though we can establish both $\vdash Pa = Pa$ and $\vdash_\Box Pa \equiv (Pa \,\&\, (Qb \vee \neg Qb))$, we may not conclude $\vdash Pa = (Pa \,\&\, (Qb \vee \neg Qb))$ because Pa is *not* a subformula of $Pa = Pa$! When the defined notation $Pa = Pa$ is expanded, the formula Pa will appear only inside complex terms and so the definition of *subformula* won't count Pa as a subformula of the formula $Pa = Pa$.

Before we explain this in detail, note that $Pa = Pa$ is a defined identity formula of the form $\Pi^0 = \Pi^0$, where Π^0 is a 0-place relation term. It is easy to show $Pa = Pa$ is a theorem: we first prove $p = p$ (67.1), then apply GEN to obtain $\forall p(p = p)$, and then use Variant Rule $\forall$E to instantiate Pa into the universal claim to obtain $Pa = Pa$ (we can do this because Pa is a 0-place relation term substitutable for p in $p = p$). Moreover, we can show $\vdash_\Box Pa \equiv (Pa \,\&\, (Qb \vee \neg Qb))$, for it is an instance of a tautology $\varphi \equiv (\varphi \,\&\, (\psi \vee \neg\psi))$, which is easily provable as a modally-strict theorem.

Notwithstanding these derivations, the fact is that Pa is not a subformula of the theorem $Pa = Pa$. To see why, expand the latter by applying definition (16.3). For then we obtain: $[\lambda y\, Pa] = [\lambda y\, Pa]$. Though this, too, is a defined formula, at this point we should suspect that already Pa is not a subformula of $[\lambda y\, Pa] = [\lambda y\, Pa]$. However, to complete our proof that it is not, we expand $[\lambda y\, Pa] = [\lambda y\, Pa]$ by definition (16.1), to obtain:

$$\Box\forall x(x[\lambda y\, Pa] \equiv x[\lambda y\, Pa])$$

The above is the result of expanding $Pa = Pa$ into primitive notation, and by applying the definition of *subformula* (8) to the above formula, we can see that Pa doesn't count as one of its subformulas.[108] Hence, we may not use the Rule of Substitution to substitute $Pa \,\&\, (Qb \vee \neg Qb)$ for Pa in $Pa = Pa$ to obtain $Pa = (Pa \,\&\, (Qb \vee \neg Qb))$. The moral here is this: one may not use the Rule of Substitution to identify necessarily equivalent propositions.

By the same token, one may not use the Rule of Substitution to identify necessarily equivalent properties. Consider the following illegitimate application of the rule, related to the one above:

- If $\vdash Px = Px$ and $\vdash_\Box Px \equiv [\lambda y\, Py \,\&\, (Qb \vee \neg Qb)]x$, then $\vdash Px = [\lambda y\, Py \,\&\, (Qb \vee \neg Qb)]x$

One can establish $\vdash Px = Px$ in the same manner that we established $\vdash Pa = Pa$. Moreover, one can establish $\vdash_\Box Px \equiv [\lambda y\, Py \,\&\, (Qb \vee \neg Qb)]x$ as follows. First, appeal to the following instance of β-Conversion (36.2):

[108] By (8.1), $\Box\forall x(x[\lambda y\, Pa] \equiv x[\lambda y\, Pa])$ is a subformula of itself; so by (8.2), $\forall x(x[\lambda y\, Pa] \equiv x[\lambda y\, Pa])$ is also a subformula; by (8.2) again, $x[\lambda y\, Pa] \equiv x[\lambda y\, Pa]$ is also a subformula; and by Metatheorem $\langle 7.3\rangle$, $x[\lambda y\, Pa]$ is also a subformula. These are the only subformulas of $\Box\forall x(x[\lambda y\, Pa] \equiv x[\lambda y\, Pa])$.

$$[\lambda y\, Py \,\&\, (Qb \vee \neg Qb)]x \equiv Px \,\&\, (Qb \vee \neg Qb)$$

Now $Px\&(Qb\vee\neg Qb)$ is provably equivalent Px; it is an instance of the tautology $(\varphi \,\&\, (\psi \vee \neg\psi)) \equiv \varphi$. Thus, it follows by a biconditional syllogism and commutativity of $\equiv$ that $Px \equiv [\lambda y\, Py \,\&\, (Qb \vee \neg Qb)]x$. Since this proof is modally-strict, we've established that:

(ϑ) $\vdash_\Box Px \equiv [\lambda y\, Py \,\&\, (Qb \vee \neg Qb)]x$

But though we have established both $\vdash Px = Px$ and (ϑ), we may not apply the Rule of Substitution to infer $\vdash Px = [\lambda y\, Py \,\&\, (Qb \vee \neg Qb)]x$. Px is not a subformula of $Px = Px$ and so we don't have a legitimate instance of the rule.

(**111**) **Metarule/Derived Rule:** Rule for Alphabetic Variants. The Rules of Necessary Equivalence not only help us to derive the Rule of Substitution, but also help us derive an important rule about alphabetically-variant formulas and terms, as these are defined in item (35):

Rule of Alphabetic Variants
$\Gamma \vdash \varphi$ if and only if $\Gamma \vdash \varphi'$, where φ' is any alphabetic variant of φ
[Variant $\varphi \dashv\vdash \varphi'$]

As a special case, when $\Gamma = \varnothing$, our rule asserts that a formula is a theorem if and only if any of its alphabetic variants is a theorem. We henceforth use the Variant form within derivations as a derived rule.

(**112**) **Theorems:** Theorems About Alphabetic Variants. It follows from the preceding rule: (.1) that alphabetically-variant formulas are equivalent, and (.2) that alphabetic-variants of logically proper descriptions can be identified:

(.1) $\varphi \equiv \varphi'$, where φ' is an alphabetic variant of φ

(.2) $\exists y(y = \imath\nu\varphi) \to \imath\nu\varphi = (\imath\nu\varphi)'$, where y doesn't occur free in φ and $(\imath\nu\varphi)'$ is any alphabetic variant of $\imath\nu\varphi$

Note that (.2) is a consequence, given (111.1), of the facts that logically proper descriptions can be instantiated into the universal claim $\forall x(x = x)$ and that formulas of the form $\imath\nu\varphi = \imath\nu\varphi$ and $\imath\nu\varphi = (\imath\nu\varphi)'$ are alphabetic variants by virtue of the fact that they differ only with respect to alphabetically-variant terms.

(**113**) **Theorems:** Additional K Theorems. The Rules of Necessary Equivalence and Substitution often simplify the proof of the following theorems:

(.1) $\Box\varphi \equiv \Box\neg\neg\varphi$

(.2) $\neg\Box\varphi \equiv \Diamond\neg\varphi$

(.3) $\Box\varphi \equiv \neg\Diamond\neg\varphi$ (Df$\Box$)

(.4) $\Box\neg\varphi \equiv \neg\Diamond\varphi$

(.5) $\Box(\varphi \to \psi) \to (\Diamond\varphi \to \Diamond\psi)$ (K$\Diamond$)

(.6) $\Diamond(\varphi \vee \psi) \equiv (\Diamond\varphi \vee \Diamond\psi)$

(.7) $(\Box\varphi \vee \Box\psi) \to \Box(\varphi \vee \psi)$

(.8) $\Diamond(\varphi \mathbin{\&} \psi) \to (\Diamond\varphi \mathbin{\&} \Diamond\psi)$

(.9) $\Diamond(\varphi \to \psi) \equiv (\Box\varphi \to \Diamond\psi)$

(.10) $\Diamond\Box\varphi \equiv \neg\Box\Diamond\neg\varphi$

(.11) $\Diamond\Diamond\varphi \equiv \neg\Box\Box\neg\varphi$

(.12) $\Box(\varphi \vee \psi) \to (\Box\varphi \vee \Diamond\psi)$

(**114**) **Theorems:** A Theorem of T. The following theorem depends on the T schema (32.2):

$$\varphi \to \Diamond\varphi \qquad (\mathrm{T}\Diamond)$$

(**115**) **Theorems:** Basic S5 Theorems. The system S5 is based on the K, T, and 5 schemata, i.e., (32.1) – (32.3). Chellas 1980 (16–18) informed the development of the following list of theorems of that system:

(.1) $\Diamond\Box\varphi \to \Box\varphi$ (5$\Diamond$)

(.2) $\Box\varphi \equiv \Diamond\Box\varphi$

(.3) $\Diamond\varphi \equiv \Box\Diamond\varphi$

(.4) $\varphi \to \Box\Diamond\varphi$ (B)

(.5) $\Diamond\Box\varphi \to \varphi$ (B$\Diamond$)

(.6) $\Box\varphi \to \Box\Box\varphi$ (4)

(.7) $\Box\varphi \equiv \Box\Box\varphi$

(.8) $\Diamond\Diamond\varphi \to \Diamond\varphi$ (4$\Diamond$)

(.9) $\Diamond\Diamond\varphi \equiv \Diamond\varphi$

(.10) $\Box(\varphi \vee \Box\psi) \equiv (\Box\varphi \vee \Box\psi)$

(.11) $\Box(\varphi \vee \Diamond\psi) \equiv (\Box\varphi \vee \Diamond\psi)$

(.12) $\Diamond(\varphi \,\&\, \Diamond\psi) \equiv (\Diamond\varphi \,\&\, \Diamond\psi)$

(.13) $\Diamond(\varphi \,\&\, \Box\psi) \equiv (\Diamond\varphi \,\&\, \Box\psi)$

(**116**) **Metarules:** The following rules are derivable with the help of the B (115.4) and B$\Diamond$ (115.5) schemata:

(.1) If $\Gamma \vdash_{\Box} \Diamond\varphi \to \psi$, then $\Box\Gamma \vdash \varphi \to \Box\psi$

(.2) If $\Gamma \vdash_{\Box} \varphi \to \Box\psi$, then $\Box\Gamma \vdash \Diamond\varphi \to \psi$

When Γ is empty and there are no premises or assumptions involved, the above reduce to:

If $\vdash_{\Box} \Diamond\varphi \to \psi$, then $\vdash \varphi \to \Box\psi$

If $\vdash_{\Box} \varphi \to \Box\psi$, then $\vdash \Diamond\varphi \to \psi$

(**117**) **Theorems:** Modal Quantification.

(.1) $\Box\forall\alpha\varphi \to \forall\alpha\Box\varphi$ (Converse Barcan Formula = CBF)

(.2) $\Diamond\exists\alpha\varphi \to \exists\alpha\Diamond\varphi$ (BF$\Diamond$)

(.3) $\exists\alpha\Diamond\varphi \to \Diamond\exists\alpha\varphi$ (CBF$\Diamond$)

(.4) $\exists\alpha\Box\varphi \to \Box\exists\alpha\varphi$ (Buridan)

(.5) $\Diamond\forall\alpha\varphi \to \forall\alpha\Diamond\varphi$ (Buridan$\Diamond$)

(.6) $\Diamond\exists\alpha(\varphi \,\&\, \psi) \to \Diamond(\exists\alpha\varphi \,\&\, \exists\alpha\psi)$

(.7) $(\Box\forall\alpha(\varphi \to \psi) \,\&\, \Box\forall\alpha(\psi \to \chi)) \to \Box\forall\alpha(\varphi \to \chi)$

(.8) $(\Box\forall\alpha(\varphi \equiv \psi) \,\&\, \Box\forall\alpha(\psi \equiv \chi)) \to \Box\forall\alpha(\varphi \equiv \chi)$

(**118**) **Theorems:** Identity, Necessity, and Possibility. The foregoing theorems and rules now allow us to prove: (.1) α and β are possibly identical if and only if they are identical; (.2) α and β are distinct if and only if they are necessarily distinct; and (.3) α and β are possibly distinct if and only if they are distinct:

(.1) $\Diamond\alpha = \beta \equiv \alpha = \beta$

(.2) $\alpha \neq \beta \equiv \Box\alpha \neq \beta$

(.3) $\Diamond\alpha \neq \beta \equiv \alpha \neq \beta$

(**119**) **Lemmas:** Logically Proper (Rigid) Descriptions and Necessity.

(.1) $\exists y(y = \imath x\varphi) \to \exists y\Box(y = \imath x\varphi)$

(.2) $\exists y(y = \imath x\varphi) \to \Box\exists y(y = \imath x\varphi)$

These theorems have modally-strict proofs. Note that (.1) does not say that: if there exists *the* x that happens to be φ then there exists something that necessarily is the x that happens to be φ. Since the description is rigid, it should be read as: if something is the x in fact such that φ then something necessarily is the x in fact such that φ. This understanding is confirmed by the theorems in item (99).

(**120**) ⋆**Lemmas:** Sufficient Conditions for the (Necessary) Existence of (Rigidly) Described Objects.

(.1) $\exists!x\varphi \to \exists y\Box(y = \imath x\varphi)$, provided y doesn't occur free in φ

(.2) $\exists!x\varphi \to \Box\exists y(y = \imath x\varphi)$, provided y doesn't occur free in φ

These theorems are not modally-strict. The first asserts that if there is a unique individual such that φ, then there exists something which, necessarily, is identical to the individual that is in fact φ.

(**121**) **Theorem:** Encoding, Modality and Actuality. We now prove a host of important *modally-strict* theorems about the status of encoding predications: (.1) an object x encodes property F iff necessarily x encodes F; (.2) possibly x encodes F iff x encodes F; (.3) possibly x encodes F iff necessarily x encodes F; (.4) xF and yG are necessarily equivalent if and only if $\Box xF$ and $\Box yG$ are materially equivalent; (.5) x fails to encode F iff necessarily x fails to encode F; (.6) possibly x fails to encode F iff x fails to encode F; (.7) possibly x fails to encode F iff necessarily x fails to encode F; and (.8) actually x encodes F iff x encodes F:

(.1) $xF \equiv \Box xF$

(.2) $\Diamond xF \equiv xF$

(.3) $\Diamond xF \equiv \Box xF$

(.4) $\Box(xF \equiv yG) \equiv (\Box xF \equiv \Box yG)$

(.5) $\neg xF \equiv \Box\neg xF$

(.6) $\Diamond\neg xF \equiv \neg xF$

(.7) $\Diamond\neg xF \equiv \Box\neg xF$

(.8) $\mathcal{A}xF \equiv xF$

(121.3) and (121.4) are especially significant theorems. (121.3) has a dual significance. Logically, it tells us that necessity and possibility collapse for encoding predications (the possible truth of xF is *equivalent* to its necessary truth), while metaphysically it tells us that an object x possibly encodes a property iff x necessarily does so. (121.4), which is a consequence of (121.3), is significant *not* because of the left-to-right condition, which is just an instance (107.6), but because of its right-to-left direction. In general, $\Box\varphi \equiv \Box\psi$ doesn't materially imply the claim $\Box(\varphi \equiv \psi)$, as we saw in the brief discussion following (107.6). But when φ and ψ are two encoding claims, such as xF and yG, the implication holds. Finally, (121.8) is interesting because it is a modally-strict theorem and so provable without an appeal to (30)★.

9.9 The Theory of Relations

In this subsection, we describe some important theorems that govern properties, relations, and propositions.

(**122**) **Theorems:** Fact about Necessary Equivalence and Complex Relations. If φ^* and ψ^* are equivalent propositional formulas, then objects $x_1,\ldots,x_n$ stand in the relation *being* $y_1,\ldots,y_n$ *such that* φ^* if and only if they stand in the relation *being* $y_1,\ldots,y_n$ *such that* ψ^*:

(.1) $(\varphi^* \equiv \psi^*) \to [\lambda y_1 \ldots y_n\, \varphi^{*}{}^{y_1,\ldots,y_n}_{x_1,\ldots,x_n}]x_1 \ldots x_n \equiv [\lambda y_1 \ldots y_n\, \psi^{*}{}^{y_1,\ldots,y_n}_{x_1,\ldots,x_n}]x_1 \ldots x_n$, provided $y_1,\ldots,y_n$ are substitutable, respectively, for $x_1,\ldots,x_n$ in φ^* and ψ^*, and don't occur free in φ^* and ψ^*

This theorem generalizes:

(.2) $\Box\forall x_1 \ldots \forall x_n(\varphi^* \equiv \psi^*) \to$ $\Box\forall x_1 \ldots \forall x_n([\lambda y_1 \ldots y_n\, \varphi^{*}{}^{y_1,\ldots,y_n}_{x_1,\ldots,x_n}]x_1 \ldots x_n \equiv [\lambda y_1 \ldots y_n\, \psi^{*}{}^{y_1,\ldots,y_n}_{x_1,\ldots,x_n}]x_1 \ldots x_n)$, provided $y_1,\ldots,y_n$ are substitutable, respectively, for $x_1,\ldots,x_n$ in φ^* and don't occur free in φ

As an example of (.2), we have:

$$\Box\forall x(\neg\Diamond E!x \equiv \Box\neg E!x) \to \Box\forall x([\lambda y\, \neg\Diamond E!y]x \equiv [\lambda y\, \Box\neg E!y]x)$$

Note that this doesn't mean that one can substitute $[\lambda y\, \Box\neg E!y]$ for $[\lambda y\, \neg\Diamond E!y]$; this theorem doesn't assert that $[\lambda y \Box\neg E!y] = [\lambda y \neg\Diamond E!y]$ whenever $\Box\forall y(\Box\neg E!y \equiv \Diamond\neg E!y]$.

(**123**) **Theorems:** Strengthened β-Conversion. Though we stated our axiom of β-Conversion (36.2) governing λ-expressions using (a) specific object-language variables in the expression $[\lambda y_1 \ldots y_n\, \varphi^*]$ and (b) exemplification predications of the form $[\lambda y_1 \ldots y_n\, \varphi^*]x_1 \ldots x_n$, we can derive β-Conversion for: (i) any λ-expressions, (ii) exemplifications involving any variables other that $y_1, \ldots, y_n$ that are bound by the λ and (iii) any variables other than $x_1, \ldots, x_n$ in the exemplification predication.

Let $\mu_1, \ldots, \mu_n$ be any distinct individual variables, and let $\nu_1, \ldots, \nu_n$ be any individual variables. Strengthened β-Conversion then states:

> $[\lambda \mu_1 \ldots \mu_n\, \varphi^*]\nu_1 \ldots \nu_n \equiv \varphi^{*\,\nu_1,\ldots,\nu_n}_{\;\mu_1,\ldots,\mu_n}$, provided the ν_i are substitutable, respectively, for the μ_i in φ^* $(1 \leq i \leq n)$.

(**124**) **Theorems:** Comprehension Principle for Properties and Relations. The following is a theorem schema derivable from β-Conversion (36.2):

(.1) $\exists F^n \Box \forall x_1 \ldots \forall x_n (F^n x_1 \ldots x_n \equiv \varphi^*)$, where $n \geq 1$ and φ^* is any propositional formula in which F^n is not free.

When $n = 1$, we have a comprehension condition on properties:

(.2) $\exists F \Box \forall x (Fx \equiv \varphi^*)$, where φ^* is any propositional formula in which F is not free.

Since (124.1) and (124.2) provide existence conditions for relations and properties, respectively, and the definitions of $F^n = G^n$ $(n \geq 2)$ and $F^1 = G^1$ in (16.2) and (16.1) provide respective identity conditions, these principles constitute a rigorous theory of relations and properties. As noted previously, not only are there two senses in which properties can be *materially equivalent*, namely $\forall x(Fx \equiv Gx)$ and $\forall x(xF \equiv xG)$, but also there are two senses in which they can be *necessarily equivalent*, namely, $\Box \forall x(Fx \equiv Gx)$ and $\Box \forall x(xF \equiv xG)$. Necessary equivalence in the former sense does not guarantee property identity, whereas necessary equivalence in the latter sense does, by (16.1).

(**125**) **Theorem:** Equivalence and Identity of Properties. The following theorem establishes that our theory of properties provides us with an *extensional* theory of intensional entities, since it shows that all we need to do to prove the identity of two properties is establish that they are materially equivalent in the encoding sense:

$$\forall x(xF \equiv xG) \rightarrow F = G$$

Thus, to establish $F = G$, it suffices to establish that F and G are encoded by the same objects.

(**126**) **Theorems:** η- and β-Conversion for 0-place Relation Terms. We have as *theorem* schemata the following η-Conversion and β-Conversion principles for any complex 0-place relation term $[\lambda\, \varphi^*]$ and propositional formula φ^*:

(.1) $[\lambda\, \varphi^*] = \varphi^*$

(.2) $[\lambda\, \varphi^*] \equiv \varphi^*$

In (.1), $[\lambda\, \varphi^*]$ and φ^* are being used as terms and are read accordingly as such: *that*-φ^* is identical to φ^*. In (.2), $[\lambda\, \varphi^*]$ and φ^* are being used as formulas and are read differently: *that*-φ^* is true if and only if φ^*, With (126.2), we have established that the propositional version of the Tarski T-Schema is a *theorem* (Zalta 2014).

(**127**) **Remark:** On the Derivability of Tautologies. The derivation of (126.1) and (126.2) allows us to finish the proof of Metatheorem ⟨9.1⟩. This metatheorem is the key lemma needed for the proof of Metatheorem ⟨9.2⟩, i.e., that all tautologies are derivable. See the proofs of these metatheorems in the Appendix to this chapter. Note that from these metatheorems, Metatheorem ⟨9.3⟩ also follows, namely, that every tautology is necessary.

(**128**) **Theorems:** Comprehension Principle for Propositions. The following comprehension principle for propositions is a consequence of (126.2), by RN and ∃I:

$\exists p \Box (p \equiv \varphi^*)$,
where φ^* is any propositional formula with no free occurrences of p.

This comprehension principle and definition (16.3) jointly offer a precise theory of propositions or, if you prefer, *states of affairs*.[109] The claim that propositions are necessarily equivalent, i.e., $\Box(p \equiv q)$, does not entail that p and q are identical. For some propositions p, q, one may consistently assert, that $\Box(p \equiv q)\,\&\, p \neq q$.

[109] We may, for present purposes, treat propositions and states of affairs as the same entities. Ontologically speaking, it is probably better to conceive of 0-place relations as *states of affairs*, since 'proposition' sometimes has a connotation on which it signifies a linguistic or mental entity. I've used the term 'proposition' mostly as a matter of economy.

In a type-theoretic version of our system, which allows one to represent beliefs as higher-order relations between individuals and 0-place relations, there is more of a reason to distinguish between propositions and states of affairs. In the type theory developed in Chapter 15 (Section 15.2), there are entities of type $\langle i, \langle\,\rangle\rangle$ that relate entities of type i (the type for individuals) to entities of type $\langle\,\rangle$ (the type for 0-place relations). One may suppose that the *de re* objects of belief are *states of affairs*. These are entities of type $\langle\,\rangle$ whose constituents are the denotations of the terms inside the belief context. By contrast, whenever the objects of belief are entities of type $\langle\,\rangle$ whose constituents are the senses of the terms inside the belief context, we may suppose that these objects of *de dicto* beliefs are propositions.

Thus, when it comes time to represent beliefs as higher-order relations, a proposition becomes an entity of type $\langle\,\rangle$ having abtract constituents (e.g., abstract individuals or abstract relations) that *represent* other relations and individuals. These abstract individuals and abstract relations serve as the senses of individual terms and relation terms, respectively. In the present 'second-order' setting, however, where higher-order relations between individuals and 0-place relations are not expressible, we may forego the distinction between propositions and states of affairs and leave it for another occasion.

(**129**) **Remark:** Identity of Relations. The theory of identity for n-place relations ($n \geq 0$) asserted in (16) takes on new significance in light of the Comprehension Principles for Relations, Properties, and Propositions. Given the present context, one can look back on the definitions in (16) and see that they are designed to answer one question in a precise way: under what conditions are relations identical? An answer to this questions tells us what we have to prove to establish that $F^n = G^n$, for any n. The theory does not take a stand on the question 'Is R identical to S?' for arbitrary relation expressions 'R' and 'S' of natural language. In some cases, one can prove that R and S are distinct, e.g., when it is provable that possibly, R and S aren't equivalent (see the next theorem). But in cases of necessarily equivalent relations, the matter is often open. For example, if one has a good reason to think that *being red and round* is identical to *being round and red*, one may, when the theory is applied, assert this identity.[110] By the same token, if one has good reason to think these properties are distinct, one may assert their distinctness, in the defined sense that there is an (abstract) object that encodes the one without encoding the other (and vice versa). Thus, we are free to answer the question, 'How fine-grained are relations?' in ways that match our intuitions. But once one adds the assertion that particular properties F and G are identical, then our theory guarantees that there are no abstract objects that encode the one without encoding the other.

(**130**) **Theorems:** Fact about Property Non-identity. Properties that might fail to be materially equivalent are distinct:

$$\Diamond\neg\forall x(Fx \equiv Gx) \rightarrow F \neq G$$

(**131**) **Term Definitions:** Definition of Relation Negation. We introduce $\overline{F^n}$ ('being non-F^n') to abbreviate *being* $y_1,\ldots,y_n$ *such that it is not the case that* $F^n y_1 \ldots y_n$ (for $n \geq 1$), and introduce $\overline{p}$ to abbreviate *that it is not the case that* p, as follows:

(.1) $\overline{F^n} =_{df} [\lambda y_1 \ldots y_n\, \neg F^n y_1 \ldots y_n]$ $\quad (n \geq 1)$

(.2) $\overline{p} =_{df} [\lambda\, \neg p]$

So, when F is a property, $\overline{F} = [\lambda y\, \neg Fy]$, and when R is a 2-place relation, $\overline{R} = [\lambda yz\, \neg Ryz]$.

(**132**) **Theorems:** Relations and their Negations.

(.1) $\overline{F^n}x_1 \ldots x_n \equiv \neg F^n x_1 \ldots x_n$ $\quad (n \geq 1)$

(.2) $\neg\overline{F^n}x_1 \ldots x_n \equiv F^n x_1 \ldots x_n$ $\quad (n \geq 1)$

[110] Indeed, if one has a strong intuition about this kind of case, one could assert: $\forall F \forall G([\lambda x\, Fx \,\&\, Gx] = [\lambda x\, Gx \,\&\, Fx])$.

(.3) $\overline{p} \equiv \neg p$

(.4) $\neg\overline{p} \equiv p$

(.5) $F^n \neq \overline{F^n}$ $(n \geq 1)$

(.6) $p \neq \overline{p}$

Note that we also have:

(.7) $\overline{p} = \neg p$

(.8) $p \neq \neg p$

(**133**) **Definitions:** Noncontingent and Contingent Relations. Let us say: (.1) a relation F^n $(n \geq 0)$ is *necessary* just in case necessarily, all objects $x_1,\ldots,x_n$ are such that $x_1,\ldots,x_n$ exemplify F^n; (.2) F^n is *impossible* just in case necessarily, all objects $x_1,\ldots,x_n$ are such that $x_1,\ldots,x_n$ fail to exemplify F^n; (.3) F^n is *noncontingent* whenever it is necessary or impossible; and (.4) F^n is *contingent* whenever it is neither necessary nor impossible:

(.1) $\mathit{Necessary}(F^n) =_{df} \Box\forall x_1 \ldots \forall x_n F^n x_1 \ldots x_n$ $(n \geq 0)$

(.2) $\mathit{Impossible}(F^n) =_{df} \Box\forall x_1 \ldots \forall x_n \neg F^n x_1 \ldots x_n$ $(n \geq 0)$

(.3) $\mathit{NonContingent}(F^n) =_{df} \mathit{Necessary}(F^n) \vee \mathit{Impossible}(F^n)$ $(n \geq 0)$

(.4) $\mathit{Contingent}(F^n) =_{df} \neg(\mathit{Necessary}(F^n) \vee \mathit{Impossible}(F^n))$ $(n \geq 0)$

(**134**) **Theorems:** Facts about Noncontingent and Contingent Properties. If we focus just on properties as opposed to n-place relations generally, the following facts can be established:

(.1) $\mathit{NonContingent}(F^1) \equiv \mathit{NonContingent}(\overline{F^1})$

(.2) $\mathit{Contingent}(F) \equiv \Diamond\exists x Fx \,\&\, \Diamond\exists x \neg Fx$

(.3) $\mathit{Contingent}(F^1) \equiv \mathit{Contingent}(\overline{F^1})$

(**135**) **Theorems:** Some Noncontingent Properties. Let $L = [\lambda x\, E!x \rightarrow E!x]$ ('being concrete if concrete'). Then we have:

(.1) $\mathit{Necessary}(L)$

(.2) $\mathit{Impossible}(\overline{L})$

(.3) $\mathit{NonContingent}(L)$

(.4) $\mathit{NonContingent}(\overline{L})$

(.5) $\exists F \exists G(F \neq G \,\&\, \mathit{NonContingent}(F) \,\&\, \mathit{NonContingent}(G))$,
i.e., there are at least two noncontingent properties.

(**136**) **Lemmas:** A Symmetry. It is possible that something exemplifies F but might not have if and only if it is possible that something does not exemplify F but might have:

$$\Diamond \exists x(Fx \,\&\, \Diamond \neg Fx) \equiv \Diamond \exists x(\neg Fx \,\&\, \Diamond Fx)$$

If we think semantically for the moment and take possible worlds as primitive entities, then this lemma tells us:

> There is a world where something both exemplifies F and, at some (other) world, fails to exemplify F
> if and only if
> there is a world where something both fails to exemplify F and, at some (other) world, exemplifies F.

So, semantically, the symmetry involves possible worlds. Algebraically, however, the symmetry is between F and its negation $\overline{F}$, since the above is equivalent to:

$$\Diamond \exists x(Fx \,\&\, \Diamond \neg Fx) \equiv \Diamond \exists x(\overline{F}x \,\&\, \Diamond \neg \overline{F}x)$$

once we apply (132.1), (132.2) and the Rule of Substitution to the right-hand condition.

(**137**) **Theorems:** $E!$ and $\overline{E!}$ are Contingent Properties.

(.1) $\Diamond \exists x(\neg E!x \,\&\, \Diamond E!x)$

(.2) $\mathit{Contingent}(E!)$

(.3) $\mathit{Contingent}(\overline{E!})$

(.4) $\exists F \exists G(F \neq G \,\&\, \mathit{Contingent}(F) \,\&\, \mathit{Contingent}(G))$,
i.e., there are at least two contingent properties.

(**138**) **Theorems:** Facts about Property Existence. Where we continue to use L to abbreviate $[\lambda x\, E!x \rightarrow E!x]$, we have the following general and specific facts about the existence of properties:

(.1) $\mathit{NonContingent}(F) \rightarrow \neg \exists G(\mathit{Contingent}(G) \,\&\, G = F)$

(.2) $\mathit{Contingent}(F) \rightarrow \neg \exists G(\mathit{NonContingent}(G) \,\&\, G = F)$

(.3) $L \neq \overline{L} \,\&\, L \neq E! \,\&\, L \neq \overline{E!} \,\&\, \overline{L} \neq E! \,\&\, \overline{L} \neq \overline{E!} \,\&\, E! \neq \overline{E!}$,
i.e., L, $\overline{L}$, $E!$, and $\overline{E!}$ are pairwise distinct

(.4) There are at least four properties.

(**139**) **Theorems:** Facts about Noncontingent and Contingent Propositions. If we focus now just on propositions, the following facts can be established:

(.1) $NonContingent(p) \equiv NonContingent(\overline{p})$

(.2) $Contingent(p) \equiv \Diamond p \,\&\, \Diamond\neg p$

(.3) $Contingent(p) \equiv Contingent(\overline{p})$

(**140**) **Theorems:** Some Noncontingent Propositions. Let $p_0 = \forall x(E!x \rightarrow E!x)$. Then we have:

(.1) $Necessary(p_0)$

(.2) $Impossible(\overline{p_0})$

(.3) $NonContingent(p_0)$

(.4) $NonContingent(\overline{p_0})$

(.5) $\exists p \exists q(p \neq q \,\&\, NonContingent(p) \,\&\, NonContingent(q))$,
i.e., there are at least two noncontingent propositions.

(**141**) **Theorems:** Some Contingent Propositions. Let $q_0 = \exists x(E!x \,\&\, \Diamond\neg E!x)$. Then axiom (32.5) becomes $\Diamond q_0 \& \Diamond\neg q_0$. The following claims regarding propositions thus become derivable:

(.1) $\exists p(\Diamond p \,\&\, \Diamond\neg p)$

(.2) $Contingent(q_0)$

(.3) $Contingent(\overline{q_0})$

(.4) $\exists p \exists q(p \neq q \,\&\, Contingent(p) \,\&\, Contingent(q))$,
i.e., there are at least two contingent propositions.

(**142**) **Theorems:** Facts about Proposition Existence. Where we continue to use p_0 to abbreviate $\forall x(E!x \rightarrow E!x)$ and q_0 to abbreviate $\exists x(E!x \,\&\, \Diamond\neg E!x)$, we have the following general and specific facts about the existence of propositions:

(.1) $NonContingent(p) \rightarrow \neg\exists q(Contingent(q) \,\&\, q = p)$

(.2) $Contingent(p) \rightarrow \neg\exists q(NonContingent(q) \,\&\, q = p)$

(.3) $p_0 \neq \overline{p_0} \,\&\, p_0 \neq q_0 \,\&\, p_0 \neq \overline{q_0} \,\&\, \overline{p_0} \neq q_0 \,\&\, \overline{p_0} \neq \overline{q_0} \,\&\, q_0 \neq \overline{q_0}$,
i.e., p_0, $\overline{p_0}$, q_0, and $\overline{q_0}$ are pairwise distinct

(.4) There are at least four propositions.

(**143**) **Theorems:** $O!$ and $A!$ Are Contingent. The properties *being ordinary* and *being abstract* are distinct, contradictory contingent properties:

(.1) $O! \neq A!$

(.2) $O!x \equiv \neg A!x$

(.3) $A!x \equiv \neg O!x$

(.4) $Contingent(O!)$

(.5) $Contingent(A!)$

Moreover, the negations of *being ordinary* and *being abstract* are distinct and contradictory contingent properties:

(.6) $\overline{O!} \neq \overline{A!}$

(.7) $\overline{O!}x \equiv \neg\overline{A!}x$

(.8) $Contingent(\overline{O!})$

(.9) $Contingent(\overline{A!})$

(**144**) **Theorems:** Further Facts about Being Ordinary and Being Abstract.

(.1) $O!x \rightarrow \Box O!x$

(.2) $A!x \rightarrow \Box A!x$

(.3) $\Diamond O!x \rightarrow O!x$

(.4) $\Diamond A!x \rightarrow A!x$

(.5) $\Diamond O!x \equiv \Box O!x$

(.6) $\Diamond A!x \equiv \Box A!x$

(.7) $O!x \equiv \mathcal{A}O!x$

(.8) $A!x \equiv \mathcal{A}A!x$

The last two theorems are interesting because they have modally-strict proofs, and so don't require an appeal to the axiom of actuality (30)★.

(**145**) **Definition:** Weakly Contingent Properties. We say that a property F is *weakly contingent* just in case F is contingent and anything that possibly exemplifies F necessarily exemplifies F:

$$WeaklyContingent(F) =_{df} Contingent(F) \,\&\, \forall x(\Diamond Fx \rightarrow \Box Fx)$$

(**146**) **Theorems:** Facts about Weakly Contingent Properties. (.1) F is weakly contingent iff $\overline{F}$ is weakly contingent; (.2) if F is weakly contingent and G is not, then F is not G:

(.1) $WeaklyContingent(F) \equiv WeaklyContingent(\overline{F})$

(.2) $(WeaklyContingent(F) \,\&\, \neg WeaklyContingent(G)) \rightarrow F \neq G$

(**147**) **Theorems:** Facts about $O!$, $A!$, $E!$, and L. If we continue to use L to abbreviate $[\lambda x\, E!x \rightarrow E!x]$ (i.e., *being concrete if concrete*), we have the following facts: (.1) *being ordinary* is weakly contingent; (.2) *being abstract* is weakly contingent; (.3) *being concrete* is not weakly contingent; (.4) *being concrete if concrete* is not weakly contingent; (.5) *being ordinary* is distinct from: $E!$, $\overline{E!}$, L, and $\overline{L}$; (.6) *being abstract* is distinct from: $E!$, $\overline{E!}$, L, and $\overline{L}$; (.7) There are at least six distinct properties:

(.1) $WeaklyContingent(O!)$

(.2) $WeaklyContingent(A!)$

(.3) $\neg WeaklyContingent(E!)$

(.4) $\neg WeaklyContingent(L)$

(.5) $O! \neq E! \,\&\, O! \neq \overline{E!} \,\&\, O! \neq L \,\&\, O! \neq \overline{L}$

(.6) $A! \neq E! \,\&\, A! \neq \overline{E!} \,\&\, A! \neq L \,\&\, A! \neq \overline{L}$

(.7) There are at least six properties.

We can show that (.5) holds for $\overline{O!}$ (i.e., $\overline{O!}$ is distinct from $E!$, $\overline{E!}$, L, and $\overline{L}$), and that (.6) holds for $\overline{A!}$. But since our axioms and definitions don't appear to imply $O! \neq \overline{A!}$ or imply $A! \neq \overline{O!}$, we can't prove that there are eight distinct properties.

(**148**) **Theorems:** Identity$_E$, Necessity, and Possibility. (.1) objects x and y are identical$_E$ if and only if they are necessarily identical$_E$; and (.2) objects x and y are possibly identical$_E$ if and only if they are identical$_E$:

(.1) $x =_E y \equiv \Box x =_E y$

(.2) $\Diamond x =_E y \equiv x =_E y$

(**149**) **Term and Formula Definitions:** Distinctness$_E$. We now introduce a new 2-place relation term *being distinct*$_E$:

(.1) $\neq_E =_{df} \overline{=_E}$

Here the definiens combines defined notation: it combines the defined negation overbar '$\overline{}$' (131.1) with the defined term $=_E$ (12). So $\neq_E$ is defined as: *being an x and y such that it is not the case that x bears $=_E$ to y*. To complement our term definition, we also introduce the following formula definition for infix notation:

(.2) $x \neq_E y \ =_{df} \neq_E xy$

These definitions play an important role when we develop the theory of Fregean numbers in Chapter 14: not only do λ-expressions such as $[\lambda x\, x \neq_E a]$ play a role in the definition of the *Predecessor* relation but also the modal properties of the $\neq_E$ relation are crucial to the derivation of the Dedekind-Peano axioms for number theory as theorems of the present theory.

It might come as a surprise just how many definitions must be unpacked in order to expand $[\lambda x\, x \neq_E a]$ into primitive notation and the defined & symbol:

$[\lambda x \ x \neq_E a]$

$= [\lambda x \ \neq_E xa]$ By (149.2)

$= [\lambda x\, \overline{=_E}xa]$ By (149.1)

$= [\lambda x\, [\lambda y_1 y_2\, \neg(=_E y_1 y_2)]xa]$ By (131.1)

$= [\lambda x\, [\lambda y_1 y_2\, \neg[\lambda xy\, O!x \,\&\, O!y \,\&\, \Box\forall F(Fx \equiv Fy)]y_1 y_2]xa]$ By (12)

$= [\lambda x\, [\lambda y_1 y_2\, \neg[\lambda xy\, [\lambda x\, \Diamond E!x]x \,\&\, [\lambda x\, \Diamond E!x]y \,\&\, \Box\forall F(Fx \equiv Fy)]y_1 y_2]xa]$ By (11.1)

$= [\lambda x\, [\lambda y_1 y_2\, \neg[\lambda xy\, [\lambda x\, \neg\Box\neg E!x]x \,\&\, [\lambda x\, \neg\Box\neg E!x]y \,\&\, \Box\forall F(Fx \equiv Fy)]y_1 y_2]xa]$ By (7.4.e)

$= [\lambda x\, [\lambda y_1 y_2\, \neg[\lambda xy\, [\lambda x\, \neg\Box\neg E!x]x \,\&\, [\lambda x\, \neg\Box\neg E!x]y \,\&\, \Box\forall F((Fx \to Fy) \,\&$
$(Fy \to Fx))]y_1 y_2]xa]$ By (7.4.c)

Fortunately, the proof of those theorems below that involve the formula $x \neq_E y$ do not require that we dig quite as deep into its chain of definitions.

(**150**) **Theorem:** Equivalence of Non-identity$_E$ and Not Identical$_E$.

$$x \neq_E y \equiv \neg(x =_E y)$$

Note that this fact is not an immediate consequence of a single formula definition, but instead requires that we cite a number of term and formula definitions, as well as principles like β-Conversion.

(**151**) **Theorems:** Non-identity$_E$, Necessity, and Possibility. (.1) objects x and y are distinct$_E$ if and only if necessarily they are distinct$_E$, and (.2) objects x and y are possibly distinct$_E$ if and only if then they are distinct$_E$:

(.1) $x \neq_E y \equiv \Box x \neq_E y$

(.2) $\Diamond x \neq_E y \equiv x \neq_E y$

(**152**) **Theorems:** Identity$_E$, Identity, and Actuality.

(.1) $x =_E y \equiv \mathcal{A}x =_E y$

(.2) $x \neq_E y \equiv \mathcal{A}x \neq_E y$

(.3) $\alpha = \beta \equiv \mathcal{A}\alpha = \beta$, where α, β are variables of the same type

(.4) $\alpha \neq \beta \equiv \mathcal{A}\alpha \neq \beta$, where α, β are variables of the same type

It is important to observe here that these modally-strict theorems, provable *without* the necessitation-averse axiom (30)★ of actuality.

(**153**) **Theorem:** A Distinguished Description Involving Identity. The following theorem is a straightforward consequence of the preceding theorem:

$$y = \imath x(x = y)$$

This asserts: y is identical to the individual that (in fact) is identical to y.

This is a modally-strict theorem and though it appears to be trivial, keep in mind that identity is a defined notion and that definite descriptions are rigid and axiomatized in terms of the actuality operator. Notice that Frege uses a somewhat more complicated version of this claim as an *axiom* governing his definite description operator. In Frege 1893 (§18), Frege asserts Basic Law VI, which we may write as: $y = \backslash\acute{\epsilon}(\epsilon = y)$. On Frege's understanding of the description operator, Law VI basically asserts that y is identical to the unique member of the extension of the concept *being identical to* y. Our theorem, by contrast asserts: y is identical to the unique object x that is identical to y. No reference to the extension of a concept is needed.

(**154**) **Definitions:** η-Variants. We now work our way towards theorems (155) and (156), which constitute the wider implications of the axiom of η-Conversion. Let ρ be a complex n-place relation term ($n \geq 0$). Then we say:

(.1) ρ is *elementary* if and only if ρ has the form $[\lambda \nu_1 \dots \nu_n\, \Pi^n \nu_1 \dots \nu_n]$, where Π^n is any n-place relation term and $\nu_1, \dots, \nu_n$ are distinct individual variables none of which occur free in Π^n.

So when $n = 0$, $[\lambda\, \Pi^0]$ is elementary. Furthermore, we say:

(.2) ρ is an *η-expansion of* Π^n if and only if ρ is the elementary λ-expression $[\lambda \nu_1 \dots \nu_n\, \Pi^n \nu_1 \dots \nu_n]$

(.3) Π^n is the *η-contraction of* ρ if and only if ρ is the elementary λ-expression $[\lambda \nu_1 \dots \nu_n\, \Pi^n \nu_1 \dots \nu_n]$

Note that ρ may have many different η-expansions, depending on the choice of $\nu_1, \ldots, \nu_n$, but an elementary λ-expression ρ can have only one η-contraction.

Before we proceed, note that we again will be using the prime symbol $'$ to help us define η-variants. Thus, we are deploying this symbol in yet a third way. Previously, we've used the symbol to help us formulate the Substitution of Identicals (25) and to talk about alphabetic variants (35). But no confusion should arise, since the context shall make it clear how the prime symbol $'$ is being used.

Next, then, where ρ and ρ' are n-place relation terms ($n \geq 0$), we say:

(.4) ρ' is an *immediate* (i.e., one-step) η-variant of ρ just in case ρ' results from ρ either (a) by replacing one n-place relation term Π^n in ρ by an η-expansion $[\lambda \nu_1 \ldots \nu_n \, \Pi^n \nu_1 \ldots \nu_n]$ or (b) by replacing one elementary λ-expression $[\lambda \nu_1 \ldots \nu_n \, \Pi^n \nu_1 \ldots \nu_n]$ in ρ by its η-contraction Π^n.

Note that this definition even applies in the case where ρ and ρ' are the 0-place propositional formulas φ^* and $\varphi^{*\prime}$, since both are relation terms. Thus, $[\lambda \, \Pi^0]$ and Π^0 are immediate η-variants of each other. Clearly, in general, if ρ' is an immediate η-variant of ρ, then ρ is an immediate η-variant of ρ'.

Finally we say, for n-place relation terms ρ and ρ' ($n \geq 0$):

(.5) ρ' *is an* η-variant of ρ whenever there is a sequence of n-place relation terms $\rho_1, \ldots, \rho_m$ ($m \geq 1$) with $\rho = \rho_1$ and $\rho' = \rho_m$ such that every member of the sequence is an immediate η-variant of the preceding member of the sequence.

Thus, the relation *is an η-variant of* is the transitive closure of the relation *is an immediate η-variant of*.[111] We now illustrate these definitions with examples:

Examples of Elementary η-Variants (Expansion/Contraction Pairs):

- $[\lambda xyz \, F^3 xyz]$ / F^3
- $[\lambda x \, [\lambda y \, \neg Fy]x]$ / $[\lambda y \, \neg Fy]$
- $[\lambda xy \, [\lambda uv \, \Box \forall F(Fu \equiv Fv)]xy]$ / $[\lambda uv \, \Box \forall F(Fu \equiv Fv)]$
- $[\lambda \, p]$ / p
- $[\lambda \, \neg Pa]$ / $\neg Pa$

Immediate η-Variants

- All of the above

[111] Intuitively, the transitive closure of R is that relation R' that relates any two elements in a chain of R-related elements. That is, for any elements x and y of a domain of R-related elements, $xR'y$ holds whenever there exist $z_0, z_1, \ldots, z_n$ such that (i) $z_0 = x$, (ii) $z_n = y$, and (iii) for all $0 \leq i < n$, $z_i R z_{i+1}$.

- $[\lambda y\, [\lambda z\, Pz]y \rightarrow Say]$ / $[\lambda y\, Py \rightarrow Say]$
- $[\lambda y\, Py \rightarrow [\lambda uv\, Suv]ay]$ / $[\lambda y\, Py \rightarrow Say]$
- $[\lambda y\, [\lambda z\, Pz]y \rightarrow [\lambda uv\, Suv]ay]$ / $[\lambda y\, [\lambda z\, Pz]y \rightarrow Say]$
- $[\lambda y\, [\lambda z\, Pz]y \rightarrow [\lambda uv\, Suv]ay]$ / $[\lambda y\, Py \rightarrow [\lambda uv\, Suv]ay]$
- $[\lambda y\, [\lambda\, p]]$ / $[\lambda y\, p]$
- $[\lambda x_1 \ldots x_n\, [\lambda\, Pa]]$ / $[\lambda x_1 \ldots x_n\, Pa]$
- $[\lambda z\, Pz]y \rightarrow Say$ / $Py \rightarrow Say$
- $[\lambda\, [\lambda z\, Pz]y]$ / $[\lambda\, Py]$

η-**Variant Pairs**

- All of the above
- $[\lambda y\, [\lambda z\, Pz]y \rightarrow [\lambda uv\, Suv]ay]$ / $[\lambda y\, Py \rightarrow Say]$
- $[\lambda z\, Pz]y \rightarrow [\lambda uv\, Suv]ay$ / $Py \rightarrow Say$
- $[\lambda y\, [\lambda z\, Pz]y \rightarrow Say]$ / $[\lambda y\, Py \rightarrow [\lambda uv\, Suv]ay]$

With a thorough grip on the notion of η-variants, we may prove the following theorems.

(**155**) **Lemmas:** Useful Facts about η-Conversion.

(.1) $[\lambda x_1 \ldots x_n\, \Pi^n x_1 \ldots x_n] = \Pi^n$, where Π^n is any n-place relation term ($n \geq 0$) in which none of $x_1, \ldots, x_n$ occur free

(.2) $[\lambda \nu_1 \ldots \nu_n\, \Pi^n \nu_1 \ldots \nu_n] = \Pi^n$, where $[\lambda \nu_1 \ldots \nu_n\, \Pi^n \nu_1 \ldots \nu_n]$ is any elementary λ-expression and $\nu_1, \ldots, \nu_n$ are any distinct individual variables none of which occur free in Π^n.[112]

(.3) $\rho = \rho'$, whenever ρ' is an *immediate* η-variant of ρ

(**156**) **Theorems:** η-Conversion for arbitrary η-Variants. Recall that ρ is a metavariable ranging over n-place relation terms. Then we have:

$\rho = \rho'$, where ρ' is any η-variant of ρ

As an instance of this theorem schema, we have following identity claim, where the λ-expression on the left-hand side of the identity symbol contains two embedded λ-expressions:

$$[\lambda y\, [\lambda z\, Pz]y \rightarrow [\lambda uv\, Suv]ay] = [\lambda y\, Py \rightarrow Say]$$

[112] This last clause is, strictly speaking, unnecessary, since we've defined *elementary* λ-expressions above so that in the elementary expression $[\lambda \nu_1 \ldots \nu_n\, \Pi^n \nu_1 \ldots \nu_n]$, none of the ν_i occur free in Π^n.

Here is a proof of this claim:

1.	$[\lambda z\, Pz] = P$	ηC
2.	$[\lambda uv\, Suv] = S$	ηC
3.	$[\lambda y\, [\lambda z\, Pz]y \rightarrow [\lambda uv\, Suv]ay] = [\lambda y\, [\lambda z\, Pz]y \rightarrow [\lambda uv\, Suv]ay]$	ReflId
4.	$[\lambda y\, [\lambda z\, Pz]y \rightarrow [\lambda uv\, Suv]ay] = [\lambda y\, Py \rightarrow [\lambda uv\, Suv]ay]$	SubId, 1, 3
5.	$[\lambda y\, [\lambda z\, Pz]y \rightarrow [\lambda uv\, Suv]ay] = [\lambda y\, Py \rightarrow Say]$	SubId, 2, 4

(**157**) **Theorem:** Propositional Equations.

$$[\lambda\, p] = [\lambda\, q] \equiv p = q$$

(**158**) **Definition:** Propositional Properties. Let us call a property F *propositional* iff for some proposition p, F is *being such that* p:

$$\mathit{Propositional}(F) \ =_{df}\ \exists p(F = [\lambda y\, p])$$

(**159**) **Theorems:** Propositional Properties. For any proposition (or state of affairs) p, the propositional property $[\lambda y\, p]$ exists.

(.1) $\forall p \exists F(F = [\lambda y\, p])$

In general, if F is *being such that* p, then necessarily, an object x exemplifies F iff p is true:

(.2) $F = [\lambda y\, p] \rightarrow \Box\forall x(Fx \equiv p)$

Propositional properties play an extremely important role in some of the applications of object theory in later chapters.

(**160**) **Theorems:** Propositional Properties, Necessity, and Possibility. The following claims about propositional properties can be established: (.1) if it is possible that F is the propositional property $[\lambda y\, p]$, then F is in fact this propositional property; (.2) if F is not a propositional property, then necessarily it isn't; (.3) if F in fact is the propositional property $[\lambda y\, p]$, then necessarily it is; and (.4) if F might not have been a propositional property, then it isn't one in fact:

(.1) $\Diamond\exists p(F = [\lambda y\, p]) \rightarrow \exists p(F = [\lambda y\, p])$

(.2) $\forall p(F \neq [\lambda y\, p]) \rightarrow \Box\forall p(F \neq [\lambda y\, p])$

(.3) $\exists p(F = [\lambda y\, p]) \rightarrow \Box\exists p(F = [\lambda y\, p])$

(.4) $\Diamond\forall p(F \neq [\lambda y\, p]) \rightarrow \forall p(F \neq [\lambda y\, p])$

(**161**) **Theorems:** Encoded Propositional Properties Are Necessarily Encoded. It is provable that: (.1) if it is possible that every property that x encodes is propositional, then in fact every property x encodes is propositional, and (.2) if every property that x encodes is propositional, then necessarily every property x encodes is propositional:

(.1) $\Diamond\forall F(xF \rightarrow \exists p(F=[\lambda y\, p])) \rightarrow \forall F(xF \rightarrow \exists p(F=[\lambda y\, p]))$

(.2) $\forall F(xF \rightarrow \exists p(F=[\lambda y\, p])) \rightarrow \Box\forall F(xF \rightarrow \exists p(F=[\lambda y\, p]))$

9.10 The Theory of Objects

(**162**) **Theorems:** The Domain of Objects is Partitioned. To show the domain of objects is partitioned, we prove two theorems. First, every object is either ordinary or abstract:

(.1) $\forall x(O!x \vee A!x)$

Second, no object is both ordinary and abstract:

(.2) $\neg\exists x(O!x \,\&\, A!x)$

These are modally-strict theorems and, hence, necessary truths.

(**163**) **Theorems:** Identity$_E$ is an Equivalence Relation on Ordinary Objects:

(.1) $O!x \rightarrow x =_E x$

(.2) $x =_E y \rightarrow y =_E x$

(.3) $(x =_E y \,\&\, y =_E z) \rightarrow x =_E z$

(**164**) **Theorem:** Ordinary Objects Obey Leibniz's Law. The identity$_E$ of indiscernible ordinary objects is a theorem:

$$(O!x \,\&\, O!y) \rightarrow (\forall F(Fx \equiv Fy) \rightarrow x =_E y)$$

This tells us that if we have two ordinary objects x, y, then we simply have to show that they exemplify the same properties to conclude they are identical$_E$. We don't have to show that they necessarily exemplify the same properties to establish their identity$_E$. By contraposition, if we know two ordinary objects x, y are distinct$_E$, then we know that there exists a property that distinguishes them.

(**165**) **Theorem:** Distinct Ordinary Objects Have Distinct Haecceities. Recall that $[\lambda y\, y =_E x]$ is the *haecceity of* x. Then it follows that ordinary objects x, z are distinct if and only if their haecceities $[\lambda y\, y =_E x]$ and $[\lambda y\, y =_E z]$ are distinct, i.e.,

$$(O!x \,\&\, O!z) \to (x \neq z \equiv [\lambda y\, y =_E x] \neq [\lambda y\, y =_E z])$$

(**166**) **Theorem:** Abstract Objects Obey a Variant of Leibniz's Law. A variant of Leibniz's Law is now derivable, namely, that whenever abstract objects x, y encode the same properties, they are identical:

$$(A!x \,\&\, A!y) \to (\forall F(xF \equiv yF) \to x = y)$$

Thus to show abstract objects x, y are identical, it suffices to prove that they encode the same properties; we don't have to show that they necessarily encode the same properties.

(**167**) **Theorem:** Ordinary Objects Necessarily Fail to Encode Properties.

$$O!x \to \Box\neg\exists F xF$$

(**168**) **Theorem:** Ordinary Objects Exist. Our system yields the existence of ordinary objects as a theorem:

$$\exists x O!x$$

It is important to note that the claim that ordinary objects exist does not imply the claim that any concrete objects exist, but only that there could be concrete objects. The claim that there exist concrete objects is an empirical claim.

(**169**) **Theorem:** Objects that Encode Properties Are Abstract. It follows by contraposing (39) that if x encodes a property, then x is abstract:

$$\exists F xF \to A!x$$

The converse fails because of there exists an abstract *null* object, which encodes no properties. See theorem (192.1) below.

(**170**) **Theorems:** Strengthened Comprehension. The Comprehension Principle for Abstract Objects (40) and the definition of identity (15) jointly imply that there is a *unique* abstract object that encodes just the properties such that φ:

$\exists! x(A!x \,\&\, \forall F(xF \equiv \varphi))$, provided x doesn't occur free in φ

The proof is simplified by appealing to (166).

(**171**) **Theorems:** Abstract Objects via Strengthened Comprehension. Strengthened Comprehension principle (170) asserts the unique existence of a number of interesting abstract objects. There exists a unique abstract object that encodes all and only the properties F such that: (.1) y exemplifies F; (.2) y and z exemplify F; (.3) y or z exemplify F; (.4) y necessarily exemplifies F; (.5) F is identical to property G; (.6) F is necessarily implied by G; and (.7) F is a propositional property constructed from a true proposition:

(.1) $\exists!x(A!x \,\&\, \forall F(xF \equiv Fy))$

(.2) $\exists!x(A!x \,\&\, \forall F(xF \equiv Fy \,\&\, Fz))$

(.3) $\exists!x(A!x \,\&\, \forall F(xF \equiv Fy \vee Fz))$

(.4) $\exists!x(A!x \,\&\, \forall F(xF \equiv \Box Fy))$

(.5) $\exists!x(A!x \,\&\, \forall F(xF \equiv F = G))$

(.6) $\exists!x(A!x \,\&\, \forall F(xF \equiv \Box\forall y(Gy \rightarrow Fy)))$

(.7) $\exists!x(A!x \,\&\, \forall F(xF \equiv \exists p(p \,\&\, F = [\lambda y\, p])))$

Many of the above objects (and others) will figure prominently in the theorems which follow.

(**172**) **Lemmas:** Actuality and Unique Existence. The actuality operator commutes with the unique existence quantifier:

(.1) $\mathcal{A}\exists!\alpha\varphi \equiv \exists!\alpha\mathcal{A}\varphi$

Furthermore, given (.1), it follows from (104.1) and (104.2), respectively, that:

(.2) $\exists y(y = \iota x\varphi) \equiv \mathcal{A}\exists!x\varphi$, provided y doesn't occur free in φ

(.3) $\exists y(y = \iota x\mathcal{A}\varphi) \equiv \mathcal{A}\exists!x\varphi$, provided y doesn't occur free in φ

(.2) is especially important. If we can establish a claim of the form $\exists!x\varphi$ by way of a modally strict proof, then the Rule of Actualization (RA) yields a modally-strict proof of $\mathcal{A}\exists!x\varphi$. Then, by (.2), we can derive $\exists y(y = \iota x\varphi)$ as a modally-strict theorem.

(**173**) **Theorems:** Some Descriptions Guaranteed to be Logically Proper. It now follows, by a modally-strict proof, that for any condition φ in which x doesn't occur free, there exists something which is identical to the individual that is both abstract and encodes just the properties such that φ:

$$\exists y(y = \iota x(A!x \,\&\, \forall F(xF \equiv \varphi))), \text{ provided } x, y \text{ don't occur free in } \varphi$$

Although the above theorem is an immediate consequence of (170) and (96)⋆, the resulting proof wouldn't be modally-strict. There is, however, a modally-strict proof that appeals to (172.2).

Since the above schema has a modally-strict proof, its necessitation follows by RN. If we think semantically for a moment, and treat possible worlds as semantically primitive entities, it becomes clear that the necessitation of our schema does *not* say that at every world $\boldsymbol{w}$, there exists something which is identical to the x such that, at $\boldsymbol{w}$, x both exemplifies being abstract and encodes all and only the properties satisfying φ (at $\boldsymbol{w}$). Rather, the necessitation says that at every possible world, there exists something identical to the x such that, at the distinguished actual world $\boldsymbol{w}_0$, x both exemplifies being abstract and encodes exactly the properties satisfying φ (at $\boldsymbol{w}_0$).

(**174**) **Metadefinitions:** Canonical Descriptions, Matrices, and Canonical Individuals. The previous theorem guarantees that descriptions of the form $\iota x(A!x \& \forall F(xF \equiv \varphi))$ are logically proper (when x isn't free in φ). We henceforth say:

> A definite description is *canonical* iff it has the form $\iota\nu(A!\nu \& \forall F(\nu F \equiv \varphi))$, for some formula φ in which the individual variable ν doesn't occur free.

We call the matrix of a canonical description a *canonical matrix*. By a slight abuse of language, we call the individuals denoted by such descriptions *canonical* individuals.

(**175**) ⋆**Theorems:** The Abstraction Principle. Canonical individuals are governed by a simple principle, albeit one that is not subject to the Rule of Necessitation. To state this principle, note that when φ is a formula in which x doesn't occur free and in which G is substitutable for F, the formula φ^G_F (i.e., the result of substituting G for every free occurrence of F in φ) asserts that G is such that φ. Here, we may think of φ^G_F as a syntactic representation of the semantic claim that G *satisfies* φ, where this particular notion of satisfaction was defined in Section 5.5.1, footnote 44. It is now provable that: the abstract object encoding exactly the properties such that φ encodes a property G if and only if G is such that φ:

> $\iota x(A!x \& \forall F(xF \equiv \varphi))G \equiv \varphi^G_F$, provided x doesn't occur free in φ and G is substitutable for F in φ

To see an example, let b be an arbitrary object. Then we have the following instance of the above theorem: $\iota x(A!x \& \forall F(xF \equiv Fb))G \equiv Gb$. This asserts: the abstract object that encodes exactly the properties that b exemplifies encodes G iff b exemplifies G. The reason we call this 'abstraction' should be clear: in the right-to-left direction, we've abstracted out, from the simple predication Gb, an encoding claim about a particular abstract object.

Intuitively, we may describe this example as follows. Since there is a unique logical pattern of predications about b of the form Gb, the Strengthened Comprehension Principle (170) simultaneously *objectifies* this pattern and asserts its unique existence. The Abstraction Principle then yields a claim to the effect that property G matching the pattern is equivalent to an encoding claim about the unique abstract object the pattern defines. This way of looking at the Abstraction Principle applies to any pattern φ of arbitrary complexity, provided it doesn't contain a free variable x.

(**176**) **Remark:** The Abstraction Principle and Necessitation. Inspection shows that the proof of the Abstraction Principle depends upon Hintikka's schema (94)★, which in turn depends on the fundamental theorem (93)★ governing descriptions. So the Abstraction Principle is not a modally strict theorem and is not subject to the Rule of Necessitation. This is an important result, and it is important to understand what underlies it. The result can be explained in either the material mode or in the formal mode. In the text that follows we give the explanation in the material and leave the explanation in the formal mode to a footnote.

To see why RN can't be applied to instances of Abstraction, suppose we've extended our theory in a natrual way way with the following facts: b exemplifies being a philosopher but b might not have been a philosopher. That is, where P is the property of *being a philosopher*, suppose we've extended our theory with:

Fact: Pb

Modal Fact: $\Diamond\neg Pb$

Since Pb implies $\Diamond Pb$, the claim Pb is contingent (indeed, contingently true). Though the above facts may be part of the data (i.e., part of the body of beliefs systematized by our theory), we have the option of either asserting them as axioms (i.e., auxiliary hypotheses) or assuming them as premises for the purposes of some reasoning. In either case, the first fact is clearly necessitation-averse!

Now consider the following instance of Abstraction given the preceding context:

(a) $\imath x(A!x \,\&\, \forall F(xF \equiv Fb))G \equiv Gb$

If one were to apply RN to (a), one would be able to derive the negation of the Modal Fact. To see how, instantiate the variable G to the property P, to obtain:

(b) $\imath x(A!x \,\&\, \forall F(xF \equiv Fb))P \equiv Pb$

It then follows from (b) and our Fact that:

(c) $\imath x(A!x \,\&\, \forall F(xF \equiv Fb))P$

Since the description in (c) is canonical, it is logically proper (173), and so may be instantiated, along with P, in to our axiom for the rigidity of encoding (37), which asserts $xF \to \Box xF$. By doing so, we may infer:

(d) $\Box\imath x(A!x \,\&\, \forall F(xF \equiv Fb))P$

Of course, the derivation of (d) fails to be modally strict, since it was derived from *two* $\star$-claims: the above contingent Fact (Pb) and (175)$\star$.[113]

Nevertheless, (d) is derivable from our Fact. Now if we could apply RN to (b), we would also obtain:

(e) $\Box(\imath x(A!x \,\&\, \forall F(xF \equiv Fb))P \equiv Pb)$

From (e), (d), and the relevant instance of (107.6) ($= \Box(\varphi \equiv \psi) \to (\Box\varphi \equiv \Box\psi)$), it would follow by MP that:

(f) $\Box Pb$

But this contradicts our Modal Fact, namely, $\Diamond\neg Pb$, i.e., $\neg\Box Pb$. Thus, we've shown what would have happened if RN were applicable to (consequences of) the Abstraction Principle — our metaphysical logic would have been incompatible with the modal facts.[114]

---

[113]When we temporarily add the above Fact Pb and Modal Fact ($\Diamond\neg Pb$) either as axioms or as premises, we should, strictly speaking, mark the Fact with $\star$. For if we add the Fact as an axiom, then it has to be labeled as a *necessitation-averse* axiom (its modal closures are *not* to be added as axioms), and (b) if we add it as a premise for the purposes of some reasoning, then its necessitation can't also be a premise (given that we've also added the Modal Fact as a premise). So no matter whether we take our Fact as an axiom or a premise, its use in any reasoning undermines the modal strictness of that reasoning.

[114]In the formal mode, the reason why we can't generally apply RN to instances of Abstraction is that there are interpretations in which necessitations of instances of Abstraction fail to be valid. To see this, let us again help ourselves for the moment to the semantically primitive notion of a possible world. The semantic counterpart of the above Fact and Modal Fact is any interpretation of our language in which Pb is true at the actual world $\boldsymbol{w}_0$ but false at some other possible world, say, $\boldsymbol{w}_1$. In such an interpretation, our consequence (b) of the Abstraction Principle fails to be necessarily true. (b) fails to be necessarily true in the left-to-right direction by the following argument. Since Pb is true at $\boldsymbol{w}_0$, $\imath x(A!x \,\&\, \forall F(xF \equiv Fb))$ encodes P at $\boldsymbol{w}_0$. Since the properties an object encodes are necessarily encoded (121.1), $\imath x(A!x \,\&\, \forall F(xF \equiv Fb))$ encodes P at $\boldsymbol{w}_1$. But, by hypothesis, Pb is false at $\boldsymbol{w}_1$. Hence, $\boldsymbol{w}_1$ is a world that is the witness to the truth of the following possibility claim:

$$\Diamond(\imath x(A!x \,\&\, \forall F(xF \equiv Fb))P \,\&\, \neg Pb)$$

So the left-to-right condition of (b) isn't necessary in this interpretation, and so the necessitation of this condition fails to be valid.

Similarly, if we consider the negation of P, namely $\overline{P}$, with respect to the interpretation described above, then the following consequence of the Abstraction Principle fails to be necessarily true in the right-to-left direction:

(g) $\imath x(A!x \,\&\, \forall F(xF \equiv Fb))\overline{P} \equiv \overline{P}b$

To see that it is possible for the right condition of (g) to be true while the left condition false, we may reason as follows. Since Pb is true at $\boldsymbol{w}_0$ in the interpretation we've described, $\overline{P}b$ is false at $\boldsymbol{w}_0$ and so $\imath x(A!x \& \forall F(xF \equiv Fb))$ fails to encode $\overline{P}$ at $\boldsymbol{w}_0$. Since the properties an abstract object fails to encode are properties it necessarily fails to encode (121.5), it follows that $\imath x(A!x \,\&\, \forall F(xF \equiv Fb))$ fails to exemplify $\overline{P}$ at $\boldsymbol{w}_1$. But, by hypothesis, Pb is false at $\boldsymbol{w}_1$ and so $\overline{P}b$ is true at $\boldsymbol{w}_1$. Hence, $\boldsymbol{w}_1$ is a witness to the truth of the following possibility claim:

$$\Diamond(\overline{P}b \,\&\, \neg\imath x(A!x \,\&\, \forall F(xF \equiv Fb))\overline{P})$$

But this shows that the right-to-left direction of (g) isn't necessarily true in this interpretation, and hence fails to be valid.

It is easy to construct examples like the above that show why instances of the Abstraction Principle are not necessary. But in the next item, we strategically place an actuality operator in the Abstraction Principle and thereby formulate a version of the Abstraction Principle that has a modally strict proof.

(**177**) **Theorems:** Actualized Abstraction. There is, however, a variant of the Abstraction Principle that is modally strict. It asserts that the abstract object encoding just the properties such that φ encodes a property G if and only if it is actually the case that G is such that φ:

> $\imath x(A!x \,\&\, \forall F(xF \equiv \varphi))G \equiv \mathcal{A}\varphi^G_F$, provided x doesn't occur free in φ and G is substitutable for F in φ

The reader is encouraged to show, without the benefit of the proof in the Appendix, that this theorem can be proved without appealing to (30)★ or any other ★-theorem.

(**178**) **Theorems:** Canonical Individuals and Conditionally Necessary Abstraction. There are modally strict proofs of the claims: (.1) if G is necessarily such that φ, then $\imath x(A!x \,\&\, \forall F(xF \equiv \varphi))$ encodes G, and (.2) if G is necessarily such that φ, then necessarily, $\imath x(A!x \,\&\, \forall F(xF \equiv \varphi))$ encodes G if and only if G is such that φ:

(.1) $\Box\varphi^G_F \to \imath x(A!x \,\&\, \forall F(xF \equiv \varphi))G$, provided x doesn't occur free in φ

(.2) $\Box\varphi^G_F \to \Box(\imath x(A!x \,\&\, \forall F(xF \equiv \varphi))G \equiv \varphi^G_F)$

The proof of (.1) in the Appendix bypasses the Abstraction Principle and uses Actualized Abstraction (177) instead. The proof of (.2) utilizes (.1).

(**179**) **Remark:** Some Examples of Encoded Properties. Let φ be the condition $F = G \vee F = H$ and consider the canonical individual that encodes a property F iff F is G or F is H, i.e.,

$$\imath x(A!x \,\&\, \forall F(xF \equiv F = G \vee F = H))$$

Then φ^G_F and φ^H_F are the following formulas:

$$\begin{array}{lcl} \varphi^G_F & = & G = G \vee G = H \\ \varphi^H_F & = & H = G \vee H = H \end{array}$$

In both cases, we can establish the antecedents of the following instances of (178.1):

$$\Box\varphi^G_F \to \imath x(A!x \,\&\, \forall F(xF \equiv F = G \vee F = H))G$$

$$\Box\varphi^H_F \to \imath x(A!x \,\&\, \forall F(xF \equiv F = G \vee F = H))H$$

For example, since it is a theorem that $G = G$ (67), it follows by $\vee$I (64.3.a) that $G = G \vee G = H$, i.e., φ^G_F. Since we've proved this claim without appeal to any $\star$-theorems, it follows by RN that $\Box\varphi^G_F$. Similarly, we can derive $\Box\varphi^H_F$. Hence from these last two results and the instances of (178.1) immediately above, we can derive the following as *modally strict* theorems:

$$\imath x(A!x \,\&\, \forall F(xF \equiv F = G \vee F = H))G$$

$$\imath x(A!x \,\&\, \forall F(xF \equiv F = G \vee F = H))H$$

Contrast these two examples with the following example, in which we again consider the canonical individual discussed in Remark (176), where 'b' denotes some particular philosopher:

$$\imath x(A!x \,\&\, \forall F(xF \equiv Fb))$$

This is the abstract object that encodes exactly the properties b exemplifies. Now let 'P' denote *being a philosopher* and suppose we have extended our theory with the facts mentioned in Remark (176), namely, the Fact that b is a philosopher (Pb) and the Modal Fact that b might not have been a philosopher ($\Diamond\neg Pb$). We saw that the Fact and an appropriate instance of the Abstraction Principle (175)$\star$ implies that $\imath x(A!x \,\&\, \forall F(xF \equiv Fb))P$, but by a proof that isn't modally strict. Note that we cannot obtain a strict proof of this encoding claim by appealing to (178.1), for that only yields the following as an instance:

(a) $\Box Pb \rightarrow \imath x(A!x \,\&\, \forall F(xF \equiv Fb))P$

Given our Modal Fact ($\Diamond\neg Pb$), i.e., $\neg\Box Pb$, we may *not* detach the consequent of (a). Clearly, in general, if φ^G_F fails to be necessary, we may not derive modally strict encoding facts about $\imath x(A!x \,\&\, \forall F(xF \equiv \varphi))$.

Now here are two more examples. Let 'B' denote *being a building*, $\overline{B}$ denote the latter's negation, and L denote the property we studied earlier, namely, $[\lambda x\, E!x \rightarrow E!x]$. Then we have the following instances of (178.1):

(b) $\Box\overline{B}b \rightarrow \imath x(A!x \,\&\, \forall F(xF \equiv Fb))\overline{B}$

(c) $\Box Lb \rightarrow \imath x(A!x \,\&\, \forall F(xF \equiv Fb))L$

Case (b) is interesting, for one might reasonably adopt an essentialist view and extend our theory with the general claim that necessarily, philosophers couldn't be buildings ($\Box\forall x(Px \rightarrow \neg\Diamond Bx)$), from which it would follow that philosopher b couldn't have been a building ($\neg\Diamond Bb$). From the latter modal claim (which is equivalent to $\Box\neg Bb$), it would follow that necessarily b fails to exemplify being a building ($\Box\neg Bb$). Hence, by the modally strict theorem (132.1) and the Rule of Substitution, we could conclude: necessarily b exemplifies *being a non-building* ($\Box\overline{B}b$). Given this derivation of the antecedent of

(b) in an extended theory, one could then further derive the consequent of (b). However, note that the resulting derivation wouldn't be modally strict! The derivation appeals to the Fact Pb, and this Fact is, as discussed previously, necessitative-averse (and so should be marked with a $\star$). Hence, derivation of $\imath x(A!x \,\&\, \forall F(xF \equiv Fb))\overline{B}$ just described fails to be modally strict.

In case (c), we have an example like the ones at the beginning of this Remark, in which the antecedent is provable (via a modally strict proof) *without* extending our theory. To see this, note first that (135.1) establishes that $Necessary(L)$, i.e., $\Box\forall xLx$. So if we apply RN to the modally strict theorem $\forall xLx \rightarrow Lb$ of classical quantification theory (79.1), we have $\Box(\forall xLx \rightarrow Lb)$. Thus, by the K schema, it follows that $\Box Lb$, which is the antecedent of (c). Hence, we can derive the consequent of (c) via a modally strict proof without any additional assumptions.

(**180**) **Metadefinitions:** Canonically Defined, and Provably Canonical, Terms and Individuals. Thus far, we've been discussing canonical descriptions and individuals, as these notions were defined in (174). But we now introduce two other kinds of canonicity. First, there are terms that have been explicitly defined by canonical descriptions:

(.1) An individual term κ is *canonically-defined with respect to* φ if and only if κ is introduced by the definition $\kappa =_{df} \imath\nu(A!\nu \,\&\, \forall F(\nu F \equiv \varphi))$, for some individual variable ν that doesn't occur free in φ.

We henceforth say that κ is *canonically-defined* whenever there is a φ such that κ is canonically-defined with respect to φ. And, by a slight abuse of language, we call the individuals denoted by such terms *canonically-defined* individuals.

A second kind of canonicity arises when one can show, by a modally strict proof, that $\kappa = \imath x(A!x \,\&\, \forall F(xF \equiv \varphi))$. As we shall see, a term κ can be canonical in this sense even if it is defined by a non-canonical description. So let us define:

(.2) An individual term κ is *provably canonical with respect to* φ if and only if $\vdash_\Box \kappa = \imath\nu(A!\nu \,\&\, \forall F(\nu F \equiv \varphi))$, for some individual variable ν that doesn't occur free in φ.

Intuitively, whenever κ is provably canonical with respect to φ, the properties F such that φ are the ones by which we identify κ. We henceforth say that κ is *provably canonical* whenever there is a φ such that κ is provably canonical with respect to φ. By a now familiar abuse of language, we call an individual denoted by such a κ *provably canonical*.

(**181**) **Remark:** Examples. A simple example of a canonically-defined individual is the logical object we shall identify as the truth value Frege dubbed *The True*. In item (214.1), we introduce the notation $\top$ for *The True* as follows:

$$\top =_{df} \iota x(A!x \,\&\, \forall F(xF \equiv \exists p(p \,\&\, F = [\lambda y\, p])))$$

This defines *The True* as the abstract individual that encodes all and only the propositional properties of the form $[\lambda y\, p]$ for which p is a true proposition. We will prove some interesting facts about $\top$ in the next chapter, but for now, it suffices to note that $\top$ is a clear example of a canonically-defined individual.

It immediately follows from (180.2) that any canonical description $\iota\nu(A!\nu \,\&\, \forall F(\nu F \equiv \varphi))$ is a provably canonical term. Moreover, it follows from both (180.1) and (180.2) that canonically-defined terms are provably canonical. For if κ is canonically-defined with respect to some formula φ, then by (.1), we know that κ is explicitly defined by a canonical description, say $\iota x(A!x \,\&\, \forall F(xF \equiv \varphi))$. Hence, by our conventions for term definitions (202.1.b), we know $\vdash_\Box \kappa = \iota x(A!x \,\&\, \forall F(xF \equiv \varphi))$. So by (180.2), κ is provably canonical with respect to φ.

To produce an example of a provably canonical term that is neither a canonical description nor canonically-defined, we preview some definitions from subsequent chapters. In Chapter 10 we define what it is for an abstract object to be an extension of a property G:

$$\mathit{ExtensionOf}(x, G) =_{df} A!x \,\&\, \forall F(xF \equiv \forall z(Fz \equiv Gz)) \tag{229.1}$$

Now if we construct an instance of Strengthened Comprehension (170) out of the definiens, we know:

$$\vdash_\Box \exists! x(A!x \,\&\, \forall F(xF \equiv \forall z(Fz \equiv Gz))$$

Thus, by definition, the previous two displayed lines imply:[115]

$$\vdash_\Box \exists! x \mathit{ExtensionOf}(x, G)$$

By (173), we therefore know that:

$$\vdash_\Box \exists y(y = \iota x \mathit{ExtensionOf}(x, G))$$

By GEN, the description is logically proper for every property G and so we may introduce the function term ϵG to denote the extension of G:

$$\epsilon G =_{df} \iota x \mathit{ExtensionOf}(x, G) \tag{234}$$

[115]Although we give a full account of the inferential role of definitions in the next section, it might be useful here to point out that by the Simple Rule of Equivalence by Definition (203.2), definition (229.1) previewed above converts to the fact:

$$\vdash_\Box \mathit{ExtensionOf}(x, G) \equiv (A!x \,\&\, \forall F(xF \equiv \forall z(Fz \equiv Gz)))$$

Then, Rule SubDefSubform (203.3) allows us to exchange the left and right sides of the above biconditional whenever one occurs as a subformula.

Clearly, the new term ϵG is neither a canonical description nor canonically-defined. But it is provably canonical, and the reasoning is straightforward. Our conventions and rules for term definitions, which we discuss later in (202.2), allow us to infer the following from (234):

(ϑ) $\vdash_\Box \epsilon G = \imath x ExtensionOf(x, G)$

From (ϑ), it follows by definition (229.1) previewed above that:[116]

(ξ) $\vdash_\Box \epsilon G = \imath x(A!x \,\&\, \forall F(xF \equiv \forall z(Fz \equiv Gz)))$

Thus, from (ξ) and (180.2), ϵG is provably canonical with respect to $\forall z(Fz \equiv Gz)$.

One might wonder here: wouldn't the term 'provably canonical' still be applicable to κ even if (180.2) required only that there be *some* proof of $\kappa = \imath\nu(A!\nu \,\&\, \forall F(\nu F \equiv \varphi))$ and not necessarily a modally strict proof? That is certainly true, but we take the *demands of canonicity* to require that we reserve the term for the case where the needed theorem is modally strict. Intuitively, a provably canonical individual is one whose canonicity can be established without appealing to any contingency or necessitation-averse principle.

(**182**) **Theorems:** Provably Canonical Individuals are Abstract. It follows by a modally strict proof that provably canonical individuals exemplify *being abstract*:

$A!\kappa$, provided κ is provably canonical.

(**183**) $\star$**Theorems:** Provably Canonical Objects Encode Their Defining Properties. Suppose κ is provably canonical with respect to φ. Although there is *not*, in general, a modally strict proof of the fact that κ encodes all and only the properties such that φ, there is a non-modally strict proof:

$\forall F(\kappa F \equiv \varphi)$, provided κ is provably canonical with respect to φ.

This is straightforwardly derived from (97.2)$\star$.

We can produce an instance of this theorem by returning to the example of ϵG previewed in (181). Since we established there that ϵG is provably canonical with respect to $\forall z(Fz \equiv Gz)$, the claim that $\forall F(\epsilon GF \equiv \forall z(Fz \equiv Gz))$ is an instance of our theorem. This claim occurs as theorem (235.2)$\star$ in Chapter 10.

(**184**) **Metadefinitions:** Strictly Canonical Terms and Individuals. Over the course of the next several items, we discuss a special group of provably canonical terms. We say:

[116] Note that (ξ) is obtained from (ϑ) by substituting a definiens for definiendum even though the latter is the matrix of a description and not a subformula of (ϑ). This is justified by Rule SubDefForm (203.5), which is thoroughly discussed in the next section.

(.1) φ is a *rigid condition* (on properties) if and only if $\vdash_{\Box} \forall H(\varphi^H_F \to \Box\varphi^H_F)$.

(.2) κ is *strictly canonical with respect to* φ if and only if κ is provably canonical with respect to φ and φ is a rigid condition.

Again, we say that κ is *strictly canonical* whenever there is a formula φ with respect to which κ is strictly canonical. And, again, we sometimes abuse language and speak of the strictly canonical individuals denoted by such terms.

(**185**) **Remark:** On Strict Canonicity. In this Remark, we show: (a) that ϵG, as previewed in (181), is not a strictly canonical term with respect to $\forall z(Fz \equiv Gz)$, even though it is provably canonical with respect to $\forall z(Fz \equiv Gz)$, and (b) that, nevertheless, examples of strictly canonical terms are easy to come by. Before embarking on this discussion, an observation is in order. Note that the concept of a *rigid condition* was just defined in terms of the existence of a modally strict proof: to show φ is rigid we have to show that there is a modally strict proof of a certain kind. That is, if we are to show that some term κ is provably but not strictly canonical with respect to φ, we have to show that φ is not a rigid condition, i.e., that there is no modally strict proof of $\forall H(\varphi^H_F \to \Box\varphi^H_F)$.

However, in what follows, we shall show φ is not a rigid condition by a different strategy, one that relies on the following:

- **Fact**: If φ is a rigid condition on properties, then $\vdash \Box\forall H(\varphi^H_F \to \Box\varphi^H_F)$.
- *Proof*. If φ is rigid, then $\vdash_{\Box} \forall H(\varphi^H_F \to \Box\varphi^H_F)$, by definition (184.1). Hence, by the Rule of Necessitation, $\vdash \Box\forall H(\varphi^H_F \to \Box\varphi^H_F)$.

Now if we can establish, for some given φ, that $\vdash \Diamond\exists H(\varphi^H_F \,\&\, \neg\Box\varphi^H_F)$, then since $\Diamond\exists H(\varphi^H_F \,\&\, \neg\Box\varphi^H_F)$ contradicts $\Box\forall H(\varphi^H_F \to \Box\varphi^H_F)$, it follows, on pain of inconsistency, that if there is a proof of the former, there is no proof of the latter. That is, if we can show $\vdash \Diamond\exists H(\varphi^H_F \,\&\, \neg\Box\varphi^H_F)$, then it follows by Modus Tollens from the above Fact that φ fails to be a rigid condition. Using the above reasoning, it would follow that ϵG is not strictly canonical with respect to $\forall z(Fz \equiv Gz)$.

Now in (181), we saw that when φ is $\forall z(Fz \equiv Gz)$, the term ϵG is provably canonical with respect to φ. Intuitively, however, there are properties G such that φ fails to be rigid condition on properties: if we let G be *creature with a heart* and K be *creature with a kidney*, then we might justifiably assert $\varphi^K_F \,\&\, \Diamond\neg\varphi^K_F$, since all and only creatures with hearts are creatures with kidneys though it is possible that there are creatures with a heart but no kidney (or vice versa). It would then follow that $\Diamond(\varphi^K_F \,\&\, \neg\Box\varphi^K_F)$. If this conclusion is derived from axioms for G and K, then φ isn't a rigid condition.

But we need not rest with an intuitive example. Without appealing to any assumptions, we can show that there are properties G whose existence is guaranteed by the system such that φ fails to be a rigid condition (on properties F).

Let G be $[\lambda x\, E!x \,\&\, \Diamond\neg E!x]$. This is the property of being contingently concrete. It is provable that for some property H, $\exists H(\varphi^H_F \,\&\, \neg\Box\varphi^H_F)$. Here is how.

By axiom (32.5), β-Conversion, and the Rule of Substitution, we know that $\Diamond\exists xGx \,\&\, \Diamond\neg\exists xGx$. That is, it follows from our axioms that possibly there are contingently concrete objects and possibly there are not. If we think semantically for the moment and take the notion of a possible world as semantically primitive, then this consequence tells us that there is a possible world, say $\boldsymbol{w}_1$, at which there are contingently concrete objects and a possible world, say $\boldsymbol{w}_2$, at which there aren't. Now let H be the property $\overline{L}$, where L was defined as $[\lambda x\, E!x \to E!x]$ (135). Thus, $\overline{L}$ is a contradictory property: necessarily, nothing exemplifies $\overline{L}$. Again, thinking semantically, this means that at w_1, $\overline{L}$ is not materially equivalent to G (since something is G there but nothing is $\overline{L}$ there), but that at $\boldsymbol{w}_2$, $\overline{L}$ is materially equivalent to G (since nothing is either G or $\overline{L}$ there). That is, in object-theoretic terms, we know $\Diamond\neg\forall x(\overline{L}x \equiv Gx) \,\&\, \Diamond\forall x(\overline{L}x \equiv Gx)$.[117] Reversing the conjuncts, we have:

$$\Diamond\forall x(\overline{L}x \equiv Gx) \,\&\, \Diamond\neg\forall x(\overline{L}x \equiv Gx)$$

Hence by (115.12):

$$\Diamond(\forall x(\overline{L}x \equiv Gx) \,\&\, \Diamond\neg\forall x(\overline{L}x \equiv Gx))$$

A fortiori:

$$\Diamond\exists H(\forall x(Hx \equiv Gx) \,\&\, \Diamond\neg\forall x(Hx \equiv Gx))$$

i.e.,

$$\Diamond\exists H(\forall x(Hx \equiv Gx) \,\&\, \neg\Box\forall x(Hx \equiv Gx))$$

i.e.,

$$(\vartheta)\quad \Diamond\exists H(\varphi^H_F \,\&\, \neg\Box\varphi^H_F)$$

Hence, our derivation of (ϑ) establishes, given the Fact proved above, that when φ is $\forall z(Fz \equiv Gz)$ and G is the property of being contingently concrete, φ fails to be a rigid condition on properties.

[117] Here is the proof within our system. We establish the conjuncts separately. (a) To establish $\Diamond\neg\forall x(\overline{L}x \equiv Gx)$, we first establish $\exists xGx \to \neg\forall x(\overline{L}x \equiv Gx)$. So assume $\exists xGx$. By definition of $\overline{L}$, we know $\neg\exists x\overline{L}x$. Hence by (86.13), it follows that $\neg\forall x(Gx \equiv \overline{L}x)$, i.e., $\neg\forall x(\overline{L}x \equiv Gx)$. Hence, by conditional proof $\exists xGx \to \neg\forall x(\overline{L}x \equiv Gx)$. Since this proof is modally strict, it follows by Rule RM (106) that $\Diamond\exists xGx \to \Diamond\neg\forall x(\overline{L}x \equiv Gx)$. But since the first conjunct of axiom (32.5) is equivalent to $\Diamond\exists xGx$, it follows that $\Diamond\neg\forall x(\overline{L}x \equiv Gx)$. (b) To establish $\Diamond\forall x(\overline{L}x \equiv Gx)$, we first establish $\neg\exists xGx \to \forall x(\overline{L}x \equiv Gx)$. So assume $\neg\exists xGx$. Again, by definition of $\overline{L}$, we know $\neg\exists x\overline{L}x$. Hence by (86.12), it follows that $\forall x(Gx \equiv \overline{L}x)$, i.e., $\forall x(\overline{L}x \equiv Gx)$. Hence, by conditional proof $\neg\exists xGx \to \forall x(\overline{L}x \equiv Gx)$. Since this proof is modally strict, it follows by Rule RM that $\Diamond\neg\exists xGx \to \Diamond\forall x(\overline{L}x \equiv Gx)$. But since the second conjunct of axiom (32.5) is equivalent to $\Diamond\neg\exists xGx$, it follows that $\Diamond\forall x(\overline{L}x \equiv Gx)$.

Since ϵG is provably canonical with respect to this particular φ, ϵG fails to be strictly canonical, by (184.2). Nevertheless, it is easy to find examples of κ and φ such that κ is strictly canonical with respect to φ, and some of these will play an important role in later chapters. In Chapter 11, we introduce two different, but related, analyses of Plato's Forms. The simpler analysis provides a *thin* conception of Plato's Forms:

$$ThinFormOf(x,G) =_{df} A!x \,\&\, \forall F(xF \equiv F = G) \tag{299}$$

$$\boldsymbol{a}_G =_{df} \imath x(ThinFormOf(x,G)) \tag{306}$$

Clearly, $\boldsymbol{a}_G$ is provably canonical; the reasoning is analogous to the case of ϵG, previewed in (181).[118] In addition, one can show that $F = G$ is a rigid condition, given any property G, as follows:

> By GEN, it suffices to show $\varphi^H_F \to \Box\varphi^H_F$. So assume φ^H_F, i.e., $H = G$. By definition (16.1), it follows that $\Box\forall x(xH \equiv xG)$. By the '4' theorem (115.6), it follows that $\Box\Box\forall x(xH \equiv xG)$, i.e., $\Box\varphi^H_F$.

Given this modally strict proof of $\forall H(\varphi^H_F \to \Box\varphi^H_F)$, it follows by (.2) that $\boldsymbol{a}_G$ is strictly canonical with respect to φ.

(**186**) **Metarule:** Rule of Encoding for Strictly Canonical Objects. The following rule is now justified: if κ is strictly canonical with respect to φ, then there is a modally strict proof that κ encodes all and only the properties F such that φ. More formally:

> **Rule of Encoding for Strictly Canonical Objects** (**Rule ESCO**)
> If κ is strictly canonical with respect to φ, then $\vdash_\Box \forall F(\kappa F \equiv \varphi)$.

As an example, let κ be $\boldsymbol{a}_G$, as this was defined at the end of (184). We established that $\boldsymbol{a}_G$ is strictly canonical with respect to $F = G$. Then since the reasoning there established that $\boldsymbol{a}_G$ is strictly canonical, Rule ESCO implies:

$$\vdash_\Box \forall F(\boldsymbol{a}_G F \equiv F = G)$$

This conclusion becomes interesting and unusual when contrasted with theorem (183)★, where it was observed that, generally speaking, one cannot give a modally-strict proof that provably canonical individuals encode their defining properties. But now we've seen that certain provably canonical individuals demonstrably encode their defining properties by such a proof.

[118] By our conventions and rules for definitions in (202), definition (306) previewed above converts to the following modally strict theorem:

$$\boldsymbol{a}_G = \imath x(ThinFormOf(x,G))$$

Hence, by definition (299), there is a modally strict proof of:

$$\boldsymbol{a}_G = \imath x(A!x \,\&\, \forall F(xF \equiv F = G))$$

So by definition (180.2), $\boldsymbol{a}_G$ is provably canonical.

(**187**) **Metadefinition:** Formulas Strictly Equivalent to Canonical Matrices. Let us now say:

- ψ is *strictly equivalent to a canonical matrix constructed from* φ if and only if $\vdash_{\Box} \psi \equiv (A!\nu \,\&\, \forall F(\nu F \equiv \varphi))$, for some individual variable ν that isn't free in φ

We've now seen two simple cases in which ψ is strictly equivalent to a canonical matrix constructed from some φ:

> $ExtensionOf(x, G)$ is strictly equivalent to a canonical matrix constructed from $\forall z(Fz \equiv Gz)$
>
> $ThinFormOf(x, G)$ is strictly equivalent to a canonical matrix constructed from $F = G$

These cases are simple because in both, ψ is defined as $A!\nu \,\&\, \forall F(\nu F \equiv \varphi)$. But more interesting cases of formulas strictly equivalent to canonical matrices arise when multiple definitions and theorems are required to establish $\vdash_{\Box} \psi \equiv (A!x \,\&\, \forall F(xF \equiv \varphi))$. We shall see such a case below, when we derive facts about the null object and the universal object.

(**188**) **Theorems:** Another Group of Logically Proper Descriptions. When ψ is strictly equivalent to a canonical matrix constructed from a formula φ, there exists something which is the x such that ψ, i.e.,

> $\exists y(y = \imath x\psi)$, provided ψ is strictly equivalent to a canonical matrix constructed from some formula φ in which x isn't free.

(**189**) **Metarule:** The Rule for Provably Canonical Individuals. The following useful metarule is now justified:

> **Rule for Provably Canonical Individuals** (**Rule PCI**)
> κ is provably canonical with respect to φ if and only if $\vdash_{\Box} \kappa = \imath\nu\psi$ for some formula ψ that is strictly equivalent to a canonical matrix constructed from φ and some individual variable ν not free in φ.

In terms of an example previewed in (181), this rule asserts:

(ϑ) ϵG is provably canonical with respect to $\forall z(Fz \equiv Gz)$ if and only if $\vdash_{\Box} \epsilon G = \imath x\psi$ for some formula ψ that is strictly equivalent to a canonical matrix constructed from $\forall z(Fz \equiv Gz)$.

In this case, (ϑ) is true because both sides of the biconditional hold: (a) we saw in (181) that ϵG is provably canonical with respect to $\forall z(Fz \equiv Gz)$ and (b) we established in (181) that $\vdash_{\Box} \epsilon G = \imath x ExtensionOf(x, G)$ and established in (187) that *ExtensionOf*(x, G) is strictly equivalent to a canonical matrix constructed from $\forall z(Fz \equiv Gz)$.

Rule PCI becomes especially useful when a series of definitions and modally-strict theorems puts one into a position to apply the rule.

(**190**) **Metarule:** Rule for Strict Equivalents to Canonical Matrices. The following rule is now justified:

> **Rule for Strict Equivalents to Canonical Matrices** (**Rule SECM**)
> If ψ is strictly equivalent to a canonical matrix constructed from some *rigid* condition in which ν doesn't occur free, then
>
> if $\vdash_\Box \kappa = \imath\nu\psi$, then $\vdash_\Box \psi^\kappa_\nu$,
>
> for any defined individual term κ.

As an example, consider the values for ψ and κ introduced in the discussion of (185). When ψ is *ThinFormOf*(x, G), κ is $\boldsymbol{a}_G$, and ν is x, then our rule asserts:

> If *ThinFormOf*(x, G) is strictly equivalent to a canonical matrix constructed from some rigid condition in which x doesn't occur free, then
> if $\vdash_\Box \boldsymbol{a}_G = \imath x\mathit{ThinFormOf}(x, G)$, then $\vdash_\Box \mathit{ThinFormOf}(\boldsymbol{a}_G, G)$.

Now we saw at the end of item (187) that *ThinFormOf*(x, G) is strictly equivalent to a canonical matrix constructed from $F = G$. Clearly, x isn't free in $F = G$. And we also established at the end of (185) that $F = G$ is a rigid condition. So we may use Rule SECM to conclude that:

> If $\vdash_\Box \boldsymbol{a}_G = \imath x\mathit{ThinFormOf}(x, G)$, then $\vdash_\Box \mathit{ThinFormOf}(\boldsymbol{a}_G, G)$

Now it follows from the definition of $\boldsymbol{a}_G$ (306), previewed in (185) above, that $\vdash_\Box \boldsymbol{a}_G = \imath x\mathit{ThinFormOf}(x, G)$. Hence *ThinFormOf*$(\boldsymbol{a}_G, G)$ is derivable by a modally strict proof. This theorem is officially presented as item (308.1).

Note that other interesting conclusions can be drawn from this example. For instance, it follows from the above reasoning that $\boldsymbol{a}_G$ is provably canonical, by Rule PCI. It also follows that $\boldsymbol{a}_G$ is strictly canonical, by definition (184.2). More generally, whenever Rule SECM is applicable, the term κ in the statement of the rule is strictly canonical.

It is important to contrast the conditional consequent of Rule SECM (i.e., if $\vdash_\Box \kappa = \imath\nu\psi$, then $\vdash_\Box \psi^\kappa_\nu$) with theorem (97.2)★, a non-modally strict theorem that asserts $y = \imath x\psi \to \psi^y_x$. The fact that a uniquely described individual satisfies the matrix of its description can only be established, in the general case, by a non-modally strict proof. But in the special case where the uniquely described individual is a strictly canonical object, there is a modally strict proof that it satisfies the matrix of its own description. As another example, see theorem (333.3), i.e., *FormOf*(Φ_G, G), discussed in Chapter 11.

We now marshal the above machinery for reasoning about provably canonical and strictly canonical objects to define, and prove facts about, the null and universal objects.

(**191**) **Definitions:** Null and Universal Objects. We say: (.1) x is a *null* object just in case x is an abstract object that encodes no properties; and (.2) x is a *universal* object just in case x is an abstract object that encodes every property:

(.1) $Null(x) =_{df} A!x \& \neg\exists F xF$

(.2) $Universal(x) =_{df} A!x \& \forall F xF$

(**192**) **Theorems:** Existence and Uniqueness of Null and Universal Objects. It is now easily established that: (.1) x is a null object if and only if x is abstract and encodes exactly the properties that are non-self-identical; and (.2) x is a universal object if and only if x is abstract and encodes exactly the properties that are self-identical:

(.1) $Null(x) \equiv (A!x \& \forall F(xF \equiv F \neq F))$

(.2) $Universal(x) \equiv (A!x \& \forall F(xF \equiv F = F))$

Note that since there is a modally strict proof of (.1), it follows by metadefinition (187) that $Null(x)$ is strictly equivalent to a canonical matrix constructed from $F \neq F$. And it similarly follows from the modally strict proof of (.2) that $Universal(x)$ is strictly equivalent to a canonical matrix constructed from $F = F$.

Consequently, it follows, by a modally strict proof (i.e., without appealing to (96)★), that (.3) there exists something that is the null object, and (.4) there exists something that is the universal object:

(.3) $\exists y(y = \imath x Null(x))$

(.4) $\exists y(y = \imath x Universal(x))$

(**193**) **Definitions:** Notation for the Null and Universal Objects. We may therefore introduce the following new terms to designate the null object and the universal object:

(.1) $\boldsymbol{a}_{\varnothing} =_{df} \imath x Null(x)$

(.2) $\boldsymbol{a}_{\boldsymbol{V}} =_{df} \imath x Universal(x)$

(**194**) **Remark:** A Strategy for Strictly Proving $Null(\boldsymbol{a}_{\varnothing})$ and $Universal(\boldsymbol{a}_{\boldsymbol{V}})$. In the series of theorems that follow next in (195), we work our way towards a modally strict proof that $Null(\boldsymbol{a}_{\varnothing})$ and $Universal(\boldsymbol{a}_{\boldsymbol{V}})$. Our strategy is now simple: if we can show that $Null(x)$ is strictly equivalent to a canonical matrix constructed from some *rigid* condition in which x doesn't occur free, then from the fact that the definition of $\boldsymbol{a}_{\varnothing}$ (193.1) converts, by convention (202.2), into $\vdash_{\Box} \boldsymbol{a}_{\varnothing} = \imath x Null(x)$, Rule SECM allows us to conclude that there is a modally strict proof of $Null(\boldsymbol{a}_{\varnothing})$. A similar strategy establishes $Universal(\boldsymbol{a}_{\boldsymbol{V}})$. (Of course, there is nothing defective about the non-modally strict proofs of these claims from the definitions of $\boldsymbol{a}_{\varnothing}$, $\boldsymbol{a}_{\boldsymbol{V}}$, and (97.2)★, but modally strict proofs of these claims should be developed when they are available.)

(**195**) **Theorems:** Facts about the Null and Universal Object. Note independently that by (118.2) it follows that (.1) every non-self-identical property is necessarily non-self-identical; and by (75) it follows that (.2) every self-identical property is necessarily self-identical:

(.1) $\forall H(H \neq H \rightarrow \Box H \neq H)$

(.2) $\forall H(H = H \rightarrow \Box H = H)$

Hence, if we set φ to $F \neq F$, then it follows from (.1) by (184.1) that φ is a rigid condition, since (.1) has the form $\forall H(\varphi^H_F \rightarrow \Box\varphi^H_F)$. Similarly, if we instead set φ to $F = F$, then by (184.1), it follows from (.2) that φ is a rigid condition, since under this new setting, (.2) has the form $\forall H(\varphi^H_F \rightarrow \Box\varphi^H_F)$.

Thus, we have the facts needed to apply Rule SECM and establish that there are modally strict proofs of:

(.3) $Null(\boldsymbol{a}_{\varnothing})$

(.4) $Universal(\boldsymbol{a}_{\boldsymbol{V}})$

(**196**) **Remark:** A Rejected Alternative. Now that we've seen how to obtain modally strict proofs of $Null(\boldsymbol{a}_{\varnothing})$ and $Universal(\boldsymbol{a}_{\boldsymbol{V}})$, one might wonder: would it not be easier to obtain these results by defining $\boldsymbol{a}_{\varnothing}$ as:

(ϑ) $\boldsymbol{a}_{\varnothing} =_{df} \imath x(A!x \,\&\, \forall F(xF \equiv F \neq F))$

Then, a modally strict proof $Null(\boldsymbol{a}_{\varnothing})$ could be given more simply as follows:

> By definition, we have to show (a) $A!\boldsymbol{a}_{\varnothing}$ and (b) $\neg\exists F(\boldsymbol{a}_{\varnothing}F)$. (a) By ($\vartheta$), we know $\boldsymbol{a}_{\varnothing} = \imath x(A!x \,\&\, \forall F(xF \equiv F \neq F))$, which in turn entails that $\boldsymbol{a}_{\varnothing}$ is provably canonical (with respect to $F \neq F$). So (182) implies $A!\boldsymbol{a}_{\varnothing}$. (b) The easily established fact that $F \neq F$ is a rigid condition means that $\boldsymbol{a}_{\varnothing}$ is also strictly canonical. So Rule ESCO yields that $\forall F(\boldsymbol{a}_{\varnothing}F \equiv F \neq F)$, from which it follows that $\neg\exists F(\boldsymbol{a}_{\varnothing}F)$. Hence $Null(\boldsymbol{a}_{\varnothing})$, by a modally strict proof.

(An similar observation could be made about $\boldsymbol{a}_{\boldsymbol{V}}$, but we henceforth consider only $\boldsymbol{a}_{\varnothing}$, since the considerations are analogous.)

The alternative definition (ϑ) is certainly legitimate, but we have a good reason for not deploying it. On definition (ϑ), $\boldsymbol{a}_\varnothing$ is the abstract object that encodes all and only non-self-identical properties; it is *not* explicitly introduced as the null object *per se*. Of course, one can then show that given (ϑ), $\boldsymbol{a}_\varnothing$ is a *unique* null object (i.e., a unique abstract object that encodes no properties) and start referencing it as 'the null object'. But this fails to take advantage of the fact that our system allows us to properly define the condition $Null(x)$ as $A!x \,\&\, \neg\exists F xF$, prove that there is a unique such object, show that the description $\imath x Null(x)$ is logically proper, and then use the well-defined description $\imath x Null(x)$ to introduce the notation $\boldsymbol{a}_\varnothing$. If we are going to use $\boldsymbol{a}_\varnothing$ to name 'the null object', the correctly motivated way to do so is to define $\boldsymbol{a}_\varnothing$ as $\imath x Null(x)$ after having shown the latter is logically proper by a modally strict proof. But once we do this, it takes a little work to show that $\boldsymbol{a}_\varnothing$ can be instantiated into its own defining matrix by a modally strict proof. Rules ESCO and SECM simplify the work that has to be done.

(**197**) **Theorems:** Further Facts about the Null and Universal Objects. If we apply the Rule of Actualization to (192.1) and (192.2), respectively, we obtain:

(.1) $\mathcal{A}(Null(x) \equiv (A!x \,\&\, \forall F(xF \equiv F \neq F)))$

(.2) $\mathcal{A}(Universal(x) \equiv (A!x \,\&\, \forall F(xF \equiv F = F)))$

Since these are modally strict theorems, it follows from (.1) and definition (193.1) by theorem (102) that:

(.3) $\boldsymbol{a}_\varnothing = \imath x(A!x \,\&\, \forall F(xF \equiv F \neq F))$

And it follows from (.2) and (193.2) by (102) that:

(.4) $\boldsymbol{a}_{\boldsymbol{V}} = \imath x(A!x \,\&\, \forall F(xF \equiv F = F))$

Again, since these are modally strict theorems, it follows from (.3) and (.4), respectively, by metadefinition (180.2) that:

(ϑ) $\boldsymbol{a}_\varnothing$ is provably canonical (with respect to $F \neq F$)

(ξ) $\boldsymbol{a}_{\boldsymbol{V}}$ is provably canonical (with respect to $F = F$)

But from (195.1), we know by (184.1) that $F \neq F$ is a rigid condition, and from this and (ϑ) it follows by (184.2) that:

- $\boldsymbol{a}_\varnothing$ is strictly canonical (with respect to $F \neq F$)

Similarly, from (195.2), we know by (184.1) that $F = F$ is a rigid condition, and from this and (ξ) it follows by (184.2) that:

- $\boldsymbol{a}_{\boldsymbol{V}}$ is strictly canonical (with respect to $F = F$)

(**198**) **Theorems:** Facts about the Granularity of Relations. The following facts govern the granularity of relational properties having abstract constituents: (.1) For any relation R, there are distinct abstract objects x and y for which *bearing R to x* is identical to *bearing R to y*, and (.2) For any relation R, there are distinct abstract objects x and y for which *being a z such that x bears R to z* is identical to *being a z such that y bears R to z*:

(.1) $\forall R \exists x \exists y (A!x \,\&\, A!y \,\&\, x \neq y \,\&\, [\lambda z\, Rzx] = [\lambda z\, Rzy])$

(.2) $\forall R \exists x \exists y (A!x \,\&\, A!y \,\&\, x \neq y \,\&\, [\lambda z\, Rxz] = [\lambda z\, Ryz])$

Moreover, for any property F, there are distinct abstract objects x, y such that *that-Fx* is identical to *that-Fy*:

(.3) $\forall F \exists x \exists y (A!x \,\&\, A!y \,\&\, x \neq y \,\&\, [\lambda\, Fx] = [\lambda\, Fy])$

(**199**) **Theorem:** Some Abstract Objects Not Strictly Leibnizian. Theorem (198.1) now implies a rather interesting theorem, namely, that there exist distinct abstract objects that exemplify exactly the same properties:[119]

$$\exists x \exists y (A!x \,\&\, A!y \,\&\, x \neq y \,\&\, \forall F (Fx \equiv Fy))$$

In other words, there are distinct abstract objects that are indiscernible with respect to the properties they exemplify.[120] To prove this claim, let R in theorem (198.1) be the impredicatively-defined relation $[\lambda xy\, \forall F(Fx \equiv Fy)]$.

Thus classical Leibnizian indiscernibility doesn't imply the identity of abstract objects. Nevertheless, by (166), a somewhat different form of the Leibnizian principle of the identity of indiscernibles applies to abstract objects: such objects are identical whenever indiscernible with respect to properties they encode.

9.11 The Metatheory of Definitions

Reasoning with definitions has played, and will continue to play, an important part in the proof of almost every theorem, since few theorems are expressed using only primitive notation or rest solely on axioms that are expressed using

[119] I am indebted to Peter Aczel for pointing out that this theorem results once we allow impredicatively-defined relations into our system.

[120] If one recalls the structure of the Aczel models described in the Introduction, this result is to be expected. When abstract objects are modeled as sets of properties, where properties are modeled as sets of urelements, then the fact that an abstract object x exemplifies a property F is modeled by ensuring that the special urelement that serves as the proxy of x is an element of F. Consequently, since distinct abstract objects must sometimes be represented by the same proxy, some distinct abstract objects exemplify the same properties.

only primitive notation. Before we go any further, it would serve well to articulate the underlying theory and inferential role of definitions in our system, so as to indicate more precisely what we mean when we say "It follows by definition that ..." within the context of a derivation or proof. With Frege being one of the few exceptions, most discussions of the theory of definition in formal systems have been framed with respect to a very simple language, namely, the language of the first-order predicate calculus, possibly extended by function terms.[121] By contrast, our language uses second-order modal syntax with 0-place relation terms and includes not only formulas that are themselves terms but also terms that result from applying variable binding operators to formulas (e.g., descriptions and λ-expressions). The theory of definitions for this language involves some interesting subtleties. In particular, we have to discuss how definitions behave with respect to complex terms, rigidly-denoting terms, non-denoting terms (i.e., some descriptions), and encoding formulas. Consequently, those familiar with the classical theory of definitions may still find the following discussion intriguing if not useful. It may be skipped, however, by those more interested in seeing how the foregoing core theorems of object theory can be further applied. But those who plan to skip this section should be aware that when we say "it follows by definition" in a proof, we are implicitly relying on one of the rules in items (202.1) – (202.4) and (203.1) – (203.6), since they constitute the principles that govern the inferential role of definitions in our system.

For the reason mentioned in Remark (19), we have adopted the position that definitions introduce new expressions into our language and are not mere metalinguistic abbreviations of our object language. Furthermore, we have assumed that in a properly formed definition, the variables having free occurrences in both the definiens and definiendum should be the same.[122] Finally, recall the distinction between *term definitions* and *formula definitions* introduced in Remark (19). Though this distinction is not mutually exclusive (since definitions in which the definiens and definiendum are propositional formulas are both term and formula definitions), we discuss the theory and inferential roles of term and formula definitions separately in what follows, in

[121] In writing this section, I consulted Frege 1879, §24; Padoa 1900; Frege 1903, §§56–67, §§139–144, and §§146–147; Frege 1914, 224–225; Suppes 1957; Mates 1972; Dudman 1973; Belnap 1993; Hodges 2008; Urbaniak & Hämäri 2012; and Gupta 2014. Hodges 2008 and Urbaniak & Hämäri 2012 provide insightful discussions of the contributions by Kotarbiński, Łukasiewicz, Leśniewski, Ajdukiewicz and Tarski to the elementary theory of definitions.

[122] Though Suppes nicely explains why a definition must never allow free variables to occur in the definiens without occurring free in the definiendum, he does allow free variables in the definiendum that don't occur free in the definiens. But he notes that, in the latter case, we can trivially get the variables to match by adding dummy clauses to the definiens. Thus, in his example (Suppes 1957, 157), the definition $Q(x,y) =_{df} x > 0$ can be turned into $Q(x,y) =_{df} x > 0 \,\&\, y = y$. In what follows, we eschew definitions in which the free variables don't match, without loss of generality.

(202) and (203), respectively. However, we preface the discussion of these two items with two important Remarks.

(**200**) **Remark:** Unique Existence Not Sufficient for a New (Individual) Term. In some systems, it is legitimate to introduce, by definition, a new individual term κ when it is established that there is a unique x such that φ. That is, in some systems, if $\exists!x\varphi$ is an axiom or theorem, one may simply introduce a new term κ to designate any entity that is uniquely such that φ. Thus, in the classical theory of definitions, when $\vdash \exists!x\varphi$, one may introduce κ by stipulating (Suppes 1957, 159–60; Gupta 2014, Section 2.4):

$$\kappa = x \ =_{df} \ \varphi$$

This definition of κ is legitimate because (i) any formula ψ^{κ}_{x} in which κ occurs can be expanded by definition into the claim that $\exists!x\varphi \mathbin{\&} \exists x(\varphi \mathbin{\&} \psi)$, and (ii) no new theorems can be proved in the language with κ that aren't already provable in the language without κ.

But such a procedure would be incorrect (indeed, disastrous) for the present system. To see why, suppose that for some formula φ we were to assert, prove, or assume:

(a) $\exists!x\varphi$

(b) $\Diamond\neg\exists x\varphi$

and, then, on the basis of (a), stipulate:

(c) $\kappa = x \ =_{df} \ \varphi$

Now since (b) immediately yields:

(d) $\neg\Box\exists x\varphi$

by (113.2), a contradiction becomes derivable. Since (a) implies $\exists x\varphi$, assume b is such an entity, so that we know φ^{b}_{x}. Now note, independently, that since κ is logically proper, we can instantiate theorem (75) to obtain:

(e) $\kappa = b \rightarrow \Box\kappa = b$

Now by substituting b for x, we obtain the following instance of (c):

(f) $\kappa = b =_{df} \varphi^{b}_{x}$

Note, independently, that we can convert (f) into an equivalence: from the tautology $\varphi^{b}_{x} \equiv \varphi^{b}_{x}$, it follows by definition (f) that $\kappa = b \equiv \varphi^{b}_{x}$. Since this last result was established by a modally strict proof, we can use it together with the Rule of Substitution to obtain the following from (e):

(g) $\varphi^{b}_{x} \rightarrow \Box\varphi^{b}_{x}$

But since we've already established φ^b_x, it follows that $\Box\varphi^b_x$. Hence, by ∃I, $\exists x\Box\varphi$. So by the Buridan formula (117.4), $\Box\exists x\varphi$, which contradicts (d).

Thus, the classical theory of definitions fails for our system because our logic assumes that terms are rigid designators, no matter whether they are simple or complex, primitive or defined. One may not define κ as $\imath x\varphi$ when $\exists!x\varphi$ *happens to be* true and, indeed, not even when $\exists!x\varphi$ is necessarily true. Intuitively, in the former case, if $\exists!x\varphi$ is a necessitation-averse axiom, a non-modally strict theorem, or a contingent premise, there may be different individuals, or no individuals, that uniquely satisfy φ at different (semantically-primitive) worlds. Moreover, in the latter case, if $\exists!x\varphi$ is necessary, there may be different individuals that are uniquely such that φ at different worlds. Consequently, if we really wanted to adapt the classical theory of definitions to the present system, we should stipulate that a definition of the form $\kappa = x =_{df} \varphi$ is permissible only when $\vdash \exists!x\Box\varphi$.[123]

But we shall not adapt the classical theory of definitions in this way. That's because we have definite descriptions in our system and the logically correct way to introduce new individual terms is by explicitly defining them using definite descriptions (we discuss the further requirement that such descriptions be logically proper in the next Remark). That is, we shall require that a new individual term, say κ, be introduced by way of a definite description, using a definition of the form:

$$\kappa =_{df} \imath x\varphi$$

Note that this introduces κ by way of a rigidly designating term (and, as noted below, one that provably designates), rather than by way of a formula that contingently asserts the unique existence of something.

Moreover, we shall require something similar for relation terms. It does not suffice to stipulate, when $\vdash \exists!\alpha\varphi^*$, that $\Pi^n = \alpha =_{df} \varphi^*$. Instead, we require that the new relation term Π^n, for $n \geq 1$, be introduced by a definition of the form:

$$\Pi^n =_{df} [\lambda x_1 \dots x_n\, \varphi^*]$$

This introduces Π^n by way of a rigidly designating λ-expression. In the case where $n = 0$, we have the option of introducing a new term Π^0 using either of the following forms:

$$\Pi^0 =_{df} [\lambda\, \varphi^*]$$

$$\Pi^0 =_{df} \varphi^*$$

[123]This blocks the case described above because one can't simultaneously assert $\exists!x\Box\varphi$ and $\Diamond\neg\exists x\varphi$. A contradiction would ensue without the mediation of any definitions of new terms. Moreover, it also blocks the case in which $\Box\exists!x\varphi$ is true though because a *different* witness uniquely satisfies φ at different worlds.

Both of these forms introduce the new term Π^0 by way of a rigidly designating expression. Note that from a semantic point of view, even though φ^* rigidly denotes a proposition, the truth value of the proposition it denotes may vary from (semantically-primitive) possible world to possible world!

Finally, note that the above constraint on the introduction of new terms is independent of our rules for reasoning with *arbitrary names*. When we reason from $\exists\alpha\varphi$ to some conclusion ψ by showing, for some new (i.e., arbitrary) constant τ, that φ^τ_α implies ψ (i.e., in accordance with the conditions laid down in (85) for Rule $\exists$E), we are not introducing any new terms into the language when we reason (we're temporarily grabbing a fresh constant already in the language). We may still reason from $\exists\alpha\varphi$ by saying "let τ be an arbitrary such entity, so that we know φ^τ_α". The arbitrary term τ is chosen from our stock of primitive constants of the appropriate type, and as long as no special assumptions about τ are used in the proof, the reasoning is valid. We may not, for example, reason as follows:

> Consider the following instance of comprehension (40): $\exists x(A!x \& \forall F(xF \equiv \varphi))$. Assume a is such an object, so that we know: $A!a \,\&\, \forall F(aF \equiv \varphi)$. Hence, by RN: $\Box(A!a \,\&\, \forall F(aF \equiv \varphi))$. So by $\exists$E: $\exists x\Box(A!x \,\&\, \forall F(xF \equiv \varphi))$.

Exercise 1. Identify what is wrong with the above reasoning. **Exercise 2**. Show that the above reasoning, would lead to modal collapse if it were valid, i.e., would allow us to prove $\varphi \equiv \Box\varphi$.[124]

(**201**) **Remark:** Constraint on Term Definitions. Recall Remark (28), in which we noted that claims having the form $\exists\beta(\beta = \tau)$ ("there exists something identical with τ") informally tell us that term τ is *logically proper*. (In the formal mode, using semantic notions, this means that term τ has a denotation.) We have been, and will continue, observing the following strict constraint on term definitions generally:

---

[124] Solution to Exercise 2: Suppose, contrary to fact, that we could legitimately reason our way to the conclusion $\exists x\Box(A!x \,\&\, \forall F(xF \equiv \varphi))$. Then let a be such an object, so that we know $\Box(A!a \,\&\, \forall F(aF \equiv \varphi))$. Since a necessary conjunction implies that the conjuncts are necessary, it follows that $\Box A!a \,\&\, \Box\forall F(aF \equiv \varphi))$. From the second conjunct of this last result, it follows both that:

(ϑ) $\forall F\Box(aF \equiv \varphi)$ by the CBF schema (117.1)

(ξ) $\forall F(aF \equiv \varphi)$ by the T schema (32.2)

If we apply $\forall$E to both, we obtain, respectively:

(ϑ') $\Box(aF \equiv \varphi)$

(ξ') $aF \equiv \varphi$

From (ϑ') it follows by (107.6) that:

(ζ) $\Box aF \equiv \Box\varphi$

Now to see how this implies modal collapse, it suffices to show $\varphi \to \Box\varphi$, since $\Box\varphi \to \varphi$ is an instance of the T schema. So assume φ. From this, it follows from (ξ') that aF. From this it follows that $\Box aF$, by (121.1). So by (ζ), it follows that $\Box\varphi$.

Constraint on Term Definitions
Only terms τ that are provably logically proper (i.e., only terms τ for which $\vdash \exists\beta(\beta = \tau)$) may serve as definientia in individual term definitions.

Note, however, that since our system guarantees that all terms other than descriptions are logically proper, the above constraint really amounts to the following: only definite descriptions that are provably logically proper may serve as definientia in individual term definitions.

At this point, it is important to introduce and discuss a group of edge cases, namely, definitions in which the definiens is a definite description that is provably logically proper but not by a modally strict proof. For example, we may sometimes want to extend our theory by adding a contingent truth of the form $\exists!x\varphi$ as an axiom or premise. For example, suppose that we want to add the claim that there exists a unique moon of the Earth as an axiom to our system. If we use 'e' as the name of the Earth and represent this axiom as $\exists!xMxe$, then we would designate the axiom as an additional *necessitation-averse* axiom and we would annotate its item number with a $\star$.[125] Hence, this axiom would become a $\star$-theorem. So by theorem (96)$\star$, it would follow that $\exists y(y = \imath xMxe)$ is a $\star$-theorem.[126] Hence, $\imath xMxe$ would be provably logically proper, though the proof would not be modally strict. Nevertheless, by the Constraint on Term Definitions, we would be entitled to introduce a name, say m, to designate this unique object, as follows:

$$m =_{df} \imath xMxe \qquad (n)\star$$

Although we shall not actually include any such definitions like the above in what follows, the reader should remember that if one were to extend our system with a definition such as the above, one should:

- annotate the definition as **Definition**$\star$ and subsequently mark the item number with a $\star$, to indicate that the definition rests on a contingency, and

- treat any derivation or proof, in which one of the definiendum or definiens

[125]Recall that a *necessitation-averse* axiom is simply one for which we are not taking modal closures. But since the claim $\exists!xMxe$ isn't just necessitation-averse but rather contingent, we could also assert $\Diamond\neg\exists!xMxe$ as an axiom. But the full assertion of contingency is not needed for the present discussion; it suffices that the axiom is marked as necessitation-averse.

[126]Of course, the application of (96)$\star$ to $\exists!x\varphi$ to obtain $\exists y(y = \imath x\varphi)$ would turn the latter into a $\star$-theorem even if $\exists!x\varphi$ had been a necessary axiom or a modally strict theorem. But in this case, we could apply the Rule of Actualization to $\exists!x\varphi$ to obtain $\mathcal{A}\exists!x\varphi$, which by (172.2) yields a modally strict proof of $\exists y(y = \imath x\varphi)$. Thus, in cases like this, a definition in which $\imath x\varphi$ is used as definien is not subject to s any of the steps discussed immediately below. We will see such definitions on occasion in what follows.

in a $\star$-definition is substituted for the other, as non-modally strict (and so mark the theorem number with a $\star$).

If the reason for this isn't already clear, then it will become clear in the next item, where we discuss the inferential role of term definitions.

(**202**) **Remark:** The Inferential Role of Term Definitions. We have often introduced term definitions under constraint (201) to extend our language with new terms and subsequently cited 'by definition' to (i) assert identities on the basis of such term definitions, and (ii) draw inferences in which we have substituted, within any formula or term, the definiens for the definiendum (or vice versa). We now explain and justify this practice. In our discussion, we won't be concerned with the fact that term definitions are independent of the choice of any free and bound variables that they may contain. We discuss this aspect of definitions in the next item (204).

Our inferential practices are grounded in the following:

(.1) **Convention for Term Definitions**
The introduction of definitions having τ as a definiendum and τ' as a definiens under constraint (201) is a convention for the following reformulation of our system:

(.a) extend the language of our system with τ as a new primitive term,

(.b) if, for some variable β not free in τ', there is a modally strict proof of $\exists\beta(\beta = \tau')$, then add the closures of the schemata

- $\exists\beta(\beta = \tau)$
- $\tau = \tau'$

as new, necessary axioms to our system; otherwise add the $\Box$-free closures of these same schemata as new, necessitation-averse axioms to our system, with item numbers that are appropriately marked with a $\star$ whenever cited.

The second part of clause (.b) covers the example described in the latter half of Remark (201). If we were to extend our theory with the contingent axiom that there exists a unique moon of the Earth ($\exists!xMxe$), then by theorem (96)$\star$, $\exists y(y = \imath xMxe)$ would become a necessitation-averse $\star$-theorem. So by the second part of clause (.b) of the above Convention for Term Definitions, the definition:

$$m =_{df} \imath xMxe$$

would be a proxy for both extending the language of our system with the individual constant m and adding the $\Box$-free closures of the following:

$$\exists y(y = m)$$

$$m = \imath x Mxe$$

as *necessitation-averse* axioms. Any derivation that depended on these axioms would fail to be modally-strict. Note, though, that by an instance of the modally-strict lemma (119), namely, $\exists y(y = \imath xMxe) \rightarrow \Box\exists y(y = \imath xMxe)$, it would follow that there is a non-modally strict proof of $\Box\exists y(y = \imath xMxe)$. This is as it should be: *given* the contingent axiom that there is a unique moon of Earth, then (i) necessarily, the thing that is in fact a moon of the Earth exists, and (ii) the proof of this necessity claim rests upon a contingency, as would the proof of the claim that necessarily, there is something that is m.

By contrast, if the definiens τ' is such that there is a modally strict proof of $\exists\beta(\beta = \tau')$, then the first part of clause (.b) in the Convention for Term Definitions applies. We know that such proofs are available in the case where τ' is any relation term, or any individual constant or variable, since such $\exists\beta(\beta = \tau')$ is then axiomatic. It remains to consider the case where the definiendum τ is introduced with the description $\imath x\varphi$ as definiens. Note that if we can produce a modally strict proof of $\exists!x\varphi$, then we can use the Rule of Actualization (RA) to get a modally strict proof and $\mathcal{A}\exists!x\varphi$, and then apply theorem (172.2) to obtain a modally strict proof of $\exists y(y = \imath x\varphi)$. This case is then governed by the first part of clause (.b) in the Convention for Term Definitions, even when there are free variables in $\exists!x\varphi$. In all of these cases, the definition becomes a convention for extending the language with τ and adding *all* the closures of $\exists\beta(\beta = \tau)$ and $\tau = \tau'$ as new, necessary axioms to our system.

Consider what the above convention implies in the interesting case where the definition has a free individual variable. Let us start with the fact that the following is a theorem of object theory (indeed, an instance of Strengthened Comprehension for abstract objects (170)), in which the variable z occurs free:

$$\exists!x(A!x \,\&\, \forall F(xF \equiv Fz))$$

By (173), we then know that there is a modally strict proof of:

$$\exists y(y = \imath x(A!x \,\&\, \forall F(xF \equiv Fz)))$$

By GEN, it follows that:

$$\forall z\exists y(y = \imath x(A!x \,\&\, \forall F(xF \equiv Fz)))$$

Hence, we've met Constraint (201) and may introduce the following function term $\boldsymbol{c}_z$, where the variable z occurs free in both definiendum and definiens:

(ϑ) $\boldsymbol{c}_z =_{df} \imath x(A!x \,\&\, \forall F(xF \equiv Fz))$

By the Convention for Term Definitions (.1), the above line is a convention for adding c_z as a primitive function term and asserting the closures of the (necessary) axiom:

$$\boldsymbol{c}_z = \imath x(A!x \,\&\, \forall F(xF \equiv Fz))$$

Hence, it is an axiom that:

$$\forall z(\boldsymbol{c}_z = \imath x(A!x \,\&\, \forall F(xF \equiv Fz)))$$

Given our axioms and rules for universal instantiation, we may not instantiate $\forall z$ in the above axiom to a description, say $\imath wPw$, if it isn't provable that $\exists y(y = \imath wPw)$. So not every substitution instance of definition (ϑ) is a *proper* instance that becomes assertible as an identity.

To look at this in another way, when $\imath wPw$ fails to denote, then the definiens:

$$\imath x(A!x \,\&\, \forall F(xF \equiv F\imath wPw))$$

also fails to denote and so *fails* Convention (201). In such a case, it may not be used to define $\boldsymbol{c}_{\imath wPw}$.[127]

Given the Convention for Term Definitions, it immediately follows by (47.1) and (47.3) that definitions become assertible as theorems having the form of (the closures of) identity statements and are thus derivable from any (possibly empty) set of premises:

(.2) **Rules of Identity by Definition**
Where one of τ, τ' is the definiendum and the other definiens in any proper instance of a term definition obeying Contraint (201):

(.a) $\Gamma \vdash_\Box \varphi$, where φ is any closure of the identity $\tau = \tau'$ and provided there is a *modally strict* proof of $\exists\beta(\beta = \tau')$.

(.b) $\Gamma \vdash \varphi$, where φ is any $\Box$-free closure of $\tau = \tau'$

Consider clause (.2.a). If there is a modally strict proof of $\exists\beta(\beta = \tau')$ and the definition $\tau =_{df} \tau'$ is just a convention that, among other things, extends our system with new, necessary axioms $\tau = \tau'$ and their closures, as described in (.1), then by (47.1), all such formulas become theorems and, by (47.3), derivable from any set of premises. So in any reasoning context, we may assert $\tau = \tau'$, for any proper instance of the definition that obeys Constraint (201). If there is no modally-strict proof of $\exists\beta(\beta = \tau')$, then clause (.2.b) tells us that

[127] It is worth observing here that if $\imath wPw$ is not logically proper, it may still appear in a proper instance of a term definition. Suppose, for example, we define the new 0-place relation term, $\overline{p}$, as $[\lambda\, \neg p]$. Then the following is a *proper* instance of this definition: $\overline{Q\imath wPw} =_{df} [\lambda\, \neg Q\imath wPw]$. Even if $\imath wPw$ isn't logically proper, the λ-expression $[\lambda\, \neg Q\imath wPw]$ is. Hence, $\overline{Q\imath wPw}$ is well-defined. We'll see further examples of this phenomenon below.