Elementary Formal Logic

Brian F. Chellas

PERRY LANE PRESS

Elementary Formal Logic

Crime is common. Logic is rare.

—Sherlock Holmes

Elementary Formal Logic

Brian F. Chellas

PERRY LANE PRESS • 1996

Second printing, with corrections, fall 1996
Third printing, with corrections, summer 1997
Fourth printing, with corrections, summer 1998
Fifth printing, with corrections, summer 1999
Sixth printing, with corrections, spring 2001

CONTENTS

PREFACE

This book is an introduction to sentential and first order predicate logic—that is, to elementary formal logic. It is meant to be used without prerequisites in undergraduate courses. Chapters 1–4, 6–8, and most of 10 can be covered in a single term; the whole text can easily be covered in two.

The semantic concept of logical consequence is illustrated and defined in chapter 1 using the ideas of models and truth. The allied concepts of validity, equivalence, and satisfiability are also presented.

After chapter 1 the book is divided into three parts. Part I is about sentential logic: chapter 2 introduces the language in terms of translation and syntax; chapter 3 explains semantics; chapter 4 presents a deductive system; chapter 5 contains a proof that the system is complete. Part II follows the pattern of part I, this time for predicate logic: chapter 6 is about translation and syntax; chapter 7 concerns models and truth; chapter 8 presents a deductive system; chapter 9 contains a completeness proof. Part III is about the logic of identity and covers the language, semantics, deduction, and completeness all in a single chapter (10). Collections of exercises are provided to illustrate and expand on points made in the sections that precede them. An appendix reviews some notions and notation from elementary set theory; another appendix provides solutions to selected exercises.

The book first appeared in 1973 in an incomplete form, and has grown as I have used it in courses at the Universities of Pennsylvania, Michigan, and Calgary. The present text is the fourth complete edition since 1993.

For instruction and inspiration I am grateful to a number of teachers and colleagues, notably Robert Beard, Solomon Feferman, Richard Jeffrey, Stig Kanger, J. J. MacIntosh, and Dana Scott. No less influential have been the students whose reactions to the book have led to improvements. Corrections and suggestions continue to be welcomed.

I dedicate the book to Merry, who has been its most logical proponent, and to the memory of Papalady, who always hoped I would be a dean.

Calgary, Alberta B.F.C.
June 1996

1

INTRODUCTION

LOGIC is the science, or art, of reasoning. *Elementary formal logic* is the study of reasoning within the realms of sentential and first order predicate logic—the subjects of this book.

We begin this chapter with a brief discussion of *sentences* and *arguments*. Then we introduce the fundamental concept of logic, *consequence*, look at examples, define the concept, and examine some of its important features. Following this, we consider some of the assumptions underlying our approach. Finally we present three further logical concepts: *validity*, *equivalence*, and *satisfiability*.

1. Arguments and sentences

The basic objects of study for us are *sentences* and collections of sentences, particularly those collections we call *arguments*.

Arguments consist of sentences structured in a certain way. For example:

> Every human is mortal. Sitting Bull is a human.
> Therefore, Sitting Bull is mortal.

This argument consists of three sentences. The first two—Every human is mortal and Sitting Bull is a human—are *premisses*, while the third—Sitting Bull is mortal—is the *conclusion*. In general, that is what an argument is: a structured collection of one or more sentences one of which is the conclusion and the others premisses. Thus by an argument we do not mean an act or event of disputation, but rather the reasoning employed on such occasions—the product, not the production, of reasoning.

In this example the word "therefore" indicates which of the sentences is the conclusion. For our purposes, we shall separate premisses and conclusion formally by means of a line. Thus we may describe the argument in this way:

$$\begin{array}{l}\text{Every human is mortal}\\ \text{Sitting Bull is a human}\\ \hline \text{Sitting Bull is mortal}\end{array}$$

A few comments on the structure of an argument. First, the order of the premisses makes no difference, and it does not matter if there are duplicates. So the argument above is the same as each of these:

$$\begin{array}{l}\text{Sitting Bull is a human}\\ \text{Every human is mortal}\\ \hline \text{Sitting Bull is mortal}\end{array} \qquad \begin{array}{l}\text{Every human is mortal}\\ \text{Sitting Bull is a human}\\ \text{Every human is mortal}\\ \hline \text{Sitting Bull is mortal}\end{array}$$

Secondly, when we speak of "a structured collection of one or more sentences" we mean some *finite* number greater than zero. So arguments always contain at least one sentence, but never have infinitely many.

Question: What if an argument has just one sentence? *Answer*: In this case the sentence is the conclusion of the argument and there are no premisses. For example:

$$\begin{array}{l}\\ \hline \text{If it's raining then it's raining}\end{array}$$

In other words, we count a "zero-premiss" argument as none the less an argument, despite its lack of premisses. Though this stretches the notion of argument a bit, it does so very usefully, as we shall see.

EXERCISES

1.1. Use the sentences in the arguments above to make new arguments. (Do not worry about whether the arguments seem cogent!)

1.2. Assume that every human is either male or female, that what females say is always true, and that what males say is always false. Consider the following dialogue.

Butch:	"Sundance is female."
Sundance:	"Etta and Butch are the same sex."
Etta:	"I'm female."

Butch, Sundance, and Etta are human. Is Etta telling the truth—i.e. is Etta female?

2. Consequence

The main question about an argument is simply whether or not its conclusion is a *consequence* of its premisses. Consider again the argument:

Every human is mortal
Sitting Bull is a human
———
Sitting Bull is mortal

The question is whether the conclusion Sitting Bull is mortal is a consequence of the premisses Every human is mortal and Sitting Bull is a human. In fact, the conclusion is a consequence; the argument is *valid*.

An important note on terminology:

> To say of an argument that the *conclusion is a consequence of the premisses* is also to say that the *argument is valid*.

Thus while consequence is a *relation between* a collection of sentences and a single sentence, validity is a *property of* an argument taken as a whole. In this book, however, we also use "valid" and "validity" for sentences of a certain kind. The idea of validity for sentences appears in section 10 below, and the connection between valid sentences and valid arguments is stated in chapter 3.

For contrast, here is an example of non-consequence:

Every bear has hair
Custer has hair
———————
Custer is a bear

In this case, as should be obvious, the conclusion, Custer is a bear, is not a consequence of the premisses, Every bear has hair and Custer has hair; in other words, the argument is *invalid*.

As we said above, consequence is a relation that holds, or fails to hold, between a collection of sentences and a single sentence. When an argument consists of a single sentence—i.e. of nothing but a conclusion—the question is whether that sentence is a consequence of the *empty set* of sentences. For example, one may ask whether If it's raining then it's raining is a consequence of the empty set, as in this argument from the preceding section:

———————
If it's raining then it's raining

We shall soon see what such a question means—and discover that much of the focus of logic is on individual sentences.

Finally, we should emphasize that the notion of consequence goes beyond validity of arguments; though we concentrate on arguments we do not restrict the study of consequence just to these finite structures. In considering the relation of consequence generally, no limits are placed on the number of sentences involved, since we are often interested in whether a sentence is a consequence of an *infinite* set of sentences. For instance, one may ask whether the sentence Wyatt could outshoot Doc is a consequence of the set of all true sentences containing the word Tombstone (this set is infinite).

EXERCISES

1.3. Using the sentences in the argument above about Custer, make a new argument in which the conclusion is a consequence of the premisses.

1.4. Make another new argument using the sentences in the argument about Custer so that the conclusion is not a consequence of the premisses.

1.5. Consider the zero-premiss argument with the conclusion If it's raining then it's raining. Is this argument valid—is the conclusion a consequence of the premisses?

1.6. Give more examples of zero-premiss arguments—some valid, some not.

3. Terminology and notation

For ease of exposition, we adopt a standard symbol for consequence. Where a sentence A is a consequence of a set of sentences Γ we use a double turnstile and write:

$$\Gamma \models A$$

In the negative case, where A is not a consequence of Γ, we write:

$$\Gamma \not\models A$$

The relation of consequence is also called *implication*. When a sentence is a consequence of a set of sentences the sentence is said to be implied by the set. Thus "$\Gamma \models A$" may be read "A is a consequence of Γ" or "Γ implies A".

When A is a consequence of a finite set of sentences $\{A_1, \ldots, A_n\}$, we write:

$$\{A_1, \ldots, A_n\} \models A$$

Or we may omit the braces and simply write:

$$A_1, \ldots, A_n \models A$$

Thus in the case of an argument considered earlier, to say that the conclusion is a consequence of the premisses we might write:

{Every human is mortal, Sitting Bull is a human} $\models$ Sitting Bull is mortal

Or just:

Every human is mortal, Sitting Bull is a human $\models$ Sitting Bull is mortal

Similarly for negative cases.

As noted, we use a line to separate premisses from conclusion in an argument:

$$\begin{array}{c} A_1 \\ \cdot \\ \cdot \\ \cdot \\ A_n \\ \hline A \end{array}$$

It is often convenient and practical to "horizontalize" the description of an argument, using a slanted line:

$$A_1, \ldots, A_n \ / \ A$$

This is a good place to emphasize the difference between referring to an argument or sentence and saying something about it. In simply referring to an argument we do not thereby assert anything about it. Thus if we refer to an argument—for example, by writing

Every human is mortal, Sitting Bull is a human / Sitting Bull is mortal

—we are not thereby saying that its conclusion is, or is not, a consequence of its premisses. Likewise, in referring to a sentence—for example, by displaying

Custer is a bear

—we are not at the same time saying that it is true (or false, for that matter).

Throughout this book we use capital roman letters—A, B, C, . . .—for sentences, and some capital greek letters—Γ, Δ, E (gamma, delta, epsilon)—for sets of sentences. We also employ a modest vocabulary from set theory, for example: $\in$ (membership), $\subseteq$ (inclusion), $\cap$ (intersection), $\cup$ (union), $\varnothing$ (the empty set). The first appendix at the end of the book has more information about these and other elementary set-theoretical concepts.

Further terminology and notation will be introduced as we go along. But note, finally, that "iff" is short for "if and only if".

EXERCISES

1.7. Coyotes tell the truth on Friday, Saturday, Sunday, and Monday. They speak falsely on the other days of the week. Roadrunners' truth days are Monday, Tuesday, Wednesday, and Thursday. Their falsity days are the others. One day Yosemite Sam met a coyote and a roadrunner. He overheard the following conversation.

Coyote:	"Yesterday was one of my falsity days."
Roadrunner:	"Yesterday was one of my falsity days."

Yosemite Sam then deduced the day of the week. What day was it? (Adapted from Raymond Smullyan's *What is the name of this book?*)

4. Truth, falsity, and consequence

What distinguishes consequence from non-consequence? To answer this question let us focus on an example:

(1)
All carrots are vegetables
No cabbages are carrots
———————————
No cabbages are vegetables

In this argument the conclusion is not a consequence of the premisses. Why?

The answer of course is that the conclusion, No cabbages are vegetables, is false, whereas both the premisses, All carrots are vegetables and No cabbages are carrots are true. The key is in the *truth values* of the sentences in the argument: *truth* for the premisses together with *falsity* for the conclusion means that the conclusion is not a consequence of the premisses. In other words, we have a sufficient condition for non-consequence: if all the sentences in Γ are true and the sentence A is false, then A is not a consequence of Γ ($\Gamma \not\models A$).

This provides an important clue in answer to our question, but it does not answer it completely. The configuration of truth values in argument (1)—truth for the premisses, falsity for the conclusion—is sufficient for non-consequence, but it is not necessary. To see this, consider this pair of arguments:

All skunks are mammals
(2) No sharks are skunks
―――――――――
No sharks are mammals

All lizards are iguanas
(3) No vertebrates are lizards
―――――――――
No vertebrates are iguanas

Neither of these arguments is valid. In (2) the sentence No sharks are mammals is not implied by the pair All skunks are mammals, No sharks are skunks—yet all three sentences are true. Likewise, in argument (3), though all three sentences are false, the conclusion is not a consequence of the premisses—No vertebrates are iguanas is not implied by the pair {All lizards are iguanas, No vertebrates are iguanas}.

So we may still ask what makes the difference between cases where consequence holds and cases where it does not, as in (1), (2), and (3)? The answer, as we shall see, lies in the idea of a *model*.

EXERCISES

1.8. Here are some arguments adapted from *Symbolic logic* by Lewis Carroll (who also wrote *Alice in wonderland*). Which are valid—i.e. in which is the conclusion a consequence of the premisses?

(1)
No fossil can be crossed in love
Any oyster can be crossed in love
———
No oysters are fossils

(2)
No emperors are dentists
All dentists are dreaded by children
———
No emperors are dreaded by children

(3)
Every eagle can fly
Some pigs cannot fly
———
Some pigs are not eagles

(4)
All these bonbons are chocolate creams
All these bonbons are delicious
———
All chocolate creams are delicious

(5)
No doctors are enthusiastic
Ned Buntline is enthusiastic
———
Someone is not a doctor

5. Models

As we have seen, in asking about consequence it is not enough to consider the actual truth values of the sentences involved, the truth values of the sentences given the meanings they in fact have. It is necessary to consider what truth values the sentences might have, or would have if their meanings were different.

To illustrate this, let us interpret some of the words in argument (2) in the preceding section in such a way that the premisses become true and the conclusion false. For example, interpret skunk to mean *carrot*, mammal to mean *vegetable*, and shark to mean *cabbage*. Then the sentences in argument (2) acquire new meanings:

- the first premiss, All skunks are mammals, means that all carrots are vegetables (true);

- the second premiss, No sharks are skunks, means that no cabbages are carrots (also true);
- the conclusion, No sharks are mammals, means that no cabbages are vegetables (false).

In the light of these truth values for the premisses and conclusion, we can refer to the sufficient condition for non-consequence. Since in this model the premisses are true and the conclusion is false, the conclusion of argument (2) is not a consequence of its premisses.

Our (re)interpretation of the expressions skunk, mammal, and shark constitutes a *model*, which we can conveniently represent by means of a table:

Expression	*Interpretation*
skunk	carrot
mammal	vegetable
shark	cabbage

In describing a model in this way we have in effect transformed argument (2) into another argument, one with true premisses and a false conclusion. Indeed, argument (2) has become argument (1) from the preceding section:

(1)
All carrots are vegetables
No cabbages are carrots

No cabbages are vegetables

Viewed in this way, argument (1) is a *counterexample* to argument (2). It is an argument having the same form as (2), but true premisses and a false conclusion. In general, consequence exists when there is no possibility of a counterexample.

We leave as an exercise the problem of finding a model in which lizard, iguana, and vertebrate are interpreted so as to make the premisses of argument (3) true and its conclusion false. (Notice that argument (3) also has the same form as (1).)

Terminological remarks. The terms *model* and *interpretation* are often used more or less interchangeably, though they are not synonymous. *Model* has a broader meaning, suggesting a structure within which an interpretation is the component that interprets a language's expressions. In the study of sentential logic, *interpretation* is more usual, since models in that context are little more than interpretations. In predicate logic *model* is more appropriate, and for this reason we prefer the term.

Model has further uses. When a sentence is true in a model we say that the model *is a model of* or *for* the sentence. Likewise for a set of sentences all of which are true in a model: the model is a model of (or for) the set. In these cases one can also say that the model *models* the sentence or set.

By contrast, to say that an interpretation interprets a sentence or set is not to say that it models the sentence or set.

Negatively speaking, a model in which a sentence is false is said to be a *countermodel to* or *for* the sentence. Note that a countermodel for a set of sentences makes at least one member of the set false—not necessarily every member. In a countermodel for an argument, moreover, all the premisses are true and the conclusion is false.

EXERCISES

1.9. Describe a model that interprets the expressions lizard, iguana, and vertebrate so that in argument (3) in section 4 the premisses are true and the conclusion is false.

1.10. Describe a model that interprets Canadian, North American, and Albertan so that in the following argument the premisses are true and the conclusion is false.

$$\frac{\begin{array}{l}\text{All Canadians are North Americans}\\ \text{Some North Americans are Albertans}\end{array}}{\text{Some Canadians are Albertans}}$$

1.11. Describe a model that interprets emperor, dentist, and dreaded by children so that in argument (2) in exercise 1.8 the premisses are true and the conclusion is false.

1.12. Describe a countermodel for argument (4) in exercise 1.8. (Interpret these bonbons, chocolate cream, and delicious.)

1.13. In creating a model it is not always necessary to interpret all the expressions in the sentences involved. For example, a countermodel for argument (2) on page 8 can be devised by (re)interpreting just one of the expressions (skunk, mammal, and shark) and letting the meanings of the others remain the same. Describe a model of this sort.

6. Consequence defined

The idea of truth in a model provides a necessary and sufficient condition for non-consequence: a sentence is not a consequence of a set of sentences if and only if in some model the sentence is false and the sentences in the set are true. Put in a positive form, this is a criterion for consequence itself: a sentence is a consequence of a set if and only if there is no model in which the sentences in the set are true while the sentence is false. This then is our definition of consequence. We state it below in these equivalent positive and negative forms.

DEFINITION 1.1. *Consequence.*

- $\Gamma \models A$ iff A is true in every model in which every sentence in Γ is true.

- Equivalently: $\Gamma \models A$ iff there is no model in which A is false while every sentence in Γ is true.

This is our first and most important definition. Notice that we can make use of the other meaning of *model* to phrase other, equally good versions of the definition: $\Gamma \models A$ if and only if every model of Γ is also a model of A; and $\Gamma \models A$ if and only if there is no model of Γ that makes A false.

Applied to arguments, the definition means:

- An argument is valid iff the conclusion is true in every model in which all the premisses are true.

- Equivalently: An argument is valid iff there is no model in which the conclusion is false while all the premisses are true.

Even more briefly: an argument is valid just in case it has no countermodels.

In the next section we set forth some important propositions about the consequence relation.

EXERCISES

1.14. Here are some more arguments adapted from Lewis Carroll's *Symbolic logic*. For each argument, describe a model that shows that the conclusion is not a consequence of the premisses.

(1)
All lions are fierce
Some lions are not coffee drinkers

Some coffee drinkers are not fierce

(2)
Some epicures are ungenerous
All my uncles are generous

None of my uncles are epicures

(3)
No professor is an ignoramus
Every ignoramus is vain

No professor is vain

(4)

Some pillows are soft
No pokers are soft

Some pokers are not pillows

(5)

No frogs are poetical
Some ducks are unpoetical

Some ducks are not frogs

1.15. Describe a countermodel for this argument:

Every European is French or Spanish
Some Libyans are Europeans
Not every Libyan is French

Some European is Spanish

7. Some properties of consequence

The relation of consequence has the fundamental properties of *reflexivity*, *transitivity*, and *augmentation*, which we set out below.

First of all, every sentence is a consequence of itself. More generally, a set of sentences implies each of its members.

PROPOSITION 1.2. *Reflexivity of consequence.*

(1) If $A \in \Gamma$ then $\Gamma \models A$.

(2) Corollary: $A \models A$.

Proof. We argue first for part (1). Consider any model that makes all of the sentences in Γ true, and suppose that A is in Γ. Then A must be true in the model, as we wished to show. (Alternatively, suppose that $A \in \Gamma$, but Γ does not imply A. This means that there is a model in which every sentence in Γ is true while A is false. But then A is both true and false in the model, which is impossible.) Part (2) now follows from part (1), since it says in particular that if $A \in \{A\}$ then $\{A\} \models A$—and it is always the case that $A \in \{A\}$. But there is no need to infer (2) from (1), since (1) simply means that every model of A is a model of A. □

Proposition 1.2 provides the foundation for the basic rule of inference called P in chapter 4. Among other things, it means that any argument is valid in which the conclusion is also one of the premisses, as for example in:

Grass is green Sitting Bull is mortal The sky is blue	Sitting Bull is mortal
Sitting Bull is mortal	Sitting Bull is mortal

This may seem puzzling, as it seems commit the so-called fallacy of "begging the question", according to which one cannot prove something by assuming that it is so. This is accurate as far as the rhetorical notion of proof goes: neither of the arguments is very persuasive. But it is off base as regards the logical notion of consequence, the relation that holds when it is impossible for a sentence to be false while all the sentences in a set are true. To achieve full generality in our understanding of this concept, we must acknowledge extreme cases such as those where an argument's conclusion is among its premisses. Any sentence in a set is implied by the set.

The next proposition says that consequence is a transitive relation. For individual sentences A, B, and C, this means that if A implies B and B implies C, then A implies C. As in the case of proposition 1.2, there are more general statements of the property as well: e.g. part (1) below says that if one set of sentences implies all the members of a second set, then the first set implies any sentence the second set does.

PROPOSITION 1.3. *Transitivity of consequence.*

(1) If $\Gamma \models B$ for every sentence $B \in \Delta$, and $\Delta \models A$, then $\Gamma \models A$.

(2) If $\Gamma \models B$ and $\Delta \cup \{B\} \models A$, then $\Gamma \cup \Delta \models A$.

(3) Corollary: If $A \models B$ and $B \models C$, then $A \models C$.

Proof. We give the proof for the corollary, and leave parts (1) and (2) as exercises. Suppose that A implies B and that B implies C. Then B is true in every model in which A is true, and likewise C is true in every model in which B is. So C is true in every model in which A is true—which means that C is a consequence of A, as we wished to show. □

It is transitivity that makes "chain reasoning" possible. For a simple example of this, note that in each of the following arguments the conclusion is a consequence of the premiss.

> It's raining
> ___
> It's raining or it's pouring

> It's raining or it's pouring
> ___
> It's raining or it's pouring or the old man is snoring

Since this is so, it follows by transitivity that the conclusion is a consequence of the premiss in this argument:

> It's raining
> ___
> It's raining or it's pouring or the old man is snoring

Next we have a proposition that affirms a principle of augmentation for consequence: A consequence of a set of sentences is also a consequence of any enlargement of that set. We state this formally as follows.

PROPOSITION 1.4. *Augmentation for consequence.*

$$\text{If } \Gamma \models A \text{ and } \Gamma \subseteq \Delta, \text{ then } \Delta \models A.$$

Proof. Suppose that A is a consequence of Γ, and all of Γ's members belong to Δ. We wish to argue that A is a consequence of Δ as well. So suppose that a model exists that makes every sentence in Δ true (to show that A is also true in this model). Then every sentence in Γ is also true in the model. But our assumption tells us that A is true in every model in which all Γ's sentences are true. So A is true in the given model. □

In our proofs of propositions 1.2, 1.3, and 1.4 we have argued for properties of consequence directly in terms of the definition of this concept. It is interesting to observe, however, that the third property, augmentation, follows from the first two. To see this, assume again that Γ implies A and that Γ is a subset of Δ. By reflexivity, Δ implies all its own members and hence every sentence in Γ. By transitivity, then, Δ implies A, as we wished to show.

Augmentation (sometimes called monotonicity) illustrates the strength of the logical notion of consequence and differentiates it from other, analogous relations. Augmentation means that once a sentence is implied by a set of sentences, no additional information—no further premisses, so to speak—will alter the situation. For example, we have agreed that the conclusion of the argument

Every human is mortal
Sitting Bull is a human

Sitting Bull is mortal

is a consequence of its premisses. By augmentation, then, the conclusion Sitting Bull is mortal is a consequence of the premisses of these arguments:

Every human is mortal	Every human is mortal
Sitting Bull is a human	Sitting Bull is a human
Grass is green	Sitting Bull is not mortal
———	———
Sitting Bull is mortal	Sitting Bull is mortal

This may seem odd, especially in the second argument, where one of the premisses explicitly contradicts the conclusion. If so, it may be reassuring to check again that there is no model in which the conclusion is false while the respective sets of premisses are all true.

Our final proposition uses the combined forces of transitivity and augmentation to express generally how arguments may be combined and chained.

PROPOSITION 1.5. *Combining and chaining.*

$$\text{If } \Gamma_1 \models A_1, \ldots, \Gamma_n \models A_n, \text{ and } A_1, \ldots, A_n \models A, \text{ then } \Gamma_1 \cup \ldots \cup \Gamma_n \models A.$$

Proof. Let us give the reasoning for the case in which $n = 2$, and leave the general argument as an exercise. When $n = 2$ we can recast the proposition simply in this way:

$$\text{If } \Gamma \models A \text{ and } \Delta \models B, \text{ and } A, B \models C, \text{ then } \Gamma \cup \Delta \models C.$$

To prove this, proposition, suppose that A is a consequence of Γ, B is a consequence of Δ, and C is a consequence of the pair A, B. By augmentation each of A and B is also a consequence of the set $\Gamma \cup \Delta$. So by transitivity it follows that C is a consequence of $\Gamma \cup \Delta$, as we wished to show. □

The converse of the principle of augmentation of course does not hold: a consequence of a set of sentences is not in general a consequence of all of its subsets. But something close to it does obtain for the elementary logics studied in this text. This is the property of *compactness*:

$$\text{If } \Gamma \models A, \text{ then there is a finite subset } \{A_1, \ldots, A_n\} \text{ of } \Gamma \text{ such that } \{A_1, \ldots, A_n\} \models A.$$

This property is obviously trivial when Γ itself is a finite set of sentences; for then $\{A_1, \ldots, A_n\}$ can be taken to be identical with Γ, since every set is a subset of itself. The interesting case is when Γ is an infinite set of sentences. In this case, when compactness holds, if A is a consequence of Γ then A is already implied by some proper subset $\{A_1, \ldots, A_n\}$ made up of finitely many sentences $A_1, \ldots, A_n$ from Γ.

The importance of compactness is that whenever a sentence A is implied by an infinite set Γ (a theory, for example) it is also implied as the conclusion of an argument

$$A_1, \ldots, A_n \ / \ A$$

in which the premisses $A_1, \ldots, A_n$ form a finite subset of Γ. To the extent that it is possible to prove consequence in such a finite case it is therefore possible to prove it in the original infinite case—perhaps by systematically searching for implying premisses $A_1, \ldots, A_n$ among the finite subsets of Γ.

Our understanding of the consequence relation relies on several assumptions, notably about truth values and about logical constants, topics we take up in the following two sections.

Note on terminology. The propositions in this section are our first theorems about logic. Starting in chapter 4 we also use *theorem* for sentences of a special kind that occur in the logics we study ("theorems of logic"). Propositions about logic are more properly called *metatheorems*. The context usually makes it clear whether we are referring to theorems *about* logic or to theorems *in* logic; where it might be unclear we say "metatheorem".

EXERCISES

1.16. Complete the proof of proposition 1.3 (transitivity) by arguing for parts (1) and (2).

1.17. Show that part (3) of proposition 1.3 is a corollary by deducing it from part (1) or part (2) directly, without using the definition of consequence.

1.18. Use the definition of consequence to prove:

If $\Gamma \models A$ then $\Gamma \cup \Delta \models A$.

In fact, this is an alternative statement of the property of augmentation for consequence (proposition 1.4). Show this by arguing (without using the definition of consequence) that the alternative statement holds if proposition 1.4 does, and that proposition 1.4 holds if the alternative does. (In one direction it helps to note that Γ is always a subset of $\Gamma \cup \Delta$.)

1.19. Appealing just to reflexivity, part (1) of transitivity, and augmentation, prove part (2) of proposition 1.3.

1.20. Of course the reverse of augmentation—diminution—does not hold: deleting a premiss from a valid argument generally results in an invalid argument. Give an example to show this.

1.21. Give the full reasoning in the proof of proposition 1.5.

1.22. It is an important (and sometimes at first puzzling) fact that in every model every sentence in the empty set, ∅, is true. Explain why. (*Hint*: Suppose otherwise—that it is not the case that in every model every sentence in ∅ is true—and argue that this leads to something impossible.)

8. Truth values

Our definition of consequence speaks of sentences' having truth values. Inasmuch as we are just commencing our study of logic, it seems appropriate to make the following simplifying assumption:

> *In any model every sentence is either true or false, but not both.*

In other words, we assume that in any model every sentence has just one truth value.

Because we stick to this assumption we must set aside the logical study of sentences that have no truth values. Imperatives and interrogatives are two kinds of sentences that come to mind as examples. Unlike sentences in the indicative mood, at least on the face of it, sentences in the imperative and interrogative moods do not have truth values. Consider Close the door! and Is the door closed? Neither of these is usually reckoned to be either true or false. This is not to say, of course, that the meaning of such sentences is unrelated to the meanings of sentences that do have truth values (such as The door is closed). And it is also not to deny that such sentences have a logic.

An example of a sentence in the indicative mood that may seem to lack a truth value is The present king of France is bald. There is no present king of France, and thus is it is not true that he is bald—but not false either. Sentences like this have captured the attention of philosophers, some arguing that they have no truth value and others saying they do. As we shall see in chapter 10, there is a way to understand such sentences as truth-valued.

One half of our simplifying assumption is that in any model each sentence has at least one truth value. The other half is that each sentence has at most one truth value. Are there sentences that have more than one—sentences that are both true and false? Consider:

This sentence is false.

If the sentence is true, then, given what it means, it is also false. On the other hand, if it is false, then, again given what it means, it is also true. So in either case the sentence is both true and false; it suffers from a surfeit of truth values.

The status of sentences like this, a version of the so-called Paradox of the Liar, has been debated for centuries. In the context of our study, however, we shall ignore these examples and assume, persistently, that no sentence is both true and false in any model.

EXERCISES

1.23. John has no children. What is the truth value of the sentence All John's children are asleep?

1.24. What students say is always true. What professors say is always false. (Never mind about deans.) Two persons, X and Y, appear before you. Each is either a student or a professor; neither is both. X says: "At least one of us, X and Y, is a professor." What is X and what is Y?

1.25. Consider the following pair of sentences.

The sentence immediately below is false.
The sentence immediately above is true.

Is the upper sentence true or false? The lower sentence?

9. Logical constants

Consider again the very first argument in this chapter:

Every human is mortal
Sitting Bull is a human
______
Sitting Bull is mortal

The conclusion is obviously a consequence of its premisses. That is to say, there is no model that will make the premisses true and the conclusion false.

Or is there? What about this:

Expression	*Interpretation*
human	prime number
mortal	even
Sitting Bull	seventeen
every	some

Under this interpretation, the argument in question looks like this:

Some prime number is even
Seventeen is a prime number
______
Seventeen is even

But here the premisses are true (two is an even prime number, seventeen is an odd one) and the conclusion is false. Something seems to have gone wrong—we have a counterexample to an evident case of consequence!

Of course the trouble lies in our treatment of the word every. By interpreting every to mean *some* we have gone too far. In the context of this argument, the meaning of every should remain constant; it should not be subject to variation through reinterpretation. The general point is that if we are allowed to (re)interpret symbols in a sentence at will, then we can always change the sentence's truth value in whatever way we wish, from truth to falsity or vice versa.

So the moral of this spurious counterexample is that we must distinguish logical constants, as they are called, from non-logical symbols. The meaning of a logical constant remains the same throughout models; we are not permitted to give it different meanings in different models. The meanings of non-logical symbols, by contrast, can change from one model to another; we may give them different meanings in different models.

Logical constants give content to the idea of *logical form*. By treating the words human, mortal, and Sitting Bull as the only non-logical symbols in the argument above, we in effect specify a logical form for the argument:

$$\begin{array}{l} \text{Every } P \text{ is } Q \\ t \text{ is a } P \\ \hline t \text{ is } Q \end{array}$$

It is important to realize that the notion of logical constant is not fixed once and for all. In chapters 2 through 5 we study logical constants corresponding to words such as "not", "and", "or", and "if". In chapters 6 through 9 we add counterparts to such words as "every" and "some". Then in chapter 10 we recognize identity, "=", as yet another logical constant. (Thus, interestingly, we do not in fact regard "every" as a logical constant until chapter 6.)

EXERCISES

1.26. Consider the following argument.

$$\begin{array}{l} \text{Emmett outlived Bob} \\ \text{Bob outlived Grat} \\ \hline \text{Emmett outlived Grat} \end{array}$$

Interpret the expressions Bob, Emmett, Grat, and outlived so that the premisses are true and the conclusion is false. Is the conclusion really not a consequence of the premisses? Is outlived a logical constant? What *are* the logical constants in this argument?

10. Validity, equivalence, and satisfiability

In this section we add to consequence three further important logical concepts and then state a number of relationships among all four.

Validity. As we remarked in section 2, sentences as well as arguments may be said to be valid or invalid.

To say that a sentence is valid is to say that it must be true—or that it cannot be false. For this reason sentential validity is sometimes called *logical truth*. Examples of valid sentences:

If it's raining then it's raining

Every lizard is a lizard

37 = 37

A little reflection will show that there is no way any of these sentences can fail to be true.

We employ the double turnstile, once again, and write

$$\models A$$

to mean that A is valid, and

$$\not\models A$$

to mean that A is not valid, or that A is invalid. The formal definition of validity, in equivalent positive and negative versions, is as follows.

DEFINITION 1.6. *Validity.*

- $\models$ A iff A is true in every model.
- Equivalently: $\models$ A iff there is no model in which A is false.

Proposition 1.7 below depends on a fact important enough to be registered as a lemma:

Lemma. In any model, every sentence in the empty set is true.

Why? Suppose otherwise, i.e. suppose that in some model there is a sentence in the empty set that is false. But this implies that there is a sentence in the empty set, which means that the empty set is not empty. This is impossible.

(This fact was mentioned earlier; see exercise 1.22 above, and compare exercise 1.28 below.)

Proposition 1.7 relates the notions of consequence and validity, and incidentally justifies using the double turnstile to express both, since it shows how we might have defined validity in terms of consequence. According to part (1), a sentence is valid just in case it is implied by the empty set; part (2) says that a sentence is valid just when is implied by every set of sentences.

PROPOSITION 1.7. *Definability of validity in terms of consequence.*

(1) $\models$ A iff $\varnothing \models$ A.

(2) $\models$ A iff for every set of sentences Γ, $\Gamma \models$ A.

Proof. For (1) from left to right, suppose that A is valid. This means that there is no model in which A is false. It follows that there is no model in which A is false while all the sentences in the empty set are true. But this means that A is a consequence of the empty set, as we wished to show. For right-to-left, suppose that the empty set implies A. This means that A is true in every model in which every the sentence in the empty set is true. By the lemma, all the sentences in the empty set are true. Hence we conclude that, like the sentences in the empty set, A is true in every model—again as we wished to show.

We use part (1) and augmentation (proposition 1.4) and to prove part (2). For left-to-right, suppose that A is valid. Then by (1), A is a consequence of the empty set, which is a subset of every set Γ. So by augmentation every set Γ implies A. For right-to-left, if A is a consequence of every set of sentences, then in particular A is a consequence of the empty set. So by part (1) again, A is valid. □

To say that proposition 1.7 asserts that validity is definable in terms of consequence is to say that we equally well could have chosen to define validity as implication by the empty set or as implication by all sets. Had we done so, the actual definition, 1.6, would have become a metatheorem.

In one direction, part (2) of proposition 1.7 means that, as the conclusion of an argument, a valid sentence is a consequence no matter what the premisses are. So, for example, because the conclusions in each of the following arguments are are valid, the arguments themselves are valid.

Every human is mortal	All cabbages are canines
Sitting Bull is a human	No cats are cabbages
If it's raining then it's raining	37 = 37

Equivalence. *Equivalence* is a relation between sentences that holds just in case they are true in all the same models (and hence false in all the same ones). We write

$$A \simeq B$$

to say that A is equivalent to B. To deny that A and B are equivalent (which is to affirm that there is a model in which they have different truth values) we write:

$$A \not\simeq B$$

DEFINITION 1.8. *Equivalence.*

$$A \simeq B \text{ iff in every model, } A \text{ and } B \text{ have the same truth value.}$$

Another way of saying that sentences are equivalent is to say that they have exactly the same models.

Here are some examples of equivalent pairs:

Some clowns are not funny, Not all clowns are funny

If it's raining then it's pouring, If it's not pouring then it's not raining

Every carrot is a carrot, If it's raining then it's raining

The last is a rather extreme case; the two sentences are equivalent because each is valid (true in every model) and hence true and false alike in every model.

Clearly, sentences are equivalent just in case they imply each other. This is the content of the following theorem.

PROPOSITION 1.9. *Definability of equivalence in terms of consequence.*

$$A \simeq B \text{ iff both } A \models B \text{ and } B \models A.$$

Proof. For left-to-right, suppose that A and B are equivalent. Then their truth values do not differ under any interpretation. So there is no model in which A is true and B is false, and there is no model in which B is true and A is false. This means that A implies B and B implies A. The reasoning for the reverse is left as an exercise. □

Equivalence, sometimes also called *logical equivalence*, gets its label from the fact that it is an equivalence relation among sentences. This means that equivalence is reflexive, symmetric, and transitive. It is worth recording formally.

PROPOSITION 1.10. *Equivalence an equivalence relation.*

(1) *Reflexivity*: $A \simeq A$.

(2) *Symmetry*: If $A \simeq B$ then $B \simeq A$.

(3) *Transitivity*: If $A \simeq B$ and $B \simeq C$, then $A \simeq C$.

Proof. We leave the proofs of reflexivity and symmetry as exercises. For transitivity, suppose that A is equivalent to B and B is equivalent to C. Then in every model the truth values of A and B agree, and likewise for B and C. So there is no model in which A and C disagree as to truth value—which means that A is equivalent to C. □

As we saw in an example above, all valid sentences are equivalent. Let us state this for the record.

PROPOSITION 1.11. *Equivalence of valid sentences.*

If $\models A$ and $\models B$, then $A \simeq B$.

Proof. We leave the reasoning as an exercise. □

Satisfiability. To say that a set Γ is satisfiable is to say that there is at least one model in which the sentences in Γ are true together—that Γ has a model. Equivalently, A set Γ has the opposite property, *un*satisfiability, just in case there is no model in which all Γ's sentences come out true together.

For satisfiability we write:

$$\mathrm{Sat}\ \Gamma$$

For unsatisfiability:

$$\mathrm{S}\cancel{\mathrm{a}}\mathrm{t}\ \Gamma$$

Though satisfiability is primarily a property of sets of sentences, when a set contains just one sentence we can affirm (or deny) that the sentence itself has the property. That is to say, in case Sat {A}, we may write

$$\text{Sat A}$$

—and similarly for the negative case.

Satisfiability is sometimes referred to as *simultaneous satisfiability*. It is also called *semantic consistency*, but because we wish to reserve the word "consistency" for another, related notion, we do not adopt this usage. In the special case where a set $\Gamma \cup \{A\}$ is satisfiable, it is sometimes said that Γ and A are compatible or that A is compatible with Γ.

Here is a formal definition, for both satisfiability and unsatisfiability.

DEFINITION 1.12. *Satisfiability (and unsatisfiability).*

- Sat Γ iff there is a model in which every sentence in Γ is true.

- S$\not{a}$t Γ iff there is no model in which every sentence in Γ is true.

Examples of satisfiable sentences and sets of sentences abound; most of the (truth-valued) things we say are probably true, individually or collectively, under some interpretation. Unsatisfiable sets and sentences are perhaps more interesting, as they correspond to what are familiarly termed *contradictions*. Examples:

{Grass is green, Grass is not green}

{Some peaches are both ripe and not ripe}

$\{37 \neq 37\}$

In the last two cases we could as well declare the sentences themselves to be unsatisfiable. For a somewhat less explicit instance of unsatisfiability, consider:

{Every human is mortal, Sitting Bull is human, Sitting Bull is not mortal}

Note that for single sentences validity and unsatisfiability are at opposite logical extremes: valid sentences are true in every model, unsatisfiable sentences ("self-contradictions") are true in none; in between lie all the sentences that are true under some interpretations and false in others.

A notable example of a satisfiable set is $\varnothing$, the empty set.

PROPOSITION 1.13. Sat $\varnothing$.

Proof. We need to show that $\varnothing$ has a model, i.e. that there is a model in which every sentence in the empty set is true. But in every model every sentence in the empty set is true. Since we agree that models or interpretations exist, it follows that at least one model makes all the empty set's sentences true, i.e. that the empty set is satisfiable. □

The following proposition connects consequence and satisfiability.

PROPOSITION 1.14. If $\mathrm{S\not{a}t}\ \Gamma$, then $\Gamma \models \mathrm{A}$ for every A.

Proof. Suppose that Γ is unsatisfiable, and let A be any sentence. By definition, there is no model in which every sentence in Γ is true. It follows that there is no model in which A is false while every sentence in Γ is true. But this means that A is a consequence of Γ. □

According to proposition 1.14, an unsatisfiable set of sentences implies every sentence whatsoever. In particular, a contradictory sentence implies every sentence. This means that in each of the following two arguments the conclusion is a consequence of the premisses.

	Every human is mortal
	Sitting Bull is a human
Grass is green and not green	Sitting Bull is not mortal
————————	————————
Sitting Bull is mortal	Grass is green

This is so even though the conclusions seem to have nothing to do with the premisses.

Later on we shall see that the converse of proposition 1.14 holds for all the languages we study—that is to say, that a set of sentences is unsatisfiable if it implies every sentence.

If a set of sentences is satisfiable, so is any part of it; likewise, if a set is unsatisfiable, so is any enlargement of it. We put these propositions as follows.

PROPOSITION 1.15. *Diminution for satisfiability; augmentation for unsatisfiability.*

(1) If Sat Γ and $\Delta \subseteq \Gamma$, then Sat Δ.

(2) Equivalently: If $\text{S}\not{\text{a}}\text{t}$ Γ and $\Gamma \subseteq \Delta$, then $\text{S}\not{\text{a}}\text{t}$ Δ.

Proof. We argue for (1) and leave (2) as an exercise. Suppose that Γ is satisfiable, so that there is some interpretation under which all Γ's sentences are true together. Such a model will make true all the sentences in any subset of Γ. So if Δ is a subset of Γ, Δ is satisfiable as well. □

As we saw above, all valid sentences are equivalent. For similar reasons, all unsatisfiable sentences are equivalent as well:

PROPOSITION 1.16. *Equivalence of unsatisfiable sentences.*

If $\text{S}\not{\text{a}}\text{t}$ A and $\text{S}\not{\text{a}}\text{t}$ B, then $A \simeq B$.

Proof. Exercise. □

This completes our survey of validity, equivalence, and satisfiability. The exercises contain some further propositions relating these ideas to each other and to that of consequence.

EXERCISES

1.27. Valid sentences imply only valid sentences; in particular, a consequence of a valid sentence is itself valid.

(1) If $\Gamma \models B$ and $\models A$, for every sentence A in Γ, then $\models B$.

(2) If A $\models$ B and $\models$ A, then $\models$ B.

Give the reasoning for (1) and explain why (2) follows.

1.28. With reference to propositions 1.7 and 1.13, explain why in every model every sentence in the empty set, $\varnothing$, is false. (Recall exercise 1.22.) Then explain why this does not mean that this set is unsatisfiable.

1.29. Give the reasoning for the right-to-left direction in proposition 1.9. Explain what it means to say that this theorem shows that equivalence is definable in terms of consequence.

1.30. Show that the relation of equivalence is reflexive and symmetric. That is, prove parts (1) and (2) of proposition 1.10.

1.31. Complete the proof of proposition 1.15.

1.32. Here are two further statements of diminution for satisfiability and augmentation for unsatisfiability:

(1) If Sat Γ then Sat $\Gamma \cap \Delta$.

(2) If $\mathrm{S\not{a}t}$ Γ then $\mathrm{S\not{a}t}$ $\Gamma \cup \Delta$.

Argue that these are indeed equivalent to the versions in proposition 1.15.

1.33. Explain why a satisfiable set implies only satisfiable sentences, or, equivalently, a set is unsatisfiable if it implies an unsatisfiable sentence:

(1) If Sat Γ and $\Gamma \models$ A, then Sat A.

(2) If $\mathrm{S\not{a}t}$ A and $\Gamma \models$ A, then $\mathrm{S\not{a}t}$ Γ.

1.34. As in the case of consequence, the reverse of augmentation for unsatisfiability does not hold. But, for the elementary logics we study here, an analogous *compactness* property obtains:

> Sat Γ if Sat $\{A_1, \ldots, A_n\}$ for every finite subset $\{A_1, \ldots, A_n\}$ of Γ.

In other words, for unsatisfiability:

> If $\text{S\not{a}t}\ \Gamma$, then $\text{S\not{a}t}\ \{A_1, \ldots, A_n\}$ for some finite subset $\{A_1, \ldots, A_n\}$ of Γ.

As in the case of consequence, this property is trivial if Γ itself is finite; it is interesting only when Γ is infinite.

Explain (1) why the two versions of compactness above are equivalent, (2) why compactness for consequence entails compactness for satisfiability, and (3) why compactness for satisfiability entails compactness for consequence.

1.35. Prove propositions 1.11 and 1.16. Give illustrations in English of the contents of these facts. Are all invalid sentences equivalent? Are all satisfiable sentences equivalent? Explain.

1.36. Explain why if sentences are equivalent then one is valid if and only if they all are—that is:

(1) If $A \simeq B$, then $\models A$ iff $\models B$.

Also explain why if sentences are equivalent then they are equally satisfiable—that is:

(2) If $A \simeq B$, then Sat A iff Sat B.

Notice that (2) means as well that if sentences are equivalent then one is unsatisfiable if and only if they all are.

1.37. True or false:

a. A contradiction implies nothing but contradictions.

b. There exist sentences of which no sentence is a consequence.

c. Every sentence is a consequence of at least one sentence.

d. For any sentences A and B, if B is not a consequence of A then either A is valid or B is not satisfiable.

e. Only valid sentences are a consequence of a valid sentence.

f. The union of satisfiable sets is itself a satisfiable set.

g. Every sentence is a consequence of the empty set.

h. A set of sentences is unsatisfiable if and only if in every model at least one sentence in the set is false.

i. Every sentence is equivalent to at least one valid sentence.

j. There is at least one set of sentences that is both satisfiable and unsatisfiable.

PART I

SENTENTIAL LOGIC

2

THE LANGUAGE OF SENTENTIAL LOGIC

IN THIS and the next three chapters we study *sentential logic*, also known as *propositional* or *truth-functional* logic. This is the logic of sentences and compounds of sentences created by means of particles such as not, and, or, and if. We shall be concerned with arguments like this one:

> If the coyote speaks the truth then today is Wednesday
> Today is not both Thursday and Wednesday
> Today is Thursday
> ―――――――――――――――――――
> The coyote does not speak the truth

This argument is valid, and its validity depends, as we shall see, only on its form and on the meanings of the logical constants if and not.

We study sentential logic—like all the logics in this book—through the medium of a *language*. As explained in chapter 1, logic depends at the bottom on the notion of a *logical constant*. In section 1 we introduce five *operators* that constitute our logical constants for sentential logic. In section 2 we specify in a preliminary way the meanings of the operators and discuss questions of translation.

In section 3 we give a precise definition of what it is to be a *sentence* of the language of sentential logic. The language's sentences are illustrated, and we look closely at their nature and form. In sections 4 and 5 we go further into the structure of the sentences, defining the ideas of *subsentence*, *proper subsentence*, and *atom of a sentence*, and then we briefly review *arguments*.

1. The logical operators

The logical constants for sentential logic are these five:

$\neg$	negation
$\wedge$	conjunction
$\vee$	disjunction

$\rightarrow$	conditionality
$\leftrightarrow$	biconditionality

The sentences of the language are built up out of these operations, or *operators*, as they are more commonly called. (They are also referred to as *connectives*, although this term is becoming outmoded.) Thus where A and B are themselves sentences, the logical operators give rise to further sentences of the following five different forms, which we list along with their formal descriptions.

$\neg$ A	the negation of A
A $\wedge$ B	the conjunction of A and B
A $\vee$ B	the disjunction of A and B
A $\rightarrow$ B	the conditional of A and B
A $\leftrightarrow$ B	the biconditional of A and B

As can be seen, negation is a one-place operator, while the others are two-place operators.

2. Translation

To get a better idea of the language of sentential logic, here are some rough equivalents to the meanings of the operators:

$\neg$	not
$\wedge$	and
$\vee$	or
$\rightarrow$	if
$\leftrightarrow$	if and only if

In terms of these and other common synonyms for negation, conjunction, disjunction, conditionality, and biconditionality, we can translate into and out of the language of sentential logic.

For ease of exposition in what follows, let us suppose that the sentence A means Snow is white and the sentence B means 7 + 5 = 12.

Negation. As its title implies, negation provides an expression of denial. So the sentence

$$\neg A$$

means

Snow is <u>not</u> white.

Negation in English is primarily marked by the presence of the word not. Sometimes not occurs as part of a prefix, as in the phrase it is not the case that. Thus Snow is white can also be negated

<u>It is not the case that</u> snow is white.

This may be regarded as a canonical form.

Negation can also be marked by a particular verb, like to fail. For example, the sentence Ben failed to eat may be thought of as Ben did not eat. Negation is frequently marked by the particle non. Thus Snow is non-white may be translated in the same way as Snow is not white. Finally, the phrase it is false that expresses negation. Snow is white may be negated It is false that snow is white; the same goes for the phrase it is not true that.

Conjunction. The conjunction operator functions much like the word and. For this reason, the sentence

$$A \wedge B$$

has the meaning

Snow is white <u>and</u> 7 + 5 = 12.

The word and is the most distinguished sign of conjunction in English sentences. Notice that it occurs not only between sentences, as in the example above, but also between nouns, verbs, adjectives, adverbs, and prepositions. Often these usages indicate sentential conjunction, as in the following pairs.

Ben and Laura ate	Ben ate and Laura ate
Ben ate and drank	Ben ate and Ben drank
Ben is hungry and thirsty	Ben is hungry and Ben is thirsty
Ben drank quickly and thirstily	Ben drank quickly and Ben drank thirstily
Ben went in and out the window	Ben went in the window and Ben went out the window

Some uses of and cannot be so easily reformed. For example, Cattle Annie and Little Britches are sisters may not be not equivalent to Cattle Annie is a sister and Little Britches is a sister. And Some burglars went in and out the window is certainly not equivalent to Some burglars went in the window and some burglars went out the window.

And has numerous counterparts, among them: but, however, moreover, though, although, even though, nevertheless, nonetheless, also, too, and, perhaps, the semicolon (;). Like and, many of these expressions turn up in positions other than between sentences. Unlike and, many of the expressions mean more than plain conjunction. Ben ate even though he was not hungry is clearly richer in meaning than Ben ate and he was not hungry. Or consider Ben took poison and died.

This last example brings up the question of *order* in translating conjunctions. Clearly, Snow is white and 7 + 5 = 12 means the same thing as 7 + 5 = 12 and snow is white. Thus it is indifferent from the point of view of meaning whether we use $A \wedge B$ or $B \wedge A$ for 7 + 5 = 12 and snow is white. On the other hand, as we shall better see later on, it may be unwise to adopt a laisser-aller attitude in this matter. So it may be preferable for some practical purposes to *transliterate* wherever possible. In this way, $A \wedge B$ transliterates Snow is white and 7 + 5 = 12, whereas $B \wedge A$ transliterates 7 + 5 = 12 and snow is white.

Disjunction. The disjunction operator corresponds closely to the word or in English—in the weak sense of the legalism and/or. So the disjunction of A and B,

$$A \vee B,$$

means

Snow is white <u>or</u> 7 + 5 = 12,

or more explicitly

Snow is white <u>or</u> 7 + 5 = 12, <u>or both</u>.

In this sense, $\vee$ is often referred to as weak or *inclusive* disjunction. *Exclusive* (strong) disjunction can be rendered by adding but not both instead of or both.

Like and and some of its fellows, or can occur between nouns, verbs, adjectives, adverbs, and prepositions. Thus Snow is white or black can be regarded as the disjunction Snow is white or snow is black. But again, as with conjunction, there is a caveat: All dogs are male or female cannot be rendered All dogs are male or all dogs are female.

Among alternatives to or there is, perhaps surprisingly, the word unless. Ben will die unless he sees a doctor means that if he does not see a doctor, Ben will die—which is to say that either Ben sees a doctor or he will die. It is tempting to read or in this last sentence as strong disjunction, but it is not warranted, since the unless-sentence does not imply that if Ben does see a doctor, he will not die. For this reason we are justified in translating unless as inclusive disjunction.

(The distinction between inclusive and exclusive disjunction is well marked in some languages, for instance, by vel and aut in Latin. The symbol $\vee$ is said to derive from the first letter of vel.)

The principle of transliteration mentioned in connection with and holds as well for or. So Snow is white or 7 + 5 = 12 is translated $A \vee B$, and not $B \vee A$. Applying the principle to unless is more difficult, since this word can stand in front of two sentences as well as between—Unless he sees a doctor, Ben will die, Ben will die unless he sees a doctor. For the sake of uniformity, let us place the sentence immediately following unless second in translation. Then the sentence Unless snow is white, 7 + 5 = 12 will be translated $B \vee A$.

Conditionality. The conditional operator is meant to correspond to a very minimal sense of the word if. Thus

$$A \to B$$

means

If snow is white then 7 + 5 = 12,

or, with if in the middle,

7 + 5 = 12 if snow is white.

The actual sense of the if corresponding to the conditional operator is more accurately given by translating $A \to B$ as

Either snow is not white or 7 + 5 = 12.

This means that a conditional is reckoned true if its antecedent (the if-part) is false or its consequent is true (or both), and it is reckoned false just in the case the antecedent is true and the consequent is false. Thus one can see that still another rendering of the meaning of $A \to B$ is

It is not the case that both snow is white and 7 + 5 ≠ 12.

This is the *truth-functional* interpretation of the conditional.

Sentences formed by the use of if (often with a companion then) are regularly translated by means of $\to$. It is notorious, however, that this rarely does justice to the meanings of sentences containing if—especially those in the subjunctive mood. For example, consider the sentence If the earth were flat then circles would be squares. This is evidently a false conditional, yet if we translate it by $P \to Q$, it is true—simply because of the falsity of its antecedent.

Even if problems of counterfactuality and mood are set aside, many indicative conditionals can be formed whose truth conditions seem at odds with those for $\to$. Take for example the indicative counterpart of the conditional mentioned above, If the earth is flat then circles are squares. In the language of sentential logic this sentence is true because its antecedent is false. Yet one may feel that, in so far as there is no connection between the flatness of the earth and the shape of circles, the sentence may have no truth value at all. Similar examples can be constructed in which the antecedent is false while the consequent is true.

The truth of the matter would seem to be that there is agreement only in the case where the antecedent is true and the consequent is false: here the conditional is surely false. As we shall see, we can make an argument to show that—so far as a truth-functional analysis is concerned—it is correct to count a conditional true except where the antecedent is true and the consequent is false (see chapter 3, page 67).

One of the most common alternative expressions for conditionality is the phrase only if. A little reflection will show that, for example, the sentences If snow is white then 7 + 5 = 12 and Snow is white only if 7 + 5 = 12 have the same truth-functional meaning, in the sense that they are false alike just in case Snow is white is true and 7 + 5 = 12 is false. Newcomers are sometimes confused, however, since if can also occur between two sentences. It is probably a good idea to straighten this out once for all, by means of this table of equivalent forms:

if p then q
p only if q
q if p
only if q, then p

One further caution: p only if q should not be construed to mean the same as p if and only if q (lest the longer phrase contain a redundancy).

Other truth-functional equivalents to if include in case, only in case, and cumbrous phrases involving the words necessary and sufficient. For example, each of the four sentences below can be regarded as equivalent to If snow is white then 7 + 5 = 12.

> In case snow is white, 7 + 5 = 12
>
> Snow is white only in case 7 + 5 = 12
>
> In order that snow is white, it is necessary that 7 + 5 = 12

It is sufficient that snow is white, in order that $7 + 5 = 12$

Biconditionality. Finally, the biconditional may be said to function like the phrase if and only if. Hence

$$A \leftrightarrow B$$

means

Snow is white if and only if $7 + 5 = 12$.

The sense of if in the phrase if and only if is truth-functional, so that $A \leftrightarrow B$ is best understood as meaning something like:

Either snow is white and $7 + 5 = 12$, or snow is not white and $7 + 5 \neq 12$.

Though subject to some degree to the problems surrounding if, the phrase if and only if is best translated by $\leftrightarrow$. There are few equivalents to the phrase; the most familiar are just in case, exactly in case, and constructions involving the words necessary and sufficient.

Here again, as with conjunction and disjunction, order is indifferent as regards meaning. We transliterate whenever possible—so that $A \leftrightarrow B$, and not $B \leftrightarrow A$, will be the translation of Snow is white if and only if $7 + 5 = 12$. Notice, too, that if and only if can occur in front of as well as between sentences. For the sake of uniformity, once again, we place the sentence immediately following if and only if second in translation. Thus $B \leftrightarrow A$ translates If and only if snow is white, then $7 + 5 = 12$.

strong Exclusive disjunction can be expressed in terms of biconditionality and negation. For example, Snow is white if and only if $7 + 5 \neq 12$ means the same as Snow is white or $7 + 5 = 12$, but not both. We return to this point later.

Combinations of operators. When the operators are used in combinations we get sentences such as these:

$$A \wedge \neg B \qquad \neg (A \vee B) \qquad \neg A \rightarrow (B \wedge C) \qquad (B \vee \neg A) \leftrightarrow \neg \neg C$$

The meanings of the first two of these are, respectively, Snow is white and $7 + 5 \neq 12$ and It is not the case that either snow is white or $7 + 5 = 12$. If we suppose that C means Pigs fly, the third example means If snow is not white then both $7 + 5 = 12$ and pigs fly. We leave the meaning of the fourth sentence as an exercise.

Notice the use of round brackets (parentheses) to indicate scope and grouping. This is very important. For example, consider the following two sentences.

$$\neg (A \vee B) \qquad\qquad \neg A \vee B$$

The first of these is a negation; the scope of $\neg$ is the whole disjunction $A \vee B$. The second is not a negation at all; rather, it is a disjunction the left disjunct of which, $\neg A$, is a negation (so the scope of $\neg$ here is just A). The sentences differ sharply in meaning: while the first says that it is not the case that either snow is white or $7 + 5 = 12$, the second says that either snow is not white or $7 + 5 = 12$. The first is false, the second true.

For another example of the importance of grouping, consider:

$$A \vee B \wedge C$$

This is not a sentence, since there is no indication of which operator, disjunction or conjunction, is dominant. To disambiguate this expression, so to speak, one must write either

$$A \vee (B \wedge C) \qquad \text{or} \qquad (A \vee B) \wedge C.$$

These are both sentences. The first says that either snow is white or both $7 + 5 = 12$ and pigs fly. This is true, since snow is white. The second says that both snow is white or $7 + 5 = 12$, and pigs fly. This is false, since pigs do not fly.

As an exercise, reckon the differences in form (and meaning) among the following three sentences.

$$\neg A \rightarrow (B \wedge C)$$

$$(\neg A \rightarrow B) \wedge C$$

$$\neg (A \rightarrow (B \wedge C))$$

It should be emphasized that difference in meaning is a sufficient, but not a necessary condition for distinguishing different sentences. As we shall see when we give a definition of the sentences in the language, the conditions of identity for sentences are a purely formal, syntactic matter.

Of course different sentences may have the same meaning, or at least truth conditions—this is, in part, what logic is about. For an example of different sentences with the same meaning, consider this pair:

$$\neg (A \vee B) \qquad\qquad \neg A \wedge \neg B$$

As we have seen, the first of these means that it is not the case that either snow is white or $7 + 5 = 12$. The second means that snow is not white and $7 + 5 \neq 12$. But then the meaning of each is simply that it is neither the case that snow is white nor the case that $7 + 5 = 12$.

Let us conclude this section by considering matters of punctuation and other subtleties.

Punctuation is not as strictly marked in natural languages as it is in the language of sentential logic. Very often important differences are indicated by the use of a comma. Witness the difference between

Snow is white, and 7 + 5 = 12 or pigs fly

and

Snow is white and 7 + 5 = 12, or pigs fly.

The first of these sentences is translated $A \wedge (B \vee C)$, whereas the second is translated $(A \wedge B) \vee C$.

Sometimes the auxiliary expressions both and either function like left parentheses to effect punctuation. Notice how the need for the comma is eliminated in these sentences:

Snow is white and either $7 + 5 = 12$ or pigs fly

Either snow is white and $7 + 5 = 12$ or pigs fly

In initial position, the expressions if, only if, and if and only if also behave like left parentheses. Consider the difference between

Snow is white and if $7 + 5 = 12$ then pigs fly

and

If snow is white and $7 + 5 = 12$ then pigs fly.

The first is translated $A \wedge (B \rightarrow C)$, the second $(A \wedge B) \rightarrow C$.

When translating a conjunctive or disjunctive series, like Snow is white, and $7 + 5 = 12$, and pigs fly, it may be tempting to use $A \wedge B \wedge C$, rather than $A \wedge (B \wedge C)$ or $(A \wedge B) \wedge C$—especially since the latter two are equivalent. Later in the book we allow such descriptions of multiple conjuncts (and disjuncts). But for now it is better to adhere to canonical forms, in which conjunctions and disjunctions are grouped in twos.

It is probably apparent that there are no mechanical rules or devices for translation. Rules of thumb and hints abound, and are about all there is to be had. Translation is best viewed as a practical matter in which the interests at hand determine what one should do.

Theoretically, we could adopt a policy of translating distinct sentences, no matter how complex, by distinct sentences. Thus we would translate the three sentences If snow is white then $7 + 5 = 12$, $7 + 5 \neq 12$, and Snow is not both heavy and white, respectively, as D, E, and F—despite the subsentences they obviously share. This policy would be a mistake from the standpoint of logic—especially if we are asking whether the third sentence is a consequence of the other two. Given this translation using D, E, and F, the third sentence is not implied by the others.

So as a practical matter we should attribute to sentences as much logical structure—in terms of not, and, or, etc.—as the situation will bear. Thus if we translate Snow is white, $7 + 5 = 12$, and Snow is heavy as $A \rightarrow B$, $\neg B$, and $\neg (C \wedge A)$—we will see that the third is indeed a consequence of the others.

We should also point out that there are situations in which different but equivalent sentences should be translated by a single sentence. For instance, All humans are mortal and Every human is mortal are equivalent yet have no distinguishing marks of negation, conjunction, etc. So the prudent course is to translate them both by a single sentence.

Much more will be said about the meanings of the logical operators $\neg$, $\wedge$, $\vee$, $\rightarrow$, and $\leftrightarrow$ in chapter 3, where we study their truth conditions, and in chapter 4, where we introduce rules of inference for them.

The purpose of the next few sections is to define the language of sentential logic and say some more about the structure of its sentences.

EXERCISES

2.1. In terms of the meanings of A, B, and C given above—Snow is white, 7 + 5 = 12, and pigs fly—what does the sentence $(B \vee \neg A) \leftrightarrow \neg\neg C$ mean? Explain the differences in form and in meaning among the sentences $\neg A \rightarrow (B \wedge C)$, $(\neg A \rightarrow B) \wedge C$, and $\neg(A \rightarrow (B \wedge C))$.

2.2. Translate the following sentences, where:

A—Arapaho rule	B—Blackfoot rule
C—Cheyenne rule	D—Dakota rule

(1) $(A \wedge C) \vee (B \wedge D)$

(2) $(A \wedge B) \wedge (C \vee D)$

(3) $A \rightarrow (B \wedge C)$

(4) $(A \wedge B) \rightarrow (C \vee D)$

(5) $\neg A \rightarrow \neg(B \vee C)$

(6) $\neg(A \wedge C) \rightarrow (B \wedge D)$

(7) $\neg((A \vee C) \rightarrow \neg(B \vee D))$

(8) $(\neg A \vee \neg C) \rightarrow (\neg B \wedge \neg D)$

(9) $(A \wedge B) \vee (\neg C \rightarrow \neg D)$

(10) $(A \rightarrow \neg C) \wedge ((\neg B \wedge \neg C) \rightarrow \neg D)$

(11) $(A \rightarrow B) \rightarrow (D \rightarrow C)$

(12) $(A \wedge B) \vee (C \leftrightarrow D)$

(13) $A \leftrightarrow (B \vee (C \vee D))$

(14) $(A \rightarrow \neg B) \rightarrow (C \vee D)$

(15) $(\neg A \wedge \neg B) \rightarrow \neg \neg (C \vee D)$

2.3. Translate the following sentences, where:

P—William Pinkerton succeeds
Q—Charles Quantrill will ride
R—Johnny Ringo will ride

(1) If it is not the case that Pinkerton succeeds, then not both Quantrill and Ringo will ride

(2) It is not the case that if Pinkerton succeeds, then neither Quantrill nor Ringo will ride

(3) It is not the case that if Pinkerton succeeds, then either Quantrill or Ringo will not ride

(4) If it is not the case that Pinkerton succeeds, then neither Quantrill nor Ringo will ride

2.4. Translate the following sentences, where:

A—Abilene wins B—Butte wins
C—Coffeyville wins D—Deadwood wins

(1) Either Abilene wins and Coffeyville wins or Butte and Deadwood do

(2) Abilene and Butte win and either Coffeyville wins or Deadwood does

(3) If Abilene wins, then so do both Butte and Coffeyville

(4) If both Abilene and Butte win, then so does either Coffeyville or Deadwood

(5) If Abilene does not win, then it is not the case that either Butte or Coffeyville wins

(6) If it is not the case that both Abilene and Coffeyville win, then both Butte and Deadwood do

(7) It is not the case that if either Abilene or Coffeyville wins, then neither Butte nor Deadwood does

(8) If Abilene or Coffeyville does not win, then both Butte and Deadwood do not win

(9) Either Abilene wins and so does Butte, or if Coffeyville does not win, then Deadwood fails to too

(10) Abilene wins only if Coffeyville does not, but if Butte and Coffeyville both fail to win, then Deadwood does likewise

(11) If, if Abilene wins then Butte does, then only if Coffeyville does, then Deadwood does

(12) Either Abilene and Butte win, or Coffeyville wins if and only if Deadwood does

(13) Abilene wins just in case Butte, Coffeyville, or Deadwood does

(14) If Abilene wins only if Butte does not, then Coffeyville wins unless Deadwood does

(15) If, along with Abilene, Butte does not win, then it is not the case that neither Coffeyville nor Deadwood wins

2.5. Translate the following sentences, where:

E—Edmund Campbell will go F—Frank Carver will go
G—George Maledon will go H—Heck Thomas will go
I—Isaac Parker will go

(1) Edmund will not go unless Frank, George, and Heck will

(2) If neither Heck nor Isaac will go, neither will Edmund

(3) Frank and George will go if and only if either Heck or Isaac, but not both, will

(4) Either exactly one of Edmund, Frank, and George will go, or all three of them will

(5) If any of the five will go, they all will

(6) At least one of the five will go, although not all of them will

(7) Provided that none of the five will go, Heck won't

(8) If none of the others will, Isaac will go

(9) Edmund or Frank or George, but not all three, will go, unless Heck or Isaac, but not both, won't

(10) It is not the case that either Edmund won't go or Frank will

3. The language

At the bottom, our language of sentential logic consists of a collection of *atomic* sentences:

$$\mathbb{P}_0, \ \mathbb{P}_1, \ \mathbb{P}_2, \ \ldots, \ \mathbb{P}_n, \ \ldots$$

These are the simplest sentences in the language; they have no logically important structure. We assume that there are infinitely many such sentences. This supposition is harmless and largely a matter of convenience. Among other benefits, it means we will never be short of atomic sentences in any practical situation.

The rest of the sentences of the language are compound, or *molecular*, in the sense that each is built up out of atomic sentences by repeated applications of five basic operators: $\neg$, $\wedge$, $\vee$, $\rightarrow$, $\leftrightarrow$. Here is an example of a molecular sentence of:

$$(\mathbb{P}_0 \wedge \neg\, \mathbb{P}_{42}) \leftrightarrow ((\mathbb{P}_{739} \vee \mathbb{P}_0) \rightarrow \mathbb{P}_{17})$$

The constituents of the various kinds of molecular sentences have labels. Here is a catalogue of these terms:

A is the *negate* of $\neg$ A.

A is the *left conjunct* of A $\wedge$ B.
B is the *right conjunct* of A $\wedge$ B.

A is the *left disjunct* of A $\vee$ B.
B is the *right disjunct* of A $\vee$ B.

A is the *antecedent* of A $\rightarrow$ B.
B is the *consequent* of A $\rightarrow$ B.

A is the *left member* of A $\leftrightarrow$ B.
B is the *right member* of A $\leftrightarrow$ B.

We presume that the atomic sentences are all distinct, that molecular sentences are distinct if any of their parts are, and that each sentence is of exactly one kind (so that, for example, no conjunction is a conditional). Given these assumptions, we may define the set of sentences simply as the smallest set containing all the atomic sentences and closed under the operations of negation, conjunction, disjunction, conditionality, and biconditionality. This is the content of the following more perspicuous definition.

DEFINITION 2.1. *Sentences.*

(1) Each atomic sentence $\mathbb{P}_n$ is a sentence.

(2) $\neg$ A is a sentence iff A is a sentence.

(3) A $\wedge$ B is a sentence iff A and B are sentences.

(4) A $\vee$ B is a sentence iff A and B are sentences.

(5) A $\rightarrow$ B is a sentence iff A and B are sentences.

(6) A $\leftrightarrow$ B is a sentence iff A and B are sentences.

Using definition 2.1, let us verify that the expression on page 51,

$$(\mathbb{P}_0 \wedge \neg\, \mathbb{P}_{42}) \leftrightarrow ((\mathbb{P}_{739} \vee \mathbb{P}_0) \rightarrow \mathbb{P}_{17}),$$

is indeed a sentence. By clause (6), this expression is a sentence if and only if both

$$\mathbb{P}_0 \wedge \neg\, \mathbb{P}_{42} \qquad\qquad \text{and} \qquad\qquad (\mathbb{P}_{739} \vee \mathbb{P}_0) \rightarrow \mathbb{P}_{17}$$

are sentences. The first of these is a sentence, according to clause (3), if and only if both

$$\mathbb{P}_0 \qquad\qquad \text{and} \qquad\qquad \neg\, \mathbb{P}_{42}$$

are sentences. The former of these is deemed a sentence by clause (1), and the latter checks out via clauses (2) and (1). The second expression,

$$(\mathbb{P}_{739} \vee \mathbb{P}_0) \rightarrow \mathbb{P}_{17},$$

is a sentence, according to clause (5), just in case both

$$\mathbb{P}_{739} \vee \mathbb{P}_0 \qquad\qquad \text{and} \qquad\qquad \mathbb{P}_{17}$$

are sentences—the latter of which is by clause (1). So, finally,

$$\mathbb{P}_{739} \vee \mathbb{P}_0$$

is a sentence, according to clause (4), if and only if both

$$\mathbb{P}_{739} \qquad\qquad \text{and} \qquad\qquad \mathbb{P}_0$$

are sentences—which they are, again by clause (1). Therefore the expression with which we began—

$$(\mathbb{P}_0 \wedge \neg\, \mathbb{P}_{42}) \leftrightarrow ((\mathbb{P}_{739} \vee \mathbb{P}_0) \rightarrow \mathbb{P}_{17})$$

—is itself a sentence.

This sentence can be described equally well as follows.

> the biconditional the left member of which is the conjunction of the atomic sentence number zero and the negation of atomic sentence number forty-two and the right member of which is the conditional whose antecedent is the disjunction of atomic sentence number seven hundred thirty-nine and atomic sentence zero and whose consequent is atomic sentence number seventeen

This description of the sentence is long and rather unwieldy. It can be summarized by means of the following *construction table*.

<table>
<tr><td colspan="5">$(\mathbb{P}_0 \wedge \neg \mathbb{P}_{42}) \leftrightarrow ((\mathbb{P}_{739} \vee \mathbb{P}_0) \rightarrow \mathbb{P}_{17})$</td></tr>
<tr><td colspan="2">$\mathbb{P}_0 \wedge \neg \mathbb{P}_{42}$</td><td colspan="3">$(\mathbb{P}_{739} \vee \mathbb{P}_0) \rightarrow \mathbb{P}_{17}$</td></tr>
<tr><td>$\mathbb{P}_0$</td><td>$\neg \mathbb{P}_{42}$</td><td colspan="2">$\mathbb{P}_{739} \vee \mathbb{P}_0$</td><td>$\mathbb{P}_{17}$</td></tr>
<tr><td></td><td>$\mathbb{P}_{42}$</td><td>$\mathbb{P}_{739}$</td><td>$\mathbb{P}_0$</td><td></td></tr>
</table>

The table shows precisely how the original sentence is built up from its atomic parts by means of the logical operators.

Remarks on terminology and notation. The terminology used above is important when it comes to giving a correct description of a sentence. It is nevertheless sometimes tempting to say, for example, "not A"—rather than "the negation of A". Or "if A then B" (or worse yet, as we shall see, "A implies B")—instead of "the conditional of A and B". While it must be admitted that these casual descriptions are often less cumbersome, use of the proper forms is urged.

On the other hand, the operators have more familiar names: $\neg$ is the *hook;* $\wedge$ is the *caret;* $\vee$ is the *wedge;* $\rightarrow$ is the *arrow;* and $\leftrightarrow$ is the *double* (or *double-headed*) *arrow*. So $\neg$ A can be called "hook A", A $\wedge$ B can be called "A caret B", and so on.

Our choice of logical constants—negation, conjunction, disjunction, conditionality, and biconditionality—is not unique or privileged, or even minimal, but the choice is traditional. The language of sentential logic is a special, "zero order" part of the language of first order logic studied beginning in chapter 6.

EXERCISES

2.6. Create a construction table for the sentence

$$\neg (\mathbb{P}_{29} \leftrightarrow (\mathbb{P}_7 \vee \neg \neg \mathbb{P}_{84})) \rightarrow (\mathbb{P}_{106} \rightarrow (\mathbb{P}_{84} \wedge \neg \mathbb{P}_{29})).$$

2.7. Let P, Q, and R be atomic sentences. Create construction tables for each of the following sentences.

(1)	$\neg P \rightarrow \neg (Q \wedge R)$	(3)	$\neg (P \rightarrow (\neg Q \vee \neg R))$
(2)	$\neg (P \rightarrow \neg (Q \vee R))$	(4)	$\neg P \rightarrow \neg (Q \vee R)$

Then describe in English the structure of each of the sentences.

4. Subsentences and atoms

As the construction table on page 54 shows, each sentence breaks down into a number of smaller *subsentences*, the smallest of which are the atomic subsentences, or *atoms*. The notion of subsentence we employ here is liberal: a sentence itself is included among its own subsentences (the others are distinguished as *proper* subsentences). Thus the set of subsentences of the sentence $(\mathbb{P}_0 \wedge \neg \mathbb{P}_{42}) \leftrightarrow ((\mathbb{P}_{739} \vee \mathbb{P}_0) \rightarrow \mathbb{P}_{17})$ is the set:

$$\{(\mathbb{P}_0 \wedge \neg \mathbb{P}_{42}) \leftrightarrow ((\mathbb{P}_{739} \vee \mathbb{P}_0) \rightarrow \mathbb{P}_{17}),\\ (\mathbb{P}_{739} \vee \mathbb{P}_0) \rightarrow \mathbb{P}_{17},\ \ \mathbb{P}_0 \wedge \neg \mathbb{P}_{42},\ \mathbb{P}_{739} \vee \mathbb{P}_0,\\ \neg \mathbb{P}_{42},\ \ \mathbb{P}_0,\ \ \mathbb{P}_{17},\ \ \mathbb{P}_{42},\ \ \mathbb{P}_{739}\}$$

In other words, the subsentences of the sentence in question are all the sentences appearing in cells of the construction table on page 54. Notice that because we are talking about the *set* of subsentences, $\mathbb{P}_0$ is listed only once even though it occurs twice in the sentence. The set of proper subsentences is formed, of course, by removing the main sentence from the set of subsentences. The atoms of the sentence are found in the set

$$\{\mathbb{P}_0, \mathbb{P}_{17}, \mathbb{P}_{42}, \mathbb{P}_{739}\}.$$

In short, the sentence's atoms are the sentences appearing at the bottom-most cells of the table.

For the sake of brevity in the exercises below, let us dub the *set of atomic sentences* "At", and similarly refer to the *set of all sentences* as "Sn".

EXERCISES

2.8. Identify the subsentences of the sentence in exercise 2.6. Identify the proper subsentences. Identify the atoms.

2.9. Identify the subsentences of the sentences in exercise 2.7. Identify the proper subsentences. Identify the atoms. Which sentences are subsentences of both sentences (2) and (4)? Which sentences are subsentences of all four of sentences?

2.10. It is possible to define the notions of subsentence, proper subsentence, and atom rigorously, as follows.

Sbsn(A)—*the set of subsentences of the sentence* A:

(1) $\mathrm{Sbsn}(\mathbb{P}_n) = \{\mathbb{P}_n\}$, for each atomic sentence $\mathbb{P}_n$.

(2) $\mathrm{Sbsn}(\neg A) = \mathrm{Sbsn}(A) \cup \{\neg A\}$.

(3) $\mathrm{Sbsn}(A \wedge B) = \mathrm{Sbsn}(A) \cup \mathrm{Sbsn}(B) \cup \{A \wedge B\}$.

(4) $\mathrm{Sbsn}(A \vee B) = \mathrm{Sbsn}(A) \cup \mathrm{Sbsn}(B) \cup \{A \vee B\}$.

(5) $\mathrm{Sbsn}(A \rightarrow B) = \mathrm{Sbsn}(A) \cup \mathrm{Sbsn}(B) \cup \{A \rightarrow B\}$.

(6) $\mathrm{Sbsn}(A \leftrightarrow B) = \mathrm{Sbsn}(A) \cup \mathrm{Sbsn}(B) \cup \{A \leftrightarrow B\}$.

Prsbsn(A)—*the set of proper subsentences of the sentence* A:

$$\text{Prsbsn}(A) = \text{Sbsn}(A) - \{A\}.$$

Atm(A)—the set of atoms (atomic subsentences) of the sentence A:

$$\text{Atm}(A) = \text{Sbsn}(A) \cap \text{At}.$$

The notions of subsentence and atom can be extended to sets of sentences, so that by Sbsn(Γ) we understand the set of all subsentences of sentences in the set Γ, and by Atm(Γ) we understand the set of all atoms of sentences in the set Γ. Verify the correctness of the following definitions of these ideas.

$$\text{Sbsn}(\Gamma) = \{A\colon A \in \text{Sbsn}(B), \text{ for some } B \in \Gamma\}.$$

$$\text{Atm}(\Gamma) = \text{Sbsn}(\Gamma) \cap \text{At}.$$

Identify each of the following sets of sentences.

(1)	Sbsn(Sn)	(6)	Atm(Sbsn(Γ))
(2)	Atm(Sn)	(7)	Sbsn(Atm(Γ))
(3)	Sbsn(At)	(8)	Atm(Atm(Γ))
(4)	Atm(At)	(9)	Prsbsn($\mathbb{P}_0$)
(5)	Sbsn(Sbsn(Γ))	(10)	Atm($\{\mathbb{P}_0\}$)

2.11. Explain why it is inappropriate to attempt a definition of the set of proper subsentences of a set of sentences—that is, give an example of a set Γ for which Prsbsn(Γ) will not make sense.

5. Arguments

As we explained in chapter 1, evaluation of arguments is one of the main concerns of logic. We characterized an argument as any finite collection of sentences one of which is the conclusion and the others of which are the premisses, and we denoted an argument with premisses $A_1, \ldots, A_n$ and a conclusion A in this way:

$$A_1, \ldots, A_n \;/\; A$$

For instance, consider the argument

$$A \to B, \;\neg B \;/\; \neg(C \wedge A).$$

In this argument, $A \to B$ and $\neg B$ are the premisses, and $\neg(C \wedge A)$ is the conclusion. Of course the argument can be described vertically. For example:

$$\begin{array}{l} A \to B \\ \neg B \\ \hline \neg(C \wedge A) \end{array}$$

Recall that there is nothing special about the order in which the premisses of an argument are listed. So the argument above is equally well described by

$$\neg B, \;A \to B \;/\; \neg(C \wedge A)$$

or by

$$\begin{array}{l} \neg B \\ A \to B \\ \hline \neg(C \wedge A) \end{array}$$

Remember that these expressions merely *describe* an argument and do not say anything about it, for example whether the conclusion is a consequence of the premisses.

EXERCISES

2.12. Let L mean The coyote speaks the truth, W mean Today is Wednesday, and T mean Today is Thursday. Translate the following argument.

If the coyote speaks the truth then today is Wednesday
Today is not both Thursday and Wednesday
Today is Thursday
___
The coyote does not speak the truth

2.13. Given your understanding of the meanings of the operators, do you think the argument described on page 58 is correct? That is, is $\neg(C \wedge A)$ a consequence of $A \rightarrow B$ and $\neg B$?

2.14. What is your opinion of the arguments in exercise 3.11 (page 77)? In which of those arguments is the conclusion a consequence of the premisses?

2.15. In your opinion, which of the arguments in exercise 3.20 (page 88) is valid?

3

SEMANTICS FOR SENTENTIAL LOGIC

THE FIRST object of this chapter is to define *model* and *truth in a model* for the sentences of the language of sentential logic. Definitions and examples occupy sections 1 and 2.

Then in section 3 we present *truth tables*, and in section 4 we give some examples of the logical notions defined in chapter 1—*consequence, validity, equivalence,* and *satisfiability*.

Section 5 is devoted to an explanation of the *reverse truth table technique*—a simple means of discovering the logical status of finite collections of sentences. This method should be mastered: it provides rapid resolution for many problems otherwise solvable only by long and tedious means.

In section 6 we record more *metatheorems* like those in chapter 1—statements of properties of and relationships among the logical notions. One of these provides a foundation for the soundness of the deductive system presented in chapter 4. Section 7 is about *compactness* and section 8 treats *replacement*.

In section 9 the topics are *truth functions, expressibility*, and the *definability of truth-functional operators*. In the final section, 10, we explain *truth-functional completeness* and see why our language of sentential logic is as expressively powerful as it can be.

1. Models

A model for a language gives meanings to the expressions in the language, in such a way that a truth value—truth or falsity—is assigned to each sentence of the language. The notion of model introduced here is narrower than this conception, in two ways.

First, models for the language of sentential logic give meanings directly only to the atomic sentences. Not all of the sentences of the language are immediately interpreted. Secondly, because we are ultimately interested only in the truth values of sentences in a model, we ignore richer notions of meaning and construe a model for the language as *just* an assignment of truth values to the atomic sentences. So the idea of meaning with respect to this language is quite meager—merely that of a sentence's having a truth value, relative to an assignment of truth values to the atomic sentences.

To facilitate a great deal of discussion in what follows, we introduce the symbols "$\top$" and "$\bot$" for the truth values *truth* and *falsity*, respectively. (As might be imagined, these symbols are sometimes called "tee" and "eet".)

DEFINITION 3.1. *Models.* A model $\mathcal{M}$ is an assignment of truth values to the atomic sentences. In other words, for each atomic sentence $\mathbb{P}_n$:

$\mathcal{M}$ assigns truth to $\mathbf{P}_n$—i.e. $\mathcal{M}(\mathbb{P}_n) = \top$; or

$\mathcal{M}$ assigns falsity to $\mathbb{P}_n$—i.e. $\mathcal{M}(\mathbb{P}_n) = \bot$.

Technically, then, a model $\mathcal{M}$ for the language is a truth-valued function on the set of atomic sentences:

$$\mathcal{M}\colon \mathrm{At} \mapsto \{\top, \bot\}$$

For example, consider the model that assigns truth values to the atomic sentences in this way:

$\mathcal{M}(\mathbb{P}_n) = \top$ if n is a prime number; and $\mathcal{M}(\mathbb{P}_n) = \bot$ otherwise (i.e. if n is not a prime number).

In this model each of the atomic sentences $\mathbf{P}_2$, $\mathbb{P}_3$, $\mathbb{P}_5$, $\mathbb{P}_7$, $\mathbb{P}_{11}$, $\mathbb{P}_{13}$, $\mathbb{P}_{17}$, . . . is true, since the index of each is a prime number, while all the other atomic sentences, those with indices that are not prime, are false.

This example of a model serves, incidentally, to illustrate the difference between our models and other, richer conceptions. For example, we might have construed a model as an assignment of *propositions* to the atomic sentences. For example, consider the model $\mathcal{N}$ defined by:

$\mathcal{N}(\mathbb{P}_n)$ = the proposition expressed by Snow is white if n is a prime number; and $\mathcal{N}(\mathbf{P}_n)$ = the proposition expressed by Pigs fly otherwise.

Provided that snow is white and that pigs do not fly, in this model the same atomic sentences, $\mathbb{P}_2$, $\mathbb{P}_3$, $\mathbb{P}_5$, $\mathbb{P}_7$, $\mathbb{P}_{11}$, $\mathbb{P}_{13}$, $\mathbb{P}_{17}$, . . . , are true, and the rest are false. Thus from the standpoint of what truth values are assigned, the models $\mathcal{M}$ and $\mathcal{N}$ come to the same thing. So, to repeat: since it is the determination of truth values in which we are primarily interested, it is enough for our purposes that models assign truth values directly to the atomic sentences.

If our language is thought to be "real", then its atomic sentences may be thought to have meanings, and hence truth values, already. Thus $\mathbb{P}_0$ might mean that the sky is blue, $\mathbb{P}_1$ that the earth is round, $\mathbb{P}_2$ that the sea is calm, and so on. From this point of view, the "real" model for the language is but one among many. This is an acceptable view so long as we recall, from chapter 1, the two assumptions that govern our treatment of models. The first is that no sentence, and hence no atomic sentence, is both truth and false in any model. The second is that every sentence, and hence every atomic sentence, is either true or false in any model. The first assumption comes of our identification of models with truth-valued functions. The second comes of our provision that these truth-valued functions are total.

EXERCISES

3.1. Define eight models, $\mathcal{M}_1$–$\mathcal{M}_8$, such that no two of them agree on all the atoms P, Q, and R of the sentences in exercise 2.7 (page 55). That is, make sure that no two of the models assign the same truth values to P, Q, and R.

2. Truth in a model

The notion of truth in a model $\mathcal{M}$ is defined for all the sentences of the language by means of a recursive procedure. We first specify the conditions under which an atomic sentence is true in $\mathcal{M}$, then the conditions under which a negation is true in $\mathcal{M}$, then the conditions under which a conjunction is true in $\mathcal{M}$, and so on for a disjunction, a conditional, and a biconditional.

Before we get to the definition itself, note an important abbreviation. To say that a sentence A is *true* in a model $\mathcal{M}$, we sometimes write:

$$\models_{\mathcal{M}} \text{A}.$$

Likewise, to say that A is *false* (not true) in $\mathcal{M}$, we may write:

$$\not\models_{\mathcal{M}} \mathrm{A}.$$

The double turnstile, "$\models$", thus has three meanings. We use it for consequence and for validity, and now, when it is subscripted with a term for a model, it means truth in the model (or, if slashed, falsity).

DEFINITION 3.2. *Truth in a model.*

(1) $\models_{\mathcal{M}} \mathbf{P}_n$ iff $\mathcal{M}(\mathbf{P}_n) = \mathsf{T}$, for each atomic sentence $\mathbf{P}_n$.

(2) $\models_{\mathcal{M}} \neg \mathrm{A}$ iff $\not\models_{\mathcal{M}} \mathrm{A}$.

(3) $\models_{\mathcal{M}} \mathrm{A} \wedge \mathrm{B}$ iff $\models_{\mathcal{M}} \mathrm{A}$ and $\models_{\mathcal{M}} \mathrm{B}$.

(4) $\models_{\mathcal{M}} \mathrm{A} \vee \mathrm{B}$ iff $\models_{\mathcal{M}} \mathrm{A}$ or $\models_{\mathcal{M}} \mathrm{B}$.

(5) $\models_{\mathcal{M}} \mathrm{A} \rightarrow \mathrm{B}$ iff if $\models_{\mathcal{M}} \mathrm{A}$ then $\models_{\mathcal{M}} \mathrm{B}$.

(6) $\models_{\mathcal{M}} \mathrm{A} \leftrightarrow \mathrm{B}$ iff $\models_{\mathcal{M}} \mathrm{A}$ if and only if $\models_{\mathcal{M}} \mathrm{B}$.

The clauses of definition 3.2 mean, in other words:

(1) An *atomic sentence* is true in a model just in case it is assigned the value truth in the model.

(2) The *negation* of a sentence is true in a model just in case the negate is not true in the model.

(3) The *conjunction* of two sentences is true in a model just in case both conjuncts are true in the model.

(4) The *disjunction* of two sentences is true in a model just in case at least one of the disjuncts is true in the model.

(5) The *conditional* of two sentences is true in a model just in case the consequent is true in the model if the antecedent is. Equivalently: just in case either the antecedent is false in the model or the consequent is true. Equivalently again: just in case it is not the case both that the antecedent is true in the model and the consequent is false in the model.

(6) The *biconditional* of two sentences is true in a model just in case the left member is true in the model if and only if the right member is. Equivalently, and perhaps more perspicuously: just in case either both members are true in the model or both members are false in the model.

The preceding remarks should help in understanding the definition of truth in a model, and also show more exactly in what sense we attribute to the five operators $\neg$, $\wedge$, $\vee$, $\rightarrow$, and $\leftrightarrow$ the meanings of not, and, or, if, and if and only if.

Notice that clause (1) is almost trivial, since it says only that an atomic sentence is true in a model just when the model assigns it truth. In the next language we study, the clause for the atomic sentences has more content, because atomic sentences there have parts that are themselves significant, and so models give truth values to atomic sentences only indirectly in terms of the meanings of their parts.

The content of the non-atomic clauses of the truth definition can be made even more precise by means of *truth tables*. These exhibit the truth value assigned to a molecular sentence as a function of the truth values of its immediate subsentences. Thus the following truth table gives the truth conditions of a negation (the number below indicates clause (2) of definition 3.2).

A	$\neg$ A
$\top$	$\bot$
$\bot$	$\top$
	(2)

Similarly, the truth conditions for conjunctions, disjunctions, conditionals, and biconditionals can be stated by means of tables (again the numbers correspond to the clauses of definition 3.2):

A	B	A $\wedge$ B	A $\vee$ B	A $\rightarrow$ B	A $\leftrightarrow$ B
$\top$	$\top$	$\top$	$\top$	$\top$	$\top$
$\top$	$\bot$	$\bot$	$\top$	$\bot$	$\bot$
$\bot$	$\top$	$\bot$	$\top$	$\top$	$\bot$
$\bot$	$\bot$	$\bot$	$\bot$	$\top$	$\top$
		(3)	(4)	(5)	(6)

Thus in the truth table for negation, $\bot$ is assigned to $\neg$ A when $\top$ is assigned to A, and vice versa, thereby showing that $\neg$ A is true in a model if and only if A is false in the model. Similarly, in the truth table for conditionality, A $\rightarrow$ B is assigned $\bot$ only when $\top$ is assigned to the antecedent A and $\bot$ is assigned to the consequent B—thereby indicating that A $\rightarrow$ B is true in a model if and only if it is not the case both that A is true in the model and that B is false in the model. Compare the other tables with their respective clauses in definition 3.2.

Now for some examples. Consider first the sentence

$$(\mathbb{P}_0 \wedge \neg\, \mathbb{P}_{42}) \leftrightarrow ((\mathbb{P}_{739} \vee \mathbb{P}_0) \rightarrow \mathbb{P}_{17}),$$

which made its first appearance in chapter 2. Let us determine the truth value of this sentence in the model $\mathcal{M}$ described in the preceding section where $\mathcal{M}(\mathbb{P}_n) = \top$ if and only if n is a prime number. The numbers 17 and 739 are prime, and the numbers 0 and 42 are not. So:

$$\mathcal{M}(\mathbb{P}_0) = \bot. \quad \mathcal{M}(\mathbb{P}_{17}) = \top. \quad \mathcal{M}(\mathbb{P}_{42}) = \bot. \quad \mathcal{M}(\mathbb{P}_{739}) = \top.$$

Therefore, according to clause (1) of definition 3.2:

$$\not\models_{\mathcal{M}} \mathbb{P}_0. \qquad \models_{\mathcal{M}} \mathbb{P}_{17}. \qquad \not\models_{\mathcal{M}} \mathbb{P}_{42}. \qquad \models_{\mathcal{M}} \mathbb{P}_{739}.$$

By clause (2) of 3.2, it follows that

$$\vDash_{\mathcal{M}} \neg \mathbb{P}_{42}.$$

And so by clause (3) of 3.2, we find that

$$\nvDash_{\mathcal{M}} \mathbb{P}_0 \wedge \neg \mathbb{P}_{42}.$$

Looking now at the right member of the biconditional, we see by clause (4) of 3.2 that

$$\vDash_{\mathcal{M}} \mathbb{P}_{739} \vee \mathbb{P}_0.$$

And so by clause (5) of 3.2,

$$\vDash_{\mathcal{M}} (\mathbb{P}_{739} \vee \mathbb{P}_0) \rightarrow \mathbb{P}_{17}.$$

Thus, finally, according to clause (6) of 3.2,

$$\nvDash_{\mathcal{M}} (\mathbb{P}_0 \wedge \neg \mathbb{P}_{42}) \leftrightarrow ((\mathbb{P}_{739} \vee \mathbb{P}_0) \rightarrow \mathbb{P}_{17}).$$

That is to say, the sentence $(\mathbb{P}_0 \wedge \neg \mathbb{P}_{42}) \leftrightarrow ((\mathbb{P}_{739} \vee \mathbb{P}_0) \rightarrow \mathbb{P}_{17})$ is false in the model $\mathcal{M}$. An even simpler way to use the definition of truth in a model to determine the truth value of the sentence in question is by means of a kind of truth table:

$\mathbb{P}_0$	$\mathbb{P}_{17}$	$\mathbb{P}_{42}$	$\mathbb{P}_{739}$	$(\mathbb{P}_0 \wedge$	$\neg \mathbb{P}_{42})$	$\leftrightarrow$	$((\mathbb{P}_{739} \vee \mathbb{P}_0)$	$\rightarrow \mathbb{P}_{17})$
$\perp$	$\top$	$\perp$	$\top$	$\perp$	$\top$	$\perp$	$\top$	$\top$
				2	**1**	**5**	**3**	**4**

In the guide column at the left we list the truth values in $\mathcal{M}$ of the four atoms of the sentence. Then truth values are assigned to the non-atomic subsentences of the sentence in accordance with the clauses of the definition of truth in a model—using either definition 3.2 or the truth table versions for molecular sentences. The numbering below the table above is not essential and only indicates the order of evaluation. Note that the order could be otherwise—e.g. **4, 3, 5, 1, 2.**

It should be pointed out that these derivations of the truth value in $\mathcal{M}$ of the sentence are, however thorough, unnecessarily long. For to see that $\mathbf{P}_0 \wedge \neg\, \mathbb{P}_{42}$ is false in $\mathcal{M}$, it is enough to notice that $\mathbf{P}_0$ is false in $\mathcal{M}$; and to see that $(\mathbb{P}_{739} \vee \mathbb{P}_0) \rightarrow \mathbf{P}_{17}$ is true in $\mathcal{M}$, it is enough to remark that $\mathbf{P}_{17}$ is true in $\mathcal{M}$. Such shortcuts mean that the truth table above could look like this:

$\mathbb{P}_0$	$\mathbb{P}_{17}$	$\mathbb{P}_{42}$	$\mathbb{P}_{739}$	$(\mathbb{P}_0 \wedge \neg\, \mathbb{P}_{42}) \leftrightarrow ((\mathbb{P}_{739} \vee \mathbb{P}_0) \rightarrow \mathbf{P}_{17})$		
$\perp$	$\top$	$\perp$	$\top$	$\perp$	$\perp$	$\top$
				1	**3**	**2**

The sentence $(\mathbb{P}_0 \wedge \neg\, \mathbb{P}_{42}) \leftrightarrow ((\mathbb{P}_{739} \vee \mathbb{P}_0) \rightarrow \mathbf{P}_{17})$ is false in the model $\mathcal{M}$, and in other models as well. For an example of a model in which the sentence is true, consider $\mathcal{M}'$, defined as follows. $\mathcal{M}'(\mathbb{P}_n) = \top$ if n is even; and $\mathcal{M}'(\mathbf{P}_n) = \perp$ if n is odd. The evaluation is left as an exercise.

The interpretation of the conditional. In chapter 2 we acknowledged the possibility of complaints about the interpretation of the conditional, according to which A $\rightarrow$ B is true if A is false or B is true, or both—that is, according to which A $\rightarrow$ B is false just in the case where A is true and B is false. We conclude this section with an argument that shows that if $\rightarrow$ is to have a truth-functional interpretation then this is the way it must be.

We take it as given that a conditional is false if its antecedent is true and its consequent is false. So the truth table for $\rightarrow$ looks at least like column (1) below.

		(1)	(2)	(3)
A	B	A $\rightarrow$ B	A $\rightarrow$ B	A $\rightarrow$ B
$\top$	$\top$		$\top$	$\top$
$\top$	$\perp$	$\perp$	$\perp$	$\perp$
$\perp$	$\top$			$\top$
$\perp$	$\perp$		$\top$	$\top$

Next we observe that any sentence of the form if p then p is always true, whether p itself is true or is false. So it would be wrong to say that a conditional is false when its antecedent and consequent are both true—or else A $\rightarrow$ A, e.g. If snow is white then snow is white, will be false. Equally, it would be wrong to say that a conditional is false when antecedent and consequent are both false—or else, again, A $\rightarrow$ A, e.g. If pigs fly then pigs fly, will be false. So truth table for $\rightarrow$ must look at least like column (2) above.

Finally, how do we fill in the third row of the truth table? Not by $\perp$, for then the conditionals A $\rightarrow$ B and B $\rightarrow$ A would have the same meaning—which in general they do not. So the third row must be filled by $\top$, yielding column (3) in the table above. But this is the same as column (5) of the truth table on page 65. Thus we conclude that under a truth-functional interpretation, this is how $\rightarrow$ should be evaluated.

EXERCISES

3.2. Verify the correctness of the truth tables for conjunctions, disjunctions, and biconditionals by comparing each with its respective clause in the definition 3.2 of truth in a model.

3.3. Show by means of a truth table that the sentence $(\mathbb{P}_0 \wedge \neg \mathbb{P}_{42}) \leftrightarrow ((\mathbb{P}_{739} \vee \mathbb{P}_0) \rightarrow \mathbf{P}_{17})$ is true in the model $\mathcal{M}$ defined by $\mathcal{M}(\mathbf{P}_n) = \top$ if n is even; and $\mathcal{M}(\mathbf{P}_n) = \perp$ if n is odd. Consider the model $\mathcal{M}'$ defined by: $\mathcal{M}'(\mathbf{P}_n) = \top$ if n is odd; and $\mathcal{M}'(\mathbf{P}_n) = \perp$ if n is even. What is the truth value in $\mathcal{M}'$ of the sentence? Define a model different from $\mathcal{M}$ in which the sentence is true.

3.4. What is the truth value of the sentence

$$\neg\,(\mathbb{P}_{29} \leftrightarrow (\mathbb{P}_7 \vee \neg\,\neg\,\mathbb{P}_{84})) \rightarrow (\mathbb{P}_{106} \rightarrow (\mathbb{P}_{84} \wedge \neg\,\mathbb{P}_{29}))$$

in the model $\mathcal{M}$ that assigns truth to just the atomic sentences whose indices are prime numbers? What is the truth value of the sentence in each of the models $\mathcal{M}$ and $\mathcal{M}'$ defined in the preceding exercise?

3.5. What is the truth value of each of the sentences $\neg\,P \rightarrow \neg\,(Q \wedge R)$, $\neg\,(P \rightarrow \neg\,(Q \vee R))$, $\neg\,(P \rightarrow (\neg\,Q \vee \neg\,R))$, and $\neg\,P \rightarrow \neg\,(Q \vee R)$ in each of the models $\mathcal{M}_1$–$\mathcal{M}_8$ defined in exercise 3.1?

3.6. Let $\mathcal{M}$ be a model in which A is true, B is false, C is true, and D is false. What is the truth value in $\mathcal{M}$ of each of the sentences in exercise 2.2 (page 48)?

3.7. Let $\downarrow$ (the dagger) be a new two-place operator, so that in addition to the others we also have sentences of the form $A \downarrow B$. The meaning of $\downarrow$ is given by the following truth table.

A	B	$A \downarrow B$
$\top$	$\top$	$\bot$
$\top$	$\bot$	$\bot$
$\bot$	$\top$	$\bot$
$\bot$	$\bot$	$\top$

What is a natural reading of $\downarrow$? That is, how should the blank be filled in the following clause for $\downarrow$ of a definition of truth in a model $\mathcal{M}$?

$$\models_{\mathcal{M}} A \downarrow B \text{ iff } \underline{\hspace{6cm}}.$$

Let $\mathcal{M}$ be a model in which X is true, Y is false, and Z is false. What is the truth value in $\mathcal{M}$ of the sentences $\neg\,(X \rightarrow (Y \downarrow Z))$ and $\neg\, X \rightarrow (Y \downarrow Z)$?

3.8. Let $\nleftrightarrow$ be a new two-place operator, so that we also have sentences of the form $A \nleftrightarrow B$. The meaning of $\nleftrightarrow$ is given by the following truth table.

A	B	$A \nleftrightarrow B$
$\top$	$\top$	$\bot$
$\top$	$\bot$	$\top$
$\bot$	$\top$	$\top$
$\bot$	$\bot$	$\bot$

Give a natural reading of $\nleftrightarrow$. That is, fill in the blank in the following clause for $\nleftrightarrow$ of a definition of truth in a model $\mathcal{M}$.

$$\models_{\mathcal{M}} A \nleftrightarrow B \text{ iff } \underline{\hspace{6cm}}.$$

What is the truth value of the sentence $(A \wedge B) \nleftrightarrow (C \leftrightarrow D)$ in a model $\mathcal{M}$ in which A is true, B is false, C is true, and D is false?

3.9. Let $\mathcal{M}$ be a model in which E is true, F is false, G is true, H is false, and I is true. What is the truth value in $\mathcal{M}$ of each of the following sentences?

(1) $\neg\, E \vee (F \wedge (G \wedge H))$

(2) $\neg\,(H \vee I) \rightarrow \neg\, E$

(3) $(F \wedge G) \leftrightarrow (H \leftrightarrow \neg\, I)$

(4) $((E \wedge \neg (F \vee G)) \vee ((F \wedge \neg (E \vee G)) \vee (G \wedge \neg (E \vee F))))$
$\vee (E \wedge (F \wedge G))$

(5) $(E \vee (F \vee (G \vee (H \vee I)))) \rightarrow (E \wedge (F \wedge (G \wedge (H \wedge I))))$

(6) $(E \vee (F \vee (G \vee (H \vee I)))) \wedge \neg (E \wedge (F \wedge (G \wedge (H \wedge I))))$

(7) $\neg (E \vee (F \vee (G \vee (H \vee I)))) \rightarrow \neg H$

(8) $\neg (E \vee (F \vee (G \vee H))) \rightarrow I$

(9) $((E \vee (F \vee G)) \wedge \neg (E \wedge (F \wedge G))) \vee (\neg H \leftrightarrow \neg \neg I)$

(10) $\neg (\neg E \vee F)$

3. Truth tables

We have already used truth tables to state truth conditions for molecular sentences—negations, conjunctions, disjunctions, conditionals, and biconditionals—and to evaluate sentences in particular models. Now we show how truth tables can be employed to give a complete semantic analysis of individual sentences and finite sets of sentences of the language of sentential logic.

Let us begin with the observation that the truth value of a sentence in a model is a function solely of the truth values assigned to its atoms in the model. For example, consider the sentence $\mathbf{P}_0 \wedge \neg \mathbf{P}_1$. Let $\mathcal{M}$ be a model in which $\mathbf{P}_0$ is true and $\mathbf{P}_1$ is false. Then $\mathbf{P}_0 \wedge \neg \mathbf{P}_1$ is true in $\mathcal{M}$, since both its conjuncts are. It should be evident that $\mathbf{P}_0 \wedge \neg \mathbf{P}_1$ is true in $\mathcal{M}$ regardless of what truth values $\mathcal{M}$ assigns to atomic sentences other than $\mathbf{P}_0$ and $\mathbf{P}_1$. It matters not to the truth value of $\mathbf{P}_0 \wedge \neg \mathbf{P}_1$ in $\mathcal{M}$ whether the model assigns truth or falsity to, say, $\mathbf{P}_{101}$. To put the point succinctly, we can say that any models that treat the atoms of $\mathbf{P}_0 \wedge \neg \mathbf{P}_1$ the same—i.e. give the same truth values respectively to $\mathbf{P}_0$ and to $\mathbf{P}_1$—must treat the sentence $\mathbf{P}_0 \wedge \neg \mathbf{P}_1$ the same.

This may already have been noticed, in connection with the sentence

$$(\mathbf{P}_0 \wedge \neg \mathbf{P}_{42}) \leftrightarrow ((\mathbf{P}_{739} \vee \mathbf{P}_0) \rightarrow \mathbf{P}_{17}),$$

considered earlier. In order to determine the truth value of this sentence in the model $\mathcal{M}$ defined there, it was enough to know the truth values in $\mathcal{M}$ of the sentence's atoms, $\mathbf{P}_0$, $\mathbf{P}_{17}$, $\mathbf{P}_{42}$, and $\mathbf{P}_{739}$. The values in that model of extraneous atomic sentences, e.g. $\mathbf{P}_{5000}$, were irrelevant.

The proposition in general is that if models assign the same truth values respectively to all the atoms of a sentence A, then—no matter what truth values they assign to the rest of the atomic sentences—A will have the same truth value in all. That is to say, A will either be true in all or false in all. Our methods in this chapter rely fundamentally on this fact—called *relevance* or *continuity* —which is stated formally and proved in chapter 5.

The property of relevance holds also for sets of sentences: So long as models agree on all the atoms of a set of sentences, they will agree as well on all the sentences in the set.

Relevance is what entitles us to use the method of *truth tables* to evaluate the logical status of sentences and finite sets of sentences, completely and effectively. As we explain below, truth tables can be used to prove or refute the logical properties of consequence, validity, equivalence, and satisfiability.

Consider again the sentence $\mathbb{P}_0 \wedge \neg \mathbb{P}_1$. Since it has but two atoms, $\mathbb{P}_0$ and $\mathbb{P}_1$, models for this sentence can agree on the values of the atoms in just four ways: they can assign truth to the atoms, they can assign falsity to them, and, in two different ways, they can make one atom true and the other false. These possibilities can be registered on a truth table, as follows.

	$\mathbb{P}_0$	$\mathbb{P}_1$	
$\mathcal{M}_1$	$\top$	$\top$	
$\mathcal{M}_2$	$\top$	$\bot$	
$\mathcal{M}_3$	$\bot$	$\top$	
$\mathcal{M}_4$	$\bot$	$\bot$	

The list in the guide column indicates the four possible combinations of truth values that models can assign to the atomic sentences $\mathbb{P}_0$ and $\mathbb{P}_1$. (The list of models itself is not essential.)

Now we place the sentence $\mathbb{P}_0 \wedge \neg \mathbb{P}_1$ to the right of the guide column of the truth table, and calculate the truth value of the sentence in terms of the values of its subsentences:

	$\mathbb{P}_0$	$\mathbb{P}_1$	$\mathbb{P}_0 \wedge \neg \mathbb{P}_1$
$\mathcal{M}_1$	T	T	$\perp\perp$
$\mathcal{M}_2$	T	$\perp$	TT
$\mathcal{M}_3$	$\perp$	T	$\perp\perp$
$\mathcal{M}_4$	$\perp$	$\perp$	$\perp$T
			2 1

From this truth table, we can see at a glance the truth values of the sentence (in column **2**) throughout its models: $\mathbf{P}_0 \wedge \neg \mathbb{P}_1$ is true in $\mathcal{M}_2$, when $\mathbf{P}_0$ is true and $\mathbf{P}_1$ is false, and it is false in $\mathcal{M}_1$, $\mathcal{M}_3$, and $\mathcal{M}_4$, when the atoms' truth values are otherwise.

As an illustration of a model for a finite collection of sentences, consider the set

$$\{\mathbb{P}_0 \rightarrow \mathbf{P}_1, \ \neg \mathbb{P}_1, \ \neg (\mathbb{P}_2 \wedge \mathbb{P}_0)\}.$$

The atoms of this set are the atomic sentences $\mathbb{P}_0$, $\mathbb{P}_1$, and $\mathbb{P}_2$. Models relevant to the set will be eight in number, since there are three atoms and hence 2^3 combinations of truth values for these atomic sentences. The sentences and their atoms are registered on the truth table below:

	$\mathbb{P}_0$	$\mathbb{P}_1$	$\mathbb{P}_2$	$\mathbb{P}_0 \to \mathbb{P}_1$	$\neg \mathbb{P}_1$	$\neg$	$(\mathbb{P}_2 \wedge \mathbb{P}_0)$
$\mathcal{M}_1$	T	T	T	T	⊥	⊥	T
$\mathcal{M}_2$	T	T	⊥	T	⊥	T	⊥
$\mathcal{M}_3$	T	⊥	T	⊥	T	⊥	T
$\mathcal{M}_4$	T	⊥	⊥	⊥	T	T	⊥
$\mathcal{M}_5$	⊥	T	T	T	⊥	T	⊥
$\mathcal{M}_6$	⊥	T	⊥	T	⊥	T	⊥
$\mathcal{M}_7$	⊥	⊥	T	T	T	T	⊥
$\mathcal{M}_8$	⊥	⊥	⊥	T	T	T	⊥
				1	**2**	**4**	**3**

Since the eight rows in the guide column exhaust the possible ways $\mathbb{P}_0$, $\mathbb{P}_1$, and $\mathbb{P}_2$ can be true and false together, the truth table contains a complete summary of the truth values of the three sentences $\mathbb{P}_0 \to \mathbb{P}_1$, $\neg\, \mathbb{P}_1$, and $\neg\, (\mathbb{P}_2 \wedge \mathbb{P}_0)$ throughout all possible models for the set of sentences. In $\mathcal{M}_1$, the first sentence is true and the others are false; in $\mathcal{M}_2$, the first and third are true and the second is false; in $\mathcal{M}_3$, the second is true and the others are false; and so on.

By now we have seen that we can use a truth table to set out systematically all the different models for a sentence or finite set of sentences and then evaluate the sentence or sentences in each model. Since the number of atoms is always finite, the number of essentially different models will always be finite too. Indeed, where n is the number of atoms, the number of models is 2^n. Thus as in the example above, there are eight (i.e. 2^3) models for the set $\{\mathbb{P}_0 \to \mathbb{P}_1, \neg\, \mathbb{P}_1, \neg\, (\mathbb{P}_2 \wedge \mathbb{P}_0)\}$ of three sentences.

The general recipe for constructing a truth table for a finite set of sentences is as follows.

> First list the atoms of the set, $A_1, \ldots, A_n$, to the left at the top of the guide column, putting the sentences in the set to the right. Below the first atom, A_1, put a column of 2^{n-1} Ts continued by 2^{n-1} ⊥s. Under the next atom, put half as many Ts and ⊥s twice as many times—that is, put 2^{n-2} Ts followed by 2^{n-2} ⊥s followed by 2^{n-2} Ts followed by 2^{n-2} ⊥s. Continue in this way until beneath the last atom, A_n, 2^{n-n} (= 1) Ts and ⊥s have been alternated a total of 2^n times.

The result of all this activity will be a truth table of 2^n rows, each row of which represents a model for the set whose atoms are $A_1, \ldots, A_n$, and which, collectively, exhaust all the possible ways of assigning truth values to $A_1, \ldots, A_n$.

For example, consider once more the truth table above for the set $\{\mathbb{P}_0 \rightarrow \mathbb{P}_1, \neg \mathbb{P}_1, \neg (\mathbb{P}_2 \wedge \mathbb{P}_0)\}$, where the number of atoms is three and the number of rows in the truth table is eight (= 2^3). The first column, under $\mathbb{P}_0$, consists of 4 (= $2^{3-1} = 2^2$) Ts followed by the same number of ⊥s. The second column, under $\mathbb{P}_1$, consists of 2 (= $2^{3-2} = 2^1$) Ts followed by the same number of ⊥s—all next to the Ts in the first column—followed by this pattern next to the ⊥s in the first column. Finally, under $\mathbb{P}_2$ the third column alternates 1 (= $2^{2-2} = 2^0$) Ts and ⊥s for the whole eight rows.

Truth tables provide an effective means of evaluating sentences and finite sets of sentences, and they constitute a decision procedure for determining the logical status of sentences and finite sets of sentences, as we shall see.

We close this section with a final example, viz. a complete evaluation in sixteen (= 2^4) models of the sentence $(\mathbb{P}_0 \wedge \neg \mathbb{P}_{42}) \leftrightarrow ((\mathbb{P}_{739} \vee \mathbb{P}_0) \rightarrow \mathbb{P}_{17})$.

$\mathbb{P}_0$	$\mathbb{P}_{17}$	$\mathbb{P}_{42}$	$\mathbb{P}_{739}$	$(\mathbb{P}_0 \wedge$	$\neg \mathbb{P}_{42})$	$\leftrightarrow$	$((\mathbb{P}_{739} \vee \mathbb{P}_0)$	$\rightarrow \mathbb{P}_{17})$
⊤	⊤	⊤	⊤	⊥	⊥	⊥	⊤	⊤
⊤	⊤	⊤	⊥	⊥	⊥	⊥	⊤	⊤
⊤	⊤	⊥	⊤	⊤	⊤	⊤	⊤	⊤
⊤	⊤	⊥	⊥	⊤	⊤	⊤	⊤	⊤
⊤	⊥	⊤	⊤	⊥	⊥	⊤	⊤	⊥
⊤	⊥	⊤	⊥	⊥	⊥	⊤	⊤	⊥
⊤	⊥	⊥	⊤	⊤	⊤	⊥	⊤	⊥
⊤	⊥	⊥	⊥	⊤	⊤	⊥	⊤	⊥
⊥	⊤	⊤	⊤	⊥	⊥	⊥	⊤	⊤
⊥	⊤	⊤	⊥	⊥	⊥	⊥	⊥	⊤
⊥	⊤	⊥	⊤	⊥	⊤	⊥	⊤	⊤
⊥	⊤	⊥	⊥	⊥	⊤	⊥	⊥	⊤
⊥	⊥	⊤	⊤	⊥	⊥	⊤	⊤	⊥
⊥	⊥	⊤	⊥	⊥	⊥	⊥	⊥	⊤
⊥	⊥	⊥	⊤	⊥	⊤	⊤	⊤	⊥
⊥	⊥	⊥	⊥	⊥	⊤	⊥	⊥	⊤
				2	**1**	**5**	**3**	**4**

Compare this truth table with the one on page 67 for this sentence. Some reduction of effort in computation could of course be achieved.

EXERCISES

3.10. Construct full truth tables for some of the sentences about the Arapaho, Blackfoot, Cheyenne, and Dakota in exercise 2.2 (page 48).

3.11. Construct truth tables for the sentences in the following arguments.

(1) $A \;/\; A$

(2)	$B \;/\; A \to B$	(8)	$A \;/\; A \vee B$
(3)	$B,\ \neg B \;/\; A$	(9)	$B \;/\; A \vee B$
(4)	$A \to B,\ A \;/\; B$	(10)	$A \vee B,\ A \to C,\ B \to C \;/\; C$
(5)	$A,\ B \;/\; A \wedge B$	(11)	$A \to B,\ B \to A \;/\; A \leftrightarrow B$
(6)	$A \wedge B \;/\; A$	(12)	$A \leftrightarrow B \;/\; A \to B$
(7)	$A \wedge B \;/\; B$	(13)	$A \leftrightarrow B \;/\; B \to A$

3.12. Construct truth tables for the following sentences.

(1) $A \to A$

(2) $\neg \neg A \to A$

(3) $A \vee \neg A$

(4) $\neg (A \wedge \neg A)$

(5) $A \leftrightarrow A$

(6) $\neg (A \leftrightarrow \neg A)$

(7) $(\neg A \to A) \to A$

(8) $A \to (B \to A)$

(9) $\neg A \to (A \to B)$

(10) $(A \wedge \neg A) \to B$

(11) $B \to (A \vee \neg A)$

(12) $A \to (B \to (A \wedge B))$

(13) $A \to (B \to (A \leftrightarrow B))$

(14) $(A \to B) \to ((A \vee B) \to B)$

(15) $(\neg B \to \neg A) \to (A \to B)$

(16) $(\neg A \to \neg B) \to ((\neg A \to B) \to A)$

(17) $((A \to B) \to A) \to A$

(18) $((A \to B) \to C) \to (B \to C)$

(19) $(A \to B) \to (\neg B \to (A \to C))$

(20) $(A \to (B \to C)) \to ((A \to B) \to (A \to C))$

3.13. Construct truth tables for the sentences in the following pairs.

(1) $A, \ \neg\neg A$

(2) $A, \ A \wedge A$

(3) $A, \ A \vee A$

(4) $A \wedge B, \ B \wedge A$

(5) $A \vee B, \ B \vee A$

(6) $A \leftrightarrow B, \ B \leftrightarrow A$

(7) $A \wedge (B \wedge C), \ (A \wedge B) \wedge C$

(8) $A \vee (B \vee C), \ (A \vee B) \vee C$

(9) $A \leftrightarrow (B \leftrightarrow C), \ (A \leftrightarrow B) \leftrightarrow C$

(10) $\neg(A \wedge B), \ \neg A \vee \neg B$

(11) $\neg(A \vee B), \ \neg A \wedge \neg B$

(12) $A \wedge (B \vee C), \ (A \wedge B) \vee (A \wedge C)$

(13) $A \vee (B \wedge C), \ (A \vee B) \wedge (A \vee C)$

(14) $A \to B, \ \neg B \to \neg A$

(15) $A \to B, \ \neg(A \wedge \neg B)$

(16) $A \to B, \ \neg A \vee B$

(17) $A \leftrightarrow B, \ \neg A \leftrightarrow \neg B$

(18) $A \leftrightarrow B, \ (A \to B) \wedge (B \to A)$

(19) $A \leftrightarrow B, \ (A \wedge B) \vee (\neg A \wedge \neg B)$

(20) $(A \wedge B) \rightarrow C, \; A \rightarrow (B \rightarrow C)$

3.14. Construct truth tables for the sentences in the following sets.

(1) $\{A, \neg A\}$

(2) $\{A \wedge \neg A\}$

(3) $\{A \wedge \neg A, B\}$

(4) $\{A \vee B, \neg A, \neg B\}$

(5) $\{A \leftrightarrow B, A \leftrightarrow \neg B\}$

(6) $\{A \leftrightarrow \neg A\}$

4. Consequence, validity, equivalence, and satisfiability

The four logical concepts in which we are interested—consequence, validity, equivalence, and satisfiability—were defined in chapter 1. In this section we review these concepts and apply them in the language of sentential logic.

Consequence. Recall that a sentence A is a consequence of a set of sentences Γ (in symbols: $\Gamma \models A$) just in case there is no model in which A is false and all the sentences in Γ are true; in other words, if and only if A is true in every model in which all the sentences of Γ are true (definition 1.1, page 12).

To illustrate the concept of consequence, consider the argument

$$\begin{array}{l} \mathbb{P}_0 \rightarrow \mathbb{P}_1 \\ \neg \mathbb{P}_1 \\ \hline \neg (\mathbb{P}_2 \wedge \mathbb{P}_0) \end{array}$$

To check whether the conclusion of this argument is a consequence of the premisses is to check whether the sentence $\neg (\mathbb{P}_2 \wedge \mathbb{P}_0)$ is true in every model in which both the sentences $\mathbb{P}_0 \rightarrow \mathbb{P}_1$ and $\neg \mathbb{P}_1$ are true. The simplest way to do this is by reference to a truth table for these three sentences:

$\mathbb{P}_0$	$\mathbb{P}_1$	$\mathbb{P}_2$	$\mathbb{P}_0 \to \mathbb{P}_1$	$\neg\,\mathbb{P}_1$	$\neg\,(\mathbb{P}_2 \wedge \mathbb{P}_0)$	
T	T	T	T	⊥	⊥	T
T	T	⊥	T	⊥	T	⊥
T	⊥	T	⊥	T	⊥	T
T	⊥	⊥	⊥	T	T	⊥
⊥	T	T	T	⊥	T	⊥
⊥	T	⊥	T	⊥	T	⊥
⊥	⊥	T	T	T	T	⊥
⊥	⊥	⊥	T	T	T	⊥

(This is the same as the truth table on page 74 above.) Here we find that the two premisses are true together only in two rows (the last two), and that in each of these the conclusion is true. In other words, there is no model in which the premisses are true and the conclusion is false. So we conclude that $\neg\,(\mathbb{P}_2 \wedge \mathbb{P}_0)$ is a consequence of the set $\{\mathbb{P}_0 \to \mathbb{P}_1, \neg\,\mathbb{P}_1\}$.

For a case of non-consequence, consider the argument

$$\begin{array}{l} \mathbb{P}_0 \to \mathbb{P}_1 \\ \neg\,\mathbb{P}_0 \\ \hline \neg\,\mathbb{P}_1 \end{array}$$

A four row truth table for the sentences in this argument shows that the conclusion is not a consequence of the premisses:

$\mathbb{P}_0$	$\mathbb{P}_1$	$\mathbb{P}_0 \to \mathbb{P}_1$	$\neg \mathbb{P}_0$	$\neg \mathbb{P}_1$
$\top$	$\top$	$\top$	$\bot$	$\bot$
$\top$	$\bot$	$\bot$	$\bot$	$\top$
$\bot$	$\top$	$\top$	$\top$	$\bot$
$\bot$	$\bot$	$\top$	$\top$	$\top$

The fact that in models represented by the third row of this truth table both the premisses of the argument are true and the conclusion is false means that the latter is not a consequence of the former. (That in the last row all three sentences are true is of course irrelevant.)

With regard to the examples above we can say, thus,

$$\mathbb{P}_0 \to \mathbb{P}_1,\ \neg \mathbb{P}_1 \ \models\ \neg (\mathbb{P}_2 \wedge \mathbb{P}_0)$$

and

$$\mathbb{P}_0 \to \mathbb{P}_1,\ \neg \mathbb{P}_0 \ \not\models\ \neg \mathbb{P}_1.$$

Validity. As we know, a sentence A is valid ($\models$ A) if and only if there is no model in which A is false—or, equivalently, A is true in every model. Recall definition 1.6 (page 25). In the context of sentential logic, a valid sentence is said to be *tautologous* or a *tautology*.

Validity in the language of sentential logic is illustrated by the sentence

$$\mathbb{P}_0 \to \mathbb{P}_0,$$

which the following two row truth table shows to be valid.

$\mathbb{P}_0$	$\mathbb{P}_0 \to \mathbb{P}_0$
$\top$	$\top$
$\bot$	$\top$

That is, because the sentence $\mathbb{P}_0 \rightarrow \mathbb{P}_0$ is true in both models, it is valid.

Another example of validity is the sentence

$$((\mathbb{P}_0 \rightarrow \mathbb{P}_1) \wedge \neg \mathbb{P}_1) \rightarrow \neg (\mathbb{P}_2 \wedge \mathbb{P}_0).$$

Here is a truth table for it:

$\mathbb{P}_0$	$\mathbb{P}_1$	$\mathbb{P}_2$	$((\mathbb{P}_0 \rightarrow \mathbb{P}_1)$	$\wedge$	$\neg \mathbb{P}_1)$	$\rightarrow$	$\neg$	$(\mathbb{P}_2 \wedge \mathbb{P}_0)$
T	T	T	T	⊥	⊥	T	⊥	T
T	T	⊥	T	⊥	⊥	T	T	⊥
T	⊥	T	⊥	⊥	T	T	⊥	T
T	⊥	⊥	⊥	⊥	T	T	T	⊥
⊥	T	T	T	⊥	⊥	T	T	⊥
⊥	T	⊥	T	⊥	⊥	T	T	⊥
⊥	⊥	T	T	T	T	T	T	⊥
⊥	⊥	⊥	T	T	T	T	T	⊥
						*		

The starred column contains the truth values of the sentence in all eight models. Because each row of this column contains T, the sentence is true in each model and hence is valid.

For a case of invalidity, consider the sentence

$$(\mathbb{P}_0 \rightarrow \mathbb{P}_1) \rightarrow (\neg \mathbb{P}_0 \rightarrow \neg \mathbb{P}_1).$$

By means of the following truth table, we can see that this sentence is invalid.

$\mathbb{P}_0$	$\mathbb{P}_1$	$(\mathbb{P}_0 \to \mathbb{P}_1) \to (\neg\, \mathbb{P}_0 \to \neg\, \mathbb{P}_1)$
$\top$	$\top$	$\top \quad \top \; \bot \quad \top \; \bot$
$\top$	$\bot$	$\bot \quad \top \; \bot \quad \top \; \top$
$\bot$	$\top$	$\top \quad \bot \; \top \quad \bot \; \bot$
$\bot$	$\bot$	$\top \quad \top \; \top \quad \top \; \top$

The third row of the truth table represents a model in which the sentence is false, which means that it is not valid.

Equivalence. Definition 1.8 (page 27) says that two sentences A and B are equivalent ($A \simeq B$) just in case there is no model in which A and B disagree as to truth value—or, in other words, the two sentences are true and false alike in every model.

A very elementary example of equivalent sentences is the pair

$$\mathbb{P}_0, \quad \neg\neg\, \mathbb{P}_0.$$

Another example is the pair

$$\mathbb{P}_0 \to \mathbb{P}_1, \quad \neg\, (\mathbb{P}_0 \wedge \neg\, \mathbb{P}_1).$$

That these last two sentences are equivalent is shown by the following truth table.

$\mathbb{P}_0$	$\mathbb{P}_1$	$\mathbb{P}_0 \to \mathbb{P}_1$	$\neg\, (\mathbb{P}_0 \wedge \neg\, \mathbb{P}_1)$
$\top$	$\top$	$\top$	$\top \quad \bot\bot$
$\top$	$\bot$	$\bot$	$\bot \quad \top\top$
$\bot$	$\top$	$\top$	$\top \quad \bot\bot$
$\bot$	$\bot$	$\top$	$\top \quad \bot\top$
		*	*

The final (starred) columns of truth values under each of the sentences are the same, i.e. the two sentences are true and false alike in all models. Therefore, they are equivalent.

A good illustration of non-equivalence is the pair of sentences

$$\mathbb{P}_0 \to \mathbb{P}_1, \ \mathbb{P}_1 \to \mathbb{P}_0.$$

Here is a truth table that shows that $\mathbb{P}_0 \to \mathbb{P}_1$ is not equivalent to $\mathbb{P}_1 \to \mathbb{P}_0$:

$\mathbb{P}_0$	$\mathbb{P}_1$	$\mathbb{P}_0 \to \mathbb{P}_1$	$\mathbb{P}_1 \to \mathbb{P}_0$
$\top$	$\top$	$\top$	$\top$
$\top$	$\bot$	$\bot$	$\top$
$\bot$	$\top$	$\top$	$\bot$
$\bot$	$\bot$	$\top$	$\top$

The columns of truth values differ in the both the second and third row, either one of which shows that the two sentences are not equivalent.

Satisfiability. A set Γ of sentences is satisfiable (Sat Γ) just when there is a model in which all the sentences in Γ are true. And to say Γ is unsatisfiable ($\text{S}\not{\text{a}}\text{t}\ \Gamma$) is to affirm that there is no model in which all the sentences in Γ are true. The definitions are on page 29.

Satisfiability is illustrated by the set of sentences

$$\{\mathbb{P}_0 \vee \mathbb{P}_1, \ \mathbb{P}_0, \neg\, \mathbb{P}_1\}.$$

Here is a truth table that establishes the satisfiability of this set:

	P_0	P_1	$P_0 \vee P_1$	P_0	$\neg P_1$
	T	T	T	T	⊥
*	T	⊥	T	T	T
	⊥	T	T	⊥	⊥
	⊥	⊥	⊥	⊥	T

For the sake of perspicuity, we have repeated P_0 and its truth values along with the other two sentences in the set. It is clear that the second (starred) row of the truth table represents a model in which all the sentences in the set are true. So the set is satisfiable. The set may seem to be "barely" satisfiable, since its three sentences are true together in only one case. But it is satisfiable nevertheless.

To illustrate unsatisfiability, consider this set of sentences:

$$\{P_0 \rightarrow P_2,\ P_1 \rightarrow P_2,\ P_0 \vee P_1,\ \neg P_2\}$$

An eight row truth table for the sentences in this set looks like this:

P_0	P_1	P_2	$P_0 \rightarrow P_2$	$P_1 \rightarrow P_2$	$P_0 \vee P_1$	$\neg P_2$
T	T	T	T	T	T	⊥
T	T	⊥	⊥	⊥	T	T
T	⊥	T	T	T	T	⊥
T	⊥	⊥	⊥	T	T	T
⊥	T	T	T	T	T	⊥
⊥	T	⊥	T	⊥	T	T
⊥	⊥	T	T	T	⊥	⊥
⊥	⊥	⊥	T	T	⊥	T

That the set of sentences is unsatisfiable is shown by the fact that "T" fails to occur under all four sentences at once in any of the eight rows of the truth table: there is no model in which all the sentences of the set are true together.

Another, very simple example of unsatisfiability is furnished by the set

$$\{\mathbb{P}_0, \neg \mathbb{P}_0\}$$

—as may be verified by means of a truth table.

Recall that it is common to call an unsatisfiable sentence a contradiction. A simple example of a contradiction is the sentence

$$\mathbb{P}_0 \wedge \neg \mathbb{P}_0.$$

This completes our introduction to the logical notions of consequence, validity, equivalence, and satisfiability in the language of sentential logic. A number of properties of and relationships among these concepts were catalogued in chapter 1, and several more will appear at the end of the chapter.

We conclude this section with some remarks about another use of truth tables.

By means of a truth table, we saw above that the sentence $\neg (\mathbb{P}_2 \wedge \mathbb{P}_0)$ is a consequence of the set of sentences $\{\mathbb{P}_0 \rightarrow \mathbb{P}_1, \neg \mathbb{P}_1\}$. We can use a truth table to show, more generally, that where A, B, and C are any three sentences, the sentence $\neg$ (C $\wedge$ A) is a consequence of the set of sentences $\{$A $\rightarrow$ B, $\neg$ B$\}$. Thus:

A	B	C	$A \to B$	$\neg B$	$\neg$	$(C \wedge A)$
T	T	T	T	⊥	⊥	T
T	T	⊥	T	⊥	T	⊥
T	⊥	T	⊥	T	⊥	T
T	⊥	⊥	⊥	T	T	⊥
⊥	T	T	T	⊥	T	⊥
⊥	T	⊥	T	⊥	T	⊥
⊥	⊥	T	T	T	T	⊥
⊥	⊥	⊥	T	T	T	⊥

In this truth table, which should be compared with the one on page 80, the sentence $\neg (C \wedge A)$ is true in every row in which both the sentences $A \to B$ and $\neg B$ are true, so that we see that consequence holds generally for any sentences A, B, and C.

We use truth tables in this way (as we have already, in some exercises) to show that consequence, validity, equivalence, or unsatisfiability holds for certain *forms* of sentences or finite sets of sentences. In truth tables like the one above, the components at the top of the guide column represent sentences of arbitrary complexity, and the rows of the guide column represent all the possible ways these components can be true and false together in a model. Thus the truth table above establishes not only that

$$\mathbb{P}_0 \to \mathbb{P}_1,\ \neg \mathbb{P}_1 \ \models \ \neg (\mathbb{P}_2 \wedge \mathbb{P}_0),$$

but also that

$$(\mathbb{P}_{150} \vee \neg \mathbb{P}_{62}) \to \neg \mathbb{P}_{2001},\ \neg \neg \mathbb{P}_{2001} \ \models \ \neg ((\neg \mathbb{P}_{13} \leftrightarrow \mathbb{P}_2) \wedge (\mathbb{P}_{150} \vee \neg \mathbb{P}_{62}))$$

—and infinitely many other such cases of consequence.

Likewise, when demonstrating the opposite of consequence, validity, equivalence, and unsatisfiability, one can regard the components at the top of the guide column as atomic, so that non-consequence, invalidity, non-equivalence, or satisfiability is established for *instances* of the forms.

In either way, truth tables thus provide a convenient method of determining the logical status of sentences and finite sets of sentences.

EXERCISES

3.15. By means of truth tables, check that:

(1) $\nvDash (\mathbb{P}_0 \rightarrow \mathbb{P}_1) \rightarrow (\neg \mathbb{P}_0 \rightarrow \neg \mathbb{P}_1)$.

(2) $\mathrm{S}\not{\mathrm{a}}\mathrm{t}\ \mathbb{P}_0 \wedge \neg \mathbb{P}_0$.

(3) $\mathbb{P}_0 \rightarrow \mathbb{P}_1 \not\simeq \mathbb{P}_1 \rightarrow \mathbb{P}_0$.

3.16. In which of the arguments in exercise 3.11 is the conclusion a consequence of the premisses?

3.17. Which of the sentences in exercise 3.12 is valid?

3.18. Which of the pairs of sentences in exercise 3.13 are equivalent?

3.19. Which of the sets in exercise 3.14 is unsatisfiable?

3.20. Use truth tables to determine in which of the following arguments the conclusion is a consequence of the premisses.

(1) $A \rightarrow B,\ B \rightarrow C\ /\ A \rightarrow C$

(2) $A \rightarrow B,\ \neg B\ /\ \neg A$

(3) $A \vee B,\ \neg A\ /\ B$

(4) $A \vee B,\ \neg B\ /\ A$

(5) $A \vee B,\ A \rightarrow C,\ B \rightarrow D\ /\ C \vee D$

(6) $\neg B \vee \neg C,\ A \rightarrow B,\ A \rightarrow C\ /\ \neg A$

(7) $\neg C \vee \neg D,\ A \rightarrow C,\ B \rightarrow D\ /\ \neg A \vee \neg B$

3.21. Use truth tables to determine in which of the following arguments the conclusion is a consequence of the premisses.

(1) $A \to B,\ A \to \neg B\ /\ \neg A$

(2) $A \to B,\ \neg A \to B\ /\ B$

(3) $A \leftrightarrow B,\ B \leftrightarrow C\ /\ A \leftrightarrow C$

3.22. Use truth tables to determine which of the following sentences are valid.

(1) $((A \to B) \wedge (B \to C)) \to (A \to C)$

(2) $(A \to \neg A) \to \neg A$

(3) $(A \to B) \to (\neg B \to \neg A)$

(4) $(A \to B) \to ((A \to \neg B) \to \neg A)$

(5) $A \to \neg \neg A$

(6) $A \to (\neg B \to \neg (A \to B))$

(7) $\neg (A \to B) \to A$

(8) $\neg (A \to B) \to \neg B$

(9) $A \vee (A \to B)$

(10) $(A \to B) \vee \neg B$

(11) $(A \to B) \vee (B \to A)$

(12) $(A \to B) \vee (B \to C)$

(13) $(A \to B) \vee (C \to A)$

(14) $\neg A \to (\neg B \to (A \leftrightarrow B))$

(15) $(A \vee B) \vee (A \leftrightarrow B)$

3.23. Use truth tables to determine which of the following pairs of sentences are equivalent.

(1) $A,\ A$

(2) $A,\ A \wedge (A \vee B)$

(3) $A,\ A \vee (A \wedge B)$

(4) $A,\ A \wedge (B \vee \neg B)$

(5) $A,\ A \vee (B \wedge \neg B)$

(6) $A,\ (A \wedge B) \vee (A \wedge \neg B)$

(7) $A,\ (A \vee B) \wedge (A \vee \neg B)$

(8) $A,\ \neg A \rightarrow A$

(9) $A,\ (B \vee \neg B) \rightarrow A$

(10) $A,\ B \leftrightarrow (A \leftrightarrow B)$

(11) $\neg A,\ A \rightarrow \neg A$

(12) $\neg A,\ A \rightarrow (B \wedge \neg B)$

(13) $\neg A,\ \neg B \leftrightarrow (A \leftrightarrow B)$

(14) $A \wedge B,\ \neg(\neg A \vee \neg B)$

(15) $A \wedge B,\ \neg(A \rightarrow \neg B)$

(16) $A \wedge B,\ A \leftrightarrow (A \rightarrow B)$

(17) $A \wedge B,\ (A \leftrightarrow B) \leftrightarrow (A \vee B)$

(18) $A \vee B,\ \neg(\neg A \wedge \neg B)$

(19) $A \vee B,\ \neg A \rightarrow B$

(20) $A \vee B,\ B \leftrightarrow (A \rightarrow B)$

(21) $A \vee B,\ (A \leftrightarrow B) \leftrightarrow (A \wedge B)$

(22) $A \rightarrow B,\ A \leftrightarrow (A \wedge B)$

(23) $A \rightarrow B,\ B \leftrightarrow (A \vee B)$

(24) $\neg (A \leftrightarrow B),\ \neg A \leftrightarrow B$

(25) $\neg (A \leftrightarrow B),\ A \leftrightarrow \neg B$

(26) $(A \rightarrow B) \rightarrow B,\ (B \rightarrow A) \rightarrow A$

(27) $A \vee (B \rightarrow C),\ B \rightarrow (A \vee C)$

(28) $A \rightarrow (B \vee C),\ (A \rightarrow B) \vee C$

(29) $A \rightarrow (B \rightarrow C),\ B \rightarrow (A \rightarrow C)$

(30) $A \rightarrow (B \wedge C),\ (A \rightarrow B) \wedge (A \rightarrow C)$

(31) $(A \vee B) \rightarrow C,\ (A \rightarrow C) \wedge (B \rightarrow C)$

(32) $(A \vee B) \wedge C,\ (A \wedge C) \vee (B \wedge C)$

(33) $(A \wedge B) \vee C,\ (A \vee C) \wedge (B \vee C)$

(34) $(A \vee B) \wedge (C \vee D), ((A \wedge C) \vee (A \wedge D)) \vee ((B \wedge C) \vee (B \wedge D))$

(35) $(A \wedge B) \vee (C \wedge D), ((A \vee C) \wedge (A \vee D)) \wedge ((B \vee C) \wedge (B \vee D))$

3.24. Use truth tables to determine which of the following sets of sentences is unsatisfiable.

(1) $\{B,\ \neg (A \rightarrow B)\}$

(2) $\{A \rightarrow B,\ A, \neg B\}$

(3) $\{A \wedge B,\ \neg A\}$

(4) $\{A,\ \neg (A \vee B)\}$

(5) $\{A \vee B,\ A \rightarrow C,\ B \rightarrow C,\ \neg C\}$

(6) $\{A \leftrightarrow B,\ \neg (A \rightarrow B)\}$

(7) $\{\neg (A \leftrightarrow B),\ \neg A \leftrightarrow \neg B\}$

3.25. The operator $\downarrow$ (the dagger), which expresses joint denial and corresponds to the words neither . . . nor, was introduced in exercise 3.7. By means of truth tables, determine which of the following pairs of sentences are equivalent.

(1) $\neg A,\ A \downarrow A$

(2) $A \wedge B,\ (A \downarrow A) \downarrow (B \downarrow B)$

(3) $A \vee B,\ (A \downarrow B) \downarrow (A \downarrow B)$

(4) $A \downarrow B,\ \neg A \wedge \neg B$

(5) $A \downarrow B,\ \neg (A \vee B)$

(6) $A \downarrow B,\ B \downarrow A$

(7) $A \downarrow (B \wedge C),\ (A \downarrow B) \vee (A \downarrow C)$

(8) $A \downarrow (B \vee C),\ (A \downarrow B) \wedge (A \downarrow C)$

(9) $A \downarrow \neg A,\ B \wedge \neg B$

(10) $\neg A,\ A \downarrow (B \wedge \neg B)$

3.26. The operator $\not\leftrightarrow$ was introduced in exercise 3.8. It expresses exclusive disjunction and corresponds to either . . . or . . . , but not both. By means of truth tables, determine which of the following pairs of sentences are equivalent.

(1) $A \not\leftrightarrow B,\ B \not\leftrightarrow A$

(2) $A \not\leftrightarrow (B \not\leftrightarrow C),\ (A \not\leftrightarrow B) \not\leftrightarrow C$

(3) $A \not\leftrightarrow A,\ B \wedge \neg B$

(4) $A \not\leftrightarrow \neg A,\ B \vee \neg B$

(5) $A,\ A \not\leftrightarrow (B \wedge \neg B)$

(6) $\neg A,\ A \not\leftrightarrow (B \vee \neg B)$

(7) $A \wedge (B \not\leftrightarrow C),\ (A \wedge B) \not\leftrightarrow (A \wedge C)$

(8) $A \not\leftrightarrow B,\ \neg A \not\leftrightarrow \neg B$

(9) $\neg (A \not\leftrightarrow B),\ \neg A \not\leftrightarrow B$

(10) $\neg (A \not\leftrightarrow B),\ A \not\leftrightarrow \neg B$

(11) $A \not\leftrightarrow B,\ \neg (A \leftrightarrow B)$

(12) $A \nleftrightarrow B, \ \neg A \leftrightarrow B$

(13) $A \nleftrightarrow B, \ A \leftrightarrow \neg B$

(14) $A \nleftrightarrow B, \ (A \wedge \neg B) \vee (B \wedge \neg A)$

(15) $A \nleftrightarrow B, \ (A \vee B) \wedge \neg (A \wedge B)$

(16) $A \vee B, \ (A \nleftrightarrow B) \nleftrightarrow (A \wedge B)$

3.27. Is the conclusion of the following argument a consequence of the premisses? (See exercise 2.12, page 59.)

$$C \rightarrow W, \ \neg (T \wedge W), \ T \ / \ \neg C$$

3.28. Which, if any, of the sentences $\neg P \rightarrow \neg (Q \wedge R)$, $\neg (P \rightarrow \neg (Q \vee R))$, $\neg (P \rightarrow (\neg Q \vee \neg R))$, and $\neg P \rightarrow \neg (Q \vee R)$ imply which of the others? (See exercise 3.5.)

5. Reverse truth table technique

To determine the logical status of a finite set of sentences one may construct a truth table of 2^n rows, where n is the number of atoms of the set. Although this is an effective procedure, it is often unwieldy, since the size of the truth table increases exponentially with the number of atomic sentences. Thus, for a simple example, to use a truth table to show that the conclusion of the argument

$$A \rightarrow B, B \rightarrow C, C \rightarrow D, D \rightarrow E, E \rightarrow F, F \rightarrow G, G \rightarrow H, H \rightarrow I, I \rightarrow J \ / \ A \rightarrow J$$

is a consequence of the premisses, a truth table of more than a thousand rows would be required (2^{10} rows, to be precise).

In this section we describe a generally quicker method for determining logical status. We call this the *reverse truth table technique* because in using it we construct a truth table—usually just part of one—backwards.

As we have seen, an assignment of truth values to the atoms of a set of sentences immediately fixes a truth value for each sentence in the set. Contrariwise, an assignment of truth values to the sentences in a set determines a *class* of models for the atoms. The models so determined are precisely those in which the set's sentences have the initially assigned truth values. From a logical point of view, it matters only whether the class of models determined by an assignment of truth values to the sentences in a set is empty or non-empty. If it is empty, then it is impossible to assign truth values in the way imagined to the sentences in the set; if it is non-empty, then such an assignment is after all possible.

Let us illustrate the reverse truth table technique to show that the following argument is valid.

$$A \rightarrow B,\ B \rightarrow \neg C\ /\ C \rightarrow \neg A$$

The conclusion of the argument is indeed a consequence of the premisses, and this means that it is impossible—coherently—to assign truth to each of the premisses and falsity to the conclusion. That such an assignment is impossible (that it determines the empty class of models) can be seen as follows.

At step **1**, each premiss $A \rightarrow B$ and $B \rightarrow \neg C$ is labeled true, and the conclusion $C \rightarrow \neg A$ is labeled false:

A	B	C	$A \rightarrow B$	$B \rightarrow \neg C$	$C \rightarrow \neg A$
			T	T	$\perp$
			1	**1**	**1**

Because the conditional $C \rightarrow \neg A$ is false, in step **2** its antecedent C is true and its consequent $\neg A$ is false:

A	B	C	$A \rightarrow B$	$B \rightarrow \neg C$	$C \rightarrow \neg A$
		T	T	T	$\perp$ $\perp$
		2	**1**	**1**	**1 2**

Since the negation $\neg A$ is false, its negate A is true—step **3**:

A	B	C	$A \to B$	$B \to \neg C$	$C \to \neg A$
T		T	T	T	⊥ ⊥
3		**2**	**1**	**1**	**1 2**

Now because the conditional $A \to B$ and its antecedent A are both true, so is the consequent B, at step **4**:

A	B	C	$A \to B$	$B \to \neg C$	$C \to \neg A$
T	T	T	T	T	⊥ ⊥
3	**4**	**2**	**1**	**1**	**1 2**

Next because the conditional $B \to \neg C$ and its antecedent B are both true, so is the consequent $\neg C$, at step **5**:

A	B	C	$A \to B$	$B \to \neg C$	$C \to \neg A$
T	T	T	T	T T	⊥ ⊥
3	**4**	**2**	**1**	**1 5**	**1 2**

Finally, given that the negation $\neg C$ is true, its negate C is false. This is recorded at step **6**:

A	B	C	$A \to B$	$B \to \neg C$	$C \to \neg A$
T	T	T	T	T T	⊥ ⊥
3	**4**	**2**	**1**	**1 5**	**1 2**
		⊥			
		6			

But now steps **2** and **6** conflict: the sentence C is both true and false. This being impossible, it follows that the initial assignment of truth values to the premisses and conclusion of the argument is impossible. Since it is thus impossible for the premisses to be true while the conclusion is false, the conclusion is a consequence of the premisses.

All this is quicker to do than to say—which is, of course, the point. In practice we do not write down a series of partial truth tables, like the six above; rather we do all the work on just one table (the last one). So the technique is generally shorter and faster than creating a full truth table.

Steps **1** through **6** above are not the only way to discover a conflicting assignment of truth values for the argument. There are usually many ways, depending on what choices are made as the development unfolds. Here, for example, is another route to the same end result:

A	B	C	$A \to B$	$B \to \neg C$	$C \to \neg A$
⊥	⊥	T	T	T ⊥	⊥ ⊥
5	**4**	**2**	**1**	**1 3**	**1 2**
					T
					6

In working through the steps above, conflicting truth values occur not at an atomic sentence but at another subsentence ($\neg$ A). All the same, it shows that it is not coherent to suppose that the premisses are true and the conclusion is false—which means that the conclusion is a consequence of the premisses.

Now let us look at an example in which assignment of truth to the premisses and falsity to the conclusion does not result in a conflict of truth values. Here is an argument in which the conclusion is not a consequence of the premisses:

$$A \to B,\ \neg (A \vee C)\ /\ \neg B$$

Any model in which this argument's premisses are true and its conclusion is false will make A false, B true, and C false, as we see from the following reverse truth table.

A	B	C	$A \to B$	$\neg$	$(A \vee C)$	$\neg B$
$\perp$	$\top$	$\perp$	$\top$	$\top$	$\perp$	$\perp$
4	**2**	**4**	**1**	**1**	**3**	**1**

The development of this table is as follows. At step **1**, we assume that the premisses are true and the conclusion is false. This means that B is true at step **2**. Falsity of the disjunction $A \vee C$ is next at step **3**, and this determines, at step **4**, that its disjuncts A and C are false. Because there are no conflicting assignments of truth values to the atomic sentences, we can be satisfied that we have arrived at a model in which the premisses are true and the conclusion is false—i.e. at a countermodel to the argument in question.

It should by now be clear that the reverse truth table technique can be used to answer questions concerning finite consequence (i.e. validity of arguments), validity of sentences, equivalence, and finite satisfiability. We summarize and illustrate the methods of application below.

Consequence. To show that a sentence A is a consequence of a (finite) set of sentences Γ, assume that a model exists in which A is false and all the sentences in Γ are true. Then show that in such a model some sentence (i.e. sentence or subsentence) involved is both true and false. This being impossible, it follows that there is no model in which A is false and all the sentences in Γ are true—so A is a consequence of Γ. If *no* conflicting assignment of truth values is reached, then there *is* a model in which A is false and all the sentences in Γ are true—so A is *not* a consequence of Γ.

Validity. To show that a sentence A is valid, assume that a model exists in which A is false. Then show that in such a model some subsentence is both true and false. This being impossible, it is thereby established that there is no model in which A is false—so A is valid. If *no* conflicting assignment of truth values is reached, then there *is* a model in which A is false—so A is *not* valid.

For example, the following table shows that the sentence $\neg(A \vee \neg B) \to B$ is valid.

A	B	$\neg$ (A $\vee$ $\neg$ B) $\rightarrow$ B
$\perp$	$\perp$	T $\perp\perp$ $\perp$
4	**2**	**2** **3 4** **1**
	T	
	5	

Following the steps, we see that a conflict appears at steps **2** and **5**, in which B is assigned both truth and falsity. Thus the initial assumption of falsity does not hold: the sentence is valid.

Equivalence. To show that two sentences A and B are equivalent is to show that A implies B and that B implies A. So assume both that a model exists in which A is true and B is false, and that a model exists in which B is true and A is false. Then show in each case that some subsentence of the sentences involved is both true and false. This being impossible, there is no model in which A is true and B is false, and no model in which B is true and A is false. Hence A implies B and B implies A—i.e. A is equivalent to B. If in either case *no* conflicting assignment of truth values is reached, then there *is* a model in which one of A and B is true while the other is false—so A is *not* equivalent to B.

Note that a conflict may obtain in one case but not the other. This means that implication holds in one direction only, and so the sentences are not equivalent.

The following two truth tables shows that the sentences A $\vee$ $\neg$ B and B $\rightarrow$ A are equivalent. The table on the left shows that A $\vee$ $\neg$ B implies B $\rightarrow$ A, and the table on the right shows that B $\rightarrow$ A implies A $\vee$ $\neg$ B.

A	B	A ∨ ¬ B	B → A
$\bot$	$\top$	$\top \bot$	$\bot$
2	**2**	**1 3**	**1**
		$\bot$	
		4	

A	B	A ∨ ¬ B	B → A
$\bot$	$\top$	$\bot \bot$	$\top$
2	**3**	**1 2**	**1**
	$\bot$		
	4		

Satisfiability. Finally, to show that a (finite) set of sentences Γ is *un*satisfiable, assume that a model exists in which all the sentences in Γ are true. Then show that some sentence or subsentence must be both true and false. This being impossible, conclude that there is no model in which all the sentences in Γ are true—so Γ is unsatisfiable. If *no* conflicting assignment of truth values is reached, then there *is* a model in which all the sentences in Γ are true—so Γ *is* satisfiable.

To illustrate the reverse truth table technique as it applies to satisfiability, consider the set {A ∨ B, A → B, ¬ B}. The following table shows this set to be unsatisfiable.

A	B	A ∨ B	A → B	¬ B
$\bot$	$\bot$	$\top$	$\top$	$\top$
3	**2**	**1**	**1**	**1**
		$\bot$		
		4		

Because a conflicting truth value assignment occurs at the sentence A ∨ B, we reject the initial assumption that all three sentences are true in some model. In other words, there is no model in which all the sentences in the set are true, and hence it is unsatisfiable.

EXERCISES

3.29. Use the reverse truth table technique to show the validity of the following argument.

$A \rightarrow B, B \rightarrow C, C \rightarrow D, D \rightarrow E, E \rightarrow F, F \rightarrow G, G \rightarrow H, H \rightarrow I, I \rightarrow J \ / \ A \rightarrow J.$

3.30. Use the reverse truth table technique to check that the conclusion is a consequence of the premisses in each of the arguments in exercise 3.11. See also exercise 3.16.

3.31. Use the reverse truth table technique to check that each of the sentences is valid in exercise 3.12. See also exercise 3.17.

3.32. Use the reverse truth table technique to check that the pairs of sentences are equivalent in exercise 3.13. See also exercise 3.18.

3.33. Use the reverse truth table technique to check that each of the sets of sentences is unsatisfiable in exercise 3.14. See also exercise 3.19.

3.34. Use the reverse truth table technique to check that the conclusion is a consequence of the premisses in each of the arguments in exercise 3.20.

3.35. Use the reverse truth table technique to check that each of the sentences is valid in exercise 3.22.

3.36. Use the reverse truth table technique to check that the pairs of sentences are equivalent in exercise 3.23.

3.37. Use the reverse truth table technique to check that each of the sets of sentences is unsatisfiable in exercise 3.24.

3.38. Use the reverse truth table technique to determine whether any of the sentences in exercises 2.2 (page 48) and 3.9 are valid or contradictory.

3.39. Use the reverse truth table technique to determine of the following sentences which is valid, which is contradictory, and which is neither.

(1)	$((A \to B) \wedge A) \to B$	(3)	$((A \to B) \to A) \to B$
(2)	$((A \to B) \wedge A) \to \neg B$	(4)	$((A \to B) \to A) \to \neg B$

6. Metatheorems

We record further important facts about consequence, validity, equivalence, and satisfiability. Several lesser propositions, some of them corollaries to those in the text, are also stated as exercises. Much of what follows is based on propositions in chapter 1, particularly *reflexivity*, *transitivity*, and *augmentation* for consequence. Readers may wish to review these before proceeding (propositions 1.2–1.4, pages 14–16).

Consequence and conditionals. We begin our survey with a proposition of great importance in argumentation:

PROPOSITION 3.3. *"Conditional proof".*

$$\text{If } \Gamma \models B, \text{ then } \Gamma - \{A\} \models A \to B.$$

Proof. Suppose that Γ implies B. We know from part (2) of exercise 3.11 (page 77) that B implies $A \to B$, so by transitivity Γ implies $A \to B$. Now to argue that $\Gamma - \{A\}$ also implies $A \to B$, let us suppose otherwise and reason to an absurdity. If $\Gamma - \{A\}$ does not imply $A \to B$, then there is a model in which all the sentences in $\Gamma - \{A\}$ are true but $A \to B$ is false. Hence in this model A is true and B is false. But then all the sentences in the set Γ—i.e. all of $\Gamma - \{A\}$ and also A—are true in the model. Since Γ implies B, B is true in the model as well. Thus B is both true and false in the model, which is absurd. □

Proposition 3.3 provides the foundation for the important rule of inference called PC in chapter 4. In a similar fashion the next proposition underlies the inference rule called MP.

PROPOSITION 3.4. *Modus ponens.*

$$\text{If } \Gamma \models A \to B \text{ and } \Delta \models A, \text{ then } \Gamma \cup \Delta \models B.$$

Proof. Suppose that Γ implies $A \to B$ while Δ implies A. By part (4) of exercise 3.11, B is a consequence of the pair $A \to B$, A. It follows by augmentation and transitivity (or proposition 1.5, page 18) that B is a consequence of $\Gamma \cup \Delta$, as we wished to show. □

Propositions 3.3 and 3.4—conditional proof and modus ponens—combine to yield the next proposition, which may be called the "semantic deduction theorem". It states a general connection between the relation of consequence and the operation of conditionality: that a set of sentences implies a conditional $A \to B$ just in case the consequent B is a consequence of the set together with the antecedent A. The corollaries state connections between valid arguments and valid conditionals—in particular, that a conditional is valid just in case its antecedent implies its consequent.

PROPOSITION 3.5. *"Semantic deduction theorem".*

(1) $\Gamma \cup \{A\} \models B$ iff $\Gamma \models A \to B$.

(2) Corollary: $A_1, \ldots, A_n \models A$ iff $\models A_1 \to (\ldots (A_n \to A) \ldots)$.

(3) Corollary: $A \models B$ iff $\models A \to B$.

Proof. Let us show how conditional proof and modus ponens yield part (1). For left-to-right, assume that $\Gamma \cup \{A\}$ implies B. By conditional proof the set $(\Gamma \cup \{A\}) - \{A\}$ implies $A \to B$. This set is the same as $\Gamma - \{A\}$. But if A is not in Γ then this set is Γ, while if A is in Γ the set $\Gamma - \{A\}$ can be augmented to Γ. In either case we have that Γ implies $A \to B$, as we wished to show. For right-to-left, assume that Γ implies $A \to B$. By reflexivity $\{A\}$ implies A. So by modus ponens $\Gamma \cup \{A\}$ implies B, as we wished to show.

The first corollary follows by repeated applications of the main principle, and the second is just a limiting case of this. □

According to proposition 3.5, premisses $A_1, \ldots, A_n$ imply a conclusion A if and only if the conditional $A_n \to A$ is a consequence of the remaining sentences $A_1, \ldots, A_{n-1}$. In other words, arguments of the following two forms are equally valid:

$$A_1, \ldots, A_n \;/\; A \qquad \text{and} \qquad A_1, \ldots, A_{n-1} \;/\; A_n \rightarrow A$$

For example, the following two arguments are equally valid (and hence equally invalid).

$$\begin{array}{l} A \rightarrow B \\ \neg B \\ \hline \neg (C \wedge A) \end{array} \qquad\qquad \begin{array}{l} A \rightarrow B \\ \hline \neg B \rightarrow \neg (C \wedge A) \end{array}$$

The second premiss of the argument on the left has become the antecedent of the conclusion of the argument on the right.

The first corollary in proposition 3.5 tells us that this procedure can be iterated till we reach an argument with no premisses at all:

$$/ \;\; A_1 \rightarrow (\ldots (A_n \rightarrow A) \ldots)$$

Thus the following arguments are equally valid.

$$\begin{array}{l} A \rightarrow B \\ \neg B \\ \hline \neg (C \wedge A) \end{array} \qquad\qquad \begin{array}{l} \\ \hline (A \rightarrow B) \rightarrow (\neg B \rightarrow \neg (C \wedge A)) \end{array}$$

So the question whether the conclusion of an argument is a consequence of its premisses can always be reduced to the question whether a certain conditional is valid.

This is just one connection between valid arguments and valid conditionals. To reach another, more useful connection, we need some brief preliminaries.

First, it may have been noticed that conjunction is *associative*: $A \wedge (B \wedge C)$ and $(A \wedge B) \wedge C$ are always equivalent (see exercise 3.13, page 78). The practical import of this fact is that it is logically indifferent which way a trio of sentences is grouped by conjunction. So in describing a conjunction of three or more sentences we may omit punctuation, describing both $A \wedge (B \wedge C)$ and $(A \wedge B) \wedge C$ as $A \wedge B \wedge C$. More generally, to describe a conjunction of n sentences $A_1, \ldots, A_n$, we may write:

$$A_1 \wedge \ldots \wedge A_n$$

Secondly, it should be obvious that a model of a finite set $\{A_1, \ldots, A_n\}$ is equally a model of a conjunction $A_1 \wedge \ldots \wedge A_n$ of the set's members—i.e.

$$\mathcal{M} \text{ is a model of } \{A_1, \ldots, A_n\} \text{ iff } \models_{\mathcal{M}} A_1 \wedge \ldots \wedge A_n.$$

This is because an n-termed conjunction is true in a model if and only if each of its conjuncts is true in the model.

So we see that the premisses of an argument can all be conjoined, i.e. that the n-premiss argument

$$A_1, \ldots, A_n \ / \ A$$

is no more and no less valid than the corresponding one-premiss argument

$$A_1 \wedge \ldots \wedge A_n \ / \ A.$$

For example, the following arguments are equally valid.

$$\begin{array}{l} A \rightarrow B \\ \neg B \\ \hline \neg (C \wedge A) \end{array} \qquad\qquad \begin{array}{l} (A \rightarrow B) \wedge \neg B \\ \hline \neg (C \wedge A) \end{array}$$

Now, using this fact and the "semantic deduction theorem" (proposition 3.5), we see that for any argument $A_1, \ldots, A_n \ / \ A$ there is a *corresponding conditional* $(A_1 \wedge \ldots \wedge A_n) \rightarrow A$. And the relationship between the two can be put simply: the argument is valid if and only if its corresponding conditional is a valid sentence. In case there are no premisses in an argument, i.e. when n is zero, we stipulate that $(A_1 \wedge \ldots \wedge A_n) \rightarrow A$ is just the sentence A. So it follows that a zero-premiss argument / A is valid if and only if its conclusion A is a valid sentence. We express it formally as follows.

PROPOSITION 3.6. *Valid arguments and valid conditionals.*

$$A_1, \ldots, A_n \models A \text{ iff } \models (A_1 \wedge \ldots \wedge A_n) \rightarrow A.$$

Proof. Let us argue for this proposition by way of a series of "if and only if"s, as follows. The sentence A is implied by the sentences $A_1, \ldots, A_n$ if and only if A is true in every model of the set $\{A_1, \ldots, A_n\}$, i.e. if and only if A is true in every model of the conjunction $A_1 \wedge \ldots \wedge A_n$, i.e. if and only if A is a consequence of the conjunction $A_1 \wedge \ldots \wedge A_n$, i.e., by part (3) of proposition 3.5, if and only if the conditional $(A_1 \wedge \ldots \wedge A_n) \rightarrow A$ is valid. □

The equivalence of the conditionals $A_1 \rightarrow (\ldots (A_n \rightarrow A) \ldots)$ and $(A_1 \wedge \ldots \wedge A_n) \rightarrow A$, which follows from propositions 3.5 and 3.6, is a simple generalization of the equivalence of the conditionals $A \rightarrow (B \rightarrow C)$ and $(A \wedge B) \rightarrow C$ (see exercises 3.13 and 3.18, pages 78 and 88).

Implication and conditionality. As we remarked above, part (3) of proposition 3.5 says that a conditional is valid just when its antecedent implies its consequent. For this reason, the conditional sign $\rightarrow$ is sometimes called an "implication" sign, and $A \rightarrow B$ is read "A implies B". This practice tends to confuse a *relation among* sentences, consequence ($\models$), with an *operation on* sentences, conditionality ($\rightarrow$). The confusion lingers on harmlessly in our own terminology, where we dub the second component of a conditional the consequent. (A conditional is valid if and only if its consequen<u>t</u> is a consequen<u>ce</u> of its antecedent.)

Consequence and negation. The next few propositions deal with negation. According to the first one, arguments of the following two forms are equally valid:

$$A_1, \ldots, A_n, A \;/\; B \qquad \text{and} \qquad A_1, \ldots, A_n, \neg B \;/\; \neg A$$

Here the negation of the conclusion is added to the premisses while the negation of one of the premisses is taken as the conclusion. The manœuver is sometimes referred to as "indirect proof". Our first proposition states the matter quite generally and in its simplest form.

PROPOSITION 3.7. *"Indirect proof".*

(1) $\Gamma \cup \{A\} \models B$ iff $\Gamma \cup \{\neg B\} \models \neg A$.

(2) Corollary: $A \models B$ iff $\neg B \models \neg A$.

Proof. Part (2) is just the case in (1) where there are no sentences in Γ, i.e. Γ is the empty set. So we argue for (1).

For left-to-right, assume that B is a consequence of $\Gamma \cup \{A\}$. To show that $\Gamma \cup \{\neg B\}$ implies $\neg A$, suppose otherwise—to reach an absurdity. This means that $\neg A$ is false in some model that makes all the sentences in $\Gamma \cup \{\neg B\}$ true. But then A is true along with all of Γ; in short, all the sentences in $\Gamma \cup \{A\}$ are true in the model. Since B is a consequence of $\Gamma \cup \{A\}$, B is also true in the model. But this is absurd, since now we have that both B and $\neg B$ are true in the model.

The reasoning for right-to-left is parallel. We assume that $\Gamma \cup \{\neg B\}$ implies $\neg A$, but (again to reach an absurdity) that B is not a consequence of $\Gamma \cup \{A\}$. So B is false in some model in which all the sentences in $\Gamma \cup \{A\}$ are true. This yields the result that every sentence in $\Gamma \cup \{\neg B\}$ is true in the model, and hence by our assumption so is $\neg A$. But now both A and $\neg A$ are true in the model, which is absurd. □

The next proposition is known as *consequentia mirabilis*—Latin for *wonderful consequence*. In effect, it says that it matters not to the validity of an argument whether the negation of the conclusion is among the premisses: arguments of the forms

$$A_1, \ldots, A_n \,/\, A \qquad \text{and} \qquad A_1, \ldots, A_n, \neg A \,/\, A$$

are equally valid. More generally (as we state it in proposition 3.8 below), a sentence is a consequence of a set if and only if it is a consequence of the set plus—or minus!—its own negation. As a special case, we see that a sentence is valid just in case it is implied by its own negation (which is thus a third way of defining validity in terms of consequence).

PROPOSITION 3.8. *Consequentia mirabilis.*

(1) $\Gamma \models A$ iff $\Gamma \cup \{\neg A\} \models A$.

(2) Equivalently: $\Gamma \models A$ iff $\Gamma - \{\neg A\} \models A$.

(3) Corollary: $\models A$ iff $\neg A \models A$.

Proof. The corollary follows from part (1) by taking Γ to be $\varnothing$. Parts (1) and (2), on the other hand, are equivalent. We prove (1) and leave (2) as an exercise.

From left to right, it is a matter of augmentation: if Γ implies A, then so does $\Gamma \cup \{\neg A\}$. For right-to-left, assume that $\Gamma \cup \{\neg A\}$ implies A, and suppose, to reach an absurdity, that A is not a consequence of Γ. Then a model exists in which A is false while all the sentences in Γ are true. So $\neg$ A holds in this model as well—which means that all of $\Gamma \cup \{\neg A\}$ is true. But then, since A is a consequence of this set, A is also true in the model—which is absurd, as we wished to show. □

Proposition 3.8 expresses a special principle of reductio ad absurdum and underlies a useful deductive strategy, as we see in chapter 4: to show that A follows from $A_1, \ldots, A_n$, it may be useful to add $\neg$ A to these premisses.

The principle of *reductio ad absurdum* itself is the content of the next proposition. In chapter 4 this principle is embodied in a rule of inference we call RAA.

PROPOSITION 3.9. *Reductio ad absurdum.*

$$\text{If } \Gamma \models B \text{ and } \Delta \models \neg B, \text{ then } (\Gamma \cup \Delta) - \{\neg A\} \models A.$$

Proof. Suppose that Γ implies B while Δ implies $\neg$ B. We know (part (3) of exercise 3.11, page 77) that any sentence A is a consequence of the pair B, $\neg$ B and hence of $\Gamma \cup \Delta$ (proposition 1.5, page 18). Now suppose, to reach an absurdity, that A is not a consequence of the set $(\Gamma \cup \Delta) - \{\neg A\}$. This means that a model exists in which A is false while all the sentences in $(\Gamma \cup \Delta) - \{\neg A\}$ are true. But then $\neg$ A is true in this model, and also all the sentences in $\Gamma \cup$

Δ are true. Since by augmentation this set implies both B and $\neg$ B, these two sentences are true in the model as well. But this means that B is both true and false in the model, which is absurd. □

Let us see that consequentia mirabilis (proposition 3.8) follows from reductio ad absurdum. We focus on the interesting direction of part (1)—from right to left. Suppose that A is a consequence of $\Gamma \cup \{\neg A\}$, to show that Γ implies A. By reflexivity, the set $\{\neg A\}$ implies $\neg$ A. In other words:

$$\Gamma \cup \{\neg A\} \models A \qquad \text{and} \qquad \{\neg A\} \models \neg A.$$

Hence, taking B to be A in reductio ad absurdum, A is a consequence of the set $(\Gamma \cup \{\neg A\} \cup \{\neg A\}) - \{\neg A\})$, which a little calculation reduces to $\Gamma - \{\neg A\}$. So A is a consequence of $\Gamma - \{\neg A\}$. Now either this set is Γ (if $\neg$ A is not in Γ) or it can be augmented to Γ (if A is in Γ), so that in either case Γ implies A, as we wished to show.

Soundness of rules of inference. The next proposition has several parts; it confirms the soundness of the rules of inference introduced in chapter 4.

PROPOSITION 3.10. *Soundness of rules of inference.*

(1) $\{A\} \models A$.

(2) If $\Gamma \models B$, then $\Gamma - \{A\} \models A \rightarrow B$.

(3) If $\Gamma \models B$ and $\Delta \models \neg B$, then $(\Gamma \cup \Delta) - \{\neg A\} \models A$.

(4) If $\Gamma \models A \rightarrow B$ and $\Delta \models A$, then $\Gamma \cup \Delta \models B$.

(5) If $\Gamma \models A$ and $\Delta \models B$, then $\Gamma \cup \Delta \models A \wedge B$.

(6) If $\Gamma \models A \wedge B$, then $\Gamma \models A$.

(7) If $\Gamma \models A \wedge B$, then $\Gamma \models B$.

(8) If $\Gamma \models A$, then $\Gamma \models A \vee B$.

(9) If $\Gamma \models B$, then $\Gamma \models A \vee B$.

(10) If $\Gamma \models A \vee B$, $\Delta \models A \rightarrow C$, and $E \models B \rightarrow C$, then $\Gamma \cup \Delta \cup E \models C$.

(11) If $\Gamma \models A \rightarrow B$ and $\Delta \models B \rightarrow A$, then $\Gamma \cup \Delta \models A \leftrightarrow B$.

(12) If $\Gamma \models A \leftrightarrow B$, then $\Gamma \models A \rightarrow B$.

(13) If $\Gamma \models A \leftrightarrow B$, then $\Gamma \models B \rightarrow A$.

Proof. Part (1) above is just reflexivity of consequence—proposition 1.2 (page 14), or part (1) of exercise 3.11 (page 77). The principle emerges in chapter 4 as the rule P. Parts (2), (3), and (4), of course, are propositions 3.3, 3.9, and 3.4, respectively—"conditional proof" for the rule PC, reductio ad absurdum for RAA, and modus ponens for MP.

The remaining parts of the proposition are proved using transitivity and augmentation (equally, proposition 1.5) and the correspondingly numbered parts of exercises 3.11 and 3.16 on pages 77 and 88. For example, for part (5)—for the rule CONJ—we reason as follows. Suppose that Γ implies A and Δ implies B. From part (5) of exercises 3.11 and 3.16 we know that the pair $\{A, B\}$ implies $A \wedge B$. So by augmentation and transitivity $\Gamma \cup \Delta$ implies $A \wedge B$.

We leave (6)–(13) as exercises. □

More on interdefinability. We saw in chapter 1 that some of the logical concepts could be used to define others—for example, that validity is definable in terms of consequence (proposition 1.7, page 25). We record some further such propositions, beginning with the following theorem, which provides a very simple explanation of equivalence in terms of validity: sentences are equivalent just when their biconditional is valid.

PROPOSITION 3.11. *Definability of equivalence in terms of validity.*

$$A \simeq B \text{ iff } \models A \leftrightarrow B.$$

Proof. To say that A and B are equivalent means that A and B have the same truth value in every model. By the definition of truth for biconditionals (clause (6) of 3.2), this means that $A \leftrightarrow B$ is true in every model. And this just means that $A \leftrightarrow B$ is valid. □

The next proposition asserts the definability of satisfiability (or unsatisfiability) in terms of consequence. Parts (1) and (2) state that every sentence is implied by an unsatisfiable set of sentences, and vice versa: a set of sentences is unsatisfiable if it implies every sentence. Parts (3) and (4) provide a more practical definition: a set of sentences is satisfiable just in case there is no sentence such that both it and its negation are implied by the set; equivalently, a set of sentences is unsatisfiable just in case there is such a pair of sentences, each of which is a consequence of the set.

PROPOSITION 3.12. *Definability of satisfiability in terms of consequence.*

(1) $\mathrm{Sat}\ \Gamma$ iff there is a sentence A such that $\Gamma \not\models A$.

(2) Equivalently: $\not{\mathrm{Sat}}\ \Gamma$ iff for every sentence A, $\Gamma \models A$.

(3) $\mathrm{Sat}\ \Gamma$ iff there is no sentence A such that $\Gamma \models A$ and $\Gamma \models \neg A$.

(4) Equivalently: $\not{\mathrm{Sat}}\ \Gamma$ iff there is a sentence A such that $\Gamma \models A$ and $\Gamma \models \neg A$.

Proof. From left to right, part (2) is just proposition 1.14 (page 30)—that an unsatisfiable set implies every sentence. For right-to-left suppose that Γ implies every sentence. Then Γ implies $\mathbf{P}_0 \wedge \neg \mathbf{P}_0$. This sentence is a contradiction—it is false in every model. Hence there is no model in which Γ's sentences are all true. So Γ is unsatisfiable. Proof that parts (1) and (2) are equivalent we leave as an exercise, along with arguments for parts (3) and (4) □

Through propositions 1.7, 3.11, and 3.12 we see that consequence suffices as a basis for the rest of the logical notions. Proposition 3.13, below, shows that consequence is definable in terms of satisfiability, and this establishes that this notion alone is a sufficient basis for all the others. For the proposition reveals that a sentence is a consequence of a set of sentences just when the set together with the negation of the sentence forms an unsatisfiable set. It follows from this that validity and equivalence are also definable in terms of satisfiability.

PROPOSITION 3.13. *Definability of consequence, validity, and equivalence in terms of satisfiability.*

(1) $\Gamma \models A$ iff $\not{\text{Sat}}\ \Gamma \cup \{\neg A\}$.

(2) Corollary: $\models A$ iff $\not{\text{Sat}}\ \neg A$.

(3) Corollary: $A \simeq B$ iff $\not{\text{Sat}}\ \neg (A \leftrightarrow B)$.

Proof. Part (1): Γ implies A if and only if there is no model in which Γ is all true and A is false. This means that there is no model in which Γ is all true and $\neg$ A is true as well. Put another way, there is no model in which $\Gamma \cup \{\neg A\}$ is all true, which just means that $\Gamma \cup \{\neg A\}$ is unsatisfiable. Part (2) is the case in (1) where $\Gamma = \varnothing$. Part (3) follows from (2) and proposition 3.11.

□

Having seen that consequence and satisfiability are each sufficient to define the other and validity and equivalence as well, readers may wonder whether this holds too for validity and equivalence. The answer is affirmative, as we discover in the next section.

EXERCISES

3.40. Referring to exercise 3.13, confirm the associativity of the conjunction operator, $\wedge$. Another part of that exercise shows that disjunction, $\vee$, is likewise associative—so that sometimes we may write $A \vee B \vee C$ rather than $A \vee (B \vee C)$ or $(A \vee B) \vee C$.

3.41. Construct "corresponding conditionals" (see before proposition 3.6) for the arguments in exercise 3.20. Which conditionals are valid?

3.42. Argue for these versions of "indirect proof" and its corollary (proposition 3.7):

(1) $\Gamma \cup \{A\} \models \neg B$ iff $\Gamma \cup \{B\} \models \neg A$.

(2) $A \models \neg B$ iff $B \models \neg A$.

3.43. The equivalence of parts (1) and (2) of consequentia mirabilis (proposition 3.8) follows from the fact that their righthand sides are themselves are equivalent, odd as this may seem. Prove this—i.e. that $\Gamma \cup \{\neg A\} \models A$ iff $\Gamma - \{\neg A\} \models A$.

3.44. We saw above that consequentia mirabilis (proposition 3.8) follows from reductio ad absurdum (proposition 3.9). Prove the reverse—that reductio ad absurdum follows from consequentia mirabilis.

3.45. Provide a characterization of equivalence in terms of validity different from proposition 3.11. (Proposition 1.9, page 27, and proposition 3.5 are helpful.)

3.46. (a) Prove that parts (1) and (2) of proposition 3.12 are equivalent. (b) Prove parts (3) and (4), by reference to the notion of truth in a model or by using parts (1) and (2).

3.47. Explain in detail why equivalence is definable in terms of satisfiability (part (3) of proposition 3.13).

3.48. For each of the odd-numbered arguments in exercise 3.11, form the set consisting of the premisses and the negation of the conclusion. Which of these sets are satisfiable? (*Hint*: exercise 3.24.) Do the same for the arguments in exercise 3.20 and say which sets are satisfiable.

3.49. Form the negation of each of the sentences in exercise 3.22. Which of these negations are satisfiable?

3.50. Explain why:

(1) $\not\text{Sat}$ Sn.

(2) $\not\text{Sat}$ A iff $\models \neg$ A. (Compare proposition 3.13.)

(3) $\not\text{Sat}$ A iff A $\models \neg$ A. (Compare proposition 3.8.)

3.51. Prove parts (6)–(13) of proposition 3.10 using transitivity and augmentation (or proposition 1.5, page 18) and exercises 3.11 and 3.16. Find the rules of inference in chapter 4 that correspond to parts (6)–(13).

3.52. Using proposition 3.10 (and perhaps others), prove:

(1) $\Gamma \models A \to B$ iff $\Gamma - \{A\} \models A \to B$.

(2) $\Gamma \models A$ iff $\Gamma - \{\neg A\} \models A$.

(3) $\models A$ iff $\neg A \models A$.

Notice that (2) and (3) appear in proposition 3.8 (consequentia mirabilis).

3.53. Prove:

(1) $\Gamma \models \neg A$, then $\Gamma - \{A\} \models A \to B$.

(2) $\Gamma \models B$ and $\Delta \models \neg B$, then $(\Gamma \cup \Delta) - \{A\} \models \neg A$.

3.54. Using results from exercise 3.20, prove:

(1) If $\Gamma \models A \to B$ and $\Delta \models B \to C$, then $\Gamma \cup \Delta \models A \to C$.

(2) If $\Gamma \models A \to B$ and $\Delta \models \neg B$, then $\Gamma \cup \Delta \models \neg A$.

(3) If $\Gamma \models A \vee B$ and $\Delta \models \neg A$, then $\Gamma \cup \Delta \models B$.

(4) If $\Gamma \models A \vee B$ and $\Delta \models \neg B$, then $\Gamma \cup \Delta \models A$.

(5) If $\Gamma \models A \vee B$, $\Delta \models A \to C$, and $E \models B \to D$, then $\Gamma \cup \Delta \cup E \models C \vee D$.

(6) If $\Gamma \models \neg B \vee \neg C$, $\Delta \models A \to B$, and $E \models A \to C$, then $\Gamma \cup \Delta \cup E \models \neg A$.

(7) If $\Gamma \models \neg C \vee \neg D$, $\Delta \models A \to C$, and $E \models B \to D$, then $\Gamma \cup \Delta \cup E \models \neg A \vee \neg B$.

3.55. Using only propositions 1.4 and 1.5 (pages 16 and 18), prove (1)–(7) of exercise 3.54.

3.56. Using only propositions 1.4 and 1.5 (pages 16 and 18), and perhaps results in exercise 3.54, prove (1) and (2) in exercise 3.53.

3.57. True or false:

a. If Γ does not imply A, then Γ implies $\neg$ A.

b. A conjunction is valid if and only if each of its conjuncts is valid.

c. Just one of these sentences is valid: $(A \to B) \to A$ and $(B \to A) \to B$.

d. A disjunction is valid if and only if at least one of its disjuncts is valid.

e. There are at least a billion negations that are equivalent to each other.

f. A conditional is unsatisfiable if and only if its antecedent is valid and its consequent is unsatisfiable.

g. Every sentence implies at least one sentence.

h. Every sentence is equivalent to some disjunction.

i. There is a sentence that is a consequence of every set of sentences.

j. A sentence is unsatisfiable if and only if it implies $A \wedge \neg A$.

k. For every sentence A, either A is valid or $\neg$ A is.

l. A conditional is valid if and only if either its antecedent is unsatisfiable or its consequent is valid.

m. Every sentence is a consequence of at least one set of sentences.

n. A conditional is valid if its consequent implies its antecedent.

o. There is a set of sentences Γ of which every sentence is a consequence.

p. The union of satisfiable sets of sentences is itself a satisfiable set.

q. If a sentence is valid, then its negation is not valid.

r. A conjunction is satisfiable if and only if each of its conjuncts is satisfiable.

s. The empty set implies a sentence if and only if its own negation does.

t. A disjunction is unsatisfiable only if its disjuncts are.

7. Compactness

As we remarked in chapter 1, *compactness for consequence* means:

> If $\Gamma \models A$, then there is a finite subset $\{A_1, \ldots, A_n\}$ of Γ such that $\{A_1, \ldots, A_n\} \models A$.

Another, important way of formulating compactness is this: whenever consequence obtains, $\Gamma \models A$, there exists a valid argument

$$A_1, \ldots, A_n \ / \ A$$

in which the premisses $A_1, \ldots, A_n$ are all members of Γ and together imply the conclusion A. Notice that the converse of compactness follows simply from augmentation: if a finite subset of Γ implies A, then by augmentation Γ implies A too. So the "if" in compactness could be strengthened to "iff".

Observe that compactness is of interest only when A is a consequence of an infinite set Γ. For when Γ is finite, Γ itself is a finite subset of Γ that implies A, so the property holds trivially.

We are not yet in a position to prove compactness; a simple proof becomes available in chapter 5 once completeness is established. We can, however, use compactness in the meantime.

In chapter 1 we saw how validity can be defined in terms of consequence (proposition 1.7, page 25). Using compactness and the idea of a corresponding conditional, we can put things the other way around and show the *definability of consequence in terms of validity of sentences*:

> $\Gamma \models A$ iff Γ contains sentences $A_1, \ldots, A_n$ such that $\models (A_1 \wedge \ldots \wedge A_n) \rightarrow A$.

To argue for this from left to right, suppose that Γ implies A. Then by compactness there is a subset $\{A_1, \ldots, A_n\}$ of Γ that implies A. Hence by proposition 3.6, the conditional $(A_1 \wedge \ldots \wedge A_n) \rightarrow A$ is valid, as we wished to show. Note that this reasoning depends on the as yet unproved statement of compactness. The argument for right-to-left does not appeal to compactness; we leave it as an exercise.

Given compactness for consequence, we can show that a set of sentences is satisfiable if all its finite subsets are satisfiable—equivalently, that every unsatisfiable set has a finite unsatisfiable subset. In other words, we have *comcactness for satisfiability*:

- If Sat Δ, for every set of sentences Δ such that Δ is finite and $\Delta \subseteq \Gamma$, then Sat Γ.

- Equivalently: If $\not{\text{Sat}}\ \Gamma$, then there is a set of sentences Δ such that Δ is finite, $\Delta \subseteq \Gamma$, and $\not{\text{Sat}}\ \Delta$.

Let us show the second version. Suppose that Γ is unsatisfiable. Then Γ implies every sentence (proposition 1.14, page 30), in particular $\mathbb{P}_0 \wedge \neg \mathbb{P}_0$. Hence by compactness for consequence there is a finite subset $\{A_1, \ldots, A_n\}$ of Γ that implies $\mathbf{P}_0 \wedge \neg \mathbf{P}_0$. But this sentence is unsatisfiable and hence $\{A_1, \ldots, A_n\}$ is unsatisfiable too (proposition 3.12). Therefore, Γ has a finite unsatisfiable subset, as we wished to show. Notice that this reasoning also depends on the as yet unproved compactness of consequence.

Proposition 3.12 explains the definability of satisfiability in terms of consequence. Thus, via the definability of consequence in terms of validity, above, one can reach characterizations of satisfiability in terms of validity. Likewise, using also proposition 1.9 (page 27), it is possible to define satisfiability in terms of equivalence. We leave these as exercises.

EXERCISES

3.58. (a) Prove the right-to-left of the proposition expressing the definability of consequence in terms of validity (page 116). (b) Characterize satisfiability in terms of validity (proposition 3.12 is also helpful).

3.59. Explain why the two versions of compactness for satisfiability are equivalent.

3.60. Show how satisfiability may be defined in terms of (a) validity and (b) equivalence.

8. Replacement

Here we present the semantic *replacement theorem*, which says that sentences alike except for equivalent parts are themselves equivalent.

PROPOSITION 3.14. *Replacement.* Let A and A′ be sentences alike except that A contains a subsentence B in places where A′ contains a subsentence B′. Then if $B \simeq B'$, then $A \simeq A'$.

Proof. Because it involves mathematical induction, we shall not prove this here. However, the proof depends essentially upon the following lemma, the argument for which does not involve induction. The lemma in effect catalogues all the possible ways equivalent sentences can occur as subsentences of sentences that are otherwise alike.

Lemma. Suppose that A and B are equivalent. Then the following pairs are also equivalent.

(i) $\neg A, \neg B$

(ii)	$A \wedge C,\ B \wedge C$	(vi)	$A \rightarrow C,\ B \rightarrow C$
(iii)	$C \wedge A,\ C \wedge B$	(vii)	$C \rightarrow A,\ C \rightarrow B$
(iv)	$A \vee C,\ B \vee C$	(viii)	$A \leftrightarrow C,\ B \leftrightarrow C$
(v)	$C \vee A,\ C \vee B$	(ix)	$C \leftrightarrow A,\ C \leftrightarrow B$

Let us prove parts (i), (ii), and (iii) and leave the rest as exercises. We suppose throughout that A and B are equivalent, which means that for every model $\mathcal{M}$, A is true in $\mathcal{M}$ if and only if B is.

For (i). It follows from the assumption that for every model $\mathcal{M}$, A is false in $\mathcal{M}$ if and only if B is. By the definition of truth for negations, then, for every model $\mathcal{M}$, $\neg A$ is true in $\mathcal{M}$ if and only if $\neg B$ is, and this means that $\neg A$ and $\neg B$ are equivalent.

For (ii). It follows from the assumption that for every model $\mathcal{M}$, both A and C are true in $\mathcal{M}$ if and only if both B and C are. By the definition of truth for conjunctions, then, for every model $\mathcal{M}$, $A \wedge C$ holds in $\mathcal{M}$ if and only if $B \wedge C$ does. This means that $A \wedge C$ is equivalent to $B \wedge C$.

For (iii). From the assumption that A and B are equivalent, we have by part (ii) of the lemma that $A \wedge C$ and $B \wedge C$ are equivalent. Because equivalence is an equivalence relation (proposition 1.10, page 28), it will suffice to show that $A \wedge C$ and $C \wedge A$ are equivalent and likewise for $B \wedge C$ and $C \wedge B$—i.e. to show that conjunction is "commutative." This can be done by means of a truth table, as in exercise 3.13 (page 78). So we may consider the proof of (iii) complete. □

We can illustrate the replacement theorem with the following pair of sentences.

(a) $(A \wedge (C \vee D)) \leftrightarrow (B \rightarrow (C \vee D))$

(b) $(A \wedge (C \vee D)) \leftrightarrow (B \rightarrow \neg(\neg C \wedge \neg D))$

Sentences (a) and (b) are alike except that (b) has an occurrence of the subsentence $\neg (\neg C \wedge \neg D)$ at one place where (a) has the subsentence $C \vee D$. Because these two subsentences are equivalent (as a truth table will show), it follows by the replacement theorem that sentences (a) and (b) are also equivalent.

The propositions in this section hold for consequence, validity, equivalence, and satisfiability not only in the language of sentential logic, but also in the more powerful languages in chapter 6 and beyond.

EXERCISES

3.61. Prove clauses (iv)–(ix) of the lemma for the replacement theorem (3.14).

3.62. Use truth tables to check the equivalence of the sentences (a) and (b) following proposition 3.14, and to check the equivalence of $C \vee D$ and $\neg (\neg C \wedge \neg D)$.

9. Truth functions, expressibility, definability

In this section we begin a discussion of the expressive power of the language of sentential logic.

Our language contains only five operators: $\neg$, $\wedge$, $\vee$, $\rightarrow$, and $\leftrightarrow$. Each is *truth-functional,* in the sense that any assignment of truth values to the sentences formed by one of these operators determines a truth value for the whole sentence.

Not every conceivable operator is truth-functional. For example, the one-place operator $\Diamond$ with the reading it is possibly true that is not truth-functional. A sentence of the form $\Diamond A$ is true if A is true; but if A is false, the truth value of $\Diamond A$ is not determined. To see this, compare the outcomes for $\Diamond A$ when A means The sky is puce and when A means Some dogs are not dogs. A is false in both cases, but $\Diamond A$ is false only when A has the second meaning: it is at least possible that the sky is puce, but it is not possible that some dogs are not dogs.

Associated with each truth-functional operator is a *truth function,* and we say that a truth-functional operator *expresses* a truth function. The truth function expressed by an operator may be represented by the column of truth values in its truth table (where we assume that the truth table is constructed in a standard way; see the remarks on page 75). Thus the truth functions expressed by the operators $\neg$, $\wedge$, $\vee$, $\rightarrow$, and $\leftrightarrow$ are, respectively:

$$\bot\top,\quad \top\bot\bot\bot,\quad \top\top\top\bot,\quad \top\bot\top\top,\quad \top\bot\bot\top$$

(These are the columns in the truth tables, turned on their sides.)

From a syntactic point of view, the language might contain any number of n-place (or n-ary) operators, for $n \geq 0$. But how many truth functions might be expressed by a given n-place operator? Put another way: How many different truth-functional n-place operators are there, for each number n?

For example, let \$ be a one-place operator. Then the truth table set-up for \$ will look like this:

A	\$(A)
$\top$	
$\bot$	

How many truth functions can \$ express? The answer is given by noting the four different ways the truth table can be filled in with $\top$ and $\bot$:

$$\top\top,\quad \top\bot,\quad \bot\top,\quad \bot\bot$$

Thus the language can contain no more than four non-equivalent truth-functional one-place operators. If there are more than four, some will be equivalent (will express the same truth functions).

In general, where \$ is an n-place operator, the number of truth functions \$ can express is:

$$2^{2^n}$$

Why? Because an n-place operator has 2^n rows in its truth table, and in each row either the value $\top$ or the value $\bot$ can appear. So there are two choices for filling in each of 2^n rows. Hence the number of possible outcomes is 2^{2^n}, and each outcome represents a different truth function expressible by \$.

Thus there are 2^{2^2}, or sixteen, truth functions expressible by two-place operators, i.e. sixteen non-equivalent truth-functional operators $\$_1, \ldots, \$_{16}$. They can be enumerated as follows.

$\$_1$	$\$_2$	$\$_3$	$\$_4$	$\$_5$	$\$_6$	$\$_7$	$\$_8$	$\$_9$	$\$_{10}$	$\$_{11}$	$\$_{12}$	$\$_{13}$	$\$_{14}$	$\$_{15}$	$\$_{16}$
⊤	⊤	⊤	⊤	⊤	⊤	⊤	⊤	⊥	⊥	⊥	⊥	⊥	⊥	⊥	⊥
⊤	⊤	⊤	⊤	⊥	⊥	⊥	⊥	⊤	⊤	⊤	⊤	⊥	⊥	⊥	⊥
⊤	⊤	⊥	⊥	⊤	⊤	⊥	⊥	⊤	⊤	⊥	⊥	⊤	⊤	⊥	⊥
⊤	⊥	⊤	⊥	⊤	⊥	⊤	⊥	⊤	⊥	⊤	⊥	⊤	⊥	⊤	⊥

The truth functions expressed by the operators $\wedge$, $\vee$, $\rightarrow$, and $\leftrightarrow$ appear in the table above: $\$_8$ expresses conjunction; $\$_2$ expresses disjunction; $\$_5$ expresses conditionality; and $\$_7$ expresses biconditionality. Or, to put the matter in another way: $\$_8(A, B)$ is equivalent to $A \wedge B$; $\$_2(A, B)$ is equivalent to $A \vee B$; $\$_5(A, B)$ is equivalent to $A \rightarrow B$; and $\$_7(A, B)$ is equivalent to $A \leftrightarrow B$.

Notice that negation, $\neg$, is expressed by two of the sixteen operators, $\$_{11}$ and $\$_{13}$—in the sense that $\$_{11}(A, B)$ is equivalent to $\neg B$, and $\$_{13}(A, B)$ is equivalent to $\neg A$.

It should also be noted that the first and last operators, $\$_1$ and $\$_{16}$, express, respectively, validity and contradiction. That is to say, any sentence of the form $\$_1(A, B)$ is equivalent to any valid sentence (e.g. $A \vee \neg A$), while any sentence of the form $\$_{16}(A, B)$ is equivalent to any unsatisfiable sentence (e.g. $A \wedge \neg A$).

We leave it as an exercise to verify these claims. We also leave it as an exercise to find sentences using only the five operators $\neg$, $\wedge$, $\vee$, $\rightarrow$, and $\leftrightarrow$ that are equivalent to sentences $\$_i(A, B)$, for $i = 3, 4, 6, 9, 10, 12, 14$, and 15.

Among the two-place truth-functional operators listed in the table above, two are worthy of note: $\$_9$ and $\$_{15}$. The first, $\$_9$, is known as the *Sheffer stroke* and is symbolized by "|". Its truth function ⊥ ⊤ ⊤ ⊤ is the the same as that for negation of conjunction. Operator $\$_{15}$ is also known as the *dagger* $\downarrow$, which we met earlier in exercises 3.7 and 3.25 (pages 69 and 91). Its truth function $\bot\,\bot\,\bot\,\top$ corresponds to negation of disjunction.

We observed above that the truth function for negation is expressed by two two-place operators. Similar results hold for the other truth functions expressed by one-place operators. To make this clear, let $¢_1$, $¢_2$, $¢_3$, and $¢_4$ be one-place operators expressing the four truth functions displayed above on page 120:

$¢_1$	$¢_2$	$¢_3$	$¢_4$
$\top$	$\top$	$\bot$	$\bot$
$\top$	$\bot$	$\top$	$\bot$

Then the following pairs are equivalent.

(i)	$¢_1(A),\ \$_1(A, B)$
(ii)	$¢_2(A),\ \$_4(A, B)$
(iii)	$¢_3(A),\ \$_{13}(A, B)$
(iv)	$¢_4(A),\ \$_{16}(A, B)$

It is left as an exercise to verify these equivalences.

This circumstance holds generally: every truth function expressible by an n-place operator is also expressed by some m-place operator, for each $m > n$. Thus among the 2^{2^3} (= 256) truth functions expressed by three-place operators—which we shall not enumerate here!—there will be those that express all that is expressed by means of one- and two-place operators. For example, the three-place operator # expressing the truth function $\top\,\top\,\top\,\top\,\top\,\bot\,\top\,\bot$ expresses disjunction, in the sense that #(A, B, C) is equivalent to $A \vee C$ (or to $\$_2(A, C)$).

To move in the opposite direction, let $n = 0$. Since

$$2^{2^0} = 2,$$

there are just two non-equivalent truth-functional zero-place operators, $\Uparrow$ and $\Downarrow$. Here are their truth tables:

$\Uparrow$	$\Downarrow$
T	$\perp$

Zero-place operators of course do not "connect" any sentences; they stand alone as sentences: $\Uparrow$ and $\Downarrow$ are like atomic sentences with fixed meanings. The former expresses validity, the latter expresses contradiction—which is to say that $\Uparrow$ is equivalent to $A \vee \neg A$, and $\Downarrow$ to $A \wedge \neg A$, for any sentence A.

So far we have spoken of truth functions as being expressed by operators. But the situation is more general, for we can speak also of truth functions as being expressed by sentences. Where a sentence has n atomic subsentences its truth table has 2^n rows, and the 2^n Ts and $\perp$s in the column beneath the main operator of the sentence expresses the truth function expressed by the whole sentence.

For example, where A, B, and C are atomic sentences (in numerical order), the sentence $B \wedge (A \rightarrow \neg C)$ expresses the truth function $\perp \mathrm{T} \perp \perp \mathrm{T} \mathrm{T} \perp \perp$, given the standard truth table set-up in which the atoms of a sentence are listed in numerical order in the guide column, from left to right, and the truth values in the guide column are arranged in the standard way:

A	B	C	$B \wedge (A \rightarrow \neg C)$
T	T	T	
T	T	$\perp$	
T	$\perp$	T	
T	$\perp$	$\perp$	
$\perp$	T	T	
$\perp$	T	$\perp$	
$\perp$	$\perp$	T	
$\perp$	$\perp$	$\perp$	

In general, a sentence A built up out of n atoms expresses a certain truth function, and A can be thought of as *defining* an n-place operator \$—in the sense that we could invent \$ to correspond to the truth function expressed by A. For example, we could define $\surd$ as a three-place operator expressing the truth function expressed by the sentence $B \wedge (A \rightarrow \neg C)$ above:

A	B	C	$\surd$(A, B, C)
T	T	T	⊥
T	T	⊥	T
T	⊥	T	⊥
T	⊥	⊥	⊥
⊥	T	T	T
⊥	T	⊥	T
⊥	⊥	T	⊥
⊥	⊥	⊥	⊥

(It is left as an exercise to verify the equivalence of the sentences $B \wedge (A \rightarrow \neg C)$ and $\surd(A, B, C)$.)

EXERCISES

3.63. Verify the claims about the truth functions expressed by the operators $\$_i$, for $i = 1, 2, 5, 7, 8, 11, 13$, and 16 (page 121).

3.64. Using only the operators $\neg$, $\wedge$, $\vee$, $\rightarrow$, and $\leftrightarrow$, construct sentences equivalent to $\$_i(A, B)$, for $i = 3, 4, 6, 9, 10, 12, 14$, and 15 (page 121). Can you think of readings for these operators?

3.65. Verify equivalences (i)–(iv) on page 122.

3.66. With regard to the operator # introduced on page 122, check that #(A, B, C) is equivalent to $A \vee C$. How many truth-functional three-place operators express binary disjunction? What are the truth functions expressed by these? How many truth-functional three-place operators express some sort of (inclusive) disjunction?

3.67. Check that $\Uparrow$ is equivalent to $A \vee \neg A$ and that $\Downarrow$ is equivalent to $A \wedge \neg A$ (see page 122).

3.68. Use a truth table to check that $B \wedge (A \rightarrow \neg C)$ is equivalent to $\surd(A, B, C)$ (see page 124).

3.69. Using clues from exercises 3.13 and 3.23, construct sentences using:

(1) Using only $\neg$ and $\wedge$, construct sentences that define each of $\vee$, $\rightarrow$, and $\leftrightarrow$.

(2) Using only $\neg$ and $\vee$, construct sentences that define each of $\wedge$, $\rightarrow$, and $\leftrightarrow$.

(3) Using only $\neg$ and $\rightarrow$, construct sentences that define each of $\wedge$, $\vee$, and $\leftrightarrow$.

3.70. Construct a sentence showing that $\neg$ is definable in terms of $\Downarrow$ and $\rightarrow$ (see page 122).

3.71. Referring to the operators in the table on page 121, write $A \leftarrow B$ for $\$_3(A, B)$, $A \nrightarrow B$ for $\$_{12}(A, B)$, and $A \nleftarrow B$ for $\$_{14}(A, B)$. Which of the following sentences define $\vee$? Which define $\wedge$?

(1) $(A \nleftarrow B) \nleftarrow B$

(2) $\neg (A \nleftarrow (\neg A \nleftarrow B))$

(3) $\neg ((A \nleftarrow B) \nleftarrow \neg A)$

(4) $(\neg A \nrightarrow B) \nrightarrow (A \nrightarrow \neg B)$

(5) $(A \nrightarrow B) \nrightarrow \neg (A \nrightarrow B)$

(6) $(A \nrightarrow \neg B) \nrightarrow (\neg A \nrightarrow B)$

(7) $(A \leftarrow B) \leftarrow A$

(8) $(A \leftarrow B) \leftarrow B$

(9) $A \leftarrow (A \leftarrow B)$

3.72. Use (just) the dagger $\downarrow$, to define $\neg$, $\wedge$, and $\vee$. (*Hint*: exercise 3.25.)

3.73. Using only $\rightarrow$, construct a sentence that defines $\vee$. (*Hint*: exercise 3.71.)

3.74. Consider a truth-functional three-place operator [, ,] yielding sentences of the form [A, B, C] and expressing the truth function $\top\ \top\ \top\ \bot\ \bot\ \bot\ \top\ \bot$. Check that [A, B, C] is equivalent to $(B \rightarrow A) \wedge (\neg B \rightarrow C)$, to see the sense in which [A, B, C] represents a "gamble" with respect to the proposition expressed by B and outcomes ("payoffs") with respect to A and C.

3.75. Suppose $\Box$, O, P, and K are one-place operators with readings as follows.

$\Box$	it is necessarily true that
O	it ought to be the case that
P	it has always been the case that
K	I know that

Are any of $\Box$, O, P, and K truth-functional? Explain why (not). What if K means God knows that?

3.76. Let B be a two-place operator with the reading because. Is B truth-functional? Explain why (not).

10. Truth-functional completeness

Given that every sentence expresses a truth function, it is natural to ask whether every truth function is expressed by some sentence. In other words: Is every truth function expressible? Since every truth function corresponds to an imaginable n-place operator, this is the same as asking: Is every truth-functional n-place operator definable in our language of sentential logic?

The answer is affirmative, and is often expressed by saying that the language, or more particularly its set of operators $\{\neg, \wedge, \vee, \rightarrow, \leftrightarrow\}$, is *truth-function-*

ally (or *expressively*) *complete*. Indeed, the situation is even better, for each of the following sets of operators is by itself truth-functionally complete.

$$\{\neg, \wedge\}, \ \{\neg, \vee\}, \ \{\neg, \rightarrow\}$$

Clearly, if a set of operators is truth-functionally complete, then any larger set is too. In particular, the set $\{\neg, \wedge, \vee\}$ is truth-functionally complete. We show this by means of a simple example.

Suppose we are given the truth function ⊥ ⊤ ⊥ ⊥ ⊤ ⊤ ⊥ ⊥. As we observed above, the sentence $B \wedge (A \rightarrow \neg C)$ expresses this function, but discovering this may well be a haphazard process. Instead we can use a *recipe*—an effective method—to design a sentence that expresses the function.

First we display the truth function on a truth table, with atomic sentences A, B, and C (again, assumed to be in numerical order):

A	B	C	
⊤	⊤	⊤	⊥
⊤	⊤	⊥	⊤
⊤	⊥	⊤	⊥
⊤	⊥	⊥	⊥
⊥	⊤	⊤	⊤
⊥	⊤	⊥	⊤
⊥	⊥	⊤	⊥
⊥	⊥	⊥	⊥

Our recipe begins by noting where the value ⊤ appears in the truth function. In this example, ⊤ appears in the truth function only three times: in the second, fifth, and sixth rows of the truth table. In the next step in the recipe, then, we form three conjunctions. For the second row, where the values of A, B, and C are, respectively, ⊤, ⊤, and ⊥, we form the conjunction

$$A \wedge B \wedge \neg C.$$

For the fifth row, where the guide column values are ⊥, T, and T, we form the conjunction

$$\neg A \wedge B \wedge C.$$

And for the sixth row, where the values are ⊥, T, and ⊥, we form the conjunction

$$\neg A \wedge B \wedge \neg C.$$

In the final step, we make a disjunction of these three conjunctions:

$$(A \wedge B \wedge \neg C) \vee (\neg A \wedge B \wedge C) \vee (\neg A \wedge B \wedge \neg C)$$

(Note that since both conjunction and disjunction are associative, there is no need to group the conjuncts and disjuncts in pairs; see exercise 3.40, page 111.)

Now the displayed disjunction expresses the truth function ⊥ T ⊥ ⊥ T T ⊥ ⊥ (and so also defines the operator √). We leave it as an exercise to check this using a truth table, but one can reason it out just as easily and see how the recipe works.

To begin, note that the first conjunction, $A \wedge B \wedge \neg C$, is true just in case A is true, B is true, and C is false—i.e. exactly in the second row of a truth table for the sentence. Likewise, the second conjunction, $\neg A \wedge B \wedge C$, is true just in case A is false, B is true, and C is true—i.e. exactly in the fifth row of a truth table. And, lastly, the third conjunction, $\neg A \wedge B \wedge \neg C$, is true just in case A is false, B is true, and C is false—i.e. exactly in the sixth row. Therefore a disjunction of these three conjunctions will be true precisely in the three alternatives just listed—i.e. in rows two, five, and six. In other words, the sentence displayed above expresses the given truth function, and it does so using only $\neg$, $\wedge$, and $\vee$.

This recipe will always (with one exception) produce a sentence that expresses a given truth function:

> For each row in the function that contains the value T, form a conjunction of atomic sentences or their negations according as they are true or false in that row. Then make a disjunction of all such conjunctions.

The sole exception is when the truth function does not contain the value T. This means that only $\perp$ appears and the recipe cannot be applied. But then what we have is contradiction. So to express the function we simply construct a contradictory sentence using $\neg$, $\wedge$, and $\vee$. For example: $A \wedge \neg A$—or if $\vee$ is wanted as well, $(A \wedge \neg A) \vee (A \wedge \neg A)$.

For that matter, if the only value in the truth function is T, then what we have is validity. So any valid sentence constructed out of $\neg$, $\wedge$, and $\vee$ will express the function. For example: $A \vee \neg A$—or if $\wedge$ is demanded as well, $(A \wedge A) \vee (\neg A \wedge \neg A)$.

Though these remarks about our recipe do not constitute a proof, they are nevertheless enough for the reader to see how any truth function can be expressed using only $\neg$, $\wedge$, and $\vee$. Thus we may state:

PROPOSITION 3.15. *Truth-functional completeness.* Every truth function is expressible. Equivalently, every truth-functional operator is definable.

Disjunctive normal form. A sentence is said to be in *disjunctive normal form*—or in DNF—when it is a disjunction of conjunctions of atomic sentences or negations of atomic sentences. (Note that this is a syntactic, not semantic, concept.)

The recipe we used to illustrate truth-functional completeness reveals that every truth function is expressible by means of a sentence that is in DNF. For the recipe generally yields a disjunction of conjunctions of atomic sentences and negations of atomic sentences.

As stated, this is not quite true—but we can deal with the exceptions so as to make it so!

The first exception is when there is just one atomic sentence in the guide column (i.e. where the truth function contains just two truth values). Then the "conjunctions" will have only one conjunct, viz. the atomic sentence A or its negation $\neg A$. But each of these can be conjoined with itself ($A \wedge A$, $\neg A \wedge \neg A$) to provide a real, equivalent conjunction.

The second exception is where the number of Ts in the truth function is one. In this case the recipe produces only a single conjunction $A_1 \wedge \ldots \wedge A_n$. But this sentence can be disjoined with itself,

$$(A_1 \wedge \ldots \wedge A_n) \vee (A_1 \wedge \ldots \wedge A_n),$$

to provide a disjunction and hence a sentence in DNF.

The third and last exception is where the number of Ts in the truth function is zero. But then the contradictory disjunction

$$(A_1 \wedge \ldots \wedge A_n \wedge \neg A_n) \vee (A_1 \wedge \ldots \wedge A_n \wedge \neg A_n)$$

both expresses the truth function and is in DNF.

As a corollary to proposition 3.15, then, we see that every sentence in the language is equivalent to some sentence in DNF.

Sheffer stroke and dagger. Among all the possible two-place operators just two are truth-functionally complete by themselves—the Sheffer stroke $\mid$ and the dagger $\downarrow$, expressing, respectively, $\perp$ T T T and $\perp \perp \perp$ T.

To prove that each of $\mid$ and $\downarrow$ is truth-functionally complete, it suffices to see that each is adequate to define all the operators in some truth-functionally complete set. For example, negation, conjunction, and disjunction may be defined using the dagger by, respectively:

$$A \downarrow A, \quad (A \downarrow A) \downarrow (B \downarrow B), \quad (A \downarrow B) \downarrow (A \downarrow B)$$

These three definitions have in effect already been checked in exercise 3.25 (page 91). We leave as an exercise the problem of defining negation, conjunction, and disjunction solely by means of the Sheffer stroke.

Truth-functional incompleteness. Each of the operators $\neg$, $\wedge$, $\vee$, $\rightarrow$, and $\leftrightarrow$ is truth-functionally incomplete; that is, each is inadequate by itself to express every truth function. Moreover, without negation, every combination of the remaining two-place operators is truth-functionally incomplete. And the combination of biconditionality with negation, $\{\neg, \leftrightarrow\}$, is also truth-functionally incomplete.

To prove some combinations of operators truth-functionally incomplete is simple. For example, the set $\{\wedge, \vee\}$ can be seen to be incomplete because the truth table for any sentence built up solely by means of $\wedge$ and $\vee$ will inevitably have the value T in the top row (and thus no truth function of the form $\perp \ldots$ will be expressible).

It should also be obvious (the explanation is left as an exercise) that no zero- or one-place operator alone is truth-functionally complete.

EXERCISES

3.77. Check that the disjunction displayed on page 128 expresses the truth function $\perp \top \perp \perp \top \top \perp \perp$.

3.78. Explain why $A_1 \wedge \ldots \wedge A_n$ is equivalent to $(A_1 \wedge \ldots \wedge A_n) \vee (A_1 \wedge \ldots \wedge A_n)$. (See page 130. *Hint*: exercise 3.13.)

3.79. Explain why $(A_1 \wedge \ldots \wedge A_n \wedge \neg A_n) \vee (A_1 \wedge \ldots \wedge A_n \wedge \neg A_n)$ is a contradiction (see page 130).

3.80. Use the operators $\neg$ and $\rightarrow$ to express the truth-function expressed by $\wedge$ or by $\vee$. Conclude that the set $\{\neg, \rightarrow\}$ is truth-functionally complete.

3.81. Referring to the operators in exercise 3.71, explain why the set $\{\rightarrow, \not\leftarrow\}$ is truth-functionally complete. What about $\{\leftarrow, \not\rightarrow\}$?

3.82. Explain why the sets $\{\Downarrow, \rightarrow\}$ and $\{\neg, [\ ,\ ,\ ]\}$ are truth-functionally complete. (See exercises 3.70 and 3.74.)

3.83. Suppose ♠, ♥, ♦, and ♣ are three-place operators expressing the following truth functions.

♠	$\perp \top \perp \perp \top \perp \perp \perp$	♥	$\perp \top \top \perp \top \perp \perp \perp$
♦	$\top \perp \top \perp \perp \top \perp \top$	♣	$\perp \perp \perp \perp \perp \top \perp \top$

(1) For each of ♠, ♥, ♦, and ♣ construct a sentence to show that it is definable in terms of $\neg$, $\wedge$, and $\vee$.

(2) Construct sentences in DNF equivalent to these sentences:

$\neg$ ♠(A, $\neg$ A, B)	$\neg$ ♥(A, B, $\neg$ B)
$\neg$ ♦(A, $\neg$ B, A $\rightarrow$ B)	$\neg$ ♣(A, ♣(A, B, A), B)

(3) Can you find simpler equivalent sentences (i.e. not necessarily in DNF) using only $\neg$, $\wedge$, $\vee$, $\rightarrow$, $\leftrightarrow$?

3.84. Let ☺ and ☹ be six-place operators expressing the following truth functions.

☺ ⊤ ⊥
⊥ ⊥
⊥ ⊥ ⊥ ⊥ ⊥ ⊤

☹ ⊤ ⊥
⊥ ⊥ ⊤ ⊤ ⊥
⊥ ⊥ ⊥ ⊥ ⊥ ⊤

Construct sentences (in DNF or not, as you please) showing that each of ☺ and ☹ is definable in terms of $\neg$, $\wedge$, and $\vee$.

3.85. A sentence is said to be in *conjunctive normal form*—in CNF—when it is a conjunction of disjunctions of atomic sentences or negations of atomic sentences. Construct a sentence in CNF expressing the truth function ⊥ ⊥ ⊤ ⊥ ⊥ ⊥ ⊤ ⊤. Describe a recipe by means of which every truth function can be shown to be expressible by means of a sentence in CNF. (This means that every sentence is equivalent to some sentence in CNF.)

3.86. Use (just) the Sheffer stroke, $|$, to define $\neg$, $\wedge$, and $\vee$ (see page 121).

3.87. Explain why the set $\{\rightarrow\}$ is truth-functionally incomplete.

3.88. Explain why no zero- or one-place operator alone is truth-functionally complete.

3.89. Explain why the Sheffer stroke ($|$) and the dagger ($\downarrow$) are the only possible truth-functionally complete two-place operators.

3.90. What considerations underlie the truth-functional incompleteness of the set $\{\neg, \leftrightarrow\}$?

3.91. Describe a truth function for three-place operator $\sharp$ that makes $\sharp$ truth-functionally complete by itself. How many truth-functionally complete three-place operators are there?

3.92. True or false:

a. Both these sentences express the truth function for disjunction: $(A \to B) \to B$ and $(C \to D) \to D$.

b. Operators expressing the truth functions $\bot\ \top\ \top\ \bot$ and $\top\ \bot\ \bot\ \top$ are truth-functionally complete by themselves.

c. A three-place operator expressing the truth function $\bot\ \top\ \top\ \top\ \top\ \top\ \top\ \top$ is truth-functionally complete by itself.

d. A three-place operator expressing the truth function $\top\ \top\ \top\ \top\ \top\ \top\ \top\ \bot$ is truth-functionally complete by itself.

e. Exactly one of these sentences expresses the truth function for disjunction: $(A \to B) \leftrightarrow B$ and $(C \to D) \leftrightarrow D$.

4

DEDUCTION IN SENTENTIAL LOGIC

In this chapter we develop a *deductive system* for sentential logic, a formal means of demonstrating consequence and other logical properties and relationships among sentences in sentential logic. Fundamental to our deductive system are the ideas of a *deduction* and *rules of inference*, in terms of which these we define four logical notions: *deducibility*, *theoremhood*, *deductive equivalence*, and *consistency*.

The deductive system is *formal* in the sense that it uses only syntactic properties of the sentences of the language. Because the rules of inference closely resemble patterns of reasoning found in natural language discourse, the deductive system is of a kind called *natural deduction*.

Our procedure in this chapter differs from that in chapter 1, where we defined each of the logical notions of consequence, validity, equivalence, and satisfiability independently in terms of the idea of truth in a model. Here, by contrast, we take one concept—deducibility—as primary, and define the other three in terms of it.

In section 1, we discuss the nature of a deductive system and the concepts of *soundness* and *completeness*. Definitions of deduction, deducibility, and theoremhood appear in section 2, with an example of a deduction. In section 3 we set out the *basic rules of inference* and give examples of their use. In section 4 we explain and illustrate some strategies for using the rules of inference. In section 5 we define deductive equivalence and consistency.

Section 6 contains a proof that the deductive system is *sound* (a proof of completeness appears in chapter 5).

In section 7 we expand the roster of inference rules by adding several others. The addition of these *derived rules* does not strengthen the deduction system—nothing more can be deduced than could be deduced before—but it does make it easier to use.

In section 8 we present a deductive version of the replacement theorem. This entitles us to introduce still more inference rules, called *rules of replacement*.

Section 9 contains a number of general *propositions* about the logical notions developed in this chapter. Finally, in section 10 we discuss the *derivability of the derived inference rules*.

1. The nature of the deductive system

Foremost among the logical notions in which we are interested is *deducibility*. To say that a sentence A is *deducible* from a set Γ, we write:

$$\Gamma \vdash A$$

Thus, like consequence, deducibility is a relation between sets of sentences and sentences. Indeed, deducibility is meant to match the semantic notion of consequence, in the sense that A should be deducible from Γ just in case A is a consequence of Γ:

$$\Gamma \models A \text{ iff } \Gamma \vdash A.$$

This condition can be split into two two properties desired for a system of deduction:

If $\Gamma \vdash A$ then $\Gamma \models A$. If $\Gamma \models A$ then $\Gamma \vdash A$.

According to the lefthand statement, the deductive system is *sound*: whenever a sentence is deducible from a set it is a consequence of the set. Conversely, according to the other statement, the system is *complete*: whenever a sentence is a consequence of a set it is deducible from the set.

Our deductive system for sentential logic is in fact sound (as we show in section 6 below): it provides a means of proving the validity of an argument, i.e. of proving that the conclusion is a consequence of the premisses.

In addition to deducibility, we define three other deductive notions: *theoremhood*, *deductive equivalence*, and *consistency*. These are the deductive analogues to the semantic concepts of validity, equivalence, and satisfiability.

Theoremhood, like validity, is a property of sentences, and we write

$$\vdash A$$

to mean that A is a *theorem*. Deductive equivalence relates sentences, and to say that sentences A and B are *deductively equivalent* we write:

$$A \sim B$$

As with satisfiability, consistency is a property of sets of sentences, and to say that Γ is *consistent* set we write:

$$\text{Con } \Gamma$$

(For *in*consistency: C$\not{o}$n Γ.)

Given soundness and completeness, theoremhood, deductive equivalence, and consistency correspond to the semantic concepts of validity, equivalence, and satisfiability. In other words, we have:

$$\models A \text{ iff } \vdash A. \qquad A \simeq B \text{ iff } A \sim B. \qquad \text{Sat } \Gamma \text{ iff Con } \Gamma.$$

2. Deductions, deducibility, and theoremhood

Let us begin with an example of a deduction:

$\{A \vee B\}$	(1)	$A \vee B$
$\{\neg B\}$	(2)	$\neg B$
$\{B\}$	(3)	B
$\{\neg B, B\}$	(4)	A
$\{\neg B\}$	(5)	$B \rightarrow A$
$\{A\}$	(6)	A
$\varnothing$	(7)	$A \rightarrow A$
$\{A \vee B, \neg B\}$	(8)	A

In a line of a deduction the set Γ is called the *premiss set of the line* or the *premiss set of the sentence* A on the line.

Intuitively, each line in a deduction represents an argument—premisses on the left, conclusion on the right. So for example we might (though we do not) write the deduction above as a sequence of arguments:

(1)	$A \vee B$	/	$A \vee B$
(2)	$\neg B$	/	$\neg B$
(3)	B	/	B
(4)	$\neg B, B$	/	A
(5)	$\neg B$	/	$B \to A$
(6)	A	/	A
(7)		/	$A \to A$
(8)	$A \vee B, \neg B$	/	A

In the notion of deduction we develop here, a deduction consists of a sequence of pairs $\langle \Gamma, A \rangle$ consisting of a set of sentences Γ (the *premiss set*) and a sentence A. Each pair in a deduction is called a *line*—where the parenthetical *line number*, (n), indicates which line in the sequence it is. The function of the rules of inference is to *justify* the lines of a deduction, i.e. to stipulate which pairs (arguments) are permitted to occur as lines in a deduction. Never mind for the moment what rules justify the lines in the deduction above; we shall see soon enough.

For the record, here is the definition of a deduction:

DEFINITION 4.1. *Deductions.* A deduction is a sequence of lines—

$$\begin{array}{ccc} \Gamma_1 & (1) & A_1 \\ & \vdots & \\ \Gamma_n & (n) & A_n \end{array}$$

—in which each line is justified by at least one of the rules of inference.

Sometimes we write the lines of a deduction horizontally: $\langle \Gamma_1, A_1 \rangle, \ldots, \langle \Gamma_n, A_n \rangle$.

As we remarked above, a sentence is said to be deducible from a set of sentences just in case there is a *deduction of the sentence from the set*—where this means that the sentence appears on the last line of a deduction with the set or one of its subsets as premiss set. Thus in the deduction above, the last line shows that the sentence A is deducible from the set $\{A \vee B, \neg B\}$ (and from any set of which $\{A \vee B, \neg B\}$ is a subset, for that matter). Formally:

DEFINITION 4.2. *Deducibility.* $\Gamma \vdash A$ iff there is a deduction of A from Γ.

In other words, A is deducible from Γ if and only if there is a deduction the last line of which is $\langle \Gamma', A \rangle$, where Γ' is a subset of Γ.

Question: Why do we not define deducibility so that A is deducible from Γ only if the set Γ itself occurs as the premiss set of A on the last line of a deduction? Why allow subsets of Γ?

Answer: Because of the nature of the rules of inference, premiss sets are always finite. So to require that A be deducible from Γ just in case the pair $\langle \Gamma, A \rangle$ itself is the last line of a deduction would yield the unwanted result that deducibility holds only between finite sets and sentences. But deducibility is meant to match consequence, and consequence can hold between infinite sets and sentences. So we must accommodate our definition of deducibility to allow for this possibility.

In the special case when a sentence is deducible from the empty set, $\varnothing$, we call the sentence a *theorem* (more precisely, a *theorem of logic*). That is to say:

DEFINITION 4.3. *Theoremhood.*

$$\vdash A \text{ iff } \varnothing \vdash A.$$

Theorems are thus sentences that appear on last lines of deductions with the empty set as premiss set.

The remaining critical idea is that of *rule of inference*, on which the notions defined above ultimately depend. The basic rules of inference are set forth in the next section.

Terminological remarks. The symbol "$\vdash$", called the single turnstile, is used to express facts about deducibility much in the same way "$\models$" is used for consequence. When it happens that a sentence B is deducible from a set $\{A_1, \ldots, A_n\}$, i.e. a set containing just the sentences $A_1, \ldots, A_n$, we may omit the braces and write:

$$A_1, \ldots, A_n \vdash B$$

In particular, where B is deducible from a set {A} containing a single sentence, we may say that B is deducible from the sentence A itself and write:

$$A \vdash B$$

When A is deducible from Γ it is often convenient to say that Γ *proves* A. The verb "to prove" provides a natural left-to-right reading of the deducibility relation.

To deny deducibility we use the slash:

$$\Gamma \nvdash A$$

Note that this means that A is not deducible from Γ, or that Γ does not prove A. It does not mean that $\Gamma \vdash \neg A$ ("Γ proves not-A").

Indeed, it should be observed that one cannot "deduce"—i.e. use a deduction to show—that a sentence A is *not* deducible from a set Γ ($\Gamma \nvdash A$) or that A is *not* a theorem ($\nvdash A$). For to say that A is not deducible from Γ is to say that no deduction exists that ends with a line A and a subset of Γ as premiss set—something that cannot be established by means of a deduction.

EXERCISES

4.1. Check (e.g. by means of truth tables) that each of the arguments in the deduction above is valid.

3. The basic rules of inference

The chief consideration prompting our choice of rules is, of course, that the deductive system be sound and complete. But we also want the rules to reflect as much as possible familiar patterns of reasoning.

It is difficult, however, to satisfy everyone's intuitions about which inference rules represent natural patterns of reasoning. Moreover, we must strike a balance between the virtues of economy and utility: if there are too few rules, deductions will be hard to find; if there are too many, a faulty or incomplete memory will make one miss otherwise obvious deductions. Our approach is to enumerate a fairly large stock of rules, but to distinguish a small group of these as *basic* and show the rest to be *derivable*.

The rules of inference state the conditions under which a pair $\langle\Gamma, A\rangle$ may occur as a line of a deduction. In general, a rule has the form

$$\frac{\Gamma_1,\ A_1\ ;\ \ldots\ ;\ \Gamma_n,\ A_n}{\Gamma,\ A}$$

—where n (the number of *hypotheses* of the rule) may be any finite number including 0. The rule says that a line (argument) of the form $\langle\Gamma, A\rangle$ may appear in a deduction provided that lines of the forms $\langle\Gamma_1, A_1\rangle, \ldots, \langle\Gamma_n, A_n\rangle$ already appear (i.e. above it) in the deduction.

There are ten basic rules of inference: P (premiss); PC (proof of a conditional); RAA (reductio ad absurdum); MP (modus ponens); CONJ (conjunction); SIMP (simplification); ADD (addition); DIL (dilemma); BI (biconditional introduction); and BE (biconditional elimination). We record them below in a formal definition (and for future reference).

DEFINITION 4.4. *Basic rules of inference for sentential logic.*

P. $$\frac{\qquad}{\{A\},\ A}$$

PC. $$\frac{\Gamma,\ B}{\Gamma - \{A\},\ A \to B}$$

RAA. $$\frac{\Gamma,\ B;\ \ \Delta,\ \neg B}{(\Gamma \cup \Delta) - \{\neg A\},\ A}$$ B = A allowed?

MP. $$\frac{\Gamma,\ A \to B;\ \ \Delta,\ A}{\Gamma \cup \Delta,\ B}$$

CONJ. $$\frac{\Gamma,\ A;\ \ \Delta,\ B}{\Gamma \cup \Delta,\ A \wedge B}$$

SIMP. $\dfrac{\Gamma,\ A \wedge B}{\Gamma,\ A}$ $\qquad$ $\dfrac{\Gamma,\ A \wedge B}{\Gamma,\ B}$

ADD. $\dfrac{\Gamma,\ A}{\Gamma,\ A \vee B}$ $\qquad$ $\dfrac{\Gamma,\ B}{\Gamma,\ A \vee B}$

DIL. $\dfrac{\Gamma,\ A \vee B;\quad \Delta,\ A \to C;\quad E,\ B \to C}{\Gamma \cup \Delta \cup E,\ C}$

BI. $\dfrac{\Gamma,\ A \to B;\quad \Delta,\ B \to A}{\Gamma \cup \Delta,\ A \leftrightarrow B}$

BE. $\dfrac{\Gamma,\ A \leftrightarrow B}{\Gamma,\ A \to B}$ $\qquad$ $\dfrac{\Gamma,\ A \leftrightarrow B}{\Gamma,\ B \to A}$

Most of the remainder of this section is devoted to examples of deductions illustrating the basic rules of inference—and proving a number of theorems along the way. In the next section we explain some strategies for using the rules to create deductions.

Premiss. Of the ten basic inference rules, only the rule P has no hypotheses. It states that the pair $\langle \{A\}, A \rangle$ may occur as a line of a deduction—for any sentence A, and no matter what lines may occur before it in the deduction. Because P is the only rule without hypotheses, every deduction must begin with a use of the rule P. Indeed, a deduction may both begin and end with a use of this rule: by the rule P, the sequence

$\{A\}$ (1) A P

is a deduction, a deduction of A from $\{A\}$.

Proof of a conditional. The rule PC has one hypothesis. It states that the pair $\langle \Gamma - \{A\}, A \to B \rangle$ may occur as a line of a deduction provided that the pair $\langle \Gamma, B \rangle$ occurs as an earlier line. The use of the rule PC is illustrated by the following deduction showing that $A \rightarrow A$ is a theorem, i.e. that $A \rightarrow A$ is deducible from the empty set.

$\{A\}$	(1)	A	P
$\varnothing$	(2)	$A \rightarrow A$	PC, 1

In this deduction, which is an extension of the first one, line (1) is again justified by the rule P, and line (2) is obtained by the rule PC from line (1): the sentence A on line (1) is brought down as consequent of a conditional on line (2), with A itself as antecedent. To construct the premiss set of line (2), we bring down the premiss set from line (1)—i.e. $\{A\}$—and remove from it the sentence A that was entered as antecedent of the conditional. The result is $\{A\} - \{A\}$, i.e. the empty set, $\varnothing$.

Notice that if in the rule PC the antecedent sentence A is not in the premiss set Γ, then the premiss set $\Gamma - \{A\}$ will just be Γ itself. Line (2) of the next deduction, showing that $A \rightarrow (B \rightarrow A)$ is a theorem, is an example of this.

$\{A\}$	(1)	A	P
$\{A\}$	(2)	$B \rightarrow A$	PC, 1
$\varnothing$	(3)	$A \rightarrow (B \rightarrow A)$	PC, 2

In going from line (1) to line (2) by PC, the antecedent B is deleted from the premiss set $\{A\}$; the result, $\{A\} - \{B\}$, is just $\{A\}$. Note that deletion of the antecedent occurs non-vacuously, however, on line (3).

Reductio ad absurdum. The rule RAA has two hypotheses and, like PC, requires removal of a sentence from the premiss set of the inferred line. The rule states that the pair $\langle(\Gamma \cup \Delta) - \{\neg A\}, A\rangle$ may occur as a line of a deduction provided that the pairs $\langle\Gamma, B\rangle$ and $\langle\Delta, \neg B\rangle$ occur as earlier lines in the deduction. (It does not matter in which order the hypotheses occur.) Use of the rule RAA is exemplified in the following deduction, showing that $\neg\neg A \rightarrow A$ is a theorem.

$\{\neg\neg A\}$	(1)	$\neg\neg A$	P
$\{\neg A\}$	(2)	$\neg A$	P
$\{\neg\neg A\}$	(3)	A	RAA, 1, 2
$\varnothing$	(4)	$\neg\neg A \rightarrow A$	PC, 3

In this deduction, lines (1) and (2) are justified by the rule P. Because a sentence on one of those lines ($\neg \neg A$) is the negation of the sentence on the other line ($\neg A$), the rule RAA permits the inference of any sentence on line (3)—in particular A. To construct the premiss set of line (3), we bring down all the contents of the premiss sets from lines (1) and (2), $\{\neg \neg A, \neg A\}$, and remove the negation of the sentence entered on line (3), $\neg A$; the result, $\{\neg \neg A, \neg A\} - \{\neg A\}$, is the set $\{\neg \neg A\}$. The deduction ends with an application of PC.

Notice that if in the rule RAA the sentence $\neg A$ is not in either of the premiss sets Γ or Δ, then the premiss set $(\Gamma \cup \Delta) - \{\neg A\}$ will just be $\Gamma \cup \Delta$. Line (3) of the following deduction is an example of this. (The deduction shows that $\neg A \to (A \to B)$, sometimes called the Law of Duns Scotus, is a theorem.)

$\{\neg A\}$	(1)	$\neg A$	P
$\{A\}$	(2)	A	P
$\{\neg A, A\}$	(3)	B	RAA, 1, 2
$\{\neg A\}$	(4)	$A \to B$	PC, 3
$\varnothing$	(5)	$\neg A \to (A \to B)$	PC, 4

In going from lines (1) and (2) to line (3) by RAA, the sentence $\neg B$ is deleted from the union of the premiss sets on lines (1) and (2); the result, $\{\neg A, A\} - \{\neg B\}$, is just $\{\neg A, A\}$.

The annotations paralleling the descriptions of the deductions above indicate in an abbreviated manner the justifications of the lines of the deductions. It must be emphasized that such annotations, while sometimes helpful, are not part of the deduction descriptions. It is not necessary in describing a deduction to annotate it in this fashion; the description of the sequence of lines is sufficient. In the deductions that follow, we generally annotate only lines to which we are paying special attention. It is left as an exercise to fill in the missing justifications.

The rules P, PC, and RAA are the "workhorses" of the deductive system; they are used over and over again in the construction of deductions. In section 4 we describe *strategies* for using PC and RAA to create deductions.

The rules PC and RAA are the only ones that stipulate deletion of sentences from premiss sets. In all the other rules, except P, the premiss set of the inferred line is simply the union of the premiss sets of the hypotheses. In short, as a cursory inspection of MP through BE will show, the instruction regarding premiss sets is "bring them all down".

Let us resume our survey of the basic rules of inference.

Modus ponens. The rule MP has two hypotheses. It states that the pair $\langle \Gamma \cup \Delta, B \rangle$ may occur as a line of a deduction provided that the pairs $\langle \Gamma, A \to B \rangle$ and $\langle \Delta, A \rangle$ occur as lines earlier in the deduction (in either order). The use of the rule MP is illustrated in the following proof that $(A \to (B \to C)) \to ((A \to B) \to (A \to C))$ is a theorem.

$\{A \to (B \to C)\}$	(1)	$A \to (B \to C)$	P
$\{A \to B\}$	(2)	$A \to B$	P
$\{A\}$	(3)	A	P
$\{A \to (B \to C), A\}$	(4)	$B \to C$	MP, 1, 3
$\{A \to B, A\}$	(5)	B	MP, 2, 3
$\{A \to (B \to C), A \to B, A\}$	(6)	C	MP, 4, 5
$\{A \to (B \to C), A \to B\}$	(7)	$A \to C$	PC, 6
$\{A \to (B \to C)\}$	(8)	$(A \to B) \to (A \to C)$	PC, 7
$\varnothing$	(9)	$(A \to (B \to C)) \to ((A \to B) \to (A \to C))$	PC, 8

In this deduction, lines (1)-(3) are justified by the rule P. On lines (1) and (3) there appear the conditional $A \to (B \to C)$ and its antecedent A; so the consequent $B \to C$ is inferred on line (4) by the rule MP, and the contents of the premiss sets on lines (1) and (3) form the premiss set of line (4). Similarly, on lines (2) and (3) there appear the conditional $A \to B$ and its antecedent A; so the consequent B is inferred on line (5) by the rule MP, and the contents of the premiss sets on lines (2) and (3) are brought down to (5). In the last use of the rule MP, the sentence C is inferred on (6) from the sentences $B \to C$ and B appearing on lines (4) and (5), and the premiss sets of the earlier lines are brought down. Three applications of the rule PC complete the deduction.

Conjunction. Like the rule MP, the rule CONJ has two hypotheses. It states that the pair $\langle \Gamma \cup \Delta, A \land B \rangle$ may occur as a line of a deduction provided that the pairs $\langle \Gamma, A \rangle$ and $\langle \Delta, B \rangle$ occur as earlier lines in the deduction. A simple use of the rule CONJ appears in the following deduction establishing the theorem $A \to (B \to (A \land B))$.

$\{A\}$	(1)	A	
$\{B\}$	(2)	B	
$\{A, B\}$	(3)	$A \wedge B$	CONJ, 1, 2
$\{A\}$	(4)	$B \rightarrow (A \wedge B)$	
$\varnothing$	(5)	$A \rightarrow (B \rightarrow (A \wedge B))$	

In this deduction, the conjunction $A \wedge B$ on line (3) is inferred from the sentences A and B on lines (1) and (2); note the construction of the premiss set. Justification of the remaining lines of the deduction is left as an exercise.

Simplification. The companion to the preceding rule of inference is the rule SIMP, which has two versions, each with one hypothesis. They can be summarized by saying that from a conjunction each of the conjuncts may be inferred. That is, if the pair $\langle \Gamma, A \wedge B \rangle$ occurs as a line of a deduction, then either (or both) of the lines $\langle \Gamma, A \rangle$ and $\langle \Gamma, B \rangle$ may occur as later lines in the deduction. Uses of both versions of the rule SIMP are exemplified in the next deduction, which shows that $(A \wedge \neg A) \rightarrow B$ is a theorem.

$\{A \wedge \neg A\}$	(1)	$A \wedge \neg A$	P
$\{A \wedge \neg A\}$	(2)	A	SIMP, 1
$\{A \wedge \neg A\}$	(3)	$\neg A$	SIMP, 1
$\{A \wedge \neg A\}$	(4)	B	RAA 2, 3
$\varnothing$	(5)	$(A \wedge \neg A) \rightarrow B$	

The uses of SIMP in the deduction above are noted to the right. The other rules used in the deduction are P, PC, and RAA; it is left as an exercise to say where and how.

Addition. The rule ADD, like SIMP, comes in two versions, each with one hypothesis. Summarized, the rule states that a disjunction may be inferred from either of its disjuncts. That is, if either of the pairs $\langle \Gamma, A \rangle$ or $\langle \Gamma, B \rangle$ occurs as a line of a deduction, then the pair $\langle \Gamma, A \vee B \rangle$ may occur as a later line. Let us illustrate both versions of the rule ADD by a deduction of a theorem known as the Law of Excluded Middle, $A \vee \neg A$.

$\{\neg (A \vee \neg A)\}$	(1)	$\neg (A \vee \neg A)$	P
$\{\neg A\}$	(2)	$\neg A$	P
$\{\neg A\}$	(3)	$A \vee \neg A$	ADD, 2

$\{\neg(A \vee \neg A)\}$	(4)	A	RAA, 1, 3
$\{\neg(A \vee \neg A)\}$	(5)	$A \vee \neg A$	ADD, 4
$\varnothing$	(6)	$A \vee \neg A$	RAA, 1, 5

The uses of the rule ADD in this deduction are in the moves from line (2) to line (3) and from line (4) to line (5). It may be surprising that the sentence $A \vee \neg A$ recurs so often in the deduction. We have noted the justifications for lines (1), (2), (4), and (6), as well as for (3) and (5). Study these carefully. We shall comment on the deduction again, later on.

Dilemma. As a companion to the rule ADD there is the rule DIL. This rule has three hypotheses. It states that the pair $\langle \Gamma \cup \Delta \cup E, C\rangle$ may occur as a line of a deduction provided that the pairs $\langle \Gamma, A \vee B\rangle$, $\langle \Delta, A \to C\rangle$, and $\langle E, B \to C\rangle$ occur as earlier lines. An example of the use of the rule DIL is afforded by the following proof of the theorem $(A \to B) \to ((A \vee B) \to B)$.

$\{A \to B\}$	(1)	$A \to B$	
$\{A \vee B\}$	(2)	$A \vee B$	
$\{B\}$	(3)	B	
$\varnothing$	(4)	$B \to B$	
$\{A \to B, A \vee B\}$	(5)	B	DIL, 1, 2, 4
$\{A \to B\}$	(6)	$(A \vee B) \to B$	
$\varnothing$	(7)	$(A \to B) \to ((A \vee B) \to B)$	

In this deduction, line (5) is inferred by the rule DIL from lines (1), (2), and (4): the sentence B is the common consequent of the conditionals $A \to B$ and $B \to B$, on lines (1) and (4); and the disjunction of the antecedents, $A \vee B$, occurs on line (2). Notice, too, the content of the premiss set of line (5): $\{A \to B, A \vee B\}$ is the union of the sets $\{A \to B\}$, $\{A \vee B\}$, and $\varnothing$—the premiss sets of lines (1), (2), and (4). Justification of the remaining lines of this deduction is left as an exercise. More examples of the use of the rule DIL are to be found in the exercises.

Biconditional introduction. The rule BI has two hypotheses, and it states that the pair $\langle \Gamma \cup \Delta, A \leftrightarrow B\rangle$ may occur as a line of a deduction provided that the pairs $\langle \Gamma, A \to B\rangle$ and $\langle \Delta, B \to A\rangle$ occur earlier in the deduction. In short, a biconditional may be inferred from its conditional halves. An amusing example of the use of the rule BI is provided in the following deduction of the theorem $A \leftrightarrow A$.

$\{A\}$	(1)	A	
$\varnothing$	(2)	$A \to A$	
$\varnothing$	(3)	$A \leftrightarrow A$	BI, 2, 2

As the annotation indicates, line (3) of the deduction above is inferred from the line just above it, (2), taken twice over. This is fair and perfectly in accord with the statement of the rule of inference.

Biconditional elimination. Finally, we come to the rule BE, the opposite number to BI. The rule has two versions and says, in summary, that each of a biconditional's conditional halves may be inferred from it. That is, if the pair $\langle \Gamma, A \leftrightarrow B \rangle$ occurs as a line of a deduction, then either (or both) of the pairs $\langle \Gamma, A \to B \rangle$ and $\langle \Gamma, B \to A \rangle$ may occur as later lines. Use of the rule BE is illustrated below, in a long and rather complicated proof of the theorem $\neg (A \leftrightarrow \neg A)$.

$\{\neg \neg (A \leftrightarrow \neg A)\}$	(1)	$\neg \neg (A \leftrightarrow \neg A)$	
$\{\neg (A \leftrightarrow \neg A)\}$	(2)	$\neg (A \leftrightarrow \neg A)$	
$\{\neg \neg (A \leftrightarrow \neg A)\}$	(3)	$A \leftrightarrow \neg A$	RAA, 1, 2
$\{\neg \neg (A \leftrightarrow \neg A)\}$	(4)	$A \to \neg A$	BE, 3
$\{\neg \neg (A \leftrightarrow \neg A)\}$	(5)	$\neg A \to A$	BE, 3
$\{\neg A\}$	(6)	$\neg A$	
$\{\neg \neg (A \leftrightarrow \neg A), \neg A\}$	(7)	A	MP, 5, 6
$\{\neg \neg (A \leftrightarrow \neg A)\}$	(8)	A	RAA, 6, 7
$\{\neg \neg (A \leftrightarrow \neg A)\}$	(9)	$\neg A$	MP, 4, 8
$\varnothing$	(10)	$\neg (A \leftrightarrow \neg A)$	RAA, 8, 9

The uses of the rule BE in the deduction above are in the moves from line (3) to line (4) and from line (3) to line (5). Except for a few obvious lines, we have supplied the justifications for the rest of the lines of the deduction. Here, again, study the deduction carefully.

Abbreviating deductions. In describing a deduction it is tedious and often cumbersome to write out the premiss sets of the lines of the deduction. For the sake of brevity, let us agree to use the line numbers of the sentences in the premiss sets instead of the sentences themselves—so that the line numbers function as proxies for such sentences. For example, the deduction on page 142 for the theorem $\neg\neg A \to A$ may be more succinctly described as follows.

{1}	(1)	$\neg\neg A$	P
{2}	(2)	$\neg A$	P
{1}	(3)	A	RAA, 1, 2
$\varnothing$	(4)	$\neg\neg A \to A$	PC, 3

In this deduction, the number 1 is a proxy for the sentence $\neg\neg A$—since that sentence appears on line number (1)—and appears in place of the sentence in the premiss sets for lines (1) and (3). Similarly, the number 2 in the premiss set on the second line is a proxy for $\neg A$, since that sentence appears on line number (2). Then in the construction of the premiss set for line (3), inferred from (1) and (2) by RAA, the line number 2 of the sentence $\neg A$ is deleted from the union {1, 2} of the premiss sets of the preceding lines, yielding the premiss set {1}. And finally, in the construction of the premiss set for line (4), inferred from (3) by PC, the line number 1 of the antecedent $\neg\neg A$ of the conditional is deleted from the premiss set {1} of the preceding line, yielding $\varnothing$ as the premiss set.

There should be little difficulty in abbreviating deduction descriptions in this way, but it is worth while to ponder the effect on the rules P, PC, and RAA (there is no important effect on any of the others).

With line numbers as proxies for sentences in premiss sets, the rules P, PC, and RAA may be stated as follows.

- The rule P: A line of a deduction may be a pair $\langle \{n\}, A \rangle$, where n is the line number of a line—ordinarily this very line (n)—with A on it.

- The rule PC: $\langle \Gamma - \{n\}, A \to B \rangle$ may be a line of a deduction provided n is the number of a line with the antecedent A on it and the line $\langle \Gamma, B \rangle$ occurs earlier in the deduction.

- The rule RAA: $\langle(\Gamma \cup \Delta) - \{n\}, \mathrm{A}\rangle$ may occur as a line of a deduction provided n is the number of a line with the sentence $\neg$ A on it and the lines $\langle\Gamma, \mathrm{B}\rangle$ and $\langle\Delta, \neg \mathrm{B}\rangle$ occur earlier in the deduction.

Sometimes the same sentence occurs on more than one line of a deduction and so has more than one line number. In such cases, one must take care in applying the rules PC and RAA to see that the right line numbers are deleted from the premiss set of the inferred line. *When in doubt, eliminate the numbers in favor of the sentences numbered.*

It is one thing to *check* to see whether or not a sequence of lines is a deduction, i.e. to determine which inference rules if any justify each line in a sequence of pairs of the form $\langle\Gamma, \mathrm{A}\rangle$. Although sometimes tedious and sometimes challenging, such checking can always be carried out in a routine fashion: there is an effective procedure for deciding whether or not a sequence of lines is in fact a deduction and, if so, what rules justify what lines. Some of the exercises that follow provide an opportunity for deduction-checking.

It is quite another thing, however, to *create* a deduction of a given sentence from a given set of sentences. This may require ingenuity or imagination, and is helped by *strategies* for using the rules.

EXERCISES

4.2. Use the basic rules of inference to justify the unjustified lines in the deductions in this section.

4.3. Rewrite the deductions in this section using numbers in the premiss sets instead of sentences.

4.4. Use the basic rules of inference to justify each of the lines of the following deductions. (Remember: If in doubt about a premiss set, eliminate the numbers in favor of the sentences numbered.)

(1)	$\{1\}$	(1)	$\mathrm{A} \to \mathrm{B}$
	$\{2\}$	(2)	$\mathrm{B} \to \mathrm{C}$
	$\{3\}$	(3)	A
	$\{1, 3\}$	(4)	B

	$\{1, 2, 3\}$	(5)	C
	$\{1, 2\}$	(6)	$A \to C$
(2)	$\{1\}$	(1)	$A \to B$
	$\{2\}$	(2)	$\neg B$
	$\{3\}$	(3)	$\neg\neg A$
	$\{4\}$	(4)	$\neg A$
	$\{3\}$	(5)	A
	$\{1, 3\}$	(6)	B
	$\{1, 2\}$	(7)	$\neg A$
(3)	$\{1\}$	(1)	$A \vee B$
	$\{2\}$	(2)	$\neg A$
	$\{3\}$	(3)	A
	$\{2, 3\}$	(4)	B
	$\{2\}$	(5)	$A \to B$
	$\{6\}$	(6)	B
	$\varnothing$	(7)	$B \to B$
	$\{1, 2\}$	(8)	B
(4)	$\{1\}$	(1)	$A \vee B$
	$\{2\}$	(2)	$\neg B$
	$\{3\}$	(3)	B
	$\{2, 3\}$	(4)	A
	$\{2\}$	(5)	$B \to A$
	$\{6\}$	(6)	A
	$\varnothing$	(7)	$A \to A$
	$\{1, 2\}$	(8)	A
(5)	$\{1\}$	(1)	$A \vee B$
	$\{2\}$	(2)	$A \to C$
	$\{3\}$	(3)	$B \to D$
	$\{4\}$	(4)	A
	$\{2, 4\}$	(5)	C

	{2, 4}	(6)	$C \vee D$
	{2}	(7)	$A \to (C \vee D)$
	{8}	(8)	B
	{3, 8}	(9)	D
	{3, 8}	(10)	$C \vee D$
	{3}	(11)	$B \to (C \vee D)$
	{1, 2, 3}	(12)	$C \vee D$

(6)	{1}	(1)	$\neg B \vee \neg C$
	{2}	(2)	$A \to B$
	{3}	(3)	$A \to C$
	{4}	(4)	$\neg \neg A$
	{5}	(5)	$\neg A$
	{4}	(6)	A
	{2, 4}	(7)	B
	{3, 4}	(8)	C
	{9}	(9)	$\neg B$
	{2, 9}	(10)	$\neg A$
	{2}	(11)	$\neg B \to \neg A$
	{12}	(12)	$\neg C$
	{3, 12}	(13)	$\neg A$
	{3}	(14)	$\neg C \to \neg A$
	{1, 2, 3}	(15)	$\neg A$

(7)	{1}	(1)	$\neg C \vee \neg D$
	{2}	(2)	$A \to C$
	{3}	(3)	$B \to D$
	{4}	(4)	$\neg \neg A$
	{5}	(5)	$\neg A$
	{4}	(6)	A
	{2, 4}	(7)	C
	{8}	(8)	$\neg C$
	{2, 8}	(9)	$\neg A$

{2, 8}	(10)	$\neg A \vee \neg B$
{2}	(11)	$\neg C \rightarrow (\neg A \vee \neg B)$
{12}	(12)	$\neg \neg B$
{13}	(13)	$\neg B$
{12}	(14)	B
{3, 12}	(15)	D
{16}	(16)	$\neg D$
{3, 16}	(17)	$\neg B$
{3, 16}	(18)	$\neg A \vee \neg B$
{3}	(19)	$\neg D \rightarrow (\neg A \vee \neg B)$
{1, 2, 3}	(20)	$\neg A \vee \neg B$

4.5. With regard to the last line of each deduction in exercise 4.4, say what sentence is shown to be deducible from what (minimal) set of sentences.

4.6. True or false: Deduction (4) in exercise 4.4 shows that A is deducible from $\{A \vee B, \neg B, C\}$.

4.7. As we have remarked, the system of deduction is sound: whenever a sentence is deducible from a set it is a consequence of the set. A full proof of the soundness of the system appears later in this chapter, but the groundwork has mostly been done.

The soundness of the deductive system depends ultimately on the soundness of the individual inference rules. To say that a rule

$$\frac{\Gamma_1,\ A_1\ ;\ \ldots\ ;\Gamma_n,\ A_n}{\Gamma,\ A}$$

is sound is just to say that the sentence A is a consequence of the set Γ in the rule's conclusion whenever in each of the hypotheses the premiss set Γ_i implies the corresponding sentence A_i ($i = 1, \ldots, n$).

For example, consider:

MP. $$\frac{\Gamma,\ A \rightarrow B;\ \ \Delta,\ A}{\Gamma \cup \Delta, B}$$

To show that MP is a sound rule, it is enough to argue that if (hypotheses) Γ implies $A \rightarrow B$ and Δ implies A then (conclusion) $\Gamma \cup \Delta$ implies B.

Referring to proposition 3.10 (page 108), check that each of the basic rules of inference is sound.

4. Strategies

As we have noted, creating a deduction often requires strategic reasoning. The purpose of this section is to explain some important strategies involved in the deductive process.

The PC strategy. Suppose that a conditional $A \rightarrow B$ is wanted as a sentence on a line of a deduction. To prove this conditional, introduce its antecedent A by the rule P and deduce (somehow) its consequent B. Then the conditional $A \rightarrow B$ itself may be inferred by the rule PC. Moreover, if the antecedent A appears in the premiss set for the sentence B (as it well may), it will be eliminated from the premiss set for the conditional $A \rightarrow B$ through the use of PC—so that A acts only as a temporary assumption in the deduction. Here is the strategy in outline:

	⋮		
$\{A\}$	( )	A	P
	⋮		
Γ	( )	B	[*deduced somehow*]
	⋮		
$\Gamma - \{A\}$	( )	$A \rightarrow B$	PC
	⋮	*desired conditional*	

The PC strategy is employed in almost all of the deductions in section 3. For example, in a deduction on page 142 the sentence A is introduced on line (1) because it is the antecedent of the desired conditional, $A \rightarrow A$. By a happy coincidence, A is the consequent to be deduced, so there is nothing more to be done than to infer A → A by the rule PC. This has the effect of voiding the premiss set brought down from line (1).

Another example of the PC strategy is on page 144. There we wish to obtain (A → (B → C)) → ((A → B) → (A → C)) on the last line of the deduction, with ∅ as premiss set. So the antecedent A → (B → C) of this conditional is entered on the first line of the deduction by the rule P, with a view to obtaining the consequent (A → B) → (A → C) on the next-to-last line. Since this conditional is now desired, its antecedent A → B is entered on line (2) by the rule P, with a view to obtaining its consequent A → C. Finally, since this conditional is now desired, its antecedent A is entered on line (3) by the rule P, with a view to obtaining its consequent C. In this way, the only ingenuity left is in the moves to lines (4), (5), and (6). When C is at last inferred on line (6), the rule PC is invoked three times in a row, to obtain the conditionals A → C, (A → B) → (A → C), and (A → (B → C)) → ((A → B) → (A → C)). And at each stage one of the sentences introduced by the rule P is eliminated from the premiss sets, until only ∅ remains.

The foregoing should suffice to show that the PC strategy is applicable at any point in the construction of a deduction; it is not restricted to the "main conclusion". (See lines (3) and (4) of the seven-line deduction on page 146 for another example.) We can outline the general form of the PC strategy as follows.

	⋮		
$\{A_1\}$	( )	A_1	P
⋮	⋮	⋮	⋮
$\{A_n\}$	( )	A_n	P
	⋮		
Γ	( )	A	[*deduced somehow*]
	⋮		
$\Gamma - \{A_n\}$	( )	$A_n \rightarrow A$	PC

$$\begin{array}{llll} \vdots & \vdots & \vdots & \vdots \\ \Gamma - \{A_1, \ldots, A_n\} & (\ ) & A_1 \rightarrow (\ldots (A_n \rightarrow A) \ldots) & \text{PC} \\ & \vdots & \nwarrow \textit{desired conditional} & \end{array}$$

The schema above may seem abstract, but a look at some of the deductions in the preceding pages will help make it more concrete.

The RAA strategy. In constructing a deduction of a sentence A, introduce its negation $\neg$ A by the rule P and deduce (somehow) a sentence B and its negation $\neg$ B (on separate lines). Then infer A by the rule RAA. If the sentence $\neg$ A appears in either of the premiss sets for the sentences B and $\neg$ B (as well it may), it will be eliminated from the premiss set for the sentence A through the use of RAA—so that $\neg$ A acts only as a temporary assumption in the deduction.

The RAA strategy may be outlined in the following way.

$$\begin{array}{rlll} & \vdots & & \\ \{\neg A\} & (\ ) & \neg A & \text{P} \\ & \vdots & & \\ \Gamma & (\ ) & B & [\textit{deduced somehow}] \\ & \vdots & & \\ \Delta & (\ ) & \neg B & [\textit{deduced somehow}] \\ & \vdots & & \\ (\Gamma \cup \Delta) - \{\neg A\} & (\ ) & A & \text{RAA} \\ & \vdots & \nwarrow \textit{desired sentence} & \end{array}$$

In the schema above B is inferred before $\neg$ B is; in practice, of course, either may appear first, possibly with one of them above $\neg$ A and possibly with one of them identical with $\neg$ A.

The RAA strategy is used in a number of the deductions above, often in combination with the PC strategy. For example, in the deduction on page 142 showing that $\neg\neg A \rightarrow A$ is a theorem, the antecedent of the conditional, $\neg\neg A$, is introduced on line (1) to begin a PC strategy. Since we then want to infer the consequent A, we introduce its negation, $\neg A$, by the rule P on line (2). The result, then, is two lines on which sentences appear one of which ($\neg\neg A$) is the negation of the other ($\neg A$). From these lines by RAA the sentence A is inferred on line (3), and $\neg A$ is removed from the union of the premiss sets on the preceding lines, so that only $\{\neg\neg A\}$ remains. This ends the RAA strategy, and the deduction closes with a use of PC, completing the PC strategy with which it was begun.

Notice that in using the RAA strategy to infer a sentence A, the conflicting sentences can be A and $\neg A$. In other words, in order to deduce A from a set Γ, take its negation $\neg A$ as a new premiss and somehow deduce A itself. Then infer A again, this time by RAA from the conflicting pair A and $\neg A$, eliminating $\neg A$ from the premiss set. This is the *special* RAA *strategy*; compare the consequentia mirabilis propositions 3.8 (above, page 107) and 4.20 (below, page 191). In outline:

	⋮		
$\{\neg A\}$	( )	$\neg A$	P
	⋮		
Γ	( )	A	[*deduced somehow*]
	⋮		
$\Gamma - \{\neg A\}$	( )	A	RAA
	⋮	↖ *desired sentence*	

It is important to recognize that, like the PC strategy, the RAA strategy can be employed at any point in a deduction. For example, consider the deduction on page 145 showing that $A \vee \neg A$ is deducible from the empty set. Here the sentence $A \vee \neg A$ is desired on the last line of a deduction, with $\varnothing$ as premiss set. So the deduction begins with the introduction of the negation $\neg(A \vee \neg A)$ by the rule P, with a view to deducing some sentence and its negation—namely, $A \vee \neg A$ on line (5), and $\neg(A \vee \neg A)$ on line (1) (already!—note that this is a use of the special RAA strategy). But there is another use of the RAA strategy in the middle of the deduction. To wit, because of a desire to obtain the sentence A, the negation $\neg A$ is introduced by the rule P on line (2), again with a view to deducing some sentence and its negation (in this case $A \vee \neg A$ on line (3) and $\neg (A \vee \neg A)$ on line (1)).

***The* DIL *strategy*.** This is useful when a deduction contains a disjunction, $A \vee B$. Suppose a sentence C is desired on a later line. Then deduce two conditionals, $A \rightarrow C$ and $B \rightarrow C$. (One or both may of course already appear in the deduction; otherwise, the PC strategy may be helpful.) Then by DIL the desired sentence C may be inferred. In outline:

	⋮		
Γ	( )	$A \vee B$	[*given somehow*]
	⋮		
Δ	( )	$A \rightarrow C$	[*deduced somehow*]
	⋮		
E	( )	$B \rightarrow C$	[*deduced somehow*]
$\Gamma \cup \Delta \cup E$	( )	C	DIL
	⋮	↖ *desired sentence*	

***Order of strategies*.** In general, *apply the* PC *strategy before the* DIL *strategy*, and *apply both the* PC *and* DIL *strategies before* the RAA *strategy*. Thus in case the desired sentence is a conditional $A \rightarrow B$, do not begin by introducing its negation $\neg (A \rightarrow B)$ by the rule P in order to deduce some sentence and its negation. It will usually be simpler and more direct to introduce the antecedent A by the rule P in order to deduce the consequent B, and then use PC. And in order to infer B it may be profitable to take its negation $\neg B$ as still another premiss, by the rule P, with an eye to obtaining some sentence and its negation, and then use RAA. Similar remarks apply to the DIL strategy. Here is a deduction that illustrates the precedence of strategies:

1	{1}	(1)	$A \vee B$	P
2	{2}	(2)	$\neg A$	P

	{3}	(3)	A
	{2, 3}	(4)	B
3	{2}	(5)	$A \to B$
	{6}	(6)	B
3	$\varnothing$	(7)	$B \to B$
2	{1, 2}	(8)	B
1	{1}	(9)	$\neg A \to B$

The deduction shows that $\neg A \to B$ is deducible from $A \vee B$. The numbering on the left in boldface indicates the order of development of the deduction: At **1** we register the assumption $(A \vee B)$ on the first line and the desired conclusion $(\neg A \to B)$ with appropriate premiss set on the last line. Since the desired conclusion is a conditional, at **2** we assume the antecedent $(\neg A)$ on line (2) and enter its consequent (B) on the next-to-last line. The presence of a disjunction on line (1) suggests a dilemma strategy, so at **3** we write the conditionals $A \to B$ (line (5)) and $B \to B$ (line (7)). The rest of the deduction uses the PC strategy (and an "unplanned" RAA) to deduce the conditionals.

It should be stressed that not every use of PC, DIL, or RAA is the result of using one of these strategies. Sometimes it just happens (as in the deductions just above, and others, e.g. on pages 142 and 143).

Creating deductions. The boldface numbering next to the deduction above also serves to emphasize that although the *order of checking* a deduction is "top down", the *order of development in creating* a deduction is usually "outside in". Thus the first lines (**1**) to be set down contain the given sentences $A \vee B$ and $\neg A \to B$ (lines (1) and (9)); the second lines (**2**) are those suggested by a PC strategy (lines (2) and (8)); the third lines (**3**) are those suggested by a DIL strategy (lines (5) and (7)). Of course the actual numbers of the lines at the bottom emerge only when the deduction is complete. For another example of "outside in" development, here again is the deduction on page 144 showing that $(A \to (B \to C)) \to ((A \to B) \to (A \to C))$ is a theorem:

2	{1}	(1)	$A \to (B \to C)$	P
3	{2}	(2)	$A \to B$	P
4	{3}	(3)	A	P
5	{1, 3}	(4)	$B \to C$	MP, 1, 3

6	{2, 3}	(5)	B	MP, 2, 3
4	{1, 2, 3}	(6)	C	MP, 4, 5
3	{1, 2}	(7)	$A \to C$	PC, 6
2	{1}	(8)	$(A \to B) \to (A \to C)$	PC, 7
1	$\varnothing$	(9)	$(A \to (B \to C)) \to ((A \to B) \to (A \to C))$	PC, 8

EXERCISES

4.8. Identify the major (PC, RAA, DIL) strategies and order of development in the deductions in section 3. Do the same for the deductions in exercise 4.4.

4.9. Here again are the valid arguments in exercise 3.11 (page 77). For each one describe a deduction of the conclusion from the premisses. (Yes, this is as easy as it seems.)

(1) $A \;/\; A$

(2)	$B \;/\; A \to B$	(8)	$A \;/\; A \vee B$
(3)	$B,\ \neg B \;/\; A$	(9)	$B \;/\; A \vee B$
(4)	$A \to B,\ A \;/\; B$	(10)	$A \vee B,\ A \to C,\ B \to C \;/\; C$
(5)	$A,\ B \;/\; A \wedge B$	(11)	$A \to B,\ B \to A \;/\; A \leftrightarrow B$
(6)	$A \wedge B \;/\; A$	(12)	$A \leftrightarrow B \;/\; A \to B$
(7)	$A \wedge B \;/\; B$	(13)	$A \leftrightarrow B \;/\; B \to A$

4.10. Here again are the valid sentences listed in exercise 3.12 (page 77). Note that (1)–(3), (5), (6), (8)–(10), (12), (14), and (20) were shown to be theorems (and hence valid) in section 3. Show the validity of the remaining sentences—(4), (7), (11), (13), and (15)–(19)—by deducing each from the empty set. (Readers may wish to begin by recreating the deductions for the theorems in section 3—without peeking, in so far as is possible.)

(1) $A \to A$

(2) $\neg\neg A \to A$

(3) $A \vee \neg A$

(4) $\neg (A \wedge \neg A)$

(5) $A \leftrightarrow A$

(6) $\neg (A \leftrightarrow \neg A)$

(7) $(\neg A \rightarrow A) \rightarrow A$

(8) $A \rightarrow (B \rightarrow A)$

(9) $\neg A \rightarrow (A \rightarrow B)$

(10) $(A \wedge \neg A) \rightarrow B$

(11) $B \rightarrow (A \vee \neg A)$

(12) $A \rightarrow (B \rightarrow (A \wedge B))$

(13) $A \rightarrow (B \rightarrow (A \leftrightarrow B))$

(14) $(A \rightarrow B) \rightarrow ((A \vee B) \rightarrow B)$

(15) $(\neg B \rightarrow \neg A) \rightarrow (A \rightarrow B)$

(16) $(\neg A \rightarrow \neg B) \rightarrow ((\neg A \rightarrow B) \rightarrow A)$

(17) $((A \rightarrow B) \rightarrow A) \rightarrow A$

(18) $((A \rightarrow B) \rightarrow C) \rightarrow (B \rightarrow C)$

(19) $(A \rightarrow B) \rightarrow (\neg B \rightarrow (A \rightarrow C))$

(20) $(A \rightarrow (B \rightarrow C)) \rightarrow ((A \rightarrow B) \rightarrow (A \rightarrow C))$

5. Deductive equivalence and consistency

As noted, deducibility is the basic deductive logical notion, the formal counterpart to the semantic notion of consequence, while theoremhood corresponds to validity. In this section we define the two companion ideas of deductive equivalence and consistency, intended to be the formal counterparts to the semantic notions of equivalence and satisfiability. As such, our choice of definitions below is guided by the semantic metatheorems characterizing equivalence and satisfiability in terms of consequence, much in the way that the definition 4.3 of theoremhood reflects part (1) of proposition 1.7 (page 25). Readers should compare definition 4.5 of deductive equivalence with proposition

1.9 (page 27), and definition 4.6 of consistency with proposition 3.12 (page 110).

Because the deductive system is sound, it provides a means of proving that sentences are equivalent and that sets of sentences are unsatisfiable.

The special symbolism for deductive equivalence ($A \sim B$) and consistency (Con Γ) was mentioned earlier in the chapter. Here are the definitions:

DEFINITION 4.5. *Deductive equivalence.*

$$A \sim B \text{ iff both } A \vdash B \text{ and } B \vdash A.$$

DEFINITION 4.6. *Consistency (and inconsistency).*

- Con Γ iff there is no sentence A such that both $\Gamma \vdash A$ and $\Gamma \vdash \neg A$.

- $\mathrm{C}\not{o}\mathrm{n}\ \Gamma$ iff there is a sentence A such that both $\Gamma \vdash A$ and $\Gamma \vdash \neg A$.

Deductive equivalence is just interdeducibility: sentences are deductively equivalent just in case each can be deduced from the other. It is often convenient to combine into a single deduction the two deductions showing deductive equivalence. For example, consider the following deduction.

{1}	(1)	$(A \wedge B) \rightarrow C$	{8}	(8)	$A \rightarrow (B \rightarrow C)$
{2}	(2)	A	{9}	(9)	$A \wedge B$
{3}	(3)	B	{9}	(10)	A
{2, 3}	(4)	$A \wedge B$	{8, 9}	(11)	$B \rightarrow C$
{1, 2, 3}	(5)	C	{9}	(12)	B
{1, 2}	(6)	$B \rightarrow C$	{8, 9}	(13)	C
{1}	(7)	$A \rightarrow (B \rightarrow C)$	{8}	(14)	$(A \wedge B) \rightarrow C$

Lines (1)–(7) above show that $(A \wedge B) \rightarrow C$ proves $A \rightarrow (B \rightarrow C)$, while (8)–(14) show that $A \rightarrow (B \rightarrow C)$ proves $(A \wedge B) \rightarrow C$. Therefore, the sentences are deductively equivalent. Notice that lines (8)-(14) really form a separate deduction, and so the longer deduction could be replaced by two separate ones. Whether to make one or two is a matter of choice.

Exercises at the end of the section will provide an opportunity to show deductive equivalences.

As in the cases of deducibility and theoremhood, one cannot "deduce" that a sentence A is *not* deductively equivalent to a sentence B (i.e. that A $\not\dashv\vdash$ B). For to say that A is not deductively equivalent to B is to say that either A does not prove B or B does not prove A, and neither of these is something that can be established by means of a deduction.

Likewise, one cannot "deduce" that a set of sentences Γ is consistent. For to say that Γ is consistent is to say that there exist no deductions ending with lines of the forms $\langle \Gamma', A \rangle$ and $\langle \Gamma'', \neg A \rangle$, for any subsets Γ′ and Γ″ of Γ—again, something that cannot be demonstrated by means of a deduction.

But deductions can be used to show that a set of sentences Γ is *in*consistent (C∅n Γ). For to show that C∅n Γ it is sufficient to establish the existence of deductions ending with lines of the forms $\langle \Gamma', A \rangle$ and $\langle \Gamma'', \neg A \rangle$, for some sentence A, where Γ′ and Γ″ are each subsets of Γ. Let us illustrate this by showing that the set $\{A, A \rightarrow B, A \rightarrow \neg B\}$ is inconsistent.

{1}	(1)	A		{1}	(1)	A
{2}	(2)	$A \rightarrow B$		{2}	(2)	$A \rightarrow \neg B$
{1, 2}	(3)	B		{1, 2}	(3)	$\neg B$

The deduction on the left shows that B is deducible from $\{A, A \rightarrow B, A \rightarrow \neg B\}$. The one on the right shows that the same set proves $\neg$ B. Therefore, $\{A, A \rightarrow B, A \rightarrow \neg B\}$ is inconsistent.

Because any initial segment of a deduction is itself a deduction, it is possible and usually more convenient to show that a set of sentences is inconsistent by the existence of a single deduction. Thus the set of sentences $\{A, A \rightarrow B, A \rightarrow \neg B\}$ in the preceding illustration is shown to be inconsistent by the following single deduction.

{1}	(1)	A
{2}	(2)	$A \rightarrow B$
{3}	(3)	$A \rightarrow \neg B$
{1, 2}	(4)	B

$\{1, 3\}$ (5) $\neg B$

Lines (1)-(4) show that $\{A, A \rightarrow B, A \rightarrow \neg B\}$ proves B; lines (1)-(5) show that the set proves $\neg$ B. So, again, $\{A, A \rightarrow B, A \rightarrow \neg B\}$ is inconsistent.

It should be emphasized that the sentences A and $\neg$ A deducible from a set Γ which show Γ to be inconsistent may either or both be in Γ already. For example, consider:

$\{1\}$ (1) $A \rightarrow \neg A$
$\{2\}$ (2) A
$\{1, 2\}$ (3) $\neg A$

The deduction consisting of lines (1) and (2) above shows that $\{A \rightarrow \neg A, A\}$ proves A (notice that A is a member of $\{A \rightarrow \neg A, A\}$); and the deduction consisting of lines (1)-(3) shows that $\{A \rightarrow \neg A, A\}$ proves $\neg$ A. Therefore, $\{A \rightarrow \neg A, A\}$ is inconsistent.

When a singleton set of sentences $\{A\}$ is consistent we may say simply that the sentence A itself is consistent and write: Con A. Similarly for inconsistency of singleton sets.

Some more strategies. We conclude this section with a few further remarks about techniques for creating deductions, attention to which may enhance readers' skills.

The first hint concerns *double negations*. Often a sentence of the form $\neg\neg A$ occurs on a line in a deduction, where A is desired on a later line. The strategy for deducing A is contained in the deduction in section 3 of $\neg\neg A \rightarrow A$ from the empty set (page 142, also page 148). In outline, it goes like this:

.
.
.

Γ ( ) $\neg\neg A$ [*deduced somehow*]

$$
\begin{array}{llll}
 & \vdots & & \\
\{\neg A\} & () & \neg A & \text{P} \\
 & \vdots & & \\
\Gamma - \{\neg A\} & () & A & \text{RAA} \\
 & \vdots & \nwarrow \textit{desired sentence} &
\end{array}
$$

That is to say, $\neg \neg A$ occurs on a line of a deduction (with some set Γ as premiss set). In order to deduce A on a later line, the sentence $\neg A$ is introduced by the rule P. This immediately provides the hypotheses for a use of the rule RAA—specifically, to infer the sentence A. Notice that $\neg A$ disappears from the premiss set for A, which is $(\Gamma \cup \{\neg A\}) - \{\neg A\}$, i.e. $\Gamma - \{\neg A\}$ (as indicated above).

Other examples of applications of this strategy for eliminating double negations appear in a deduction on page 147, in parts (1), (6), and (7) of exercise 4.4, and in this deduction of $\neg(A \wedge \neg A)$ from the empty set (from exercise 4.10):

$$
\begin{array}{lll}
\{1\} & (1) & \neg\neg(A \wedge \neg A) \\
\{2\} & (2) & \neg(A \wedge \neg A) \\
\{1\} & (3) & A \wedge \neg A \\
\{1\} & (4) & A \\
\{1\} & (5) & \neg A \\
\varnothing & (6) & \neg(A \wedge \neg A)
\end{array}
$$

As has doubtless been observed, the use of the strategy for eliminating double negations often arises in connection with an application of the RAA strategy—to wit, when, for the sake of an RAA strategy, the negation introduced on a line of a deduction turns out to be a double negation. A number of instances of this occur in exercises at the end of the section.

Secondly, there is a technique for handling *negated disjunctions*. When a sentence of the form $\neg(A \vee B)$ occurs on a line of a deduction and the negation of one of the disjuncts—let us say $\neg A$—is desired on a later line, its deduction may be outlined as follows.

$$
\begin{array}{rcll}
 & \vdots & & \\
\Gamma & (\ ) & \neg(A \vee B) & [\textit{deduced somehow}] \\
 & \vdots & & \\
\{\neg\neg A\} & (\ ) & \neg\neg A & \text{P} \\
\{\neg A\} & (\ ) & \neg A & \text{P} \\
\{\neg\neg A\} & (\ ) & A & \text{RAA} \\
\{\neg\neg A\} & (\ ) & A \vee B & \text{ADD} \\
\Gamma - \{\neg\neg A\} & (\ ) & \neg A & \text{RAA} \\
 & \vdots & \nwarrow \textit{desired sentence} &
\end{array}
$$

Essentially, the idea above is to obtain A on a line of the deduction, then $A \vee B$ by ADD, and then the desired $\neg A$ by RAA. But the dominant strategy above is RAA: In order to obtain $\neg A$, its negation $\neg\neg A$ was introduced by the rule P (whence A is obtained via the strategy for reducing double negations).

A strategy for deducing $\neg B$ from $\neg(A \vee B)$ can also be described. But it is symmetric with the one above and so is probably obvious.

One use of this strategy for dealing with negated disjunctions appears in the following deduction showing that $A \vee \neg A$ is a theorem (but compare the shorter deduction on page 145).

$$
\begin{array}{ccl}
\{1\} & (1) & \neg(A \vee \neg A) \\
\{2\} & (2) & \neg\neg A \\
\{3\} & (3) & \neg A \\
\{2\} & (4) & A \\
\{2\} & (5) & A \vee \neg A \\
\{1\} & (6) & \neg A \\
\{1\} & (7) & A \vee \neg A \\
\varnothing & (8) & A \vee \neg A
\end{array}
$$

Strategies for deducing A and $\neg B$ from a *negated conditional* $\neg(A \rightarrow B)$ can also be described. They will be discovered in the course of constructing deductions in the exercises that follow.

It should be apparent by now that more rules of inference are required if our deductive system is to be of practical value. In section 7 we introduce a number of further rules. As a result, deductions are easier to create. In section 8 we present a rule of replacement, which completes our stock of rules of inference for the language of sentential logic.

EXERCISES

4.11. Justify the lines of the deductions in this section.

4.12. Find the uses of the strategy for eliminating double negations in the deduction showing that $\neg (A \leftrightarrow \neg A)$ is a theorem (page 147), and in parts (1), (6), and (7) of exercise 4.4.

4.13. Find the use of the strategy for dealing with a negated disjunction in the last deduction in the section (showing that $A \vee \neg A$ is a theorem).

4.14. Outline a strategy for deducing $\neg B$ from $\neg (A \vee B)$.

4.15. Here again are the valid sentences listed in exercise 3.22 (page 89). By means of deductions demonstrate that each is a theorem.

(1) $((A \rightarrow B) \wedge (B \rightarrow C)) \rightarrow (A \rightarrow C)$

(2) $(A \rightarrow \neg A) \rightarrow \neg A$

(3) $(A \rightarrow B) \rightarrow (\neg B \rightarrow \neg A)$

(4) $(A \rightarrow B) \rightarrow ((A \rightarrow \neg B) \rightarrow \neg A)$

(5) $A \rightarrow \neg \neg A$

(6) $A \rightarrow (\neg B \rightarrow \neg (A \rightarrow B))$

(7) $\neg (A \rightarrow B) \rightarrow A$

(8) $\neg (A \rightarrow B) \rightarrow \neg B$

(9) $A \vee (A \rightarrow B)$

(10) $(A \rightarrow B) \vee \neg B$

(11) $(A \to B) \lor (B \to A)$

(12) $(A \to B) \lor (B \to C)$

(13) $(A \to B) \lor (C \to A)$

(14) $\neg A \to (\neg B \to (A \leftrightarrow B))$

(15) $(A \lor B) \lor (A \leftrightarrow B)$

4.16. Here again are the pairs listed in exercise 3.13 (page 78). For (2)–(7) and (15)–(18) describe deductions showing deductive equivalence (and hence equivalence).

(1) $A,\ \neg\neg A$

(2) $A,\ A \land A$

(3) $A,\ A \lor A$

(4) $A \land B,\ B \land A$

(5) $A \lor B,\ B \lor A$

(6) $A \leftrightarrow B,\ B \leftrightarrow A$

(7) $A \land (B \land C),\ (A \land B) \land C$

(8) $A \lor (B \lor C),\ (A \lor B) \lor C$

(9) $A \leftrightarrow (B \leftrightarrow C),\ (A \leftrightarrow B) \leftrightarrow C$

(10) $\neg(A \land B),\ \neg A \lor \neg B$

(11) $\neg(A \lor B),\ \neg A \land \neg B$

(12) $A \land (B \lor C),\ (A \land B) \lor (A \land C)$

(13) $A \lor (B \land C),\ (A \lor B) \land (A \lor C)$

(14) $A \to B,\ \neg B \to \neg A$

(15) $A \to B,\ \neg(A \land \neg B)$

(16) $A \to B,\ \neg A \lor B$

(17) $A \leftrightarrow B,\ \neg A \leftrightarrow \neg B$

(18) $A \leftrightarrow B,\ (A \rightarrow B) \wedge (B \rightarrow A)$

(19) $A \leftrightarrow B,\ (A \wedge B) \vee (\neg A \wedge \neg B)$

(20) $(A \wedge B) \rightarrow C,\ A \rightarrow (B \rightarrow C)$

4.17. Here again are the sets in exercise 3.24 (page 91). Describe deductions showing that they are inconsistent and therefore unsatisfiable.

(1) $\{B,\ \neg (A \rightarrow B)\}$

(2) $\{A \rightarrow B,\ A, \neg B\}$

(3) $\{A \wedge B,\ \neg A\}$

(4) $\{A,\ \neg (A \vee B)\}$

(5) $\{A \vee B,\ A \rightarrow C,\ B \rightarrow C,\ \neg C\}$

(6) $\{A \leftrightarrow B,\ \neg (A \rightarrow B)\}$

(7) $\{\neg (A \leftrightarrow B),\ \neg A \leftrightarrow \neg B\}$

4.18. Here again are the unsatisfiable sets in exercise 3.14 (page 79). Describe deductions showing that the sets are inconsistent.

(1) $\{A,\ \neg A\}$

(2) $\{A \wedge \neg A\}$

(3) $\{A \wedge \neg A,\ B\}$

(4) $\{A \vee B,\ \neg A,\ \neg B\}$

(5) $\{A \leftrightarrow B,\ A \leftrightarrow \neg B\}$

(6) $\{A \leftrightarrow \neg A\}$

4.19. Here again are the arguments in exercise 3.20 (page 88). Describe deductions showing the conclusions to be deducible from—and thus consequences of—the premisses. (No peeking.)

(1) $A \rightarrow B,\ B \rightarrow C\ /\ A \rightarrow C$

(2) $A \rightarrow B,\ \neg B\ /\ \neg A$

(3) $A \vee B,\ \neg A\ /\ B$

(4) $A \vee B,\ \neg B\ /\ A$

(5) $A \vee B,\ A \rightarrow C,\ B \rightarrow D\ /\ C \vee D$

(6) $\neg B \vee \neg C,\ A \rightarrow B,\ A \rightarrow C\ /\ \neg A$

(7) $\neg C \vee \neg D,\ A \rightarrow C,\ B \rightarrow D\ /\ \neg A \vee \neg B$

4.20. Here again are the arguments in exercise 3.21 (page 89). Describe deductions showing them to be valid.

(1) $A \rightarrow B,\ A \rightarrow \neg B\ /\ \neg A$

(2) $A \rightarrow B,\ \neg A \rightarrow B\ /\ B$

(3) $A \leftrightarrow B,\ B \leftrightarrow C\ /\ A \leftrightarrow C$

6. Soundness of the deductive system

When a sentence A is deducible from a set of sentences Γ in our deductive system ($\Gamma \vdash A$), we may assert that A is also a consequence of Γ ($\Gamma \models A$). This is to say that the deductive system is *sound*. Indeed, the deductive system would be rather pointless if it allowed for deductions of sentences that were not consequences of their premiss sets.

In this section we demonstrate that our system of deduction does not let us go wrong, i.e. that

$$\text{if } \Gamma \vdash A \text{ then } \Gamma \models A.$$

The argument for soundness reduces to showing that each line of a deduction is itself sound, in the sense that the sentence on the line is a consequence of its premiss set. In other words, we wish to prove this:

Lemma. For each line $\langle \Gamma, A \rangle$ in a deduction, $\Gamma \models A$.

Our argument for this lemma is inductive. We show (for the *basis* of the induction) that the first line of any deduction is sound, and then (for the *induction* step) that any further line is sound if all the lines that precede it are sound. As we see below, the reasoning in each case boils down to the fact that each rule of inference preserves consequence, which is what we proved in proposition 3.10 (page 108).

For the basis, note that the first line of a deduction is always simply of the form $\langle \{A\}, A\rangle$, justified by the rule P. But part (1) of proposition 3.10, or proposition 1.2 (page 14), tells us that every sentence implies itself, i.e. that A is a consequence of $\{A\}$. So the lemma holds in this case.

Now for the induction. As an *inductive hypothesis* we assume that up to some given line (n) of a deduction the sentence on each line is a consequence of its premiss set. Then we argue by cases that no matter which rule of inference justifies line (n), the sentence on this line is a consequence of its premiss set.

Case 1. Line (n) is of the form $\langle \{A\}, A\rangle$ and is justified by the rule P. Then as we saw in the argument for the basis, $\{A\}$ implies A.

Case 2. Line (n) has the form $\langle \Gamma - \{A\}, A \rightarrow B\rangle$ and comes from an earlier line $\langle \Gamma, B\rangle$ by the rule PC. From the inductive hypothesis, B is a consequence of Γ. So by proposition 3.3 (page 101; or part (2) of 3.10), $A \rightarrow B$ is a consequence of $\Gamma - \{A\}$.

Case 3. Line (n) is $\langle (\Gamma \cup \Delta) - \{\neg A\}, A\rangle$ and comes from two earlier lines $\langle \Gamma, B\rangle$ and $\langle \Delta, \neg B\rangle$ by the rule RAA. From the inductive hypothesis, it follows that Γ implies B and Δ implies $\neg$ B. So by proposition 3.9 (page 107; or part (3) of 3.10), $(\Gamma \cup \Delta) - \{\neg A\}$ implies A.

Case 4. Line (n) is $\langle \Gamma \cup \Delta, B\rangle$ and comes from two earlier lines $\langle \Gamma, A \rightarrow B\rangle$ and $\langle \Delta, A\rangle$ by the rule MP. From the inductive hypothesis, $A \rightarrow B$ is a consequence of Γ, and A is a consequence of Δ. So by proposition 3.4 (page 101; or part (4) of 3.10), B is a consequence of $\Gamma \cup \Delta$.

Cases 5–10—in which line (n) comes by the rules CONJ, SIMP, ADD, DIL, BI, and BE. The arguments are similar to those above and use proposition 3.10, parts (5)–(13). Exercises.

This concludes our proof of the lemma. We have shown that in every line of a deduction the sentence on the line is a consequence of its premiss set. Note that in particular the sentence on the last line of a deduction is always a consequence of its premiss set.

PROPOSITION 4.7. *Soundness*. The deductive system for sentential logic is sound. That is, if $\Gamma \vdash A$ then $\Gamma \models A$.

Proof. Assume that A is deducible from Γ. Then A lies on the last line of some deduction with a subset Γ' of Γ as premiss set; i.e. there is a deduction that ends with a line $\langle \Gamma', A \rangle$, where Γ' is a subset of Γ (of course Γ' may be Γ itself). So by the lemma, A is a consequence of Γ'. Hence by augmentation, A is a consequence of Γ. □

Proposition 4.7 has three important corollaries, which we register (together with the proposition itself) as follows.

PROPOSITION 4.8. *Soundness and corollaries.*

(1) If $\Gamma \vdash A$ then $\Gamma \models A$.

(2) If $\vdash A$ then $\models A$.

(3) If $A \sim B$ then $A \simeq B$.

(4) If Sat Γ then Con Γ.

Proof. Parts (2), (3), and (4) are corollaries to part (1).

For (2). From the soundness theorem it follows that if A is deducible from $\varnothing$, then A is also a consequence of $\varnothing$. So by the definition of theoremhood (4.3) and proposition 1.7 (definability of validity in terms of consequence, page 25), if A is a theorem A is valid.

For (3). Exercise.

For (4). Note that this is equivalent, contrapositively, to: Every inconsistent set of sentences is unsatisfiable. So suppose that Γ is inconsistent. This means that Γ proves some sentence A and its negation $\neg A$. Hence by soundness, both these sentences are a consequence of Γ. Clearly, then, Γ is unsatisfiable □

EXERCISES

4.21. Complete the proof of the lemma for proposition 4.7 (cases 5–10).

4.22. Complete the proof of proposition 4.8 (part (3)).

4.23. Show that part (1) of proposition 4.8 follows from part (4). (Challenging, at this stage. Use proposition 4.30, part (1).)

7. Further rules of inference

In this section we introduce seven additional rules of inference. Three are alternate versions of the basic rules of inference PC, RAA, and DIL. The other four are new: HS (hypothetical syllogism); MT (modus tollens); DS (disjunctive syllogism); and CDIL (complex dilemma).

Each of the additional rules exemplifies a familiar pattern of inference, one that can already be accounted for in terms of the ten basic rules of inference. This is to say that each of the new rules is *derivable*: given a deduction containing lines corresponding to the hypotheses of the rule, the conclusion can be deduced by means of the basic rules of inference alone.

That the seven new rules are derivable from the basic rules is the subject of section 10. Just now it is more important to become acquainted with the rules and their uses in creating deductions.

DEFINITION 4.9. *Derived rules of inference for sentential logic.*

PC (2). $$\frac{\Gamma,\ \neg A}{\Gamma - \{A\},\ A \rightarrow B}$$

RAA (2). $$\frac{\Gamma,\ B; \quad \Delta,\ \neg B}{(\Gamma \cup \Delta) - \{A\},\ \neg A}$$

HS. $$\frac{\Gamma,\ A \rightarrow B; \quad \Delta,\ B \rightarrow C}{\Gamma \cup \Delta,\ A \rightarrow C}$$

MT. $$\frac{\Gamma,\ A \rightarrow B; \quad \Delta,\ \neg B}{\Gamma \cup \Delta,\ \neg A}$$

DS. $\dfrac{\Gamma,\ A \vee B;\ \ \Delta,\ \neg A}{\Gamma \cup \Delta,\ B}$ $\qquad$ $\dfrac{\Gamma,\ A \vee B;\ \ \Delta,\ \neg B}{\Gamma \cup \Delta,\ A}$

DIL (2). $\dfrac{\Gamma,\ \neg B \vee \neg C;\ \ \Delta,\ A \rightarrow B;\ \ E,\ A \rightarrow C}{\Gamma \cup \Delta \cup E,\ \neg A}$

CDIL (1). $\dfrac{\Gamma,\ A \vee B;\ \ \Delta,\ A \rightarrow C;\ \ E,\ B \rightarrow D}{\Gamma \cup \Delta \cup E,\ C \vee D}$

CDIL (2). $\dfrac{\Gamma,\ \neg C \vee \neg D;\ \ \Delta,\ A \rightarrow C;\ \ E,\ B \rightarrow D}{\Gamma \cup \Delta \cup E,\ \neg A \vee \neg B}$

Remarks are in order on PC (2) and RAA (2); the meanings of the rest of the new rules should be apparent.

The second version of PC permits the introduction of a conditional $A \rightarrow B$ on a line of a deduction in case the negation of the antecedent, $\neg A$, occurs on an earlier line. The construction of the premiss set of the inferred line is the same as in the original version of PC: the antecedent A of the conditional is deleted from the premiss set of the hypothesis. (In practice it will rarely happen that the premiss set of the hypothesis contains the antecedent, and so most deletions will occur vacuously in this version of PC. We have stated the rule this way for the sake of uniformity and, it is hoped, memorability.)

The rule PC (2) effectively shortens many deductions. For example, to show that $\neg A \rightarrow (A \rightarrow B)$ is a theorem, the following deduction suffices.

$\{1\}$	(1)	$\neg A$	
$\{1\}$	(2)	$A \rightarrow B$	PC, 1
$\varnothing$	(3)	$\neg A \rightarrow (A \rightarrow B)$	

Compare this with the five-line deduction of $\neg A \rightarrow (A \rightarrow B)$ on page 143.

According to the second version of RAA, a line of the form $\langle (\Gamma \cup \Delta) - \{A\}, \neg A\rangle$ may appear in a deduction if it contains earlier lines of the forms $\langle \Gamma, B\rangle$

and $\langle \Delta, \neg B \rangle$. Thus the content of the inferred line in this version of RAA is the reverse of that in the original version. The use of this new version of RAA also serves to reduce the length of deductions. For example, consider the following deduction, which shows that the sentence $(A \to B) \to (\neg B \to \neg A)$ is a theorem.

$\{1\}$	(1)	$A \to B$	
$\{2\}$	(2)	$\neg B$	
$\{3\}$	(3)	A	
$\{1, 3\}$	(4)	B	
$\{1, 2\}$	(5)	$\neg A$	RAA, 2, 4
$\{1\}$	(6)	$\neg B \to \neg A$	
$\varnothing$	(7)	$(A \to B) \to (\neg B \to \neg A)$	

In the deduction above, line (5) is inferred from lines (2) and (4) by means of RAA (2). The premiss set of line (5) is accordingly $\{1, 2, 3\} - \{3\}$ (3 being the line number of the negate A of the inferred sentence $\neg A$), i.e. $\{1, 2\}$. Compare this deduction with the solution to part (2) of exercise 4.15; that deduction has at least nine lines.

The second version of the rule RAA yields another RAA strategy. To deduce $\neg A$, introduce its negate A by the rule P and deduce (somehow) a sentence B and its negation $\neg B$. Then infer $\neg A$ by the new RAA rule and delete A from the premiss set if it appears. In outline:

	⋮		
$\{A\}$	( )	A	P
	⋮		
Γ	( )	B	[*deduced somehow*]

⋮

Δ	( )	$\neg$ B	*[deduced somehow]*
	⋮		
$(\Gamma \cup \Delta) - \{A\}$	( )	$\neg$ A	RAA
	⋮	↖ *desired sentence*	

Of course A, B, and $\neg$ B may appear in any order, and one of B and $\neg$ B may be identical with A. For example, in the deduction above showing that (A $\rightarrow$ B) $\rightarrow$ ($\neg$ B $\rightarrow$ $\neg$ A) is a theorem, after A $\rightarrow$ B and $\neg$ B are introduced on lines (1) and (2) (with an eye to PC strategies), the sentence A is introduced on line (3) for an RAA strategy. The strategy succeeds when, on line (4), the sentence B appears, for then lines (2) and (4) contain a sentence and its negation; and then the desired sentence $\neg$ A is inferred on line (5) by RAA—and the influence of the sentence A on line (3) ends as it is deleted from the premiss set of line (5). Two applications of PC complete the deduction.

EXERCISES

4.24. Justify the unjustified lines in the two deductions beginning on page 173.

4.25. Use derived (as well as basic) rules to prove the validity of the following sentences (from exercise 3.22, page 89, and exercise 4.15).

(1) $((A \rightarrow B) \wedge (B \rightarrow C)) \rightarrow (A \rightarrow C)$

(2) $(A \rightarrow \neg A) \rightarrow \neg A$

(3) $(A \rightarrow B) \rightarrow (\neg B \rightarrow \neg A)$

(4) $(A \rightarrow B) \rightarrow ((A \rightarrow \neg B) \rightarrow \neg A)$

(5) $A \rightarrow \neg \neg A$

(6) $A \rightarrow (\neg B \rightarrow \neg (A \rightarrow B))$

(7) $\neg (A \rightarrow B) \rightarrow A$

(8) $\neg (A \rightarrow B) \rightarrow \neg B$

(9) $A \vee (A \rightarrow B)$

(10) $(A \rightarrow B) \vee \neg B$

(11) $(A \rightarrow B) \vee (B \rightarrow A)$

(12) $(A \rightarrow B) \vee (B \rightarrow C)$

(13) $(A \rightarrow B) \vee (C \rightarrow A)$

(14) $\neg A \rightarrow (\neg B \rightarrow (A \leftrightarrow B))$

(15) $(A \vee B) \vee (A \leftrightarrow B)$

4.26. Describe deductions showing equivalence for each of pairs (1), (8), (10)–(14), and (19)–(20) in exercise 4.16 (also in exercise 3.13, page 78).

8. Replacement

In section 7 we enlarged the roster of rules of inference for the language of sentential logic. In this section we introduce yet another rule—the *rule of replacement*. The following *replacement theorem*, the deductive analogue of the semantic replacement theorem (proposition 3.14, page 117) in chapter 3, prepares the way.

PROPOSITION 4.10. *Replacement.* Let A and A′ be sentences alike except that A contains a subsentence B in one or more places where A′ contains a subsentence B′. Then $A \sim A'$ if $B \sim B'$.

Proof. Because the proof of this proposition involves mathematical induction, like the one for proposition 3.14, we omit most of it. The proof depends, however, on the following lemma (compare the lemma on page 118), which does not itself use mathematical induction.

Lemma. Suppose A and B are deductively equivalent. Then the following pairs are also deductively equivalent.

(i) $\neg A, \neg B$

(ii)	$A \wedge C,\ B \wedge C$	(vi)	$A \rightarrow C,\ B \rightarrow C$
(iii)	$C \wedge A,\ C \wedge B$	(vii)	$C \rightarrow A,\ C \rightarrow B$
(iv)	$A \vee C,\ B \vee C$	(viii)	$A \leftrightarrow C,\ B \leftrightarrow C$
(v)	$C \vee A,\ C \vee B$	(ix)	$C \leftrightarrow A,\ C \leftrightarrow B$

The clauses of the lemma cover all the possible ways equivalent sentences can occur as subsentences of otherwise alike sentences (except the case in which the subsentence is identical with the sentence itself—which is trivial).

We prove clauses (i), (ii), (iv), (vi), and (viii) of the lemma, and leave the rest as exercises. We suppose throughout that A and B are deductively equivalent, which means that there exist two deductions ending with these lines:

$$\{A\} \quad (\,) \quad B \qquad\qquad \{B\} \quad (\,) \quad A$$

For (i). By the rule P, continue the deductions in these ways:

$$\{\neg B\} \quad (\,) \quad \neg B \qquad\qquad \{\neg A\} \quad (\,) \quad \neg A$$

Then by RAA on each we have:

$$\{\neg B\} \quad (\,) \quad \neg A \qquad\qquad \{\neg A\} \quad (\,) \quad \neg B$$

Deductions ending with these lines show that $\neg B$ proves $\neg A$ and that $\neg A$ proves $\neg B$; and this means that $\neg A$ and $\neg B$ are deductively equivalent, as we wished to show.

For (ii). By the rule PC, continue the initial two deductions thus:

$$\varnothing \quad (\,) \quad A \rightarrow B \qquad\qquad \varnothing \quad (\,) \quad B \rightarrow A$$

Then by P and two applications of SIMP add the following lines.

$\{A \wedge C\}$	$(\,)$	$A \wedge C$	$\{B \wedge C\}$	$(\,)$	$B \wedge C$
$\{A \wedge C\}$	$(\,)$	A	$\{B \wedge C\}$	$(\,)$	B
$\{A \wedge C\}$	$(\,)$	C	$\{B \wedge C\}$	$(\,)$	C

By MP from the first and third lines of the continuation, we obtain the lines:

$$\{A \wedge C\} \quad (\,) \quad B \qquad\qquad \{B \wedge C\} \quad (\,) \quad A$$

And finally, by CONJ, we have:

$$\{A \wedge C\} \quad (\,) \quad B \wedge C \qquad\qquad \{B \wedge C\} \quad (\,) \quad A \wedge C$$

Deductions ending with these lines show that $A \wedge C$ and $B \wedge C$ are deductively equivalent.

For (iv). Once more by PC, continue the initial two deductions thus:

$$\varnothing \quad (\,) \quad A \rightarrow B \qquad\qquad \varnothing \quad (\,) \quad B \rightarrow A$$

Then by P and PC add the following lines.

$$\{C\} \quad (\,) \quad C \qquad\qquad \{C\} \quad (\,) \quad C$$
$$\varnothing \quad (\,) \quad C \rightarrow C \qquad\qquad \varnothing \quad (\,) \quad C \rightarrow C$$

Next, by P, add these lines:

$$\{A \vee C\} \quad (\,) \quad A \vee C \qquad\qquad \{B \vee C\} \quad (\,) \quad B \vee C$$

Then by the rule CDIL the following lines may be introduced.

$$\{A \vee C\} \quad (\,) \quad B \vee C \qquad\qquad \{B \vee C\} \quad (\,) \quad A \vee C$$

These lines show that $A \vee C$ is deductively equivalent to $B \vee C$.

For (vi). Once again, continue the two initial deductions thus, by the rule PC:

$$\varnothing \quad (\,) \quad A \rightarrow B \qquad\qquad \varnothing \quad (\,) \quad B \rightarrow A$$

Then by the rule P add the lines:

$$\{B \rightarrow C\} \quad (\,) \quad B \rightarrow C \qquad\qquad \{A \rightarrow C\} \quad (\,) \quad A \rightarrow C$$

The rule HS then entitles us to infer the lines:

$$\{B \rightarrow C\} \quad (\,) \quad A \rightarrow C \qquad\qquad \{A \rightarrow C\} \quad (\,) \quad B \rightarrow C$$

Since these lines end deductions, $A \rightarrow C$ and $B \rightarrow C$ are deductively equivalent.

For (viii). For this demonstration, let us concatenate the two initial deductions—schematically:

$$\begin{array}{lll} & \vdots & \\ \{A\} & () & B \\ & \vdots & \\ \{B\} & () & A \end{array}$$

Two applications of the rule PC extend this deduction thus:

$$\begin{array}{lll} \varnothing & () & A \to B \\ \varnothing & () & B \to A \end{array}$$

Add to this deduction the following line, by the rule P.

$$\begin{array}{lll} \{A \leftrightarrow C\} & () & A \leftrightarrow C \end{array}$$

Two applications of the rule BE to this last line give:

$$\begin{array}{lll} \{A \leftrightarrow C\} & () & A \to C \\ \{A \leftrightarrow C\} & () & C \to A \end{array}$$

A couple of uses of HS yield these two lines:

$$\begin{array}{lll} \{A \leftrightarrow C\} & () & B \to C \\ \{A \leftrightarrow C\} & () & C \to B \end{array}$$

And from these last two lines, by BI:

$$\begin{array}{lll} \{A \leftrightarrow C\} & () & B \leftrightarrow C \end{array}$$

This line shows that $A \leftrightarrow C$ proves $B \leftrightarrow C$. For the reverse, continue the deduction above thus, by the rule P:

$$\begin{array}{lll} \{B \leftrightarrow C\} & () & B \leftrightarrow C \end{array}$$

This yields the following two lines, by BE.

$$\begin{array}{lll} \{B \leftrightarrow C\} & () & B \to C \\ \{B \leftrightarrow C\} & () & C \to B \end{array}$$

Again, a couple of applications of HS give the lines:

$\{B \leftrightarrow C\}$ ( ) $A \rightarrow C$

$\{B \leftrightarrow C\}$ ( ) $C \rightarrow A$

Finally, BI delivers the line:

$\{B \leftrightarrow C\}$ ( ) $A \leftrightarrow C$

This shows that $B \leftrightarrow C$ proves $A \leftrightarrow C$. Therefore, $A \leftrightarrow C$ is deductively equivalent to $B \leftrightarrow C$, as we wished to show.

This completes the proofs of the selected five clauses of the lemma for proposition 4.10. So we may consider the proof of the whole complete. □

The theorem itself can be illustrated by the same pair of sentences used to illustrate the semantic version of the proposition:

(a) $(A \wedge (C \vee D)) \leftrightarrow (B \rightarrow (C \vee D))$

(b) $(A \wedge (C \vee D)) \leftrightarrow (B \rightarrow \neg (\neg C \wedge \neg D))$

That is to say, given that $C \vee D$ and $\neg (\neg C \wedge \neg D)$ are deductively equivalent—something left to be verified—it follows that the sentences (a) and (b) are deductively equivalent.

Let us move now to a formulation of a *rule of replacement*, REP. The rule itself consists of twenty pairs of equivalences grouped under nine headings: DN (double negation); EXP (exportation, sometimes called importation); RED (redundancy, sometimes called tautology); DEM (DeMorgan's Laws); DIST (distributivity); COM (commutativity); ASSOC (associativity); COND (conditionality); and BICOND (biconditionality).

DEFINITION 4.11. *Rule of replacement.*

REP. $$\frac{\Gamma,\ X}{\Gamma,\ X'}$$

where X and X′ are alike except that X contains a subsentence Y in one or more places where X′ contains a subsentence Y′, and Y and Y′ are drawn from the following pairs.

DN. $A,\ \neg\neg A$

EXP. $(A \wedge B) \rightarrow C,\ A \rightarrow (B \rightarrow C)$

RED. $A,\ A \wedge A$
$A,\ A \vee A$

DEM. $\neg(A \wedge B),\ \neg A \vee \neg B$
$\neg(A \vee B),\ \neg A \wedge \neg B$

DIST. $A \wedge (B \vee C),\ (A \wedge B) \vee (A \wedge C)$
$A \vee (B \wedge C),\ (A \vee B) \wedge (A \vee C)$

COM. $A \wedge B,\ B \wedge A$
$A \vee B,\ B \vee A$
$A \leftrightarrow B,\ B \leftrightarrow A$

ASSOC. $A \wedge (B \wedge C),\ (A \wedge B) \wedge C$
$A \vee (B \vee C),\ (A \vee B) \vee C$
$A \leftrightarrow (B \leftrightarrow C),\ (A \leftrightarrow B) \leftrightarrow C$

COND. $A \rightarrow B,\ \neg B \rightarrow \neg A$
$A \rightarrow B,\ \neg(A \wedge \neg B)$
$A \rightarrow B,\ \neg A \vee B$

	$A \leftrightarrow B,\ \neg A \leftrightarrow \neg B$
BICOND.	$A \leftrightarrow B,\ (A \rightarrow B) \wedge (B \rightarrow A)$
	$A \leftrightarrow B,\ (A \wedge B) \vee (\neg A \wedge \neg B)$

That this rule of replacement is sound depends—in light of proposition 4.10—entirely on whether the sentences in each of the listed pairs are deductively equivalent. That this is so has mostly been shown in exercises 4.16 and 4.26. (The only exception to this statement is the proof of the associativity of the biconditional, part (9) in exercise 4.16. Caution: deductions for this are long, complicated, and tedious.)

Use of the replacement rules RED and DIST is illustrated in the following deduction, which shows that A is a consequence of $A \vee (A \wedge B)$.

{1}	(1)	$A \vee (A \wedge B)$	
{1}	(2)	$(A \vee A) \wedge (A \vee B)$	DIST, 1
{1}	(3)	$A \vee A$	
{1}	(4)	A	RED, 3

In the deduction above each use of a rule of replacement involves application to the whole sentence on the line—not to a proper subsentence. In the following deduction, which shows that $A \wedge B$ is implied by $\neg (\neg A \vee \neg B)$, notice the uses of DN.

{1}	(1)	$\neg (\neg A \vee \neg B)$	
{1}	(2)	$\neg \neg A \wedge \neg \neg B$	DEM, 1
{1}	(3)	$A \wedge \neg \neg B$	DN, 2
{1}	(4)	$A \wedge B$	DN, 3

It should be noted that it is not permissible to infer line (4), above, directly from line (2) by two uses of DN at once. Such a move is not licensed by the statement of the rule of replacement. On the other hand, the inference from line (1) to line (2) in the simple deduction below is allowed by replacement (DN).

{1}	(1)	$\neg \neg A \vee (B \rightarrow \neg \neg \neg A)$	
{1}	(2)	$A \vee (B \rightarrow \neg A)$	DN, 1

The sentence $\neg\neg A$ occurs above twice as a subsentence of the sentence on line (1). Thus the rule DN permits replacement at both its occurrences by the equivalent sentence A; hence line (2).

The following deduction, showing that $\neg A \to B$ implies $A \vee B$, also illustrates a use of DN that cannot be circumvented.

{1}	(1)	$\neg A \to B$	
{1}	(2)	$\neg\neg A \vee B$	COND, 1
{1}	(3)	$A \vee B$	DN, 2

The point here is that the sentence on line (3) is not directly inferable from that on line (1) by the rule COND.

Finally, here is a deduction that shows that $\neg(A \wedge \neg B)$ proves $\neg A \vee B$:

{1}	(1)	$\neg(A \wedge \neg B)$	
{1}	(2)	$A \to B$	COND, 1
{1}	(3)	$\neg A \vee B$	COND, 2

The point in this case is that although each of $\neg(A \wedge \neg B)$ and $\neg A \vee B$ is given as equivalent to $A \to B$ by the rule COND, they are not given as directly equivalent to each other—and so line (3) is not directly inferable from line (1).

The exercises that follow provide an opportunity to employ the rules of replacement in the construction of deductions.

EXERCISES

4.27. Prove clauses (iii), (v), (vii), and (ix) of the lemma on page 176 for the replacement theorem.

4.28. Justify the unjustified lines of the deductions in this section.

4.29. Prove the equivalences below (from exercise 3.23, page 90) by showing that the sentences in each pair are deductively equivalent.

(1) A, A

(2) A, $A \wedge (A \vee B)$

(3) $A, \; A \vee (A \wedge B)$

(4) $A, \; A \wedge (B \vee \neg B)$

(5) $A, \; A \vee (B \wedge \neg B)$

(6) $A, \; (A \wedge B) \vee (A \wedge \neg B)$

(7) $A, \; (A \vee B) \wedge (A \vee \neg B)$

(8) $A, \; \neg A \rightarrow A$

(9) $A, \; (B \vee \neg B) \rightarrow A$

(10) $A, \; B \leftrightarrow (A \leftrightarrow B)$

(11) $\neg A, \; A \rightarrow \neg A$

(12) $\neg A, \; A \rightarrow (B \wedge \neg B)$

(13) $\neg A, \; \neg B \leftrightarrow (A \leftrightarrow B)$

(14) $A \wedge B, \; \neg(\neg A \vee \neg B)$

(15) $A \wedge B, \; \neg(A \rightarrow \neg B)$

(16) $A \wedge B, \; A \leftrightarrow (A \rightarrow B)$

(17) $A \wedge B, \; (A \leftrightarrow B) \leftrightarrow (A \vee B)$

(18) $A \vee B, \; \neg(\neg A \wedge \neg B)$

(19) $A \vee B, \; \neg A \rightarrow B$

(20) $A \vee B, \; B \leftrightarrow (A \rightarrow B)$

(21) $A \vee B, \; (A \leftrightarrow B) \leftrightarrow (A \wedge B)$

(22) $A \rightarrow B, \; A \leftrightarrow (A \wedge B)$

(23) $A \rightarrow B, \; B \leftrightarrow (A \vee B)$

(24) $\neg(A \leftrightarrow B), \; \neg A \leftrightarrow B$

(25) $\neg(A \leftrightarrow B), \; A \leftrightarrow \neg B$

(26) $(A \rightarrow B) \rightarrow B, \; (B \rightarrow A) \rightarrow A$

(27) $A \vee (B \rightarrow C),\ B \rightarrow (A \vee C)$

(28) $A \rightarrow (B \vee C),\ (A \rightarrow B) \vee C$

(29) $A \rightarrow (B \rightarrow C),\ B \rightarrow (A \rightarrow C)$

(30) $A \rightarrow (B \wedge C),\ (A \rightarrow B) \wedge (A \rightarrow C)$

(31) $(A \vee B) \rightarrow C,\ (A \rightarrow C) \wedge (B \rightarrow C)$

(32) $(A \vee B) \wedge C,\ (A \wedge C) \vee (B \wedge C)$

(33) $(A \wedge B) \vee C,\ (A \vee C) \wedge (B \vee C)$

(34) $(A \vee B) \wedge (C \vee D), ((A \wedge C) \vee (A \wedge D)) \vee ((B \wedge C) \vee (B \wedge D))$

(35) $(A \wedge B) \vee (C \wedge D), ((A \vee C) \wedge (A \vee D)) \wedge ((B \vee C) \wedge (B \vee D))$

4.30. Without using any rules of replacement, describe deductions showing that $C \vee D$ and $\neg(\neg C \wedge \neg D)$ are equivalent—so that, by replacement, so are:

$$(A \wedge (C \vee D)) \leftrightarrow (B \rightarrow (C \vee D)),\quad (A \wedge (C \vee D)) \leftrightarrow (B \rightarrow \neg(\neg C \wedge \neg D))$$

9. Metatheorems

Like their semantic counterparts, deducibility and the other deductive concepts satisfy principles like those for consequence—e.g. reflexivity, transitivity, and augmentation. In this section, we set out and prove some of the more significant of these. We begin with some basic facts about deductions, stated as lemmas.

Lemma 1. Deductions can be concatenated to form new deductions.

In other words, any two deductions—

$$\begin{array}{ccc} & \vdots & \\ \Gamma & (\ ) & A \end{array} \qquad\qquad \begin{array}{ccc} & \vdots & \\ \Delta & (\ ) & B \end{array}$$

—can always be "stacked up" to form a single deduction:

$$\begin{array}{ccc} & \vdots & \\ \Gamma & (\,) & A \\ & \vdots & \\ \Delta & (\,) & B \end{array}$$

The point is simply that if each line in the sequences ending $\langle \Gamma, A \rangle$ and $\langle \Delta, B \rangle$ is justified by a rule of inference so is each line in the longer sequence. We have tacitly appealed to this fact already, e.g. in proving part (viii) of the lemma in the proof of replacement (proposition 4.10).

Lemma 2. Any sentence can be added to any premiss set.

To see that this is so, let us show that to the premiss set of any line $\langle \Gamma, A \rangle$ we can add any sentence X we wish. Suppose we have a deduction containing the line $\langle \Gamma, A \rangle$:

$$\begin{array}{ccc} & \vdots & \\ \Gamma & (\,) & A \\ & \vdots & \end{array}$$

Extend this by the following three lines.

$$\begin{array}{rcll} \{X\} & (\,) & X & \qquad\qquad \text{P} \\ \Gamma \cup \{X\} & (\,) & A \wedge X & \qquad\qquad \text{CONJ} \\ \Gamma \cup \{X\} & (\,) & A & \qquad\qquad \text{SIMP} \end{array}$$

Thus we have a deduction in which X has been added to the original premiss set for A.

The moral of this lemma is that hereafter we may tacitly expand premiss sets to include additional sentences as necessary. We take advantage of this frequently in what follows.

Lemma 3. Every premiss set in a deduction is finite.

This fact was remarked earlier, and is easy to prove. First note that deductions always begin by the rule P, which admits just one sentence into the premiss set. So the first line of a deduction has a finite premiss set. Secondly, observe that if all premiss sets up to a given line (n) are finite, then the premiss set on (n) itself will be finite too: either (n) is justified by P (in which case its premiss set is finite) or by some other rule. But in the other rules the inferred line's premiss set is at most the union of premiss sets from earlier lines (possibly reduced, if the rule is PC or RAA) and will therefore be finite. In short, no matter which line of a deduction is considered, its premiss set is finite.

With these facts at hand, let us begin our survey of properties of deducibility and the other deductive logical concepts.

Recall that consequence is *compact*—whenever a sentence is a consequence of a set it is a consequence as well of a finite subset of the set—although we are still not in a position to demonstrate this. Deducibility has this property too—a sentence deducible from a set is also deducible from a finite subset—and we can prove this now.

PROPOSITION 4.12. *Compactness for deducibility.*

$$\text{If } \Gamma \vdash \text{A, then } \Gamma' \vdash \text{A for some finite subset } \Gamma' \text{ of } \Gamma.$$

Proof. Suppose A is deducible from Γ. This means that there is a deduction ending with a line $\langle \Gamma', \text{A} \rangle$, where the premiss set Γ′ is a subset of Γ. But by lemma 3, Γ′ is finite. So the existence of the line $\langle \Gamma', \text{A} \rangle$ means that A is deducible from a finite subset Γ′ of Γ. □

PROPOSITION 4.13. *Reflexivity of deducibility.*

(1) If $\text{A} \in \Gamma$ then $\Gamma \vdash \text{A}$.

(2) Corollary: $\text{A} \vdash \text{A}$.

Proof. The one-line deduction $\langle \{\text{A}\}, \text{A} \rangle$ shows that A is both deducible from itself and deducible from any set that contains it. □

PROPOSITION 4.14. *Transitivity of deducibility.*

(1) If $\Gamma \vdash B$ for every sentence $B \in \Delta$, and $\Delta \vdash A$, then $\Gamma \vdash A$.

(2) If $\Gamma \vdash B$ and $\Delta \cup \{B\} \vdash A$, then $\Gamma \cup \Delta \vdash A$.

(3) Corollary: If $A \vdash B$ and $B \vdash C$, then $A \vdash C$.

Proof. We give the argument for the corollary, and leave those for (1) and (2) as exercises. Suppose that A proves B and B proves C. This means that deductions exist ending with the lines $\langle \{A\}, B \rangle$ and $\langle \{B\}, C \rangle$. (Here we rely on lemma 2 to insure that the premiss sets actually do contain A and B.) Given lemma 1, we can form a single deduction—

	⋮	
$\{A\}$	( )	B
	⋮	
$\{B\}$	( )	C

—and then continue it with these lines:

$\varnothing$	( )	$B \to C$	PC
$\{A\}$	( )	C	MP

Thus we have a deduction ending with $\langle \{A\}, C \rangle$, which means that C is deducible from A, as we wished to show. □

PROPOSITION 4.15. *Augmentation for deducibility.*

If $\Gamma \vdash A$ and $\Gamma \subseteq \Delta$, then $\Delta \vdash A$.

Proof. Suppose that Γ proves A and $\Gamma \subseteq \Delta$. By definition, A is on the last line of a deduction with a subset of Γ, which is thus a subset of Δ. So Δ proves A too. □

As a corollary to propositions 4.14 and 4.15 we have the following counterpart to proposition 1.5 (page 18) for consequence.

PROPOSITION 4.16. *Combining and chaining.*

$$\text{If } \Gamma_1 \vdash A_1, \dots, \Gamma_n \vdash A_n, \text{ and } A_1, \dots, A_n \vdash A, \text{ then } \Gamma_1 \cup \dots \cup \Gamma_n \vdash A.$$

Proof. Exercise. □

Here is an alternative, if not very useful, definition of theoremhood (compare part (2) of proposition 1.7, page 25):

PROPOSITION 4.17. *Alternative definition of theoremhood.*

$$\vdash A \text{ iff for every set of sentences } \Gamma, \Gamma \vdash A.$$

Proof. For left-to-right, suppose that A is a theorem. This means that A is deducible from the empty set. Since the empty set is a subset of every set, it follows that A is deducible from every set of sentences. In the other direction, note that if every set of sentences proves a sentence then the empty set does, and by definition this means that the sentence is a theorem. □

Next we come to the deductive version of proposition 3.5 (page 102; corollary (4) corresponds to proposition 3.6, page 104).

PROPOSITION 4.18. *"Deduction theorem".*

(1) $\Gamma \cup \{A\} \vdash B$ iff $\Gamma \vdash A \to B$.

(2) Corollary: $A_1, \dots, A_n \vdash A$ iff $\vdash A_1 \to (\dots (A_n \to A) \dots)$.

(3) Corollary: $A \vdash B$ iff $\vdash A \to B$.

(4) Corollary: $A_1, \dots, A_n \vdash A$ iff $\vdash (A_1 \wedge \dots \wedge A_n) \to A$.

Proof. First for part (1). From left to right, assume that B is deducible from $\Gamma \cup \{A\}$. Then some deduction ends with a line of the form $\langle \Delta, B \rangle$, where Δ is a subset of $\Gamma \cup \{A\}$. By PC, this deduction can be extended by the line $\langle \Delta - \{A\}, A \rightarrow B \rangle$. Since $\Delta - \{A\}$ is a subset of $\Gamma - \{A\}$, which in turn is a subset of Γ, this means that $A \rightarrow B$ is deducible from Γ, as we wished to show. For the reverse, assume that Γ proves $A \rightarrow B$. By reflexivity, the set $\Gamma \cup \{A\}$ proves A. By MP, B is deducible from $\{A \rightarrow B, A\}$. Hence by proposition 4.16, B is deducible from $\Gamma \cup \Gamma \cup \{A\}$, i.e. from $\Gamma \cup \{A\}$.

Part (2) follows from (1) by n repetitions, and for part (3) take Γ to be the empty set in (1). Part (4) follows from (2) (and vice versa) by n applications of the rule EXP. □

PROPOSITION 4.19. *"Indirect proof"*.

(1) $\Gamma \cup \{A\} \vdash B$ iff $\Gamma \cup \{\neg B\} \vdash \neg A$.

(2) Corollary: $A \vdash B$ iff $\neg B \vdash \neg A$.

Proof. Part (2) is just the case in (1) where there are no sentences in Γ, i.e. Γ is the empty set. So we argue for (1). By the deduction theorem, $\Gamma \cup \{A\}$ proves B if and only if Γ proves $A \rightarrow B$, and $\Gamma \cup \{\neg B\}$ proves $\neg A$ if and only if Γ proves $\neg B \rightarrow \neg A$. But by COND $A \rightarrow B$ proves $\neg B \rightarrow \neg A$, and vice versa. Part (1) then follows by transitivity. □

We have seen already, in connection with the special RAA strategy (page 156), that a sentence is deducible from a set if it is deducible from the set plus its own negation. The reverse is also the case, and we record this important fact here under the same rubric as its semantic counterpart (proposition 3.8, page 107):

PROPOSITION 4.20. *Consequentia mirabilis.*

(1) $\Gamma \vdash A$ iff $\Gamma \cup \{\neg A\} \vdash A$.

(2) Equivalently: $\Gamma \vdash A$ iff $\Gamma - \{\neg A\} \vdash A$.

(3) Corollary: $\vdash A$ iff $\neg A \vdash A$.

Proof. The left-to-right of (1) is by augmentation, as is the right-to-left of (2). The other directions are left as an exercise. For the corollary, take Γ to be the empty set. □

The corollary in proposition 4.20 is yet another way of defining theoremhood in terms of deducibility (compare proposition 3.8, page 107).

Next we see how deducibility might have been defined had we begun with theoremhood as our basic idea.

PROPOSITION 4.21. *Definability of deducibility in terms of theoremhood.*

$\Gamma \vdash A$ iff Γ contains sentences $A_1, \ldots, A_n$ such that $\vdash (A_1 \wedge \ldots \wedge A_n) \rightarrow A$.

Proof. Exercise. □

Similarly, here is a way of defining deductive equivalence in terms of theoremhood (compare proposition 3.11, page 109):

PROPOSITION 4.22. *Definability of deductive equivalence in terms of theoremhood.*

$A \sim B$ iff $\vdash A \leftrightarrow B$.

Proof. Exercise (use proposition 4.18). □

Like its semantic counterpart, deductive equivalence is indeed an equivalence relation (compare proposition 1.10, page 28):

PROPOSITION 4.23. *Deductive equivalence an equivalence relation.*

(1) *Reflexivity*: $A \sim A$.

(2) *Symmetry*: If $A \sim B$ then $B \sim A$.

(3) *Transitivity*: If $A \sim B$ and $B \sim C$, then $A \sim C$.

Proof. Exercise. □

PROPOSITION 4.24. *Deductive equivalence of theorems.*

$$\text{If } \vdash A \text{ and } \vdash B, \text{ then } A \sim B.$$

Proof. If A and B are both theorems, and hence deducible from the empty set, then by augmentation each is deducible from the other, which means that they are deductively equivalent. (Compare proposition 1.11, page 28.) □

PROPOSITION 4.25. Con $\varnothing$.

Proof. Proposition 1.13 (page 30) tells us that the empty set is satisfiable, and a corollary to the soundness of the deductive system (proposition 4.8) tells us that every satisfiable set is consistent. Therefore the empty set is consistent. □

The following proposition shows another way we could have defined consistency (compare proposition 3.12, page 110).

PROPOSITION 4.26. *Alternative definition of consistency (inconsistency).*

(1) Con Γ iff there is a sentence A such that $\Gamma \nvdash A$.

(2) Equivalently: C$\not{o}$n Γ iff for every sentence A, $\Gamma \vdash A$.

Proof. Let us show (2). Suppose, for left-to-right, that Γ is inconsistent. This means that Γ proves both A and $\neg$ A, for some sentence A. By RAA, every sentence is deducible from $\{A, \neg A\}$. Hence by transitivity (proposition 4.14), Γ proves every sentence. Conversely, suppose that every sentence is deducible from Γ. Then, in particular, Γ proves both A and $\neg$ A, for any sentence A. This means that Γ is inconsistent, as we wished to show. □

We saw in proposition 1.15 (page 31) that satisfiable sets can be diminished satisfiably (and likewise that unsatisfiable sets can be augmented unsatisfiably). Following is the deductive analogue.

PROPOSITION 4.27. *Diminution for consistency; augmentation for inconsistency.*

(1) If Con Γ and $\Delta \subseteq \Gamma$, then Con Δ.

(2) Equivalently: If C$\not o$n Γ and $\Gamma \subseteq \Delta$, then C$\not o$n Δ.

Proof. For the (2), suppose that Γ is inconsistent and is a subset of Δ. Then Γ proves A and $\neg$ A, for some sentence A, and hence by augmentation (proposition 4.15) so does Δ. □

Like satisfiability, consistency is compact. We state the proposition in two versions:

PROPOSITION 4.28. *Compactness for consistency.*

(1) Con Γ, if Con Γ' for every finite subset Γ' of Γ.

(2) Equivalently: If C$\not o$n Γ then Γ has a finite subset Γ' such that C$\not o$n Γ'.

Proof. We show (2). Suppose that Γ is inconsistent, so that Γ proves both A and $\neg$ A for some sentence A. Then deductions exist ending with lines $\langle \Gamma', A \rangle$ and $\langle \Gamma'', \neg A \rangle$, where both Γ' and Γ'' are finite subsets of Γ. So Γ' proves A and Γ'' proves $\neg$ A, from which it follows by augmentation (proposition 4.15) that $\Gamma' \cup \Gamma''$ also proves both A and $\neg$ A. Therefore, $\Gamma' \cup \Gamma''$ is an inconsistent finite subset of Γ. □

PROPOSITION 4.29. *Deductive equivalence of inconsistent sentences.*

If $C\not{o}n$ A and $C\not{o}n$ B, then $A \sim B$.

Proof. This has a counterpart in proposition 1.16 (page 31). Exercise. □

Parts (1) and (2) of the next proposition show how deducibility and theoremhood may be defined in terms of (in)consistency. Compare proposition 3.13 (page 111).

PROPOSITION 4.30.

(1) $\Gamma \vdash A$ iff $C\not{o}n\ \Gamma \cup \{\neg A\}$.

(2) Corollary: $\vdash A$ iff $C\not{o}n\ \neg A$.

(3) $Con\ \Gamma \cup \{A\}$ iff $\Gamma \nvdash \neg A$.

(4) Corollary: Con A iff $\nvdash \neg A$.

Proof. For (1): For left-to-right, assume that A is deducible from Γ. Then, by reflexivity, both A and $\neg A$ are deducible from $\Gamma \cup \{\neg A\}$, which means that this set is inconsistent. For right-to-left, assume that $\Gamma \cup \{\neg A\}$ is inconsistent. Then by proposition 4.26, $\Gamma \cup \{\neg A\}$ proves every sentence, in particular, A. It follows from this that, given RAA, Γ itself proves A—as we wished to show. For parts (2) and (4) take Γ to be the empty set in (1) and (3). Part (3) is left as an exercise. □

EXERCISES

4.31. Prove parts (1) and (2) of proposition 4.14. The proof of (1) uses compactness, among other propositions.

4.32. Use the definition of deducibility to prove:

If $\Gamma \vdash A$ then $\Gamma \cup \Delta \vdash A$.

This is an alternative statement of the property of augmentation for deducibility (proposition 4.15). Show this by arguing that the alternative holds if proposition 4.15 does, and vice versa. (Note that Γ is always a subset of $\Gamma \cup \Delta$.)

4.33. Prove proposition 4.16.

4.34. Theorems prove only theorems; in particular, a sentence deducible from a theorem is itself a theorem.

(1) If $\Gamma \vdash$ B and $\vdash$ A for every sentence A in Γ, then $\vdash$ B.

(2) If A $\vdash$ B and $\vdash$ A, then $\vdash$ B.

Give the reasoning for part (2).

4.35. Prove parts (2) and (3) of proposition 4.18.

4.36. Complete the proof of proposition 4.20.

4.37. Prove proposition 4.21.

4.38. Prove proposition 4.22.

4.39. Prove proposition 4.23.

4.40. Explain why a consistent set proves only consistent sentences, or, equivalently, a set is inconsistent if an inconsistent sentence is deducible from it:

(1) If Con Γ and $\Gamma \vdash$ A, then Con A.

(2) Equivalently: If C$\emptyset$n A and $\Gamma \vdash$ A, then C$\emptyset$n Γ.

4.41. Complete the proof of proposition 4.27.

4.42. Here are two further statements of diminution for consistency and augmentation for inconsistency:

(1) If Con Γ then Con $\Gamma \cap \Delta$.

(2) Equivalently: If C$\not{o}$n Γ then C$\not{o}$n $\Gamma \cup \Delta$.

Argue that these are equivalent to the versions in proposition 4.27.

4.43. Prove proposition 4.29.

4.44. Explain why if sentences are deductively equivalent then one is a theorem if and only if they all are—that is:

(1) If $A \sim B$, then $\vdash A$ iff $\vdash B$.

Also explain why if sentences are deductively equivalent then one is consistent if and only if they all are—that is:

(2) If $A \sim B$, then Con A iff Con B.

Notice that (2) means as well that if sentences are deductively equivalent then one is inconsistent if and only if the other is.

4.45. Prove part (2) of proposition 4.30.

4.46. Explain why C$\not{o}$n Sn.

4.47. Prove that $\Gamma \vdash A \rightarrow B$ iff $\Gamma - \{A\} \vdash A \rightarrow B$.

4.48. Prove that $\Gamma \vdash A$ iff $\Gamma - \{\neg A\} \vdash A$.

4.49. Explain why an initial segment of a deduction is itself a deduction. (An initial segment is what results when lines are removed starting from the bottom—up to but not including the first!)

4.50. Use the result in exercise 4.49 to prove this alternative definition of deducibility: $\Gamma \vdash A$ iff there is a deduction in which A appears on some line with Γ or a subset of Γ as a premiss set.

4.51. True or false:

a. There is a sentence that is deducible from every other sentence.

b. Every sentence is deducible from some finite set of sentences.

c. If a sentence is consistent, then its negation is not consistent.

d. Every sentence is deductively equivalent to some conjunction.

e. At least 9,876,543,210 deductions end with a line justified by the rule P.

f. A sentence is a theorem if and only if in some deduction it appears on a line with the empty set as premiss set.

g. In a complete deductive system every inconsistent set of sentences is unsatisfiable.

h. Every sentence is deducible from some infinite set of sentences.

i. A disjunction is a theorem if and only if at least one of its disjuncts is a theorem.

j. An infinite number of deductions end with a line justified by the rule PC.

k. Every sentence is deductively equivalent to at least one theorem.

l. Every sentence is deducible from the empty set.

m. Both of these sentences are theorems: $(A \to B) \vee (B \to C)$ and $(D \to E) \vee (E \to F)$.

n. For every sentence A, if A is not a theorem then $\neg A$ is.

o. In a sound deductive system, a sentence is not deducible from a set of sentences unless it is a consequence of the set.

p. No deduction begins with a line justified by the rule RAA.

q. A conjunction is consistent if and only if its conjuncts are consistent.

r. If the rule P were dropped, the deductive system would no longer be complete.

s. No disjunction is deductively equivalent to any biconditional.

t. The union of consistent sets of sentences is itself a consistent set.

u. Every sentence is deducible from every infinite set of sentences.

v. A sentence is a theorem if and only if it is deducible from its own negation.

w. In a sound deductive system every valid sentence is a theorem.

x. No sentence is deducible from itself.

y. If the rule P were dropped, the deductive system would no longer be sound.

10. Derivability of the derived rules

As we remarked earlier, to say that a derived rule of inference is *derivable* means that in any deduction containing lines corresponding to hypotheses of the rule a line corresponding to the rule's conclusion can be deduced using only the basic rules of inference. We close this chapter with a proposition about the derivability of the derived rules of inference.

PROPOSITION 4.31. *Derivability of the derived rules.* Each of the rules of inference PC (2), RAA (2), HS, MT, DS, DIL (2), CDIL (1), CDIL (2), and REP is derivable using only the basic rules of inference for sentential logic.

Proof. The important parts of the demonstrations of derivability have appeared at various places earlier in the chapter, notably in the deductions in exercise 4.4 (page 149) and in section 8 on the rule of replacement. We give the reasoning for just three—PC(2), HS, and MT—and leave the others as exercises.

For PC(2): We wish to show that from $\langle \Gamma, \neg A \rangle$ one may infer $\langle \Gamma - \{A\}, A \rightarrow B \rangle$—i.e. that if a deduction contains the first line then the second can be deduced using only the basic rules of inference. So suppose we have a deduction containing the line

Γ	( )	$\neg A$

Then using only basic rules we add these lines to the deduction:

$\{A\}$	( )	A	P
$\Gamma \cup \{A\}$	( )	B	RAA
$\Gamma - \{A\}$	( )	$A \rightarrow B$	PC

The premiss set on the next-to-last line should of course be $(\Gamma \cup \{A\}) - \{\neg B\}$. But in case $\neg B$ is a member of Γ we may assume that we have restored it by the method outlined in the proof of lemma 2 on page 186. Note too that the premiss set on the last line is the same as $(\Gamma \cup \{A\}) - \{A\}$. In any case, using only basic inference rules we have arrived at a deduction of $A \rightarrow B$ from $\Gamma - \{A\}$. Hence the rule PC(2) is derivable.

For HS: Suppose we have a deduction containing lines corresponding to the hypotheses of the rule:

Γ	( )	$A \rightarrow B$
Δ	( )	$B \rightarrow C$

Then, again using only basic rules, we extend this deduction by the following lines.

$\{A\}$	( )	A	P
$\Gamma \cup \{A\}$	( )	B	MP
$\Gamma \cup \Delta \cup \{A\}$	( )	C	MP
$\Gamma \cup \Delta$	( )	$A \rightarrow C$	PC

The final premiss set should be $(\Gamma \cup \Delta) - \{A\}$. But if A is a member of either Γ or Δ we would wish to maintain it. Since we have a deduction ending with the desired line, $\langle \Gamma \cup \Delta, A \to C \rangle$, the rule HS is derivable. Compare deduction (1) in exercise 4.4.

For MT: Suppose a deduction exists that has lines corresponding to the rule's hypotheses:

Γ	( )	$A \to B$
Δ	( )	$\neg B$

Then we continue with the following lines.

$\{\neg\neg A\}$	( )	$\neg\neg A$	P
$\{\neg A\}$	( )	$\neg A$	P
$\{\neg\neg A\}$	( )	A	RAA
$\Gamma \cup \{\neg\neg A\}$	( )	B	MP
$\Gamma \cup \Delta$	( )	$\neg A$	RAA

The first use of RAA is merely to produce A on the third-to-last line. Notice that the premiss set on the last line should properly be $(\Gamma \cup \Delta) - \{\neg\neg A\}$. But we would wish to keep $\neg\neg A$ in the premiss set if Γ or Δ contains it. So the last line is, as desired, $\langle \Gamma \cup \Delta, \neg A \rangle$, and thus MT is derivable. Compare deduction (2) in exercise 4.4. □

EXERCISES

4.52. Prove the rest of proposition 4.31—i.e. that the inference rules RAA (2), DS, DIL (2), CDIL (1), and CDIL (2) are derivable using only the basic rules of inference for sentential logic.

4.53. Remove RAA from the set of basic rules and replace it by the rules MT and DN. Then show that RAA is derivable given this revised set of basic rules. Experiment similarly with alternatives to other basic rules.

5

COMPLETENESS FOR SENTENTIAL LOGIC

IN THIS chapter we prove that the deductive system for sentential logic set out in chapter 4 is *complete*, i.e. that whenever a sentence A is a consequence of a set of sentences Γ it is also deducible from the set:

$$\text{if } \Gamma \models A \text{ then } \Gamma \vdash A.$$

Note that in chapter 4 we established that the system is *sound* (proposition 4.7, page 170), i.e. that

$$\text{if } \Gamma \vdash A \text{ then } \Gamma \models A.$$

So once completeness is proved we will have shown that the semantic concept of consequence precisely matches the deductive concept of deducibility:

$$\Gamma \models A \text{ iff } \Gamma \vdash A.$$

To pave the way toward completeness, we present in section 1 an important alternative statement of this property, and we indicate briefly how our proof will proceed in terms of this equivalent proposition. Then in sections 2 and 3 we introduce the concept of a *maximal* set of sentences and prove an important result, known as *Lindenbaum's lemma*, about such sets. The completeness theorem itself appears in section 4. In section 5 we recall the idea of compactness and note an important corollary of soundness and completeness. Finally, in section 6 we return to the idea of relevance and fulfill a promise made in chapter 3, and in section 7 we touch on the idea of a minimal set of rules.

The ideas in this chapter will be useful again in chapters 9 and 10.

1. Preliminaries to completeness

Once again, completeness is the proposition that whatever is a consequence of a set of sentences is also deducible from it:

$$\text{if } \Gamma \models A \text{ then } \Gamma \vdash A.$$

Our approach to the proof of completeness is indirect, by way of the equivalent proposition that every consistent set of sentences is satisfiable—i.e.:

$$\text{if Con } \Gamma \text{ then Sat } \Gamma.$$

Put contrapositively, this means that every unsatisfiable set of sentences is inconsistent:

$$\text{if } \mathrm{S\cancel{a}t}\ \Gamma \text{ then } \mathrm{C\cancel{o}n}\ \Gamma.$$

To see that these statements of completeness are indeed equivalent, assume the first—that any consequence of a set of sentences is also deducible from the set—and let us show the second in its contrapositive guise. To this end, suppose further that Γ is an unsatisfiable set of sentences. Then by proposition 1.14 (page 30), Γ implies every sentence. Hence by our assumption, every sentence is deducible from Γ. But this means that Γ is inconsistent (by proposition 4.26, page 192). So the first statement of completeness implies the second. For the reverse, assume the second statement—that every unsatisfiable set of sentences is inconsistent. Now suppose that a sentence A is a consequence of a set Γ. By proposition 3.13 (page 111), the set $\Gamma \cup \{\neg A\}$ is unsatisfiable. From this, by our assumption, it follows that the set $\Gamma \cup \{\neg A\}$ is inconsistent. But then by proposition 4.30 (page 194; part (1)), A is deducible from Γ. So the second statement of completeness implies the first, and hence the two statements are equivalent.

We shall prove completeness (proposition 5.5) by way of two further propositions, which play the role of lemmas. The first of these (5.3) is the proposition that every consistent set of sentences is included in what is called a maximal set; the second (5.4) affirms that every maximal set of sentences is satisfiable. Given these results, we can argue simply that every consistent set of sentences is satisfiable: for if Γ is consistent, then it is a subset of a maximal set, and since the maximal set is satisfiable, so is its subset Γ.

Proposition 5.3—Lindenbaum's lemma—is proved in section 3; proposition 5.4 appears in section 4.

2. Maximality

A set of sentences is said to be *maximal* just in case, first, it is consistent, and, secondly, the addition to it of any new sentence yields an inconsistent set. Intuitively, maximal sets of sentences are consistent sets that are as full of sentences as they can be without being inconsistent.

DEFINITION 5.1. *Maximality.* Max Γ iff (i) Con Γ, and (ii) for every A, if Con $\Gamma \cup \{A\}$, then $A \in \Gamma$.

The following proposition lists a number of important properties of maximal sets of sentences.

PROPOSITION 5.2. Let Γ be a maximal set of sentences.

(1) $A \in \Gamma$ iff $\Gamma \vdash A$.

(2) If $\Delta \subseteq \Gamma$ and $\Delta \vdash A$, then $A \in \Gamma$.

(3) If $\vdash A$ then $A \in \Gamma$.

(4) $\neg A \in \Gamma$ iff $A \notin \Gamma$.

(5) $A \wedge B \in \Gamma$ iff both $A \in \Gamma$ and $B \in \Gamma$.

(6) $A \vee B \in \Gamma$ iff either $A \in \Gamma$ or $B \in \Gamma$.

(7) $A \rightarrow B \in \Gamma$ iff if $A \in \Gamma$ then $B \in \Gamma$.

(8) $A \leftrightarrow B \in \Gamma$ iff $A \in \Gamma$ if and only if $B \in \Gamma$.

Proof. We assume that Γ is a maximal set of sentences.

For (1). This says that the members of a maximal set are exactly the sentences deducible from the set. Left-to-right is just proposition 4.13 (page 187). For the other direction, assume—to reach a contradiction—that A is deducible from Γ, but is not a member of Γ. Because Γ is maximal, this means that the set $\Gamma \cup \{A\}$ is inconsistent. So by proposition 4.30 (page 194; part (2)), $\neg A$ is deducible from Γ. Hence by definition 4.6 of consistency (page 161), Γ itself is inconsistent—which contradicts the maximality of Γ.

For (2). If A is deducible from a subset Δ of Γ, then by augmentation of deducibility (proposition 4.15, page 188), A is also deducible from Γ. So it follows from part (1) that A is in Γ.

For (3). To show that a maximal set contains every theorem, suppose that a sentence A is a theorem. Then by definition 4.3 (page 138) of theoremhood, A is deducible from $\varnothing$. But then, since $\varnothing$ is a subset of Γ, A is in Γ by part (2).

For (4). Let us split this in two: (i) Not both A and $\neg$ A are in Γ. (ii) Either A or $\neg$ A is in Γ. To show (i), assume to the contrary that both A and $\neg$ A are in Γ. Then by part (1) both A and $\neg$ A are deducible from Γ, which means that Γ is inconsistent. This contradicts the maximality of Γ. To show (ii), suppose to the contrary that neither A nor $\neg$ A is in Γ—i.e. that both A and $\neg$ A are outside of Γ. By definition of maximality, then, both the set $\Gamma \cup \{A\}$ and the set $\Gamma \cup \{\neg A\}$ are inconsistent. By proposition 4.30 (page 194) it follows that the sentences $\neg$ A and A are both deducible from Γ. And from this it follows by definition that Γ is inconsistent, which again contradicts the maximality of the set.

For (5). For left-to-right, suppose that Γ contains the conjunction $A \wedge B$, so that by part (1), $A \wedge B$ is deducible from Γ. Of course, each of A and B is deducible from $A \wedge B$. So by transitivity of deducibility (proposition 4.14, 188), each of A and B is deducible from Γ. Therefore, by (1) again, A is a member of Γ and so is B. The argument for right-to-left is left as an exercise.

For (6). For left-to-right, suppose that $A \vee B$ is in Γ, so that by (1), $A \vee B$ is deducible from Γ. Now suppose that neither A nor B is in Γ. Then by part (4), both $\neg$ A and $\neg$ B are in Γ. We leave it as an exercise for the reader to reach a contradiction from here. For right-to-left, suppose that either A or B is in Γ. By (1), A is deducible from Γ or B is. But $A \vee B$ is deducible from A and also from B. So in either case $A \vee B$ is deducible from Γ, and hence, by (1), $A \vee B$ is a member of Γ.

For (7). Another way of putting the left-to-right direction is that a maximal set of sentences is closed under the rule of modus ponens (MP). To prove it, suppose that both $A \rightarrow B$ and A are members of Γ. So $\{A \rightarrow B, A\}$ is a subset of Γ. Since B is deducible from $\{A \rightarrow B, A\}$, B is deducible from Γ as well—which by (1) means that B is in Γ. The reasoning from right to left is left as an exercise.

For (8). This too is left as an exercise. Note, however, that $A \leftrightarrow B$ is deductively equivalent to (i.e. interdeducible with) $(A \rightarrow B) \wedge (B \rightarrow A)$.

This concludes the proof of proposition 5.2. □

EXERCISES

5.1. Complete the proof of proposition 5.2 (parts (5)–(8)).

5.2. *An alternative definition of maximality.* Definition 5.1 characterizes a maximal set as a consistent set of sentences that contains every sentence that is consistent with it (or, contrapositively, that becomes inconsistent upon the addition of a sentence not already in it). An equivalent way of defining maximality characterizes maximal sets as those sets of sentences that are consistent but without consistent proper extensions, i.e.: Max Γ iff (i) Con Γ, and (ii) C$\not{o}$n Δ whenever $\Gamma \subset \Delta$. Prove that this definition is just as good as 5.1.

5.3. Let Γ and Δ be maximal sets of sentences. Prove that $\Gamma = \Delta$ if and only if $\Gamma \subseteq \Delta$.

5.4. We say that a set of sentences is *negation complete* just in case it satisfies the condition that for every sentence A, either $A \in \Gamma$ or $\neg A \in \Gamma$. Prove that every consistent negation complete set of sentences is maximal.

5.5. A set of sentences is said to be *deductively closed* just in case it contains every sentence deducible from it. Prove that a set is maximal if it is deductively closed and has the property of containing the negation of a sentence A if and only if it does not contain A itself.

5.6. Show that the set of sentences true in a model $\mathcal{M}$ is maximal.

5.7. Let us say that a set of sentences is *finitely satisfiable*—Fat, for short—iff all its finite subsets are satisfiable. Suppose a Fat set Γ is also negation complete (see exercise 5.4). Prove that Γ behaves just like a maximal set with respect to the truth-functional operators. That is to say, prove:

(1) $\neg A \in \Gamma$ iff $A \notin \Gamma$.

(2) $A \wedge B \in \Gamma$ iff both $A \in \Gamma$ and $B \in \Gamma$.

(3) $A \vee B \in \Gamma$ iff either $A \in \Gamma$ or $B \in \Gamma$.

(4) $A \rightarrow B \in \Gamma$ iff if $A \in \Gamma$ then $B \in \Gamma$.

(5) $A \leftrightarrow B \in \Gamma$ iff $A \in \Gamma$ if and only if $B \in \Gamma$.

3. Lindenbaum's lemma

This important proposition states that every consistent set of sentences has a maximal extension.

PROPOSITION 5.3. *Lindenbaum's lemma.* If Con Γ, then there exists a set of sentences Δ such that (i) $\Gamma \subseteq \Delta$, and (ii) Max Δ.

Proof. The proof of Lindenbaum's lemma involves several definitions and lemmas. We begin with a set of sentences Γ which we suppose to be consistent; that is, we assume throughout that Con Γ. The plan then is to define a set of sentences Δ and prove that Δ is a maximal extension of Γ.

For the duration of the proof we assume we have a fixed enumeration of the set of sentences:

$$A_1, A_2, A_3, \ldots$$

That is, we suppose that each sentence in the language occurs at least once in the sequence $A_1, A_2, A_3, \ldots$: for each sentence A in the language, there is a positive integer n such that A is the sentence A_n in the enumeration. (It does not matter if a sentence shows up at more than one place in the sequence—only that it show up at least once.)

In terms of the set Γ and this enumeration of the sentences, we define an infinite sequence of sets of sentences:

$$\Delta_0, \Delta_1, \Delta_2, \ldots$$

The definition is inductive. First we define the initial set Δ_0 in the sequence. Then we specify how any other set in the sequence, Δ_n, is to be defined in terms of its immediate predecessor, Δ_{n-1}. Here is the definition:

Definition 1.

(i) $\Delta_0 = \Gamma$.

$$\text{(ii)}\quad \Delta_n = \begin{cases} \Delta_{n-1} \cup \{A_n\}, \text{ if Con } \Delta_{n-1} \cup \{A_n\} \\ \Delta_{n-1}, \text{ if C}\not{o}\text{n } \Delta_{n-1} \cup \{A_n\} \end{cases} \qquad n > 0.$$

In plainer language: Δ_0 is the consistent set Γ with which we began; and, for $n > 0$, the set Δ_n in the sequence is formed by adding the nth sentence A_n in the enumeration to the preceding set Δ_{n-1}, if that addition is consistent. If that addition is not consistent, then nothing is added, and Δ_n is the same as its predecessor Δ_{n-1}.

It should be obvious from the definition that the sets $\Delta_0, \Delta_1, \Delta_2, \ldots$ are all consistent: the first set in the sequence is consistent by hypothesis, and the rest are formed only by making consistent additions to their immediate predecessors. We state this fact as a lemma (and leave its proper proof as an exercise in proof by induction).

Lemma 2. Con Δ_n, for each $n \geq 0$.

Reflection on the construction of the sets $\Delta_0, \Delta_1, \Delta_2, \ldots$ also reveals that each of them includes it immediate predecessor and hence all its predecessors. Let us state this, too, as a lemma (without further proof).

Lemma 3. $\Delta_i \subseteq \Delta_j$, whenever $i \leq j \geq 0$.

We are ready now to define the set Δ advertised earlier. Intuitively, there is a "largest" set in the sequence $\Delta_0, \Delta_1, \Delta_2, \ldots$—one that contains the contents of all the rest. Actually, of course, there is no such set in the sequence. But the contents of all the sets in the sequence can be collected into a single set Δ, by constructing the (infinite) union of the sets in the sequence. That is to say, we may define the desired set Δ to be this collection.

Definition 4. $\Delta = \Delta_0 \cup \Delta_1 \cup \Delta_2 \cup \ldots$.

(A fancier, more formal definition might read: $\Delta = \bigcup_{n=0}^{\infty} \Delta_n$. But our formulation is good enough.)

Clearly, the set Δ includes each of the sets $\Delta_0, \Delta_1, \Delta_2, \ldots$. We state this as a lemma.

Lemma 5. $\Delta_n \subseteq \Delta$, for each $n \geq 0$.

We shall also need to make use of the fact that any sentence A_k that is in the set Δ is also a member of the corresponding set Δ_k in the sequence.

Lemma 6. $A_k \in \Delta_k$ whenever $A_k \in \Delta$, for each $k > 0$.

This may not be obvious, so let us prove it before we go on. Suppose the sentence A_k to be in the set Δ. Then A_k must be a member of some set Δ_n in the sequence, where $n \geq 0$. Now if $n \leq k$, we have our conclusion, since by lemma 3, Δ_n is a subset of Δ_k. So let us assume A_k is in Δ_n for $n > k$. By lemma 2, Δ_n is consistent, and by the principle of diminution for consistency (proposition 4.27, page 193) so is any subset of Δ_n. In particular, then, the subset $\Delta_{k-1} \cup \{A_k\}$ is consistent (this union *is* a subset of Δ_n). So by definition, Δ_k is this set, and hence Δ_k contains the sentence A_k.

With lemma 6 at hand, we can argue that every finite subset of the set Δ is a subset of one of the sets in the sequence $\Delta_0, \Delta_1, \Delta_2, \ldots$. The argument: Suppose Δ' to be a finite subset of Δ, and let A_n be the sentence in Δ' with largest integer index n. Now we show that Δ' is a subset of the set Δ_n. Let A be any sentence in Δ'. Since A occurs somewhere in the enumeration $A_1, A_2, A_3, \ldots$, A is A_k, for some $k \leq n$. Since A ($= A_k$) is in Δ', it is in Δ. So applying lemma 6, A ($= A_k$) is in the set Δ_k. Using lemma 3, we see that Δ_k is a subset of Δ_n. Hence A is in the set Δ_n, as we wished to show. We record this fact as another lemma.

Lemma 7. For every finite subset Δ' of Δ, $\Delta' \subseteq \Delta_n$ for some $n \geq 0$.

We are ready at last to prove that Δ is a maximal extension of Γ. By lemma 5, Δ includes the set Δ_0, otherwise known as Γ:

Lemma 8. $\Gamma \subseteq \Delta$.

With Δ thus established as an extension of Γ, it remains only for us to show that Δ is maximal. We state this as our last lemma, and then prove it.

Lemma 9. Max Δ.

To say Max Δ means, of course, that:

(a) Con Δ

and

(b) For every sentence A, if Con $\Delta \cup \{A\}$, then $A \in \Delta$.

To prove (a), it is sufficient in virtue of compactness for consistency (proposition 4.28, page 193) to prove that every finite subset of Δ is consistent. So let Δ' be a finite subset of Δ. By lemma 7, Δ' is a subset of one of the sets in the sequence $\Delta_0, \Delta_1, \Delta_2, \ldots$. Since each of these sets is consistent (lemma 1) so is Δ' (by diminution for consistency, proposition 4.27). This ends the proof for part (a).

To prove (b), assume that the set $\Delta \cup \{A\}$ is consistent, for some sentence A. We wish to show that A is in Δ. Since A occurs somewhere in the enumeration $A_1, A_2, A_3, \ldots$, A is A_n, for some n. (If A occurs more than once in the enumeration, let A_n be its first occurrence.) Thus our assumption can be equally well expressed: $\Delta \cup \{A_n\}$ is consistent; and our desired conclusion can be equally well expressed: A_n is in Δ. Consider now the set Δ_n in the sequence $\Delta_0, \Delta_1, \Delta_2, \ldots$. By clause (ii) of definition 1, Δ_n is the set $\Delta_{n-1} \cup \{A_n\}$ if this set is consistent—which it is, since it is a subset of $\Delta \cup \{A_n\}$ (proposition 4.27 once again). So A_n is in Δ_n. But since Δ_n is a subset of Δ (lemma 5), we have that A_n is in Δ. This ends the proof of part (b).

With lemmas 8 and 9 thus established, the proof of Lindenbaum's lemma is complete: Every consistent set of sentences is a subset of a maximal set of sentences. Note that in view of proposition 4.27 the result can be strengthened: A set of sentences is consistent if and only if it has a maximal extension. □

It is important to remark that in the proof of Lindenbaum's lemma we appealed to various propositions which themselves depend ultimately on the rules of inference that characterize the fundamental notion of deducibility and its allied concepts—theoremhood, consistency, deductive equivalence, and maximality. Thus it should be apparent that Lindenbaum's lemma holds good for any deductive system in which the rules P, PC, RAA, MP, etc., appear (as basic or derived). The usefulness of this fact will become apparent later, in chapters 8 and 9, when we need to know that the proof of Lindenbaum's lemma can be carried out for logics stronger than the present deductive system.

EXERCISES

5.8. Prove lemmas 2 and 3 in the proof of Lindenbaum's lemma (proposition 5.3).

5.9. Prove as a corollary to Lindenbaum's lemma that a sentence is deducible from a set of sentences if (and only if) the sentence belongs to every maximal extension of the set; i.e.:

$$\Gamma \vdash A \text{ iff } A \in \Delta, \text{ for every Max } \Delta \text{ such that } \Gamma \subseteq \Delta.$$

(The "only if" half is trivial; it is the "if" part that requires Lindenbaum's lemma.) Then prove—as a corollary to the corollary—that a sentence is a theorem precisely when it is a member of every maximal set of sentences; i.e.:

$$\vdash A \text{ iff } A \in \Delta, \text{ for every Max } \Delta.$$

5.10. The (first) corollary in exercise 5.9 is in fact an alternative statement of Lindenbaum's lemma. Prove this by proving that Lindenbaum's lemma follows from the corollary.

5.11. Let $A_1, A_2, A_3, \ldots$ be an enumeration of the sentences, and let Γ be a consistent set of sentences. Define the sequence $\Delta_0, \Delta_1, \Delta_2, \ldots$ of sets of sentences thus:

(i) $\Delta_0 = \Gamma$.

$$\text{(ii)}\ \Delta_n = \begin{cases} \Delta_{n-1} \cup \{A_n\}, \text{ if } \Delta_{n-1} \vdash A_n \\ \Delta_{n-1}, \text{ otherwise} \end{cases} \qquad n > 0.$$

Show by an inductive proof that Con Δ_n for each $n \geq 0$. Could the definition above replace definition 1 in the proof of (Lindenbaum's lemma?

4. Completeness

Together with Lindenbaum's lemma, the following theorem provides the key to our proof of the completeness of sentential logic. It is a simple version of its counterpart in chapter 9, proposition 9.9, due to Leon Henkin. It asserts that every maximal set of sentences is satisfiable.

PROPOSITION 5.4. If Max Γ, then Sat Γ.

Proof. Let Γ be a maximal set of sentences. We define the model $\mathcal{M}$ in terms of the set Γ, as follows.

Definition. $\mathcal{M}(\mathbb{P}_n) = \mathsf{T}$ iff $\mathbb{P}_n \in \Gamma$, for each atomic sentence $\mathbb{P}_n$.

That is, an atomic sentence is stipulated to be true in the model $\mathcal{M}$ just in case it appears in the maximal set Γ. The model $\mathcal{M}$ is sometimes referred to as the *canonical model* with respect to Γ.

We prove now that, in general, just those sentences are true in the model $\mathcal{M}$ as are members of the set Γ. That is:

Lemma. $\models_{\mathcal{M}} A$ iff $A \in \Gamma$, for every sentence A.

This result is sometimes called the *truth lemma*. Its proof is inductive. We show first that it holds of all the simplest sentences in the language (the atomic sentences), and then that it holds for any complex (molecular) sentence given that it holds for all sentences of less complexity ("shorter" sentences, as we shall say). Actually, we shall treat only the cases in which a sentence A is (a) atomic, (b) a negation, and (c) a conditional. The other cases are left as exercises.

For (a). When A is an atomic sentence, $\mathbb{P}_n$, the result is trivial. We show it by a sequence of biconditionals:

$$\vDash_{\mathcal{M}} \mathbb{P}_n \text{ iff } \mathcal{M}(\mathbb{P}_n) = \top$$
—definition 3.2, part (1);
$$\text{iff } \mathbb{P}_n \in \Gamma$$
—definition of $\mathcal{M}$.

So the lemma holds when A is atomic.

The next two cases are inductive, so we assume as an inductive hypothesis that the lemma holds for all sentences shorter than A.

For (b). Suppose A is a negation, $\neg$ B. Since B is shorter than A, the inductive hypothesis tells us that:

$$\vDash_{\mathcal{M}} \mathrm{B} \text{ iff } \mathrm{B} \in \Gamma.$$

Again we argue via a sequence of biconditionals:

$$\vDash_{\mathcal{M}} \neg \mathrm{B} \text{ iff } \nvDash_{\mathcal{M}} \mathrm{B}$$
—definition 3.2, part (2);
$$\text{iff } \mathrm{B} \notin \Gamma$$
—inductive hypothesis;
$$\text{iff } \neg \mathrm{B} \in \Gamma$$
—proposition 5.2, part(4).

So the lemma holds when A is a negation.

For (c). Suppose A is a conditional, B $\rightarrow$ C. Since B and C are shorter than A, the inductive hypothesis tells us that:

$$\vDash_{\mathcal{M}} \mathrm{B} \text{ iff } \mathrm{B} \in \Gamma \quad \text{and} \quad \vDash_{\mathcal{M}} \mathrm{C} \text{ iff } \mathrm{C} \in \Gamma.$$

Thus:

$$\vDash_{\mathcal{M}} \mathrm{B} \rightarrow \mathrm{C} \text{ iff if } \vDash_{\mathcal{M}} \mathrm{B} \text{ then } \vDash_{\mathcal{M}} \mathrm{C}$$
—definition 3.2, part (5);
$$\text{iff if } \mathrm{B} \in \Gamma \text{ then } \mathrm{C} \in \Gamma$$
—inductive hypothesis;
$$\text{iff } \mathrm{B} \rightarrow \mathrm{C} \in \Gamma$$
—proposition 5.2, part (7).

So the lemma holds when A is a conditional, and our proof of the lemma is ended.

Thus all the sentences in the maximal set Γ are true in the model $\mathcal{M}$. It follows that Γ is satisfiable—which is what we set out to prove. The proof of proposition 5.4 is therefore complete. □

Lindenbaum's lemma and the truth lemma—propositions 5.3 and 5.4—together get us to completeness:

PROPOSITION 5.5. *Completeness.* The deductive system for sentential logic is complete. That is, if $\Gamma \models A$ then $\Gamma \vdash A$.

Proof. Recall that in section 1 we showed that this statement of completeness had an equivalent form: every consistent set of sentences is satisfiable. So we may give the proof for this proposition.

Suppose that Γ is a consistent set of sentences. Then by Lindenbaum's lemma, Γ has a maximal extension Δ. By proposition 5.4, moreover, the set Δ is satisfiable. Therefore, by diminution for satisfiability (proposition 1.15, page 31), Γ is also satisfiable. □

As in the case of soundness, we can state some corollaries to the completeness theorem.

PROPOSITION 5.6. *Completeness and corollaries.*

(1) If $\Gamma \models A$ then $\Gamma \vdash A$.

(2) If $\models A$ then $\vdash A$.

(3) If $A \simeq B$ then $A \sim B$.

(4) If Con Γ then Sat Γ.

Proof. We have already established the equivalence of parts (1) and (4). Parts (2) and (3) we leave as exercises. □

EXERCISES

5.12. Complete the proof of the lemma in the proof of proposition 5.4 by proving the inductive cases in which A is (a) a conjunction, $B \wedge C$, (b) a disjunction, $B \vee C$, and (c) a biconditional, $B \leftrightarrow C$.

5.13. Prove parts (2) and (3) of proposition 5.6.

5. Compactness, and another corollary

We learned in propositions 4.12 and 4.28 (pages 187 and 193) that compactness holds for deducibility and consistency—i.e.:

$$\text{if } \Gamma \vdash A, \text{ then } \Gamma' \vdash A \text{ for some finite subset } \Gamma' \text{ of } \Gamma,$$

and

$$\text{Con } \Gamma, \text{ if Con } \Gamma' \text{ for every finite subset } \Gamma' \text{ of } \Gamma.$$

In virtue of our soundness and completeness results, we have a corresponding pair of compactness results for the semantic concepts of consequence and satisfiability:

PROPOSITION 5.7. *Compactness for consequence and satisfiability in sentential logic.*

(1) If $\Gamma \models A$, then $\Gamma' \models A$ for some finite subset Γ' of Γ.

(2) Sat Γ, if Sat Γ' for every finite subset Γ' of Γ.

Proof. By soundness and completeness, consequence and deducibility are the same relation, and likewise satisfiability and consistency hold of exactly the same sentences. So (1) and (2) follow at once from the results cited in the paragraph above. □

As we have seen, soundness and completeness come in more than one guise; indeed, they come in several. The foregoing results about compactness enable us to recognize some further expressions of soundness and completeness, which we register as follows.

PROPOSITION 5.8. Each of statements (2)–(4) in the *Soundness* column of the table below is equivalent to the statement (1) of soundness itself and hence to each of the others. In like fashion, each of statements (2)–(4) in the *Completeness* column is equivalent to the statement (1) of completeness and hence to each of the others. Finally, therefore, each statement in the *Both* column expresses a statement of soundness and completeness.

	Soundness	*Completeness*	*Both*
(1)	$\Gamma \vdash A \Rightarrow \Gamma \models A$	$\Gamma \models A \Rightarrow \Gamma \vdash A$	$\Gamma \models A$ iff $\Gamma \vdash A$
(2)	$\vdash A \Rightarrow \models A$	$\models A \Rightarrow \vdash A$	$\models A$ iff $\vdash A$
(3)	$A \sim B \Rightarrow A \simeq B$	$A \simeq B \Rightarrow A \sim B$	$A \simeq B$ iff $A \sim B$
(4)	Sat $\Gamma \Rightarrow$ Con Γ	Con $\Gamma \Rightarrow$ Sat Γ	Con Γ iff Sat Γ

Proof. We leave it for the reader as an exercise to work out the implications and necessary arguments. Compactness will be useful, along with various propositions in chapters 3 and 4. □

EXERCISES

5.14. Prove proposition 5.8.

5.15. Recall Fat (finitely satisfiable) and negation complete sets of sentences from exercises 5.4 and 5.7. Given a Fat negation complete set Γ, define a model $\mathcal{M}$ by the condition that in it truth is assigned just to the atomic sentences in Γ. Show that the sentences true in $\mathcal{M}$ are exactly those in Γ. If the argument is inductive, prove the proposition for the atomic, negation, and disjunction cases.

6. Relevance, and decidability

We come now to the statement and proof promised in chapter 3 of *relevance* for the language of sentential logic. An important corollary to this is the *decidability* of sentential logic.

PROPOSITION 5.9. *Relevance for sentential logic.* The only thing relevant to the truth value of a sentence in a model are the atoms of the sentence. In other words, models agree on the truth value of a sentence A whenever they agree on the truth values they assign to the atoms of A. Likewise, they agree on all the sentences in a set Γ if they agree on the atoms of Γ.

Proof. Let A be a sentence of sentential logic, and suppose $\mathcal{M}$ and $\mathcal{M}^*$ are models that give the same truth value to each atomic subsentence $\mathbb{P}_n$ of A. Then what we wish to show is that $\mathcal{M}$ and $\mathcal{M}^*$ agree on the truth value of A itself:

$$\models_{\mathcal{M}} \mathrm{A} \text{ iff } \models_{\mathcal{M}^*} \mathrm{A}.$$

The proof is by induction on A, and we give it for the cases in which A is (a) atomic, (b) a negation, and (c) a conditional, and leave the others as exercises.

For (a). If A is an atomic sentence $\mathbb{P}_n$, the result is trivial:

$$\begin{aligned} \models_{\mathcal{M}} \mathbb{P}_n \text{ iff } & \mathcal{M}(\mathbb{P}_n) = \mathsf{T} \\ & \text{—definition 3.2, part (1);} \\ \text{iff } & \mathcal{M}^*(\mathbb{P}_n) = \mathsf{T} \\ & \text{—assumption;} \\ \text{iff } & \models_{\mathcal{M}^*} \mathbb{P}_n \\ & \text{—definition 3.2, part (1).} \end{aligned}$$

Now we assume as an inductive hypothesis that the proposition holds for all sentences shorter than A.

For (b). Suppose A is a negation, $\neg$ B. Since B is shorter than A, and B's atoms are the same as those of A, i.e. of $\neg$ B, the inductive hypothesis tells us that:

$$\models_{\mathcal{M}} B \text{ iff } \models_{\mathcal{M}^*} B.$$

Hence:

$$\begin{aligned}
\models_{\mathcal{M}} \neg B \text{ iff } & \not\models_{\mathcal{M}} B \\
& \quad\text{—definition 3.2, part (2);} \\
\text{iff } & \not\models_{\mathcal{M}^*} B \\
& \quad\text{—inductive hypothesis;} \\
\text{iff } & \models_{\mathcal{M}^*} \neg B \\
& \quad\text{—definition 3.2, part (2).}
\end{aligned}$$

For (c). Suppose A is a conditional, $B \rightarrow C$. Since B and C are less complex than A, and their atoms are among those of A, i.e. of $B \rightarrow C$, the inductive hypothesis tells us that:

$$\models_{\mathcal{M}} B \text{ iff } \models_{\mathcal{M}^*} B \quad \text{and} \quad \models_{\mathcal{M}} C \text{ iff } \models_{\mathcal{M}^*} C.$$

Thus:

$$\begin{aligned}
\models_{\mathcal{M}} B \rightarrow C \text{ iff } & \text{if } \models_{\mathcal{M}} B \text{ then } \models_{\mathcal{M}} C \\
& \quad\text{—definition 3.2, part (5);} \\
\text{iff } & \text{if } \models_{\mathcal{M}^*} B \text{ then } \models_{\mathcal{M}^*} C \\
& \quad\text{—inductive hypothesis;} \\
\text{iff } & \models_{\mathcal{M}^*} B \rightarrow C \\
& \quad\text{—definition 3.2, part (5).}
\end{aligned}$$

This completes the proof. □

As we saw in chapter 3, in virtue of relevance we can always restrict models to sentences or to sets of sentences. Thus $\mathcal{M}$ will be a model restricted to A or Γ just in case $\mathcal{M}$ is an assignment of truth values to the atoms of A or Γ, and our actual conception (in definition 3.1, page 61) is the limiting case of the set Sn of all sentences.

One implication of this is that logical questions about finite numbers of sentences are always *decidable*: for the method of truth tables provides, as we have seen, a *yes* or *no* answer in principle to questions about validity and equivalence, and to those about consequence and satisfiability where only finitely many sentences are involved. We state this as the final theorem of the chapter.

PROPOSITION 5.10. *Decidability in sentential logic.* Every question of validity, equivalence, finite consequence, and finite satisfiability in sentential logic is decidable.

So, for example, to decide whether or not a sentence A is a consequence of sentences $A_1, \ldots, A_n$, it is enough to evaluate these sentences on a truth table that enumerates the finitely many relevant models (finitely many since the number of atoms is finite). Since construction of the truth table is a finite, effective matter, as is the evaluation, a *yes* or *no* answer will always be forthcoming.

EXERCISES

5.16. Give the proof of proposition 5.9 for the inductive cases in which A is (a) a conjunction, (b) a disjunction, and (c) a biconditional.

5.17. Use relevance (proposition 5.9) to prove that a set containing all the atomic subsentences of a sentence A either implies A or implies A's negation. In other words, assume that Γ contains the atoms of A and show that either $\Gamma \models A$ or $\Gamma \models \neg A$. Conclude as a corollary that if Sat $\Gamma \cup \{A\}$ then $\Gamma \models A$.

5.18. Use relevance to prove that if A and B have no atomic sentences in common, then if $A \models B$ then either S$\not{\text{a}}$t A or $\models B$. (This means that if a satisfiable sentence implies an invalid sentence, then they have at least one atomic sentence in common.) Likewise show that if A and B are satisfiable sentences that have no atomic sentences in common, then the set $\{A, B\}$ is also satisfiable.

5.19. Use the results in exercise 5.17 and completeness (proposition 5.5) to prove that either A or its negation is deducible from a set of sentences containing all A's atomic subsentences—i.e. that if Γ contains all the atomic subsentences of A, then either $\Gamma \vdash A$ or $\Gamma \vdash \neg A$. Then conclude as a corollary that if Con $\Gamma \cup \{A\}$ then $\Gamma \vdash A$.

5.20. Without appealing to the results in exercises 5.17 or 5.19, show that either A or its negation is deducible from a set of sentences containing all A's atomic subsentences. Give the proof by induction on the complexity of A. Give it for the cases in which A is (a) atomic, (b) a negation, and (c) a conditional.

11. Minimality

Our completeness proof appeals ultimately only to the basic rules of inference—P, PC, RAA, MP, CONJ, SIMP, ADD, DIL, BI, and BE. This set of rules is *minimal* for sentential logic, in the sense that if any of the rules is dropped it cannot be derived using only those that remain. In other words, the deductive system is incomplete in the absence of any one of the basic rules: some consequences will no longer be deducible and, indeed, there will be valid sentences not deducible as theorems. This is something we shall not prove.

The basic rules are not unique, however. Other minimal sets of rules exist. One such is hinted at in exercise 4.53.

PART II

PREDICATE LOGIC

6

THE LANGUAGE OF PREDICATE LOGIC

CONSIDER the argument with which we began this book:

> Every human is mortal. Sitting Bull is a human.
> Therefore, Sitting Bull is mortal.

This argument is valid—the conclusion is a consequence of the premisses. But if we translate the argument into the language of sentential logic, we get something like

$$\frac{\begin{array}{c}\mathbb{P}_0\\ \mathbb{P}_1\end{array}}{\mathbb{P}_2}$$

—which is clearly invalid (as is shown by any model that assigns $\top$ to $\mathbb{P}_0$ and $\mathbb{P}_1$ but $\bot$ to $\mathbb{P}_2$).

What has gone wrong? Perhaps a more subtle translation is needed, one that reveals the logical structure of the sentences Every human is mortal, Sitting Bull is a human, and Sitting Bull is mortal so as to make the argument valid. But within the resources of the language of sentential logic, there is no further structure to be revealed: at least on the surface, none of the three sentences contains anything corresponding to the logical operators of negation, conjunction, disjunction, conditionality, or biconditionality.

What this example shows is that sentential logic is not adequate for an analysis of consequence as it applies to the sentences in the argument above, Every human is mortal, Sitting Bull is a human, and Sitting Bull is mortal. Something more is required.

In this chapter we begin the study of *predicate logic*—the logic of predicates, terms, and quantification. The language of predicate logic is an extension of the language of sentential logic. In section 1 We introduce its *logical constants*, notably the *universal* and *existential quantifiers*. In sections 2 through 7 we discuss *translation*, proceeding from simpler to more complex forms. Then in sections 8, 9, and 10 we formulate a precise account of the syntax of the language, including several syntactic ideas that will be of use in subsequent chapters.

1. The logical operators

The logical constants for predicate logic include the five already familiar from sentential logic:

$\neg$	negation
$\wedge$	conjunction
$\vee$	disjunction
$\rightarrow$	conditionality
$\leftrightarrow$	biconditionality

In addition, there are two new logical constants—operators called *quantifiers*:

$\forall$	universal quantification
$\exists$	existential quantification

As before, the operators $\neg$, $\wedge$, $\vee$, $\rightarrow$, and $\leftrightarrow$ give rise to formulas:

$\neg A$	the negation of A
$A \wedge B$	the conjunction of A and B
$A \vee B$	the disjunction of A and B
$A \rightarrow B$	the conditional of A and B
$A \leftrightarrow B$	the biconditional of A and B

The quantifiers $\forall$ and $\exists$ go together with *variables* to make formulas of two new forms:

$\forall xA$	the universal quantification of A (with respect to the variable x)
$\exists xA$	the existential quantification of A (with respect to the variable x)

Thus each quantifier-plus-variable forms a one-place operator.

Very roughly, the quantifiers have these meanings:

$\forall$	every
$\exists$	some

The remaining expressions in the language of predicate logic are non-logical and come in two categories—*names* and *predicates*.

Notice that we speak above of formulas. This is necessary since formulas are more general structures of which sentences are a special case.

The other novelty in the language of predicate logic is the atomic sentences and atomic formulas, which deserve an account of their own.

2. Atomic sentences and formulas

Atomic sentences in the language of predicate logic differ in general from their sentential logic counterparts by having internal structure: an atomic sentence consists of a predicate plus names. For example, the sentence

Wyatt Earp is a human

may be rendered in by

$$Ha$$

—where H is a predicate translating is a human and the name a translates Wyatt Earp. Similarly, we can write

$$Ma$$

to express the proposition that Wyatt Earp is mortal.

Thus we use uppercase letters for predicates and lowercase letters for names—somewhat the reverse of the practice in English.

Predicates come in *degrees*: H and M above are one-place, or monadic, predicates. Here is an example of a sentence using a two-place (dyadic) predicate:

Tab

If T means taught and b means Doc Holliday, then the displayed sentence translates Wyatt Earp taught Doc Holliday. The order of the names following the predicate is important: Tba means that Doc Holliday taught Wyatt Earp.

Similar to names, but different in function, are *variables*. These appear, as we have seen, with quantifiers ($\forall$x, $\exists$y, . . .); they can also appear in *atomic formulas*.

All atomic sentences are also atomic formulas, but the class of atomic formulas is broader. So in addition to atomic formulas like those above—which are also sentences—we have atomic formulas such as:

Hx, My, Txb, Txy

Atomic formulas containing variables are in a certain sense meaningless, since they are not true or false relative to a model. But variables contribute systematically to the meanings of quantificational sentences, as we shall see.

The general form of an *atomic formula* is thus

$$\mathbb{P}\tau_1 \ldots \tau_n,$$

where $\mathbb{P}$ is a predicate of degree n (an n–adic predicate), and $\tau_1, \ldots, \tau_n$ are any n terms, i.e. names or variables. When all the terms in an atomic formula are names—so no variables occur—the expression is an atomic sentence.

In practice, we mostly deal with atomic formulas containing one-place and two-place predicates, but it is well to note that the language contains predicates of every finite degree $n = 0, 1, 2, \ldots$. If need be, the degree of an n–place predicate can be indicated by a superscript, $\mathbb{P}^n$; but usually the degree of the predicate can be told from the number of terms accompanying it in a sentence.

Simple predicates of degree greater than 2 seem rare in natural languages. Perhaps there is one in a sentence such as Eric, Ginger, and Jack formed a trio. We may adopt the viewpoint that a predicate in English results whenever names are extracted from a sentence.

The language also contains predicates of degree 0. A zero-place predicate forms an atomic formula without any names or variables: $\mathbb{P}$. Since there are no variables, such an atomic formula is in fact an atomic sentence, and thus is exactly like an atomic sentence in sentential logic.

EXERCISES

6.1. Using the translation key, translate the sentences below.

a—Wild Bill Hickok	b—Calamity Jane
F—is a schoolteacher	G—is a scout

R—loves

(1) Wild Bill is a schoolteacher

(2) Calamity Jane is a scout

(3) Wild Bill loves Calamity Jane

(4) Calamity Jane does not love Wild Bill

(5) Calamity Jane is loved by Wild Bill

3. Some simple quantificational sentences

Quantificational sentences express generality. For example, where H again means is a human, the sentence

$$\forall xHx$$

means that everything is human, while (with M for is mortal) the sentence

$$\exists xMx$$

means that something is mortal—i.e. that there is at least one thing that is mortal.

Notice that in the sentences $\forall xHx$ and $\exists xMx$ neither the atomic formula Hx nor the atomic formula Mx is a sentence. Variables thus contribute to the meanings of quantificational sentences both by occurring with quantifiers and as structural place-holders in atomic formulas.

Let us translate $\forall xHx$ by the stilted but accurate sentence:

Everything is such that it is human

Then a helpful analogy between the quantificational sentence $\forall xHx$ and this translation is provided by the following table.

$\forall\underline{x}$	$H\underline{x}$
Everything is such that	it is human

The underlining shows the relationship between the occurrences of the variable x, on the one hand, and the words thing and it, on the other. The gapped phrase every . . . is such that may be thought of as corresponding to the universal quantifier.

A similar analogy holds between $\exists xMx$ and Something is such that it is mortal or There is at least one thing such that it is mortal:

$\exists\underline{x}$	$M\underline{x}$
Something is such that	it is mortal

$\exists\underline{x}$	$M\underline{x}$
There is at least one thing such that	it is mortal

Here the gapped phrases some . . . is such that and there is at least one . . . such that represent the existential quantifier.

Universal and existential quantification provide a choice when it comes to expressions of negative generality, such as Nothing is mortal. We can write either

$$\forall x \neg Mx$$

(Everything is not mortal), or

$$\neg \exists x Mx$$

(It is not the case that something is mortal). The two forms have the same meaning.

The last two sentences should not be mistaken for:

$$\neg \forall x Mx \qquad \exists x \neg Mx$$

These two say the same thing, but mean that not everything is mortal, or, equivalently, that something is not mortal.

Thus in general we have the following equivalent forms.

$$\neg \forall x A \quad \text{and} \quad \exists x \neg A$$

$$\neg \exists x A \quad \text{and} \quad \forall x \neg A$$

We will see the equivalence of these pairs in chapter 7 and again in chapter 8; they are useful in understanding alternative translations.

Before proceeding, this is a good place to observe that the simple sentences we have been considering can equally well be formulated using a variable other than x. For example, for $\forall xHx$ each of these will do as well:

$$\forall uHu, \ \forall vHv, \ \forall wHw, \ \forall yHy, \ \forall zHz$$

A note on domains. In the translations above we have rendered universal and existential quantification by means of the phrases everything is such that and something is such that. What is the background against which we evaluate sentences such as Everything is a human and Something is mortal? What is the range of "things" that are thus said to be humans or to be mortal?

When translating informally, the range of the quantifiers often does not matter and need not be specified. But in some contexts it does matter. For example, if the quantifiers range over all animals, then the sentence $\forall xHx$—Everything is a human—is false, whereas if the range is the set of individuals who read this book, then the sentence is (probably) true. Similarly, if the quantifiers range over all animals, then ∃xMx is true; but if they range over, say, numbers, then the sentence is false. Thus different ranges make for differences in truth value and, in that sense, meaning.

In models for the language of predicate logic, as we shall see in chapter 7, the set of things over which the quantifiers range is called the *domain of discourse* or, simply, the *domain*.

Sometimes a domain is implicit. For example, everyone, someone, anyone, and the phrase no one indicate that the domain is a set of human beings—as do everybody, somebody, anybody, and nobody.

Later on we pay strict attention to the question of domains. In the context of informal translation, as in the present chapter, we do so only as necessary.

EXERCISES

6.2. Translate the following sentences, where:

a—Billy the Kid	F—can dance the fandango
G—will go to the ball	*domain*: humans

(1) ∀xFx → Fa

(2) Fa → ∃xFx

(3) ∀xFx → ∃xFx

(4) (∀xFx ∧ ∀xGx) → (Fa ∧ Ga)

(5) (Fa ∨ Ga) → (∃xFx ∨ ∃xGx)

6.3. Translate the following sentences, where:

a—Bill Miner F—is an outlaw
domain: humans

(1) If Bill is an outlaw then everyone is an outlaw

(2) If someone is an outlaw then Bill is an outlaw

(3) If someone is an outlaw then everyone is an outlaw

(4) If someone is not an outlaw then no one is an outlaw

(5) Either everyone is an outlaw or no one is an outlaw

4. More quantificational sentences

Consider now the major premiss of the argument at the beginning of the chapter:

Every human is mortal

We translate this into the language of predicate logic by means of universal quantification and conditionality:

$$\forall x(Hx \rightarrow Mx)$$

In other words, we regard Every human is mortal as meaning

Everything is such that if it is a human then it is mortal

Compare these structures as we did those in section 3:

$\forall \underline{x}$	$(H\underline{x} \rightarrow M\underline{x})$
Everything is such that	if it is human then it is mortal

There are other ways of expressing universal quantification in English, using such words as each, any, and all. Thus we will count each of

Each human is mortal

Any human is mortal

All humans are mortal

as equivalent and translatable as $\forall x(Hx \rightarrow Mx)$.

Such expressions as only, none but, only a, and nothing but a can also be used to translate sentences of this form. Suppose for a moment that H means has hair and M means is a mammal. Then the sentences

Only mammals have hair

None but mammals have hair

Only a mammal has hair

Nothing but a mammal has hair

all say that everything that has hair is a mammal, and so may be translated $\forall x(Hx \rightarrow Mx)$.

To translate the sentence

Some human is mortal,

we use existential quantification and conjunction:

$$\exists x(Hx \wedge Mx)$$

Among the numerous equivalents in English:

Some humans are mortal

There is at least one human that is mortal

There exists at least one human that is mortal

There are humans that are mortal

Something is both a human and mortal

As before, note the analogy:

$\exists\underline{x}$	$(H\underline{x} \wedge M\underline{x})$
Something is such that	it is human and it is mortal

For a "negative existential", such as No human is mortal (equally: No humans are mortal), there are two obvious choices:

$$\forall x(Hx \rightarrow \neg Mx) \qquad \neg \exists x(Hx \wedge Mx)$$

These are equivalent. The first says that every human is not mortal (everything is such that if it is a human then it is not mortal), the second that it is not the case that some human is mortal (it is not the case that something is both a human and mortal).

Finally, note that the sentence

$$\exists x(Hx \wedge \neg Mx)$$

means that some human is not mortal (not that no human is mortal).

Another note on domains. In translating sentences such as Every human is mortal, we have not been concerned with questions of quantificational range. But observe that if the domain is restricted to human beings then Every human is mortal can be translated simply:

$$\forall xMx$$

That is to say, with the domain assumed to be confined to humans, Every human is mortal just means that everything is mortal. Similarly for other such sentences—Some human is mortal, No human is mortal, and Some human is not mortal.

EXERCISES

6.4. Using the translation keys supplied in each case, translate the premisses and conclusions of the following arguments.

(1) $\forall x(Fx \to Gx),\ \exists x(Hx \land Fx)\ /\ \exists x(Hx \land Gx)$

F—is a cat G—understands French H—is a chicken

(2) $\forall x(Fx \to Gx),\ \exists x(Hx \land \neg Fx)\ /\ \exists x(Hx \land \neg Gx)$

F—is a soldier G—marches well H—is a baby

(3) $\forall x(Fx \to Gx),\ \forall x(Hx \to \neg Gx)\ /\ \forall x(Fx \to \neg Hx)$

F—is a pig G—is fat H—is fed on barley-water

(4) $\forall x(Fx \to Gx),\ \exists x(Fx \land \neg Hx)\ /\ \exists x(Hx \land \neg Gx)$

F—is a lion G—is fierce H—drinks coffee

(5) $\exists x(Fx \land Gx),\ \forall x(Gx \to \neg Hx)\ /\ \exists x(Fx \land \neg Hx)$

F—is an oyster G—is silent H—is amusing

(6) $\forall x(Fx \to Gx),\ \forall x(Hx \to Gx)\ /\ \forall x(Fx \to Hx)$

F—is clever G—is popular H—is obliging

(7) $\forall x(Fx \to \neg Gx),\ \forall x(Hx \to Gx)\ /\ \forall x(Fx \to \neg Hx)$

F—is a professor G—is ignorant H—is vain

(8) $\exists x(Fx \land Gx),\ \forall x(Hx \to Gx)\ /\ \exists x(Hx \land Fx)$

F—is a geranium G—is red H—is a flower

(9) $\exists x(Fx \wedge Gx),\ \forall x(Hx \rightarrow \neg Gx)\ /\ \exists x(Fx \wedge \neg Hx)$

F—is a dream G—is terrible H—is a lamb

(10) $\forall x(Fx \rightarrow \neg Gx),\ \forall x(Fx \rightarrow \neg Hx)\ /\ \forall x(Gx \rightarrow \neg Hx)$

F—is a baby G—is studious H—is a good violinist

6.5. Which of the arguments in exercise 6.4 are valid?

6.6. Translate There are old pilots and there are bold pilots, but there are no old bold pilots using each of the translation keys (1) and (2):

(1) O—old B—bold *domain*: pilots

(2) O—old B—bold P—pilot *domain*: humans

5. Categorical propositions and "squares of opposition"

In traditional logic, sentences of the four forms discussed above—Every human is mortal, Some human is mortal, No human is mortal, and Some human is not mortal—were said to express *categorical propositions* and were often represented in a *square of opposition*:

Every human is mortal	No human is mortal
Some human is mortal	Some human is not mortal

Propositions expressed at the top of the square were called *universal*; those at the bottom, *particular*. Those on the left side were called *affirmative*; those on the right, *negative*. Thus the propositions divide into four categories: universal affirmative, universal negative, particular affirmative, and particular negative.

It was held in traditional logic that sentences diagonally opposite each other in the square were contradictories (exactly one of them was true), and that the sentences at the top implied those below. From this it follows that the top sentences cannot be true together (otherwise, by contradictoriness, Every human is mortal would be true and Some human is mortal false, so that implication fails). Likewise, the bottom sentences cannot be false together (otherwise, by contradiction the top two could be true together).

Corresponding to the traditional square of opposition, we have in the language of predicate logic:

$\forall x(Hx \rightarrow Mx)$	$\forall x(Hx \rightarrow \neg Mx)$
$\exists x(Hx \wedge Mx)$	$\exists x(Hx \wedge \neg Mx)$

In this square only the relation of contradictoriness along the diagonals still holds. The top sentences do not imply the bottom ones, the tops can be true together, and the bottoms can be false together.

Sometimes the failure of top-to-bottom implication is explained by saying that universal sentences in the language of predicate logic do not have "existential import". But to say that in the new square the tops do not imply the corresponding bottoms does *not* mean that in predicate logic *every* does not imply *some*. On the contrary, any universal quantification

$$\forall xA$$

implies its existential counterpart

$$\exists xA$$

—since we assume that there is always something in the domain of any model. But in the new square of opposition this yields only that $\forall x(Hx \rightarrow Mx)$ implies $\exists x(Hx \rightarrow Mx)$. The latter sentence does not mean that some human is mortal; it means only that something is such that it is mortal if it is a human, which is quite a different matter. The situation is parallel with respect to $\forall x(Hx \rightarrow \neg Mx)$ and $\exists x(Hx \rightarrow \neg Mx)$.

This is a good place to point out that in order to translate Every human is mortal it will not do to write:

$$\forall x(Hx \wedge Mx)$$

This sentence says that everything is both a human and mortal, which is too strong.

EXERCISES

6.7. Here are some sentences adapted from Lewis Carroll's *Symbolic logic*. Express each in the language of predicate logic, using the translation keys. (Two-place predicates are indicated by superscripts.)

(1) Some holidays are tiresome.

H—is a holiday T—is tiresome

(2) Some greyhounds are not fat.

F—is fat G—is a greyhound

(3) No photograph of a lady is satisfactory.

P—is a photograph of a lady S—is satisfactory

(4) Some cats cannot whistle.

C—is a cat W—can whistle

(5) Some things that are not umbrellas should be left behind on a journey.

J—should be left behind on a journey U—is an umbrella

(6) Some brave persons get their deserts.

B—is brave D—gets his or her deserts P—is a person

(7) Some rich persons are not Esquimaux.

E—is an Esquimau P—is a person R—is rich

(8) Some black rabbits are not old.

B—black O—is old R—is a rabbit

(9) Some thin persons are not cheerful.

C—is cheerful P—is a person T—is thin

(10) No antelopes fail to delight the eye.

A—is an antelope D—delights the eye

(11) No one looks poetical unless he or she is phlegmatic.

F—is phlegmatic L—looks poetical P—is a person

(12) Some judges do not exercise self-control.

J—is a judge E^2—exercises control of

(13) No photograph of a lady is satisfactory.

L—is a lady P^2—is a photograph of S—is satisfactory

(14) No one who forgets a promise fails to do mischief.

F^2—forgets M—does mischief P—is a person
R—is a promise

(15) Logic is unintelligible.

I—is intelligible L—is logic

6.8. Translate the premisses and conclusion of the argument in exercise 1.15 (page 14), where:

E—is European	F—is French
S—is Spanish	L—is Libyan

domain: humans

6. Multiple quantification

Let the name o denote the number zero and the predicate G mean is greater than or equal to. If we suppose that the quantifiers range over natural numbers (0, 1, 2, . . .), then the quantificational sentences

$$\forall xGxo$$

and

$$\exists xGox$$

mean, respectively, that every number is greater than or equal to zero and that zero is greater than or equal to some number (i.e. that there is a number such that zero is greater than or equal to it).

We can iterate quantifiers to form multiple quantifications. For example:

$$\forall y\exists xGxy$$

This says that for every number there is at least one that is greater than or equal to it.

Notice that the order of quantification generally makes a difference: If we reverse the quantificational prefixes in the sentence above—to get

$$\exists x\forall yGxy$$

—then the sentence means that at least one number is greater than or equal to every number. This is clearly different in meaning: $\exists x \forall y Gxy$ is false (there is no maximal number), whereas $\forall y \exists x Gxy$ is true (since every number is greater than or equal to itself).

Multiple quantifications sometimes seem to resist simple translation, particularly when negation is mixed in. Take for example the sentence

$$\exists x \neg \forall y \neg Gxy$$

This appears to say something like:

> There is at least one number such that it is not the case that every number is such that it is not the case that the first is greater than or equal to the second.

But we can use the equivalences between the forms $\neg \forall xA$ and $\exists x \neg A$, and between the forms $\neg \exists xA$ and $\forall x \neg A$, to rewrite $\exists x \neg \forall y \neg Gxy$ successively as

$$\exists x \neg \neg \exists y Gxy, \quad \neg \forall x \neg \exists y Gxy, \quad \text{and} \quad \neg \neg \exists x \exists y Gxy$$

—thus reaching a readable form that says (ignoring the double negation) that some number is greater than or equal to some number.

A final note on domains. As we observed above, with G interpreted to mean greater than or equal to and relative to the domain $\{0, 1, 2, \ldots\}$ of natural numbers, the sentence $\exists x \forall y Gxy$ says that there is a maximal number, and hence is false. Observe that if we continue to let G mean greater than or equal to, then we can still change the truth value of $\exists x \forall y Gxy$ by changing the domain. For example, take the domain to be the set $\{0, \ldots, 9\}$ of numbers from zero to nine. Then, since nine is greater than or equal to every number in the set, the sentence $\exists x \forall y Gxy$ is true. This a simple but dramatic illustration of the point that altering the domain may vary the truth value of a sentence even when the meanings of its non-logical constants (the predicate G in this case) are assumed to be fixed.

EXERCISES

6.9. Translate the following sentences, where:

a — Sacajawea R — loves
domain: humans

(1)	Raa	(16)	$\neg Raa$
(2)	$\forall xRax$	(17)	$\forall x\neg Rax$
(3)	$\forall xRxa$	(18)	$\forall x\neg Rxa$
(4)	$\exists xRax$	(19)	$\exists x\neg Rax$
(5)	$\exists xRxa$	(20)	$\exists x\neg Rxa$
(6)	$\forall xRxx$	(21)	$\forall x\neg Rxx$
(7)	$\exists xRxx$	(22)	$\exists x\neg Rxx$
(8)	$\forall x\forall yRxy$	(23)	$\forall x\forall y\neg Rxy$
(9)	$\forall x\forall yRyx$	(24)	$\forall x\forall y\neg Ryx$
(10)	$\exists x\exists yRxy$	(25)	$\exists x\exists y\neg Rxy$
(11)	$\exists x\exists yRyx$	(26)	$\exists x\exists y\neg Ryx$
(12)	$\forall x\exists yRxy$	(27)	$\forall x\exists y\neg Rxy$
(13)	$\forall x\exists yRyx$	(28)	$\forall x\exists y\neg Ryx$
(14)	$\exists x\forall yRxy$	(29)	$\exists x\forall y\neg Rxy$
(15)	$\exists x\forall yRyx$	(30)	$\exists x\forall y\neg Ryx$

6.10. Using the translation key in the preceding exercise, translate the following sentences.

(1)	$\forall x \neg \forall y Rxy$	(5)	$\forall x \neg \exists y Rxy$
(2)	$\forall x \neg \forall y Ryx$	(6)	$\forall x \neg \exists y Ryx$
(3)	$\exists x \neg \exists y Rxy$	(7)	$\exists x \neg \forall y Rxy$
(4)	$\exists x \neg \exists y Ryx$	(8)	$\exists x \neg \forall y Ryx$

7. Postscript on *any*, *some*, *if*, and *every*

We said above that in many contexts every, each, and any are interchangeable as expressions of universal quantification—as in Every human is mortal, Each human is mortal, and Any human is mortal. The particle any can be used in yet another way to express universal quantification:

If anything is a human then it is mortal

In a similar manner, some—which is usually associated with existential quantification—can be used like any to express universal quantification, as in:

If something is a human then it is mortal

Both of these sentences are translated:

$$\forall x(Hx \to Mx)$$

Notice, moreover, that neither sentence is translatable as $\forall xHx \to Mx$ or as $\exists xHx \to Mx$—for these formulas are not even sentences.

Now consider sentences such as:

If anything comes in the mail then Buffalo Bill will be pleased

If something comes in the mail then Buffalo Bill will be pleased

These may both be translated

$$\forall x(Cx \rightarrow Pa)$$

—where C means comes in the mail, P means is pleased, and a translates Buffalo Bill. But in this situation, any and some can also be rendered by means of a conditional with an antecedent existential quantification,

$$\exists xCx \rightarrow Pa$$

—which corresponds closely in form to If something comes in the mail then Buffalo Bill will be pleased.

The difference is that in this case there is no variable in the consequents of the conditionals in $\forall x(Cx \rightarrow Pa)$ and $\exists xCx \rightarrow Pa$, just as in the consequents of the English sentences there is no relative pronoun it.

The moral of all this is that under the influence of the word if, expressions such as any and some can behave existentially, but the forms if any . . . and if some . . . can nevertheless always be translated as every . . . if.

Finally, note the apparent shift in the meaning of any under the influence of negation. Whereas the sentence Any human is mortal may be rendered

$$\forall x(Hx \rightarrow Mx),$$

the sentence Not any human is mortal means

$$\neg \exists x(Hx \wedge Mx),$$

i.e. No human is mortal—and not

$$\neg \forall x(Hx \rightarrow Mx),$$

i.e. not Not every human is mortal.

EXERCISES

6.11. Translate the premisses and conclusions of the following arguments (adapted from Lewis Carroll's *Symbolic logic*).

(1) No misers are unselfish. None but misers save eggshells. Therefore, no unselfish people save eggshells.

E—saves eggshells M—is a miser S—is selfish *domain*: humans

(2) A prudent person shuns hyenas. No banker is imprudent. Therefore, no banker fails to shun hyenas.

B—is a banker P—is prudent S—shuns hyenas *domain*: humans

(3) If anything is a baby it is illogical. No one is despised who can manage a crocodile. Illogical persons are despised. Therefore, babies cannot manage crocodiles.

B—is a baby C—can manage a crocodile D—is despised L—is logical *domain*: humans

(4) No ducks waltz. No officers ever decline to waltz. All my poultry are ducks. Therefore, my poultry are not officers.

D—is a duck O—is an officer P—is one of my poultry W—waltzes *domain*: animals

(5) Everyone who is sane can do logic. No lunatics are fit to serve on a jury. None of *your* sons can do logic. Therefore, none of your sons is fit to serve on a jury.

J—is fit to serve on a jury L—can do logic S—is sane [is not a lunatic] Y—is one of *your* sons *domain*: humans

(6) No terriers wander among the signs of the zodiac. Nothing that does not wander among the signs of the zodiac is a comet. Only a terrier has a curly tail. Therefore, no comet has a curly tail.

C—is a comet T—is a terrier U—has a curly tail W—wanders among the signs of the zodiac *domain*: animals and celestial beings

(7) Nothing takes in the *Times* unless it is well educated. No hedgehogs can read. Those who cannot read are not well educated. Therefore, no hedgehog takes in the *Times*.

E—is well educated H—is a hedgehog R—can read T—takes in the *Times* *domain*: animals

(8) Nothing that really appreciates Beethoven fails to keep silent while the Moonlight Sonata is being played. If something is a guinea pig then it is hopelessly ignorant of music. Nothing that is hopelessly ignorant of music ever keeps silent while the Moonlight Sonata is being played. Therefore, guinea pigs never really appreciate Beethoven.

B—really appreciates Beethoven G—is a guinea pig I—is hopelessly ignorant of music S—keeps silent while the Moonlight Sonata is being played *domain*: animals

(9) No kitten that loves fish is unteachable. No kitten without a tail will play with a gorilla. Kittens with whiskers always love fish. No teachable kitten has green eyes. No kittens have tails unless they have whiskers. Therefore, no kitten with green eyes will play with a gorilla.

F—loves fish G—has green eyes K—is a kitten L—has a tail P—will play with a gorilla T—is teachable W—has whiskers *domain*: animals

(10) No shark ever doubts that it is well fitted out. A fish that cannot dance a minuet is contemptible. No fish is quite certain that it is well fitted out, unless it has three rows of teeth. All fishes except sharks are kind to children. No heavy fish can dance a minuet. A fish with three rows of teeth is not to be despised. Therefore, no heavy fish is unkind to children.

(Predicates and domain: left to taste.)

6.12. Relative to a model in which R means robs and the domain is the set of female humans, match each sentence (a)–(e) with its translation in (1)–(5).

(a) Everyone who robs herself robs everyone who robs her

(b) Everyone who robs everyone who robs her robs herself

(c) Everyone who robs herself robs everyone who robs herself

(d) Everyone who robs everyone who robs herself robs herself

(e) Everyone who robs herself robs only her who robs herself

(1) $\forall x(Rxx \to \forall y(Rxy \to Ryy))$

(2) $\forall x(\forall y(Ryx \to Rxy) \to Rxx)$

(3) $\forall x(Rxx \to \forall y(Ryx \to Rxy))$

(4) $\forall x(\forall y(Ryy \to Rxy) \to Rxx)$

(5) $\forall x(Rxx \to \forall y(Ryy \to Rxy))$

6.13. For each of arguments (2), (4), (6), (8), and (10) in exercise 6.4, translate the indicated predicates so that the premisses are true and the conclusion is false. (Be sure to give a domain in each case.)

6.14. Relative to a model in which S means shaves and the domain is the set of barbers, translate the following sentences.

(1) There is a barber who shaves all those barbers who do not shave themselves.

(2) There is a barber who shaves only those barbers who do not shave themselves.

(3) There is a barber who shaves all and only those barbers who do not shave themselves.

8. The language

We turn now from questions of translation to a formal description of the language of predicate logic.

Names, variables, and terms. The names are infinite in number and are officially listed as:

$$a_0, \ a_1, \ a_2, \ . \ . \ . \ , \ a_n, \ . \ . \ .$$

Unofficially, we use the first twenty lowercase roman letters:

$$a, \ b, \ c, \ . \ . \ . \ , \ t$$

Names are sometimes called *individual constants*—misleadingly, since as non-logical symbols their values change from model to model. To talk generally about names we use lowercase alpha, beta, and gamma (α, β, γ).

The variables (also called *individual variables*) are likewise infinite in number:

$$x_0, \ x_1, \ x_2, \ . \ . \ . \ , \ x_n, \ . \ . \ .$$

Unofficially, we use the last six lowercase roman letters (and usually the last three of these) for variables:

$$u, \ v, \ w, \ x, \ y, \ z$$

We also use these letters when speaking generally about variables.

Any name or variable is a *term* (*individual term*), and we use lowercase tau, sigma, and rho (τ, σ, ρ) when we wish to speak of arbitrary terms.

Predicates. The language contains infinitely many *predicates* of each degree n. So the totality of predicates can be pictured in a two-dimensional array:

$$
\begin{array}{cccccc}
\mathbb{P}^0_0 & \mathbb{P}^0_1 & \mathbb{P}^0_2 & \ldots & \mathbb{P}^0_i & \ldots \\
\mathbb{P}^1_0 & \mathbb{P}^1_1 & \mathbb{P}^1_2 & \ldots & \mathbb{P}^1_i & \ldots \\
\mathbb{P}^2_0 & \mathbb{P}^2_1 & \mathbb{P}^2_2 & \ldots & \mathbb{P}^2_i & \ldots \\
\cdot & \cdot & \cdot & & \cdot & \\
\cdot & \cdot & \cdot & & \cdot & \\
\cdot & \cdot & \cdot & & \cdot & \\
\mathbb{P}^n_0 & \mathbb{P}^n_1 & \mathbb{P}^n_2 & \ldots & \mathbb{P}^n_i & \ldots \\
\cdot & \cdot & \cdot & & \cdot & \\
\cdot & \cdot & \cdot & & \cdot & \\
\cdot & \cdot & \cdot & & \cdot &
\end{array}
$$

Even when using this "official" notation, we usually omit the subscripts, and we often leave off the superscript where a predicate's degree is evident. Unofficially, we use ordinary uppercase roman letters for predicates.

Formulas and sentences. As we have seen, sentences in the language of predicate logic are a species of a more general class of structures called formulas. The "smallest" of the formulas are the *atomic* formulas,

$$\mathbb{P}\tau_1 \ldots \tau_n,$$

in which each term τ_i ($i = 1, \ldots, n$) is either a name or a variable, and $\mathbb{P}$ is an n-place predicate.

For all their structure and possible length, atomic formulas are the "smallest" formulas in the language. There are infinitely many of them.

The rest of the formulas in the language are built from the atomic formulas by means of the logical operators $\neg$, $\wedge$, $\vee$, $\rightarrow$, $\leftrightarrow$, $\forall$, and $\exists$. In a formula $\forall$xA or $\exists$xA, $\forall$x and $\exists$x are called the *prefix* of the quantification, while the quantified formula A itself is referred to as the *matrix*. Symbols in the matrix A of $\forall$xA (or $\exists$xA) are said to be in the *scope* of the quantifier $\forall$ (or $\exists$).

In terms of atomic formulas and the logical operators, the class of all formulas is defined formally as follows.

DEFINITION 6.1. *Formulas.*

(1) Each atomic formula $\mathbb{P}\tau_1 \ldots \tau_n$ is a formula.

(2) $\neg$ A is a formula iff A is a formula.

(3) A $\wedge$ B is a formula iff A and B are formulas.

(4) A $\vee$ B is a formula iff A and B are formulas.

(5) A $\rightarrow$ B is a formula iff A and B are formulas.

(6) A $\leftrightarrow$ B is a formula iff A and B are formulas.

(7) $\forall$xA is a formula iff x is a variable and A is a formula.

(8) $\exists$xA is a formula iff x is a variable and A is a formula.

Free and bound variables. A variable in a formula is said to be *free* just when it is not in the scope of a quantifier using that variable; otherwise the occurrence is *bound*. Thus an occurrence of a variable x is free in a formula A if and only if that occurrence is not within the scope of either $\forall$x or $\exists$x. For example, in the formula

$$\mathrm{Fx} \rightarrow \exists \mathrm{xGx}$$

the variable x has three occurrences; the first is free and the second and third are bound (by $\exists$).

Every occurrence of a variable in a prefix $\forall$x (or $\exists$x) is bound by the quantifier $\forall$ (or $\exists$), and in a quantificational formula $\forall$xA (or $\exists$xA) all free occurrences of x in the matrix A are bound by the occurrence of the quantifier $\forall$ (or $\exists$) in the prefix. Thus in the formula

$$\forall \mathrm{x}(\mathrm{Fx} \rightarrow \mathrm{Gx})$$

the first occurrence of x is bound by $\forall$ since it is part of the prefix $\forall x$, and the other two occurrences of x are also bound by $\forall$ since they are both free in the matrix $Fx \rightarrow Gx$. On the other hand, in

$$\forall x(Fx \rightarrow \exists xGx),$$

while the first and second occurrences of x are bound by $\forall$, the third and fourth are not bound by $\forall$, since they are not free in the matrix $Fx \rightarrow \exists xGx$ (indeed, they are bound by the occurrence of $\exists$).

Notice that in the formula

$$\forall xFx \rightarrow Gx$$

the first two occurrences of x are bound (by $\forall$), but the last is not, since Gx is not in the scope of the quantifier.

A sentence is a formula in which there are no free variables, i.e. in which either there are no variables, or all variables are bound. Thus it is useful to have a definition of the free variables in a formula A. We denote the set of free variables in A by Fr(A), and define this as follows.

DEFINITION 6.2. *Free variables of a formula.*

(1) $\mathrm{Fr}(\mathbb{P}\tau_1 \ldots \tau_n)$ = the set of variables among $\tau_1 \ldots \tau_n$.

(2) $\mathrm{Fr}(\neg A) = \mathrm{Fr}(A)$.

(3) $\mathrm{Fr}(A \wedge B) = \mathrm{Fr}(A) \cup \mathrm{Fr}(B)$.

(4) $\mathrm{Fr}(A \vee B) = \mathrm{Fr}(A) \cup \mathrm{Fr}(B)$.

(5) $\mathrm{Fr}(A \rightarrow B) = \mathrm{Fr}(A) \cup \mathrm{Fr}(B)$.

(6) $\mathrm{Fr}(A \leftrightarrow B) = \mathrm{Fr}(A) \cup \mathrm{Fr}(B)$.

(7) $\mathrm{Fr}(\forall xA) = \mathrm{Fr}(A) - \{x\}$.

(8) $\mathrm{Fr}(\exists xA) = \mathrm{Fr}(A) - \{x\}$.

In other words: the free variables of an atomic formula are just the variables that accompany the predicate in the formula; the free variables of formulas formed by means of the operators $\neg$, $\wedge$, $\vee$, $\rightarrow$, and $\leftrightarrow$ are just the free variables in the operands, and the free variables of a quantification are just the free variables of the matrix with the exception of the variable that occurs in the quantifier prefix.

Open and closed formulas. In terms of the idea of a free variable we distinguish *open* and *closed* formulas. Open formulas are those in which at least one variable has a free occurrence; closed formulas are those in which either there are no variables or every occurrence of a variable is bound. Formally:

DEFINITION 6.3. *Open and closed formulas.*

- A is open iff $\mathrm{Fr}(A) \neq \varnothing$.
- A is closed iff $\mathrm{Fr}(A) = \varnothing$.

Sentences. Since a sentence is just a formula in which there are no free variables, we arrive at last at this definition:

DEFINITION 6.4. *Sentences.* A is a sentence iff A is a closed formula, i.e. iff $\mathrm{Fr}(A) = \varnothing$.

Vacuous quantification. If there are no free occurrences of x in A, then the quantifiers $\forall$ and $\exists$ are said to be *vacuous* in $\forall xA$ and in $\exists xA$. For example, $\forall$ occurs vacuously in the formula $\forall xFy$, and $\exists$ occurs vacuously in $\exists xRay$, since in neither case does the quantifier bind any variables in its scope. The same goes for these examples: $\forall xFa$, $\exists x\forall xRxb$. In each of these the matrix (Fa, $\forall xRxb$) is a sentence, and therefore so are $\forall xFa$ and $\exists x\forall xRxb$. (The first might read Everything is such that Abe is a human; the second might be translated Something is such that everything is greater than or equal to zero. Both are weird but grammatically acceptable sentences: the first means that Abe is a human; the second means that everything is greater than or equal to zero.)

EXERCISES

6.15. Define, for a formula A, the sets of (i) subformulas of A (Sbfm(A)), (ii) proper subformulas of A (Prsbfm(A)), and (iii) atoms of A (Atm(A)). (See section 4 of chapter 2.)

6.16. Identify the free and bound occurrences of the variables in this formula:

$$(Rxa \vee \exists xRxy) \leftrightarrow \neg \forall y(Rxy \rightarrow Rya)$$

9. Replacement, instances, and substitution

In this section we explain the ideas of replacement, instances of a quantification, and substitution. All are useful in subsequent chapters.

Replacement. If in the sentence Rbb we replace all occurrences of the name b by the name a, we obtain the sentence Raa. Likewise if we replace all free occurrences of the variable x in Rxx by the name a, we also obtain the sentence Raa. If we use a to replace all free occurrences of x in Fx → ∃xGx, however, we get Fa → ∃xGx. We can also use the idea of replacement of terms by terms within terms themselves. Thus if in the name a we replace the name b by the variable x, we get either x (if a and b are the same name) or just a (if a and b are different names).

In general, we are interested in formulas that result when all occurrences of a name or all free occurrences of a variable in a formula are replaced by a name or variable. To get at this notion, we first define the idea of replacement of a name or free variable in a term:

DEFINITION 6.5. *Replacement of a name or variable in a term.* $\rho(\sigma/\tau)$—"ρ with σ out for τ"—is the term that results when the term σ is replaced in the term ρ by the term τ. In other words:

$$\rho(\sigma/\tau) = \tau \text{ if } \rho \text{ and } \sigma \text{ are the same;}$$

and

$$\rho(\sigma/\tau) = \rho \text{ if } \rho \text{ and } \sigma \text{ are different.}$$

In the chapters that follow, we will need two ideas of replacement in formulas: replacement of a name by a name, and replacement of a free variable by a name. See how definition 6.5 yields these definitions:

DEFINITION 6.6. *Replacement of a name in a formula by a name.* $A(\beta/\alpha)$—"A with β out for α"—is the formula that results when every occurrence of the name β in the formula A is replaced by the name α. In other words:

(1) $(\mathbb{P}\tau_1 \ldots \tau_n)(\beta/\alpha) = \mathbb{P}\tau_1(\beta/\alpha) \ldots \tau_n(\beta/\alpha)$.

(2) $(\neg A)(\beta/\alpha) = \neg (A(\beta/\alpha))$.

(3) $(A \wedge B)(\beta/\alpha) = A(\beta/\alpha) \wedge B(\beta/\alpha)$.

(4) $(A \vee B)(\beta/\alpha) = A(\beta/\alpha) \vee B(\beta/\alpha)$.

(5) $(A \rightarrow B)(\beta/\alpha) = A(\beta/\alpha) \rightarrow B(\beta/\alpha)$.

(6) $(A \leftrightarrow B)(\beta/\alpha) = A(\beta/\alpha) \leftrightarrow B(\beta/\alpha)$.

(7) $(\forall x A)(\beta/\alpha) = \forall x(A(\beta/\alpha))$.

(8) $(\exists x A)(\beta/\alpha) = \exists x(A(\beta/\alpha))$.

Let us illustrate definitions 6.5 and 6.6 by showing how they yield the correct result when every occurrence of the name b in the sentence $\forall x(Fb \rightarrow Rxb)$ is replaced by the name a:

$$
\begin{aligned}
(\forall x(Fb \rightarrow Rxb))(b/a) &= \forall x((Fb \rightarrow Rxb)(b/a)) \\
&\quad \text{—definition 6.6, part (7);} \\
&= \forall x((Fb)(b/a) \rightarrow (Rxb)(b/a)) \\
&\quad \text{—definition 6.6, part (5);} \\
&= \forall x(Fb(b/a) \rightarrow Rx(b/a)b(b/a)) \\
&\quad \text{—definition 6.6, part (1);} \\
&= \forall x(Fa \rightarrow Rxa) \\
&\quad \text{—definition 6.5.}
\end{aligned}
$$

Thus $(\forall x(Fb \rightarrow Rxb))(b/a)$ is $\forall x(Fa \rightarrow Rxa)$, as it should be.

The definition below of replacement of a free variable by a name is like the one just above, except when it comes to the quantificational parts:

DEFINITION 6.7. *Replacement of a free variable in a formula by a name.* $A(x/\alpha)$—"A with free x out for α"—is the formula that results when every free occurrence of the variable x in the formula A is replaced by the name α. In other words:

(1) $(\mathbb{P}\tau_1 \ldots \tau_n)(x/\alpha) = \mathbb{P}\tau_1(x/\alpha) \ldots \tau_n(x/\alpha)$.

(2) $(\neg A)(x/\alpha) = \neg (A(x/\alpha))$.

(3) $(A \wedge B)(x/\alpha) = A(x/\alpha) \wedge B(x/\alpha)$.

(4) $(A \vee B)(x/\alpha) = A(x/\alpha) \vee B(x/\alpha)$.

(5) $(A \rightarrow B)(x/\alpha) = A(x/\alpha) \rightarrow B(x/\alpha)$.

(6) $(A \leftrightarrow B)(x/\alpha) = A(x/\alpha) \leftrightarrow B(x/\alpha)$.

(7) (a) $(\forall yA)(x/\alpha) = \forall yA$ if x and y are the same;

(b) $(\forall yA)(x/\alpha) = \forall y(A(x/\alpha))$ if x and y are different.

(8) (a) $(\exists yA)(x/\alpha) = \exists yA$ if x and y are the same;

(b) $(\exists yA)(x/\alpha) = \exists y(A(x/\alpha))$ if x and y are different.

The point of parts (7)(a) and (8)(a) is that there is no replacement of x by α if in fact x is the variable in the quantificational prefix, since in that case the variable x has no free occurrences in the formula.

Here is an example of how definitions 6.5 and 6.7 yield the correct result when every free occurrence of the variable x in the formula $\exists xFx \to Rxb$ is replaced by the name a:

$$
\begin{aligned}
(\exists xFx \to Rxb)(x/a) &= (\exists xFx)(x/a) \to (Rxb)(x/a) \\
&\quad \text{—definition 6.7, part (5);} \\
&= \exists xFx \to Rx(x/a)b(x/a) \\
&\quad \text{—definition 6.7, parts (1) and (8);} \\
&= \exists xFx \to Rab \\
&\quad \text{—definition 6.5.}
\end{aligned}
$$

Thus $(\exists xFx \to Rxb)(x/a)$ is $\exists xFx \to Rab$—as it should be, since the first and second occurrences of x in $\exists xFx \to Rxb$ are bound, while the third is free.

It is important to realize that a given formula with multiple occurrences of a given name arises in more than one way by virtue of replacement that uses the name. Thus for example, the sentence Raa is the result of replacing all occurrences of x in Rxx by a. In other words:

$$Raa = (Rxx)(x/a)$$

But Raa is also the result of replacing all occurrences of x in Rxa by a:

$$Raa = (Rxa)(x/a)$$

In like manner, we have:

$$Raa = (Rax)(x/a)$$

And even:

$$Raa = (Raa)(x/a)$$

Instances. An *instance* of a quantification ∀xA or ∃xA is any formula of the form $A(x/\alpha)$. Note that if a quantification is vacuous then it has just one instance, since $A(x/\alpha)$ will be A. The idea of an instance is important later on, especially when we come to the quantificational rules of inference.

Substitution. In replacement, *every* occurrence of a name or free variable is replaced by a name. When it comes to *substitution* it is not required that every occurrence of a name or free variable be replaced. Simply put, substitution is selective replacement—replacement that is total, partial, or even nil.

For example, the sentence Rab is the result of *substituting* the name b for one occurrence of a in the sentence Raa, whereas if b *replaces* a in Raa then the result is Rbb.

We can use the idea of replacement to define substitution. Observe that since the name c does not occur in Raa, this sentence can be described as (Rac)(c/a). Indeed, the sentence Raa can be described in four different ways in terms of replacing c by a:

(Rcc)(c/a)

(Rca)(c/a)

(Rac)(c/a)

(Raa)(c/a)

Each of these ways of describing Raa gives rise to a different way of substituting b for a in Raa:

Rbb, i.e. (Rcc)(c/b)

Rba, i.e. (Rca)(c/b)

Rab, i.e. (Rac)(c/b)

Raa, i.e. (Raa)(c/b)

It is the third of these—(Rac)(c/b)—that produces the sentence Rab in the example above. Substitution is vacuous in the last case since replacement is nil.

Thus we see that we can use replacement to give a formal definition of substitution. Because we will be interested only in names-for-name substitution, we ignore other sorts (name-for-variable, etc.).

DEFINITION 6.8. *Substitution.* $A(\gamma/\beta)$ is the result of substituting the name β for zero or more occurrences of the name α in $A(\gamma/\alpha)$.

Notice that in the limiting case where α and γ are the same name, $A(\gamma/\alpha) = A(\alpha/\alpha) = A$. In this circumstance $A(\gamma/\beta) = A(\alpha/\beta)$—i.e. β is substituted for every occurrence of α in A.

EXERCISES

6.17. What is the result of replacing all free occurrences of the variable x in the formula $\exists xRxa \vee \neg (Fy \wedge Rbx)$ by the name a? In other words, what is the formula $(\exists xRxa \vee \neg (Fy \wedge Rbx))(x/a)$?

6.18. The sentence Raa contains two occurrences of the name a. Describe the sentence in four different ways as arising by replacement of the name b by the name a.

6.19. Here is a sentence containing two occurrences of the name a:

$$Rab \rightarrow \exists yRay$$

Describe the sentence in four different ways as arising by replacement of a free variable by the name a. Also describe the sentence in four different ways as arising by replacement of some name by the name a.

6.20. Suppose A is a formula containing no occurrence of the name β. Explain why it follows that the formula $A(x/\beta)$ contains occurrences of β in exactly the places where A contains free occurrences of the variable x. (This is not to say that A contains any free occurrences of x—just that if it does then β occurs in $A(x/\beta)$ just where x is free in A.) What is the relationship between the formula $A(x/\alpha)$ and the formula $A(x/\beta)(\beta/\alpha)$ in this case?

10. Newness, fresh instances, universal closures

When it happens that a term does not occur in a formula A, we say that it is *new* to A. Thus for example, the name a is new to the formula Rxb. Notice especially that a name is new to a quantification $\forall xA$ or $\exists xA$ if and only if it is new to its matrix A. Thus, for example, a is equally new to $\forall xRxb$, $\exists xRxb$, and Rxb. We also say that a name is new to a set of sentences Γ just in case it is new to all the sentences in Γ.

By a *fresh instance* of a formula A we mean any result of replacing all free occurrences of variables in A by names that are new to A, putting distinct names for distinct variables. In other words, if $x_1, \ldots, x_n$ are the (distinct) free variables in a formula A, and $\alpha_1, \ldots, \alpha_n$ are distinct names all new to A, then the sentence

$$A(x_1/\alpha_1) \ldots (x_n/\alpha_n)$$

is a fresh instance of A. (Note that fresh instances always are sentences.) For example, $Rab \rightarrow \exists xRxb$ is a fresh instance of the formula $Rxy \rightarrow \exists xRxy$, since the free occurrence of x has been replaced by a while the free occurrences of y have been replaced by b.

Finally, let A be a formula containing the free variables $x_1, \ldots, x_n$ (in numerical order). Then by the *universal closure* of A we mean the result of universally quantifying A's free variables (in order):

$$\forall x_1 \ \ldots \ \forall x_n A$$

For example, the universal closure of the formula $Rxa \rightarrow Rxy$ is the sentence $\forall x \forall y(Rxa \rightarrow Rxy)$. (By definition, universal closures always are sentences.)

EXERCISES

6.21. Define the existential closure of a formula.

7

SEMANTICS FOR PREDICATE LOGIC

THE PRINCIPAL ideas in this chapter are those of a *model* and *truth in a model* for the language of predicate logic. The first is defined and illustrated in section 1, and in section 2 we discuss *covered models*. Questions of consequence, validity, equivalence, and satisfiability occupy section 3, where it becomes apparent that models are in effect formal structures for translating sentences in the language. In section 4 we introduce the notion of a *variant* of a model and use this in the definition of truth in section 5. Section 6 returns briefly to the logical properties of consequence, etc. Finally, in section 7 we present some propositions that underpin the soundness of the quantificational rules of inference found in chapter 8.

1. Models

As always, a model for a language gives meanings to the expressions of the language in such a way that meanings and hence truth values are determined for the sentences of the language.

In the case of predicate logic, unlike that of sentential logic, a model does not just assign a meaning (truth value) directly to each of the atomic sentences. Rather, a model for the language provides meanings for the parts of such sentences—i.e. for the names and predicates of the language.

In predicate logic a model consists of a non-empty set, the *domain* of the model, together with an assignment of *values* to each of the names and predicates in the language. The value of a name in a model is an element of the domain, and the value of an n-place predicate is an n-place relation among elements of the domain. In terms of a model, truth values are determined in a natural way for each of the atomic sentences of the language and, ultimately, for all the others.

For example, suppose the domain of a model is the set $\boldsymbol{N}$ of natural numbers, $\{0, 1, 2, \ldots\}$, suppose that to the names a and b the model assigns the numbers 0 and 1, and suppose that to the two-place predicate R the model assigns the less-than relation. (We may disregard the other assignments made by the model.) Then, relative to this model, the atomic sentence Rab says that zero is less than one. In other words, the sentence Rab is true in the model if and only if $0 < 1$.

Here is the formal definition of a model:

DEFINITION 7.1. *Models.* $\mathcal{M}$ is a model iff $\mathcal{M}$ assigns values to the names and predicates relative to a non-empty set D (the domain of $\mathcal{M}$) as follows.

(i) $\mathcal{M}(\alpha)$ is an element of D—i.e. $\mathcal{M}(\alpha) \in D$—for each name α.

(ii) $\mathcal{M}(\mathbb{P}^n)$ is an n-place relation in D—i.e. $\mathcal{M}(\mathbb{P}^n) \subseteq D^n$—for each n-place predicate $\mathbb{P}^n$.

Note that the only condition on the domain of a model is that it not be empty, i.e that it contain at least one element.

The *size* (or *cardinality*) of a model is the number of elements in its domain. A model is said to be *finite* just in case its domain is, i.e. just in case there are only a finite number of things in the set; otherwise the model is *infinite.*

It is not required of a model that it assign different elements of the domain to different names. So, for example, a model like the one above might assign 1 both to b and to the name c. (What then does Rbc mean?) Indeed, assignment of a single individual to more than one name will have to occur whenever the domain of a model is finite, since there are infinitely many names.

By the same token, it is not necessary that every individual in the domain of a model be assigned to some name. For example, in a model having as domain the set of natural numbers, all the names can be made to take just 0 as value.

We specify n-place ("n-ary") relations as values of n-place predicates so as to achieve maximum generality. In the case where $n = 1$, a one-place relation, or set of one-tuples, can be regarded simply a set. For example, the one-place relation $\{\langle \text{Butch} \rangle, \langle \text{Sundance} \rangle, \langle \text{Etta} \rangle\}$ amounts to the set $\{\text{Butch, Sundance, Etta}\}$. When $n = 0$, i.e. where zero-place predicates are concerned, there are just two possibilities, $\{\langle\ \rangle\}$ and $\varnothing$—a point to which we return in section 5 below.

The relations assigned to predicates are regarded *extensionally*, even though this may seem to blur distinctions in meaning. For example, suppose the domain of a model to be the set of living things, and that we wish to interpret the one-place predicates K and H to mean, respectively, has a kidney and has a heart. Intuitively, these two phrases have different meanings. Their *extensions*, however, are identical. Every creature with a kidney is a creature with a heart, and vice versa. So the set of creatures having kidneys is the same as the set of creatures having hearts. And so the model assigns the same sets to K and H.

Another example of the extensionality is afforded by centaurs and unicorns. Being a centaur is different from being a unicorn. But inasmuch as creatures of neither kind exist, the set of centaurs is the same as the set of unicorns — the empty set. Thus with domain the set of living things, a model of one-place predicates C and U as centaurs and unicorns might as well be expressed more succinctly as $\varnothing$.

These illustrations of models of predicates should make it clear that it is not required of a model that distinct predicates be assigned distinct relations as values. Sometimes it happens in subtle ways. For example, let the domain of a model be the set of natural numbers, and suppose the two-place predicates R and S to be given the values $\{\langle d, e\rangle: d < e\}$ and $\{\langle d, e\rangle: e > d\}$, respectively. Then R and S have been assigned the same relations, although the value of one is characterized in terms of the less-than relation while that of the other is characterized in terms of greater-than.

Finally, it is not required of a model that every possible relation in the domain be assigned to some predicate or other. To take the most extreme example, suppose the value of each n-place predicate is the empty n-place relation in the domain of a model, i.e. the value of each predicate $\mathbb{P}^n$ is the set $\varnothing$. Then since the domain itself cannot be empty, there will always be a relation left unassigned to any predicate. Notice, too, that it will be impossible to assign each relation in a domain to some predicate if the domain is countably infinite or larger. There are uncountably many subsets of such a domain, whereas there are only countably many predicates.

EXERCISES

7.1. For each of arguments (2), (4), (6), (8), and (10) in exercise 6.4 (page 233), describe a model in which the premisses are true and the conclusion is false. (Done already? See exercise 6.13, page 246.)

7.2. Describe a model that shows that the conclusion is not a consequence of the premisses in the following argument.

$$\forall x(Ex \rightarrow (Fx \vee Sx)),\ \exists x(Lx \wedge Ex),\ \neg \forall x(Lx \rightarrow Fx)\ \ /\ \ \exists x(Ex \wedge Sx)$$

(Compare exercise 6.8, page 239.)

7.3. For each odd-numbered sentence in exercise 6.9 (page 241), describe a model that verifies it; for each even-numbered sentence, describe a model that falsifies it. Do the same for the sentences in exercise 6.10 (page 242).

7.4. The following sentences express properties of binary relations.

	Sentence	*Property*
(a)	$\forall xRxx$	reflexivity
(b)	$\forall x \neg Rxx$	irreflexivity
(c)	$\forall x\forall yRxy$	universality; totality
(d)	$\forall x\forall y \neg Rxy$	emptiness; nullity
(e)	$\forall x\exists yRxy$	seriality
(f)	$\forall x\forall y(Rxy \rightarrow Ryx)$	symmetry
(g)	$\forall x\forall y(Rxy \rightarrow \neg Ryx)$	asymmetry
(h)	$\forall x\forall y(Rxy \vee Ryx)$	strong connectedness; dichotomy
(i)	$\forall x\forall y\forall z((Rxy \wedge Ryz) \rightarrow Rxz)$	transitivity
(j)	$\forall x\forall y\forall z((Rxy \wedge Ryz) \rightarrow \neg Rxz)$	intransitivity
(k)	$\forall x\forall y\forall z((Rxy \wedge Rxz) \rightarrow Ryz)$	euclideanness

(l) $\forall x \forall y(Rxy \rightarrow \exists z(Rxz \wedge Rzy))$ weak density

For each property (a)–(l), give an example of a relation having that property by describing a model in which the corresponding sentence is true. For example, if the domain of $\mathcal{M}$ is the set of natural numbers and $\mathcal{M}$(R) is the less-than relation, sentence (b) is true and the relation is irreflexive.

2. Covered models

A covered model is simply one in which each element in the domain has a name, i.e. each element in the domain is the value of at least one name. We also say of such a model that its domain is covered.

DEFINITION 7.2. *Covered model.* A model $\mathcal{M}$ is covered iff each d in the domain of $\mathcal{M}$ is the value in $\mathcal{M}$ of some name α (i.e. for every d in the domain of $\mathcal{M}$ there is a name α such that $\mathcal{M}(\alpha) = d$).

Examples come easily to mind. For instance, where the domain consists of just one thing (the first person to set foot on the moon, say), the model is perforce covered, since in it that one individual is named by every name. At the opposite extreme, suppose the domain is the set of natural numbers and $\mathcal{M}(a_n) = n$ for every $n \geq 0$. Then the model $\mathcal{M}$ is covered—though barely, since this time each element of the domain is named by only one name.

Of course not every model is covered (such models may be said to be *uncovered*). Indeed, some individuals will have to go nameless in any model in which the domain is uncountably infinite—e.g. set $\boldsymbol{R}$ of real numbers—since there is only a countable infinity of names. In other words, the existence of domains more numerous than the natural numbers means that there are models that cannot be covered. (See section 3 of the appendix on set theory for more about countable and uncountable infinities.)

But a model need not be infinite, or even large, in order to be uncovered. For example, let the domain of $\mathcal{M}$ be the set $\{0, 1\}$ and suppose $\mathcal{M}(\alpha) = 0$ for every name α. Then 1 is unnamed in $\mathcal{M}$ and so the model is uncovered.

One interesting thing about covered models is that in them the truth conditions for quantificational sentences can be reduced to the truth conditions of their instances. Thus if $\mathcal{M}$ is covered the universal quantification $\forall x Fx$ is true if and only if each of its instances Fa_0, Fa_1, Fa_2, . . . is true, while the existential quantification $\exists xFx$ is true if and only if at least one of the instances Fa_0, Fa_1, Fa_2, . . . is true. In general, where $\mathcal{M}$ is a covered model we have the following propositions (which we prove in chapter 9).

- $\models_{\mathcal{M}} \forall xA$ if and only if for every name α, $\models_{\mathcal{M}} A(x/\alpha)$.

- $\models_{\mathcal{M}} \exists xA$ if and only if for some name α, $\models_{\mathcal{M}} A(x/\alpha)$.

EXERCISES

7.5. Give an example of an *un*covered model whose domain consists of the set of natural numbers.

7.6. Consider the set $\{\exists xFx, \neg Fa_0, \neg Fa_1, \neg Fa_2, \ldots\}$. Is it satisfiable? Does the set have a model that is covered?

3. Consequence, validity, equivalence, and satisfiability

The logical concepts of consequence, validity, equivalence, and satisfiability were introduced in chapter 1 and discussed further in chapter 3. We continue the conventions used before, where "$A \models B$" abbreviates "$\{A\} \models B$", and so on.

The content of these concepts has greatly increased. The notion of a model in predicate logic differs vastly from that in sentential logic, and as a result the catalogue of consequences, etc., extends that for sentential logic.

An important feature of predicate logic is that in general the method of truth tables is not available to test sentences and sets of sentences for their logical properties. To show that a sentence or set of sentences has or does not have this-or-that logical property, other means must be employed.

On the other hand, when it comes to demonstrating *non*-consequence, *in*validity, *non*-equivalence, or satisfiability in predicate logic, we look for a model. That is, to show that a sentence A is not a consequence of a set Γ, it is sufficient to describe a model in which A is false while all of Γ is true; and similarly for the other notions. Intuition and ingenuity are often needed.

Exercises for the preceding two sections have already illustrated this, as have, in effect, some of the translation exercises in chapter 6. In order to fix ideas and formalize practice, we give a couple of illustrations here.

First, let us describe a model that shows that ∃xFx does not imply ∀xFx. A counterexample in this case is easy to find. Let $\mathcal{M}$ be a model whose domain D is the set $\mathbf{N}$ of natural numbers and where the value of F in $\mathcal{M}$ is the set of odd numbers. In other words, $D = \{0, 1, 2, \ldots\}$ and $\mathcal{M}(\mathrm{F}) = \{\langle n\rangle: n \text{ is odd}\}$. Now we see that ∃xFx is true in $\mathcal{M}$, since relative to this model what the sentence says is that at least one natural number is odd, whereas ∀xFx is false in $\mathcal{M}$ inasmuch as it says that every natural number is odd.

Especially for the purposes of the exercises to follow, readers should note that to establish non-consequence, etc., it is enough simply to describe a model that does the job; explanatory comment is unnecessary. For example, the following display is enough for the purpose of showing that ∃xFx does not imply ∀xFx.

> Let $\mathcal{M}$ be a model with domain D, where $D = \{0, 1, 2, \ldots\}$ and $\mathcal{M}(\mathrm{F}) = \{1, 3, 5, \ldots\}$.

(Notice that, technically, we should write "$\{\langle 1\rangle, \langle 3\rangle, \langle 5\rangle, \ldots\}$", and not "$\{1, 3, 5, \ldots\}$". The difference is insignificant.)

Another illustration is provided by the following counterexample to the claim that ∀x(Fx → Gx) is a consequence of {∀x(Fx → Hx), ∀x(Gx → Hx)}.

> Let $\mathcal{M}$ be a model with domain D, where D = the set of animals, $\mathcal{M}(\mathrm{F})$ = the set of cats, $\mathcal{M}(\mathrm{G})$ = the set of dogs, and $\mathcal{M}(\mathrm{H}) = D$ (i.e. the set of animals).

(All cats are animals; so are all dogs. But not all cats are dogs.)

EXERCISES

7.7. For each of the following sentences, describe a model showing it to be invalid.

(1) Fa → ∀xFx

(2) $\exists xFx \rightarrow Fa$

(3) $\forall x(Fx \vee Gx) \rightarrow (\forall xFx \vee \forall xGx)$

(4) $(\exists xFx \wedge \exists xGx) \rightarrow \exists x(Fx \wedge Gx)$

(5) $\exists xFx \rightarrow \forall xFx$

(6) $\exists x(Fx \wedge Gx) \rightarrow \exists x(Fx \wedge \neg Gx)$

(7) $\forall x(Fx \rightarrow Gx) \rightarrow \exists x(Fx \wedge Gx)$

(8) $(\forall xFx \rightarrow \forall xGx) \rightarrow \forall x(Fx \rightarrow Gx)$

(9) $\exists x(Fx \rightarrow Gx) \rightarrow \exists x(Fx \wedge Gx)$

(10) $(\exists xFx \rightarrow \exists xGx) \rightarrow \forall x(Fx \rightarrow Gx)$

(11) $\forall x(Fa \rightarrow Fx)$

(12) $\forall x(Fx \rightarrow Fa)$

(13) $\forall x\forall y(Fx \rightarrow Fy)$

(14) $\forall x(Fx \rightarrow \forall xFx)$

(15) $\forall x(\exists xFx \rightarrow Fx)$

(16) $\forall xRxx \rightarrow \forall x\forall yRxy$

(17) $\exists x\exists yRxy \rightarrow \exists xRxx$

(18) $\forall x\exists yRxy \rightarrow \exists y\forall xRxy$

(19) $\forall x\forall y\forall z(Rxy \rightarrow Rxz)$

(20) $\forall x\forall y\forall z\forall w(Rxz \rightarrow Ryw)$

7.8. Recall from exercise 7.4 the properties of binary relations and the sentences (a)–(l) that express them. Use models for each of the following problems.

(1) Give an example of a reflexive transitive relation. That is, show that {(a), (i)} is satisfiable.

(2) Give an example of an irreflexive transitive relation. That is, show that {(b), (i)} is satisfiable.

(3) Give an example of a symmetric transitive relation. That is, show that {(f), (i)} is satisfiable.

(4) Give an example of an asymmetric transitive relation. That is, show that {(g), (i)} is satisfiable.

(5) Give an example of a reflexive symmetric relation. That is, show that {(a), (f)} is satisfiable.

(6) Give an example of a symmetric asymmetric relation. That is, show that {(f), (g)} is satisfiable.

(7) Give an example of a transitive intransitive relation. That is, show that {(i), (j)} is satisfiable.

(8) Give an example of an *equivalence* relation (reflexive, symmetric, and transitive). That is, show that {(a), (f), (i)} is satisfiable.

(9) Give an example of a strongly connected equivalence relation. That is, show that {(a), (f), (h), (i)} is satisfiable.

(10) Give an example of an irreflexive serial transitive relation. That is, show that {(b), (e), (i)} is satisfiable.

(11) Give an example of a reflexive euclidean relation. That is, show that {(a), (k)} is satisfiable.

(12) Give an example of a non-reflexive transitive relation. That is, show that (i) does not imply (a).

(13) Give an example of a non-irreflexive transitive relation. That is, show that (i) does not imply (b).

(14) Give an example of a non-symmetric transitive relation. That is, show that (i) does not imply (f).

(15) Give an example of a non-asymmetric transitive relation. That is, show that (i) does not imply (g).

(16) Give an example of a non-reflexive symmetric relation. That is, show that (f) does not imply (a).

(17) Give an example of a non-irreflexive symmetric relation. That is, show that (f) does not imply (b).

(18) Give an example of a non-symmetric reflexive transitive relation. That is, show that {(a), (i)} does not imply (f).

(19) Give an example of a non-reflexive symmetric transitive relation. That is, show that {(f), (i)} does not imply (a).

(20) Give an example of a non-transitive reflexive symmetric relation. That is, show that {(a), (f)} does not imply (i).

(21) Give an example of a non-reflexive weakly dense relation. That is, show that (l) does not imply (a).

(22) Give an example of a non-connected equivalence (reflexive, symmetric, transitive) relation. That is, show that {(a), (f), (i)} does not imply (h).

(23) Give an example of a non-transitive euclidean relation. That is, show that (k) does not imply (i).

(24) Give an example of a non-euclidean transitive relation. That is, show that (i) does not imply (k).

(25) Give an example of a non-universal strongly connected relation. That is, show that (h) does not imply (c).

7.9. Use models for each of the following problems.

(1) Show that $\{\exists x Fx, \exists x \neg Fx\}$ is satisfiable.

(2) Show that $\{\forall x(Fx \rightarrow Gx), \forall x(Fx \rightarrow \neg Gx)\}$ is satisfiable.

(3) Show that $\{\exists x \neg Fx, Fa_0, Fa_1, Fa_2, \ldots\}$ is satisfiable (compare exercise 7.6).

(4) Show that $\{Fa_0, Fa_1, Fa_2, \ldots\}$ does not imply $\forall x Fx$.

7.10. Consider again the three sentences in part (10) of exercise 7.8. How many elements are there in the domain of the model you described for these sentences? What is the smallest possible number of elements in the domain of any model for these sentences?

4. Variants of a model

We introduce now an idea that plays an important role in stating the truth conditions of quantificational sentences—that of a variant of a model.

Let $\mathcal{M}$ be a model with its domain and assignment of values to the names and predicates. Then by $\mathcal{M}^{\alpha}_{d}$ we mean the model just like $\mathcal{M}$ except possibly that $\mathcal{M}^{\alpha}_{d}$ assigns to the name α the element d of the domain. The model $\mathcal{M}^{\alpha}_{d}$ is called the *α-variant* of $\mathcal{M}$ in which α's value is the element d in $\mathcal{M}$'s domain.

For example, let $\mathcal{M}$ be a model that has the set **N** of natural numbers as domain and evaluates the names thus: $\mathcal{M}(a_n) = n$, for each natural number n. So in $\mathcal{M}$ the value of a_0 is 0, the value of a_1 is 1, the value of a_2 is 2, and so on. For this example, suppose that the value of the predicate F is the set $\{1, 3, 5, \ldots\}$ of odd natural numbers, and the value of all the other predicates is the empty relation, $\varnothing$.

Consider now the model $\mathcal{M}^{a_0}_{1}$, the a_0-variant of $\mathcal{M}$ in which a_0 takes the value 1. This variant model agrees completely with the model $\mathcal{M}$ with respect to the values of all the predicates and of all the names except for a_0. Whereas the value of a_0 in $\mathcal{M}$ is the number 0, the value of a_0 in $\mathcal{M}^{a_0}_{1}$ is the number 1—i.e. $\mathcal{M}(a_0) = 0$, but $\mathcal{M}^{a_0}_{1}(a_0) = 1$.

To see the effect on a sentence of changing the value of a name in a model, consider Fa_0. Relative to the model $\mathcal{M}$, this sentence means that 0 is an odd number—i.e. that $\mathcal{M}(a_0)$ is in $\mathcal{M}(F)$—which is false. But relative to the model $\mathcal{M}^{a_0}_{1}$, the sentence means that 1 is an odd number—i.e. that $\mathcal{M}^{a_0}_{1}(a_0)$ is in $\mathcal{M}^{a_0}_{1}(F)$—which is true. Notice that for every n other than 0 the sentence Fa_n has the same truth value in both models $\mathcal{M}$ and $\mathcal{M}^{a_0}_{1}$. Indeed, it should be clear that so long as a sentence does not contain the name a_0, the sentence will have the same truth value in $\mathcal{M}$ as it does in $\mathcal{M}^{a_0}_{1}$.

It should be emphasized that an α-variant of a model $\mathcal{M}$ does not have to alter the value in $\mathcal{M}$ of the name α. If $\mathcal{M}$ assigns d to α already, then the model $\mathcal{M}^{\alpha}_{d}$ does not differ at all from $\mathcal{M}$ itself. That is to say, if $\mathcal{M}(\alpha) = d$, then $\mathcal{M}^{\alpha}_{d} = \mathcal{M}$.

To see the usefulness of variant models, consider this statement: for some $d \in \boldsymbol{N}$, Fa is true in $\mathcal{M}^{\mathrm{a}}_{d}$. This means there is at least one natural number that can be assigned to the name a with the result that the sentence Fa is true. Intuitively, this is just what the sentence ∃xFx means with respect to the model $\mathcal{M}$. So the displayed statement precisely gives the truth conditions of the sentence ∃xFx in the model $\mathcal{M}$. In other words:

∃xFx is true in the model $\mathcal{M}$ iff for some $d \in \boldsymbol{N}$, Fa is true in $\mathcal{M}^{\mathrm{a}}_{d}$.

Likewise, consider: for every $d \in \boldsymbol{N}$, Fa is true in $\mathcal{M}^{\mathrm{a}}_{d}$. According to this, no matter what natural number is assigned to the name a, the sentence Fa is true. The displayed statement gives precisely the truth conditions of the universal quantification ∀xFx in the model $\mathcal{M}$. In other words:

∀xFx is true in the model $\mathcal{M}$ iff for every $d \in \boldsymbol{N}$, Fa is true in $\mathcal{M}^{\mathrm{a}}_{d}$.

This is the principal use to which we put the notion of variance: to formulate the truth conditions of quantificational sentences. Here is the definition:

DEFINITION 7.3. *Variants of a model.* The model $\mathcal{M}^{\alpha}_{d}$ is the α-variant of the model $\mathcal{M}$ in which the name α is assigned the value d (in $\mathcal{M}$'s domain). That is, the models $\mathcal{M}^{\alpha}_{d}$ and $\mathcal{M}$ have the same domain and assign the same values to all predicates; and for every name β:

$\mathcal{M}^{\alpha}_{d}(\beta) = d$ if α and β are the same name;

and

$\mathcal{M}^{\alpha}_{d}(\beta) = \mathcal{M}(\beta)$ if α and β are different names.

To see that this definition accords with the explanation of variance above, compare the models $\mathcal{M}$ and $\mathcal{M}^{\mathrm{a}_0}_{1}$ as defined above. Then for every name a_n:

$$\mathcal{M}^{\mathrm{a}_0}_{1}(\mathrm{a}_n) = 1 \text{ if } \mathrm{a}_0 = \mathrm{a}_n\text{; and } \mathcal{M}^{\mathrm{a}_0}_{1}(\mathrm{a}_n) = \mathcal{M}(\mathrm{a}_n) \text{ (i.e. } n) \text{ if } \mathrm{a}_0 \neq \mathrm{a}_n.$$

Note a curious form of variance, models like:

$$\mathcal{M}^{\beta}_{\mathcal{M}(\alpha)}$$

The model $\mathcal{M}^{\beta}_{\mathcal{M}(\alpha)}$ is the β–variant of $\mathcal{M}$ that assigns to β whatever it is that $\mathcal{M}$ assigns to α. In short: $\mathcal{M}^{\beta}_{\mathcal{M}(\alpha)}(\beta) = \mathcal{M}(\alpha)$. (Notice that the remark above about vacuous variance can be expressed even more succinctly by saying that $\mathcal{M}^{\alpha}_{\mathcal{M}(\alpha)} = \mathcal{M}$.) We need variant models of this sort in connection with some semantic theorems.

Finally, observe that formation of variant models can be iterated. For instance, from the variant model $\mathcal{M}^{\alpha}_{d}$ one can form the further variant $\mathcal{M}^{\alpha,\beta}_{d,e}$. This is the β-variant of $\mathcal{M}^{\alpha}_{d}$ in which β takes the value e, where $\mathcal{M}^{\alpha}_{d}$ itself, of course, is the α-variant of the model $\mathcal{M}$ in which α takes the value d. (It should be noted that d and e need not be distinct individuals, and α and β may be the same name.)

EXERCISES

7.11. Describe a–variants of the models in your solutions for exercise 7.3.

5. Truth in a model

The definition of the notion of truth in a model should by now be easy to understand. A few preliminary comments may nevertheless be useful.

To begin, as we define it, the notion of truth applies exclusively to *sentences*. (Open formulas—those containing free variables—are neither true nor false relative to a model: they simply are not truth-valued.)

As before, in chapter 3, the definition of truth in a model is recursive. Truth values of atomic sentences are given initially, in terms of the values of their names and predicates, and truth conditions of complex sentences are determined by the truth conditions of smaller sentences.

The truth conditions of molecular sentences—negations, conjunctions, etc.—are specified exactly the way they were in sentential logic. Indeed, for such sentences evaluation by truth tables still suffices.

It is the atomic sentences and the universal and existential quantifications that get special treatment. We indicated in section 1 how the meaning of an atomic sentence is to be accounted for in terms of the meanings of its predicate and names, and we do so again formally in the definition below.

We want to state the truth conditions of a quantification in terms of those of a smaller sentence. Since the matrix A of a quantification ∀xA or ∃xA is not in general a sentence (it usually contains free occurrences of x), it is necessary to make the matrix A into a sentence $A(x/\alpha)$ by replacing all free occurrences of x by a new name α, i.e. one that does not already occur in A. Then we consider the truth conditions of the sentence $A(x/\alpha)$ with respect to possible values of α, while of course keeping fixed the values of the remaining names and predicates.

For example, to determine the truth values of the universal quantification ∀xFx in a model $\mathcal{M}$, we first take the matrix Fx and replace all free occurrences of the variable x by a name a. The result is (Fx)(x/a), i.e. Fa. Then we consider the truth values of Fa in the model for all possible values of the name a; that is, we check whether Fa is true in the model $\mathcal{M}^{\alpha}_{d}$ for every element *d* in the domain of $\mathcal{M}$. In brief, the sentence ∀xFx is true in a model $\mathcal{M}$ just in case the sentence Fa is true in $\mathcal{M}^{\alpha}_{d}$ for every *d* in the domain of $\mathcal{M}$.

In a similar fashion, the existential quantification ∃xFx is accounted true in $\mathcal{M}$ if and only if Fa is true in $\mathcal{M}^{\alpha}_{d}$ for some *d* in the domain of $\mathcal{M}$.

Here now is the definition of truth in a model. We continue to use the double turnstile, "$\models$", as we did in chapter 3.

DEFINITION 7.4. *Truth in a model.* Let $\mathcal{M}$ be a model with a domain D, let $\mathbb{P}$ be an n-place predicate, and let A and B be formulas. In clauses (2)–(6) below, A and B are assumed to be sentences. In (7) and (8), A has at most the variable x free and the name α is new to A.

(1) $\models_{\mathcal{M}} \mathbb{P}\alpha_1 \ldots \alpha_n$ iff $\langle \mathcal{M}(\alpha_1), \ldots, \mathcal{M}(\alpha_n)\rangle \in \mathcal{M}(\mathbb{P})$.

(2) $\models_{\mathcal{M}} \neg A$ iff $\not\models_{\mathcal{M}} A$.

(3) $\models_{\mathcal{M}} A \wedge B$ iff $\models_{\mathcal{M}} A$ and $\models_{\mathcal{M}} B$.

(4) $\models_{\mathcal{M}} A \vee B$ iff $\models_{\mathcal{M}} A$ or $\models_{\mathcal{M}} B$.

(5) $\models_{\mathcal{M}} A \rightarrow B$ iff if $\models_{\mathcal{M}} A$ then $\models_{\mathcal{M}} B$.

(6) $\models_{\mathcal{M}} A \leftrightarrow B$ iff $\models_{\mathcal{M}} A$ if and only if $\models_{\mathcal{M}} B$.

(7) $\models_{\mathcal{M}} \forall xA$ iff for every $d \in D$, $\models_{\mathcal{M}^{\alpha}_{d}} A(x/\alpha)$.

(8) $\models_{\mathcal{M}} \exists xA$ iff for some $d \in D$, $\models_{\mathcal{M}^{\alpha}_{d}} A(x/\alpha)$.

For the sake of clarity, let us also state the conditions for the *falsity* of universal and existential quantifications:

(7′) $\not\models_{\mathcal{M}} \forall xA$ iff for some $d \in D$, $\not\models_{\mathcal{M}^{\alpha}_{d}} A(x/\alpha)$.

(8′) $\not\models_{\mathcal{M}} \exists xA$ iff for every $d \in D$, $\not\models_{\mathcal{M}^{\alpha}_{d}} A(x/\alpha)$.

As in the case of sentential logic, the definition of truth can be viewed as a formal scheme for translation. Let us illustrate this by applying the definition to the sentence $\exists xFx$, relative to a model $\mathcal{M}$ in which the domain is the set $\boldsymbol{N}$ of natural numbers and the value of F is the set of odd natural numbers, i.e. $\{\langle n\rangle: n \text{ is odd}\}$. Thus, relative to this model, the sentence means that some natural number is odd (compare the example in the preceding section).

- Clause (8) of definition 7.4 tells us that the sentence $\exists xFx$ is true in the model $\mathcal{M}$ if and only if for some d in the domain of $\mathcal{M}$, the sentence $(Fx)(x/a)$ is true in the model $\mathcal{M}^{a}_{d}$.

- Because the domain is $\boldsymbol{N}$ and that $(Fx)(x/a)$ is the sentence Fa, we can simplify this to the statement that $\exists xFx$ is true in the model $\mathcal{M}$ if and only if for some natural number n, Fa is true in $\mathcal{M}^{a}_{n}$.

- Now we apply clause (1) of the definition. Thus $\exists xFx$ is true in the model $\mathcal{M}$ if and only if for some natural number n, $\langle \mathcal{M}^{a}_{n}(a)\rangle \in \mathcal{M}^{a}_{n}(F)$.

- Since in $\mathcal{M}^{a}_{n}$ the value of a is n and the value of F is $\{\langle n\rangle: n \text{ is odd}\}$, we conclude that $\exists xFx$ is true in the model $\mathcal{M}$ if and only if for some natural number n, $\langle n\rangle$ is in $\{\langle n\rangle: n \text{ is odd}\}$.

• In plainer language, we arrive at what we may call *the final result of the truth conditions* of the sentence: ∃xFx is true in the model $\mathcal{M}$ if and only if for some natural number n, n is odd.

Note that the final result of the truth conditions of a sentence is simply a *translation* of the sentence relative to the model in question.

Here is another illustration, using the sentence ∀x∃yRxy and the model $\mathcal{M}$ with the set of natural numbers as domain. We take the value of R to be the less-than relation (i.e. $\mathcal{M}(\mathrm{R}) = \{\langle m, n\rangle: m < n\}$). Relative to this model, then, the sentence means that every natural number is less than some natural number or other.

• By clause (7) of definition 7.4, ∀x∃yRxy is true in the model $\mathcal{M}$ if and only if for every d in the domain of $\mathcal{M}$, the sentence (∃yRxy)(x/a) is true in $\mathcal{M}^{\mathrm{a}}_{d}$.

• Noting that the domain is $\boldsymbol{N}$ and that (∃yRxy)(x/a) is the sentence ∃yRay, we can simplify this to the statement that ∀x∃yRxy is true in the model $\mathcal{M}$ if and only if for every natural number m,∃yRay is true in $\mathcal{M}^{\mathrm{a}}_{m}$.

• Now we use clause (8) of definition 7.4 to see that ∀x∃yRxy is true in the model $\mathcal{M}$ if and only if for every natural number m, for some d in the domain, the sentence (Ray)(y/b) is true in $\mathcal{M}^{\mathrm{a}\ \mathrm{b}}_{m,\,d}$.

• This result in turn can be expressed equally well by saying that ∀x∃yRxy is true in the model $\mathcal{M}$ if and only if for every natural number m, for some natural number n, the sentence Rab is true in $\mathcal{M}^{\mathrm{a}\ \mathrm{b}}_{m,\,n}$.

• Finally, we apply clause (1) of the definition. Thus ∀x∃yRxy is true in the model $\mathcal{M}$ if and only if for every natural number m, for some natural number n, $\langle \mathcal{M}^{\mathrm{a}\ \mathrm{b}}_{m,\,n}(\mathrm{a}), \mathcal{M}^{\mathrm{a}\ \mathrm{b}}_{m,\,n}(\mathrm{b})\rangle \in \mathcal{M}^{\mathrm{a}\ \mathrm{b}}_{m,\,n}(\mathrm{R})$.

• But in $\mathcal{M}^{a\ b}_{m\ n}$ the value of a is m, the value of b is n, and the value of R is $\{\langle m, n\rangle: m < n\}$. So we conclude that ∀x∃yRxy is true in the model $\mathcal{M}$ if and only if for every natural number m, for some natural number n, the pair $\langle m, n\rangle$ is in $\{\langle m, n\rangle: m < n\}$.

• In plainer language, we reach the final result of the truth conditions of the sentence: ∀x∃yRxy is true in the model $\mathcal{M}$ if and only if for every natural number m, for some natural number n, $m < n$—which is how we translated the sentence to begin with.

We said in section 1 that a set of one-tuples of individuals can be regarded as just the set of individuals themselves. Notice how this works out for the interpretation of sentences containing one-place predicates. In the example above with the model $\mathcal{M}$ and predicate F the truth condition emerges as

for some natural number n, $\langle n\rangle$ is in $\{\langle n\rangle: n \text{ is odd}\}$,

which is tantamount to

for some natural number n, n is in $\{n: n \text{ is odd}\}$.

In any case each yields the truth condition: n is odd.

Postscript on zero-tuples and truth values. What about when the truth definition is applied to a sentence that is simply a zero-place predicate? Let P be such a sentence. By the definition of a model, $\mathcal{M}$(P) is either the set $\{\langle\ \rangle\}$ or the empty set ∅. Let us say the former. Then:

$_{\mathcal{M}}$P iff $\langle\ \rangle \in \mathcal{M}(\text{P})$
—definition 7.4, part (1);
iff $\langle\ \rangle \in \{\langle\ \rangle\}$
—definition of $\mathcal{M}$.

In other words, P is true in $\mathcal{M}$—since the zero-tuple $\langle\ \rangle$ is a member of the set $\{\langle\ \rangle\}$. By contrast, suppose the value of P in $\mathcal{M}$ is not $\{\langle\ \rangle\}$ but the empty set. Then the truth conditions for P work out this way:

$$\begin{aligned}\models_{\mathcal{M}} \mathrm{P} \text{ iff } & \langle\ \rangle \in \mathcal{M}(\mathrm{P}) \\ & \quad \text{—definition 7.4, part (1);} \\ \text{iff } & \langle\ \rangle \in \varnothing \\ & \quad \text{—definition of } \mathcal{M}.\end{aligned}$$

In other words, P is false in $\mathcal{M}$, since the zero-tuple $\langle\ \rangle$ is not a member of the empty set.

These truth conditions show why it is plausible to identify the zero-place predicates with the atomic sentences in sentential logic. They also suggest why we might identify the set $\{\langle\ \rangle\}$ with the truth value *truth* ($\top$) and, by the same token, the empty set with *falsity* ($\bot$).

EXERCISES

7.12. As above, let $\mathcal{M}$ be a model that has the set of natural numbers as domain and interprets the predicate F as the set of odd natural numbers. Use the definition of truth to work through to the final result of the truth conditions of the sentence $\forall$xFx. (Compare the example in the preceding section.)

7.13. Where, as above, $\mathcal{M}$ has the set of natural numbers as domain and $\mathcal{M}$(R) is the less-than relation, state the final result of the truth conditions of the sentence $\exists$y$\forall$xRxy.

7.14. Let $\mathcal{M}$ be a model in which the domain is the set of natural numbers, $\mathcal{M}$(a) is 0, $\mathcal{M}$(F) is the set of even numbers, $\mathcal{M}$(G) is the set of odd numbers, and $\mathcal{M}$(R) is the less-than relation. State the final result of the truth conditions of the following sentences.

(1) $\forall$x(Fx $\rightarrow$ Rxa)

(2) $\exists$x(Gx $\wedge$ Rax)

(3) $\forall$x(Fx $\rightarrow$ $\exists$y(Gy $\wedge$ Rxy))

(4) $\forall x(Fx \rightarrow \exists y(Gy \wedge Ryx))$

(5) $\exists x(Fx \wedge \forall y(Gy \rightarrow \neg Ryx))$

7.15. Where the domain of a model $\mathcal{M}$ is the set of barbers and $\mathcal{M}(S)$ relates two individuals just in case the first shaves the second, state the final result of the truth conditions in $\mathcal{M}$ of the following sentences.

(1) $\exists x\forall y(\neg Syy \rightarrow Sxy)$

(2) $\exists x\forall y(Sxy \rightarrow \neg Syy)$

(3) $\exists x\forall y(Sxy \leftrightarrow \neg Syy)$

(Compare exercise 6.14, page 246.)

7.16. Where $\mathcal{M}$ is any model, use the definition of truth to work through the truth conditions of the following sentences.

(1) $\forall xFx \rightarrow Fa$ (2) $Fa \rightarrow \exists xFx$

Is there a model in which either (1) or (2) is false?

6. Consequence, validity, equivalence, and satisfiability—again

We said above in section 3 that, generally, in predicate logic there is nothing like the truth table method test sentences and sets of sentences for their logical properties. (Truth tables can be used in some cases, but by no means always.)

Let us show here just for example that $\forall xFx$ implies $\exists xFx$, using the definition of truth in a model. We argue as follows.

Suppose that $\forall xFx$ is true in a model $\mathcal{M}$. By definition 7.4, this means that for every element d of the domain of the model $\mathcal{M}$, Fa is true in $\mathcal{M}_d^{\hat{a}}$. Because the domain of a model is non-empty, we infer that there is an element d of the domain of the model $\mathcal{M}$ such that Fa is true in $\mathcal{M}_d^{\hat{a}}$. By definition 7.4 again, this means that $\exists xFx$ is true in $\mathcal{M}$ itself. Thus we have shown that if $\forall xFx$ is true in $\mathcal{M}$ then $\exists xFx$ is too. But since $\mathcal{M}$ was picked arbitrarily, i.e. without any conditions on it, we are entitled to conclude generally that for every model $\mathcal{M}$, if ∀xFx is true in $\mathcal{M}$ then so is ∃xFx—i.e. every model of ∀xFx is a model of ∃xFx. Therefore ∀xFx implies ∃xFx, as we wished to show.

To demonstrate that a sentence is a consequence in predicate logic of a set of sentences, or that a sentence is valid in the language, or that sentences are equivalent, or that a set of sentences is *un*satisfiable, it is usually easiest to use a sound deductive system, such as the one developed in chapter 8. For example, here is a simple, three-line deduction showing that ∀xFx implies ∃xFx; compare it with the rather more complicated argument in the preceding paragraph.

{1}	(1)	∀xFx	P
{1}	(2)	Fa	UI, 1
{1}	(3)	∃xFx	EG, 2

This of course goes beyond the proper subject of this chapter to the deductive system in chapter 8.

EXERCISES

7.17. Explain why the following two sentences are valid (compare exercise 7.16).

(1) ∀xFx → Fa (2) Fa → ∃xFx

7.18. Here for the record are a number of examples of consequence (valid arguments), valid sentences, and pairs of equivalent sentences. Each item illustrates a general principle. For example, (1) exemplifies the fact that any argument of the form ∀xA / A(x/α) is valid.

Select some of the items and use the definition of truth to substantiate the relevant claims. (In (21)–(30), P is a zero-place predicate.)

(1) ∀xFx / Fa

(2) Fa / ∃xFx

(3) ∀x∀yRxy / Rab

(4) $\forall x \forall y Rxy$ / Raa

(5) Rab / $\exists x \exists y Rxy$

(6) Raa / $\exists x \exists y Rxy$

(7) $\forall x(Fx \wedge Gx)$, $\forall xFx \wedge \forall xGx$

(8) $\exists x(Fx \vee Gx)$, $\exists xFx \vee \exists xGx$

(9) $\forall xFx \vee \forall xGx$ / $\forall x(Fx \vee Gx)$

(10) $\exists x(Fx \wedge Gx)$ / $\exists xFx \wedge \exists xGx$

(11) $\forall x(Fx \to Gx)$ / $\forall xFx \to \forall xGx$

(12) $\forall x(Fx \to Gx)$ / $\exists xFx \to \exists xGx$

(13) $\exists x(Fx \to Gx)$, $\forall xFx \to \exists xGx$

(14) $\forall xFx \wedge \exists xGx$ / $\exists x(Fx \wedge Gx)$

(15) $\forall x(Fx \vee Gx)$ / $\exists xFx \vee \forall xGx$

(16) $\forall x \forall y Rxy$, $\forall y \forall x Rxy$

(17) $\exists x \exists y Rxy$, $\exists y \exists x Rxy$

(18) $\forall x \forall y Rxy$ / $\forall x Rxx$

(19) $\exists x Rxx$ / $\exists x \exists y Rxy$

(20) $\exists x \forall y Rxy$ / $\forall y \exists x Rxy$

(21) $\forall xP$, P

(22) $\exists xP$, P

(23) $\forall x(Fx \wedge P)$, $\forall xFx \wedge P$

(24) $\exists x(Fx \wedge P)$, $\exists xFx \wedge P$

(25) $\forall x(Fx \vee P)$, $\forall xFx \vee P$

(26) $\exists x(Fx \vee P)$, $\exists xFx \vee P$

(27) $\forall x(Fx \to P)$, $\exists xFx \to P$

(28) $\exists x(Fx \rightarrow P),\ \forall xFx \rightarrow P$

(29) $\forall x(P \rightarrow Fx),\ P \rightarrow \forall xFx$

(30) $\exists x(P \rightarrow Fx),\ P \rightarrow \exists xFx$

(31) $\forall x(Fx \rightarrow Fx)$

(32) $\exists x(Fx \rightarrow Fx)$

(33) $\forall x(Fx \rightarrow \exists xFx)$

(34) $\exists x(Fx \rightarrow \forall xFx)$

(35) $\forall x(\forall xFx \rightarrow Fx)$

(36) $\exists x(\exists xFx \rightarrow Fx)$

(37) $\exists x(Fx \rightarrow \exists xFx)$

(38) $\exists x(\forall xFx \rightarrow Fx)$

(39) $\exists x(\forall xFx \leftrightarrow Fx)$

(40) $\exists x(\exists xFx \leftrightarrow xFx)$

(41) $\forall x\exists y(Fx \rightarrow Fy)$

(42) $\forall x\exists y(Fy \rightarrow Fx)$

(43) $\exists x\exists y(Fx \rightarrow Fy)$

(44) $\exists x\forall y(Fx \rightarrow Fy)$

(45) $\exists x\forall y(Fy \rightarrow Fx)$

(46) $\forall x(Rxx \rightarrow Rxx)$

(47) $\forall x\forall y(Rxy \rightarrow Rxy)$

(48) $\forall x\forall y\exists z(Rxy \rightarrow Rxz)$

(49) $\forall x\exists y\forall z(Rxz \rightarrow Ryz)$

(50) $\forall x\exists y\forall z\exists w(Rxz \rightarrow Ryw)$

(51) $\forall xFx,\ \neg \exists x \neg Fx$

(52) $\exists xFx,\ \neg \forall x \neg Fx$

7. Metatheorems

In chapter 8 we introduce several new rules of inference for dealing with universal and existential quantifications. The object of this section is to pave the way for these rules by way of propositions to the effect that the rules to come are sound, i.e. that they preserve consequence.

As we have seen in various examples, a universal quantification implies any of its instances. For example, ∀xFx implies Fa, Fb, Fc, and so on. In short, ∀xFx implies Fα, no matter which name α is. In like fashion, an existential quantification is a consequence of any one of it instances, so that each of Fa, Fb, Fc, . . . implies ∃xFx. We can state these facts generally:

(a) $\forall x A \models A(x/\alpha)$, for every name α.

(b) $A(x/\alpha) \models \exists x A$, for every name α.

Proofs of these propositions seem hardly necessary when one considers examples such as those above. If ∀xFx is true in a model, then the predicate F is, so to speak, true of each individual in the domain, and hence each instance Fa, Fb, Fc, . . . is also true in the model. And similarly for the implications from the instances to ∃xFx.

Propositions (a) and (b) depend on a fundamental property called *compositionality*, which says that two sentences have the same truth value if they differ from each other only in occurrences of names that have the same value.

An example in English will illustrate this. Consider:

Robert LeRoy Parker led the Wild Bunch

Butch Cassidy led the Wild Bunch

These sentences differ only in that one has the name Robert LeRoy Parker where the other has Butch Cassidy. These names refer to the same person, so the two sentences have the same truth value, i.e. both are true or both are false.

Compositionality can be stated formally in this way:

$$\models_{\mathcal{M}} A(\beta/\alpha) \text{ iff } \models_{\mathcal{M}^{\beta}_{\mathcal{M}(\alpha)}} A.$$

Despite the seeming complexity of its formulation, the idea is simple. First, since $A(\beta/\alpha)$ is a sentence containing no occurrences of the name β, A itself is a sentence that results from substituting the name α for β (in zero or more places). (For this reason compositionality is sometimes called *substitutivity*.) Next, α denotes $\mathcal{M}(\alpha)$ in the model $\mathcal{M}$ (this is trivial) and β denotes *the same individual*, $\mathcal{M}(\alpha)$, in the model $\mathcal{M}^{\beta}_{\mathcal{M}(\alpha)}$. Therefore the sentence $A(\beta/\alpha)$ will have the same truth value in model $\mathcal{M}$ as does the sentence A in model $\mathcal{M}^{\beta}_{\mathcal{M}(\alpha)}$. It may help the reader if we state compositionality more elaborately—but equivalently—as follows.

$$\models_{\mathcal{M}^{\alpha}_{\mathcal{M}(\alpha)}} A(\beta/\alpha) \text{ iff } \models_{\mathcal{M}^{\beta}_{\mathcal{M}(\alpha)}} A(\beta/\beta).$$

We prove compositionality itself in chapter 9, but we can use the principle here to prove (a). (The reasoning for (b) is similar and is left as an exercise.)

1. Suppose $\forall xA$ is true in a model $\mathcal{M}$.
2. This means that, where β is a name not in A, the sentence $A(x/\beta)$ is true in $\mathcal{M}^{\beta}_{d}$ for every d in the domain.
3. So, where α is any name, $A(x/\beta)$ is true in the model $\mathcal{M}^{\beta}_{\mathcal{M}(\alpha)}$.
4. Hence by compositionality, the sentence $A(x/\beta)(\beta/\alpha)$ is true in $\mathcal{M}$ itself.
5. But $A(x/\beta)(\beta/\alpha)$ is the same sentence as $A(x/\alpha)$, since β is new to A.
6. In other words, where α is any name, $A(x/\alpha)$ is true in $\mathcal{M}$—as we wished to show.

It follows from (a) that if Γ implies $\forall xA$ then Γ implies $A(x/\alpha)$ for every name α. Similarly, from (b) we see that, for every name α, if Γ implies $A(x/\alpha)$ then Γ implies $\exists xA$. Thus we have proved parts (1) and (2) of proposition 7.5 below.

Two further propositions license two further new rules:

(c) If $\Gamma \models A(x/\alpha)$ then $\Gamma \models \forall xA$—provided that the name α is new to $\forall xA$ and to Γ.

(d) If $\Gamma \models \exists xA$ and $\Delta \cup \{A(x/\alpha)\} \models B$, then $\Gamma \cup \Delta \models B$—provided that the name α is new to $\exists xA$, to B, and to Δ.

An example makes it is easy to see why (c) is correct: The sentence Raa ("Amy rubs Amy") is a consequence of $\forall x\forall yRxy$ ("Everyone rubs everyone"), and the name a forms no part of $\forall x\forall yRxy$. As a is likewise not in the sentence $\forall xRxx$, its appearance is in a certain sense arbitrary, and hence $\forall xRxx$ ("Everyone is a self-rubber") is also a consequence of $\forall x\forall yRxy$. (Note that this does *not* mean that Raa itself implies $\forall xRxx$, which it does not.)

Propositions (c) and (d) depend on the property of *relevance*, which says that models give the same truth values to a sentence A whenever they have the same domain and give the same values to all the names and predicates in A. In other words, where $\mathcal{M}$ and $\mathcal{M}^*$ are two models,

$$\models_{\mathcal{M}} A \text{ iff } \models_{\mathcal{M}^*} A$$

—provided that $\mathcal{M}$ and $\mathcal{M}^*$ share a common domain and agree on the values they assign to all the names and predicates in the sentence A. To put the point negatively, so long as $\mathcal{M}$ and $\mathcal{M}^*$ share the same domain it is *ir*relevant what meanings they give to names and predicates that are not in A. (Recall the analogous property for sentential logic discussed in chapters 3 and 5.)

Relevance is proved in chapter 9 (proposition 9.2). Meanwhile we can use it to argue for (c) and (d). For example, for (c):

1. Suppose that Γ implies a sentence $A(x/\alpha)$, where α is not in $\forall xA$ and not in any sentence in Γ.
2. Now suppose—to reach an absurdity—that Γ does not imply $\forall xA$.
3. So there is a model $\mathcal{M}$ in which all of Γ is true, but $\forall xA$ is false.
4. This means that for at least one d in the domain of $\mathcal{M}$, $A(x/\alpha)$ is false in the variant model M^{α}_{d}. (Note that we may use α for the instance since it is new to $\forall xA$.)
5. Observe that all the sentences in Γ are true in $\mathcal{M}^{\alpha}_{d}$, since none of them contain α, and the models $\mathcal{M}$ and $\mathcal{M}^{\alpha}_{d}$ have the same

domain and agree on the values of all other names and predicates.

6. So by our first assumption, $A(x/\alpha)$ is also true in $\mathcal{M}_d^\alpha$, which contradicts what we found in step 4.

Similar reasoning, using relevance, establishes (d) as well. This is left as an exercise.

We noticed in chapter 6 that the universal and existential quantifiers are interdefinable using negation, i.e. that the following equivalences hold.

(e) $\forall xA \simeq \neg\exists x\neg A.$

(f) $\exists xA \simeq \neg\forall x\neg A.$

These principles underlie the soundness of (the two parts of) another inference rule (see parts (5) and (6) of proposition 7.5 below). Let us demonstrate one of them, (e), now.

We let $\mathcal{M}$ be a model with domain D. Except for one step the reasoning is wholly by way of the definition of truth (7.4):

$\models_{\mathcal{M}} \forall x\, A$ iff for every $d \in D$, $\models_{\mathcal{M}_d^\alpha} A(x/\alpha)$
—definition 7.4, part (7);
iff it is not the case that for some $d \in D$, $\not\models_{\mathcal{M}_d^\alpha} A(x/\alpha)$
—[reasoning!];
iff it is not the case that for some $d \in D$, $\models_{\mathcal{M}_d^\alpha} \neg A(x/\alpha)$
—definition 7.4, part (2);
iff it is not the case that $\models_{\mathcal{M}_d^\alpha} \exists\neg A(x/\alpha)$
—definition 7.4, part (8);
iff $\models_{\mathcal{M}_d^\alpha} \neg\exists\neg A(x/\alpha)$
—definition 7.4, part (2).

The reasoning for (f) imitates that for (e), and we set it as an exercise.

PROPOSITION 7.5. *Soundness of the quantificational rules.*

(1) If $\Gamma \models \forall xA$, then for every name α, $\Gamma \models A(x/\alpha)$.

(2) If $\Gamma \models A(x/\alpha)$, for some name α, then $\Gamma \models \exists xA$.

(3) If $\Gamma \models A(x/\alpha)$, then $\Gamma \models \forall xA$—provided that the name α is new to Γ and to A.

(4) If $\Gamma \models \exists xA$ and $\Delta \cup \{A(x/\alpha)\} \models B$, then $\Gamma \cup \Delta \models B$—provided that the name α is new to $\exists xA$, to B, and to Δ.

(5) $\neg \forall xA \simeq \exists x \neg A$.

(6) $\neg \exists xA \simeq \forall x \neg A$.

Proof. The proofs for parts (1) and (3) were given above, with those for (2) and (4) left as exercises. For part (5), observe that from (e) we get immediately that

$$\neg \forall xA \simeq \neg \neg \exists x \neg A,$$

from which it follows that

$$\neg \forall xA \simeq \exists x \neg A.$$

Likewise, from (f) we get that $\neg \exists xA \simeq \forall x \neg A$; we leave the details as an exercise. □

EXERCISES

7.19. Check in the proof of (a) above—$\forall xA \models A(x/\alpha)$, for every name α—that the sentences $A(x/\beta)(\beta/\alpha)$ and $A(x/\alpha)$ are the same (compare exercise 6.20, page 258).

7.20. Prove part (2) of proposition 7.5 by using compositionality to argue for (b) above—i.e. that $A(x/\alpha) \models \exists xA$ for every name α.

7.21. Use relevance to argue for (d) above (equivalently, part (4) of proposition 7.5), i.e. that if $\Gamma \models \exists xA$ and $\Delta \cup \{A(x/\alpha)\} \models B$, then $\Gamma \cup \Delta \models B$—provided that the name α is new to $\exists xA$, to B, and to Δ.

7.22. Using the idea of truth in a model, argue for (f) above, i.e. that $\exists xA \simeq \neg \forall x \neg A$. Then show part (6) of proposition 7.5, i.e. that $\neg \exists xA \simeq \forall x \neg A$.

7.23. True or false:

a. The sentence $(\forall xFx \rightarrow \forall xGx) \rightarrow \forall x(Fx \rightarrow Gx)$ is false in a model in which F means is female, G means is a mother, and the domain is the set of humans.

b. If every instance $A(x/\alpha)$ of a universal quantification $\forall xA$ is true in a model, then the universal quantification is true in the model.

c. Every model of the set $\{\exists x(Fx \wedge Gx), \exists x(Fx \wedge \neg Gx)\}$ has at least two elements in its domain.

d. The sentence $\exists xFx \vee \exists x \neg Fx$ is valid.

e. The sentence $\forall x(\exists yFy \rightarrow Fx)$ is true in every model.

f. The sentences $\forall x \exists yRxy$ and $\exists y \forall xRxy$ are equivalent.

g. If any instance $A(x/\alpha)$ of an existential quantification $\exists xA$ is true in a model, then the existential quantification is true in the model.

h. The set of sentences $\{\forall x \neg Fx, Fa_0, Fa_1, Fa_2, \ldots\}$ is satisfiable.

i. If an existential quantification $\exists xA$ is true in a model, then at least one instance $A(x/\alpha)$ is true in the model.

j. If $A(x/\alpha)$ is valid and the name α is not in $\forall xA$, then $\forall xA$ is valid.

k. There is a formula A such that $\exists xA$ is valid and $\forall xA$ is not valid.

l. Every universal quantification is equivalent to at least one negation.

m. An existential quantification $\exists xA$ is valid only if at least one instance $A(x/\alpha)$ is valid.

n. The sentence $(\exists xFx \vee \exists xGx) \rightarrow \exists x(Fx \vee Gx)$ is false in a model in which F means is odd, G means is even, and the domain is the set of natural numbers.

o. There are unsatisfiable sentences that are false in at least one infinite model.

8

DEDUCTION IN PREDICATE LOGIC

IN THIS chapter we return to the formal side of logic and examine the language of predicate logic from the standpoint of a *deductive system*. As before, we introduce a set of *rules of inference* that characterize the system; the rules for predicate logic modestly but powerfully extend the collection in chapter 4.

The ideas of *deducibility* ($\Gamma \vdash A$), *theoremhood* ($\vdash A$), *deductive equivalence* ($A \sim B$), and *consistency* (Con Γ) are defined as they were earlier. What is different is their content: the relation of deducibility is far larger than it was, and likewise for the other logical notions.

Also as before, our goal is a system of inference that is both *sound* and *complete*. Thus we design the system so that:

(1) $\Gamma \models A$ iff $\Gamma \vdash A$.

(2) $\models A$ iff $\vdash A$.

(3) $A \simeq B$ iff $A \sim B$.

(4) Sat Γ iff Con Γ.

Readers may find it helpful to re-read the introduction and sections 1 and 2 of chapter 4.

In section 1 we review the ideas of a deduction and the definitions of deducibility, theoremhood, deductive equivalence, and consistency, and we introduce a new rule of inference for sentential logic (TT) and some quantificational *rules of inference*: UI, EG, UG, E, and QN. Then we illustrate the use of the rules—UI and EG (section 2), UG and E (section 3), and QN (section 4). In section 5 we discuss *strategies* for using the new rules, again with a number of examples. In section 6 we establish the *soundness* of the deductive system (completeness is demonstrated in chapter 9). So as in the case of sentential logic the system provides a means of proving consequence, validity, equivalence, and unsatisfiability, this time in predicate logic.

The quantificational rules are *generalized* in section 7, in section 8 we generalize the *rule of replacement*, and we end the chapter in section 9 with a discussion of the *interderivability of the quantificational rules*.

EFL 289 • 26 Jun 96

1. Deductions, deducibility, and rules of inference

The key notions will be familiar from chapter 4: A deduction is a finite sequence of lines of the form $\langle \Gamma, A \rangle$ determined by the rules of inference (definition 4.1, page 137), and a deduction of A from Γ is a deduction ending with a line $\langle \Gamma', A \rangle$, where Γ includes Γ'.

The concepts of deducibility, theoremhood, deductive equivalence, and consistency are also the same: A is deducible from Γ ($\Gamma \vdash A$) if and only if there is a deduction of A from Γ; A is a theorem ($\vdash A$) just when A is deducible from the empty set; A and B are deductively equivalent ($A \sim B$) just in case each is deducible from the other; and a set of sentences Γ is consistent (Con Γ) if and only if there is no sentence A such that both A and $\neg$ A are deducible from Γ. In other words, the definitions of these concepts in chapter 4 remain correct (pages 138 and 161, definitions 4.2, 4.3, 4.5, and 4.6).

What changes in moving from sentential to predicate logic is the rules of inference, in terms of which the concepts above are defined. The deductive system takes over all the rules of inference of sentential logic and adds further rules that deal with quantificational idioms.

Because we wish to concentrate on quantification we shall abbreviate as much as possible the rules of inference that have to do only with $\neg$, $\wedge$, $\vee$, $\rightarrow$, and $\leftrightarrow$. We do this by means of a new sentential rule of inference, TT, that covers most combinations of the rules for sentential logic in chapter 4. The rule is formulated in this way:

DEFINITION 8.1. *The rule TT.*

TT. $$\frac{\Gamma_1,\ A_1\ ;\ \ldots\ ;\Gamma_n,\ A_n}{\Gamma_1 \cup \ldots \cup \Gamma_n,\ A}$$ where A is deducible in sentential logic from $A_1, \ldots, A_n$ ($n \geq 0$)

The rule TT allows the inference of a sentence A from sentences $A_1, \ldots, A_n$ provided that A already follows from these sentences by means of the inference rules of sentential logic. Since in sentential logic deducibility is the same as consequence (i.e. the deductive system for sentential logic is sound and complete), the proviso that A is deducible from $A_1, \ldots, A_n$ in sentential logic can also be stated:

A is a consequence of $A_1, \ldots, A_n$ in sentential logic.

So application of the rule TT depends solely on whether A is a consequence of $A_1, \ldots, A_n$ in sentential logic—i.e. relative to models that assign truth values to the propositional atoms of the sentences involved. Since there will always be only a finite number of sentences involved in an application of the rule TT, the question of consequence can in principle always be settled, affirmatively or negatively, by means of a truth table. Hence the name "TT", for "truth tables".

In practical terms, the rule TT makes it possible to "telescope" inferences that would ordinarily require the use of several sentential rules. For an example, consider the following deduction of the theorem $A \vee (A \rightarrow B)$.

{1}	(1)	$\neg (A \vee (A \rightarrow B))$	P
{1}	(2)	C	TT, 1
{1}	(3)	$\neg C$	TT, 1
$\varnothing$	(4)	$A \vee (A \rightarrow B)$	RAA, 2, 3

A truth table suffices to show that the sentence C on line (2) is a consequence of (and hence deducible from) the sentence $\neg (A \vee (A \rightarrow B))$ on line (1); likewise for the inference from line (1) to $\neg C$ on line (3). (Of course, $\neg (A \vee (A \rightarrow B))$ implies both C and $\neg C$ because it is a contradiction.) Hence both inferences, from (1) to (2) and from (1) to (3), are licensed by the rule TT. Line (4) then follows in the usual way by RAA.

Let us not speculate here on what combination of rules of inference of sentential logic would be required to infer each of C and $\neg C$ from $\neg (A \vee (A \rightarrow B))$—a deduction in chapter 4 affords some clues. The point is that in making the inferences above by means of the rule TT, we need only to check that the sentences on line (2) and (3) are "truth table" consequences of the sentence on line (1).

It should be emphasized that the particular form of the sentences involved in the deduction above is immaterial. An exact reproduction of the deduction above will show that this sentence is a theorem:

$$\forall x \neg Fx \vee (\forall x \neg Fx \rightarrow \exists y Ryy)$$

Thus:

{1}	(1)	$\neg (\forall x \neg Fx \vee (\forall x \neg Fx \to \exists y Ryy))$	P
{1}	(2)	C	TT, 1
{1}	(3)	$\neg$ C	TT, 1
$\varnothing$	(4)	$\forall x \neg Fx \vee (\forall x \neg Fx \to \exists y Ryy)$	RAA, 2, 3

Many uses of TT appear further along, in connection with deductions illustrating the rules for quantificational idioms. Meanwhile, notice the meaning of the rule TT when $n = 0$. In this case, the rule takes the form:

$$\frac{\qquad}{\Gamma,\ A} \quad \text{where A is a theorem of sentential logic}$$

In short, the rule TT permits the introduction of any theorem (valid sentence) of sentential logic on a line of a deduction, with $\varnothing$ as premiss set. Thus an even shorter deduction of the theorem $A \vee (A \to B)$ above is:

$\varnothing$	(1)	$A \vee (A \to B)$	TT

Finally, note that TT covers the rule REP of replacement too, although we generalize this rule at the end of the chapter.

It is important to understand that while the rule TT can be used to combine sentential rules, it is not meant to supplant them. In other words, we are free to use the rules MP, CONJ, SIMP, etc., if and as we wish (they are all special cases of TT in any case); there is no requirement that every sentential inference be reduced to TT.

Likewise, it is important to note that TT does not cover the full versions of the rules P, PC, and RAA. This is because each of these three rules contains special instructions about the content of the premiss set of the inferred line, and thus more than a truth table is involved in checking that the inference is correct.

Now let us turn to the quantificational rules of inference: UI (universal instantiation); EG (existential generalization); UG (universal generalization); E (exemplar); and QN (quantifiers and negation).

DEFINITION 8.2. *Rules of inference for predicate logic.* In what follows, A is a formula with at most the variable x free, and B and B′ are sentences.

UI. $$\frac{\Gamma,\ \forall xA}{\Gamma,\ A(x/\alpha)}$$

EG. $$\frac{\Gamma,\ A(x/\alpha)}{\Gamma,\ \exists xA}$$

UG. $$\frac{\Gamma,\ A(x/\alpha)}{\Gamma,\ \forall xA}$$ where α is new to ∀xA and to Γ

E. $$\frac{\Gamma,\ \exists xA;\ \Delta \cup \{A(x/\alpha)\},\ B}{\Gamma \cup \Delta,\ B}$$ where α is new to ∃xA, to B, and to Δ

QN. $$\frac{\Gamma,\ B}{\Gamma,\ B'}$$ where B and B′ are alike except that B contains ¬ ∀xA (¬ ∃xA) in one or more places in which B′ contains ∃x¬ A (∀x¬ A), or vice versa

For the sake of convenience in the use of the quantificational rules, we may regard them all as basic. In fact, however (and as we shall show later on), each of the rules can be derived given others in the set. For example, the QN rules are derivable given UI, EG, UG, and E.

In the next three sections we illustrate the use of the quantificational rules, and also TT. Following these, we take up the question of strategies for using the rules to create deductions.

EXERCISES

8.1. Re-read sections 1 and 2 of chapter 4.

2. The rules UI and EG

Universal instantiation. The rule UI is the simplest of the lot. By UI one may infer a line of the form ⟨Γ, A(x/α)⟩, where α is any name, from one of the form ⟨Γ, ∀xA⟩. In short, one may *instantiate* a universal quantification. A use of the rule UI is illustrated in the following deduction showing that ∀xFx → Fa is a theorem.

{1}	(1)	∀xFx	P
{1}	(2)	Fa	UI, 1
∅	(3)	∀xFx → Fa	PC, 2

In line (2) above, Fa—i.e. (Fx)(x/a)—is inferred from ∀xFx on line (1), and the premiss set of line (1) is brought down to line (2).

The next deduction illustrates the situation in which there are multiple occurrences of universal quantifier prefixes; it shows that Rab is deducible from—and hence a consequence of—∀x∀yRxy.

{1}	(1)	∀x∀yRxy	P
{1}	(2)	∀yRay	UI, 1
{1}	(3)	Rab	UI, 2

In line (2) above, ∀yRay, i.e. (∀yRxy)(x/a), is inferred from ∀x∀yRxy on line (1). Then on line (3), Rab, i.e. (Ray)(y/b), is inferred from ∀yRay on line (2). At each step the premiss set of the hypothesis is brought down.

The following deduction is a variation on the one just above; it shows that ∀x∀yRxy implies Raa.

{1}	(1)	∀x∀yRxy	P
{1}	(2)	∀yRay	UI, 1
{1}	(3)	Raa	UI, 2

Here lines (1) and (2) are as before; but on line (3), Raa, i.e. (Ray)(y/a), is inferred by UI rather than Rab as before. This example should serve to emphasize the point that in instantiating a universal quantification, *any* name may replace the free occurrences of the variable in the matrix.

One more example of UI: it shows that ∀x(∃xFx → Gx) proves ∃xFx → Ga.

{1}	(1)	$\forall x(\exists xFx \to Gx)$	P
{1}	(2)	$\exists xFx \to Ga$	UI, 1

The point of this example is that it is only *free* occurrences of the variable that may be replaced by a name in an application of UI. In line (2) above, it is $\exists xFx \to Ga$, i.e. $(\exists xFx \to Gx)(x/a)$, that is inferred from $\forall x(\exists xFx \to Gx)$—and *not*, of course, something like $\exists xFa \to Ga$.

Finally, let us mention two possible pitfalls. In the pseudo-deduction

{1}	(1)	$\forall xFx$	P
{1}	(2)	Fx	ERROR

the formula Fx is erroneously entered on the second line—erroneously, in this case, because it is not a sentence and only sentences (closed formulas) are involved in deductions, never open formulas. *Moral*: Instantiate to names, never to variables.

Now consider this pseudo-deduction:

{1}	(1)	$\forall x\forall yRxy$	P
{1}	(2)	$\forall xRxa$	ERROR

Here the sentence $\forall xRxa$ is erroneously entered on line (2)—erroneously in this case because line (2) is not justified by any rule of inference—in particular, not by UI. An application of UI to the sentence $\forall x\forall yRxy$ on line (1) would yield a sentence of the form $\forall yR\alpha y$, for some name α. *Moral*: Instantiate universal quantifiers in order of their occurrence—"from outside in".

Existential generalization. According to the rule EG, a line of the form $\langle\Gamma, \exists xA\rangle$ may be inferred from $\langle\Gamma, A(x/\alpha)\rangle$, where α is any name. In short, one may *generalize* to an existential quantification. A simple use of the rule is found in the following deduction showing that $Fa \to \exists xFx$ is a theorem.

{1}	(1)	Fa	P
{1}	(2)	$\exists xFx$	EG, 1
$\varnothing$	(3)	$Fa \to \exists xFx$	PC, 2

In line (2) above, $\exists xFx$ is inferred from Fa, i.e. $(Fx)(x/a)$, on line (1), and the premiss set of line (1) is brought down to line (2).

It may not be immediately obvious that in applying the rule EG to a name in a sentence the name need not be replaced at every occurrence by the existentially quantified variable. But consider the four uses of EG in the following deduction.

{1}	(1)	Raa	P
{1}	(2)	∃xRxx	EG, 1
{1}	(3)	∃xRax	EG, 1
{1}	(4)	∃xRxa	EG, 1
{1}	(5)	∃xRaa	EG, 1

Each use of EG above on each of lines (2)–(5) is justified by regarding the sentence Raa on line (1), respectively, as (Rxx)(x/a), (Rax)(x/a), (Rxa)(x/a), and (Raa)(x/a). (Recall section 9 of chapter 6.) Thus each of the sentences on lines (2)–(5) is indeed a consequence of the sentence Raa, and so each should be deducible from Raa.

The following deduction shows that ∀xFx → ∃xFx is a theorem, and it uses both the rule UI and the rule EG.

{1}	(1)	∀xFx	P
{1}	(2)	Fa	UI, 1
{1}	(3)	∃xFx	EG, 2
∅	(4)	∀xFx → ∃xFx	PC, 3

Different occurrences of a single name may be existentially generalized with respect to different variables in separate uses of the rule EG. This is exemplified in the following deduction, which shows that Raa proves ∃x∃yRxy.

{1}	(1)	Raa	P
{1}	(2)	∃yRay	EG, 1
{1}	(3)	∃x∃yRxy	EG, 2

In line (2) above, ∃yRay is inferred from Raa on line (1). In this case it is the second occurrence of the name a in Raa that is generalized upon. In line (3), ∃x∃yRxy is inferred from ∃yRay by generalization on the remaining occurrence of a.

There are two mistakes to be avoided in using the rule EG. Consider, first, the following pseudo-deduction.

{1}	(1)	∀xRxa	P
{1}	(2)	∃x∀xRxx	ERROR

In this example, the sentence ∃x∀xRxx is erroneously entered on line (2)—erroneously because in attempting to existentially generalize on the name a in the sentence ∀xRxa on line (1), the variable x is introduced in such a way as to become bound by the inner occurrence of the universal quantifier in ∃x∀xRxx. To appreciate the mistake fully, notice that ∃x∀xRxx is equivalent simply to ∀xRxx and that this sentence is not a consequence of ∀xRxa. *Moral*: Existentially generalize so as to avoid prior bondage of the variable in the prefix.

Secondly, consider this pseudo-deduction:

{1}	(1)	∃xRxa	P
{1}	(2)	∃x∃yRxy	ERROR

Here the sentence ∃x∃yRxy is erroneously entered on the second line. Line (2) is not justified by EG. An application of EG to the sentence ∃xRxa on line (1) would yield a sentence of the form ∃y∃xRxy, for some variable y. *Moral*: Existentially generalize on the outside, never the inside.

EXERCISES

8.2. Use the rules from sentential logic (including TT!) and UI and EG to justify each of the lines of the following deductions.

(1) (Compare the argument at the beginning of chapters 1 and 6.)

{1}	(1)	∀x(Hx → Mx)
{2}	(2)	Hs
{1}	(3)	Hs → Ms
{1, 2}	(4)	Ms

(2) (Compare argument (5) in exercise 1.8, page 8.)

{1}	(1)	$\forall x(Dx \rightarrow \neg Ex)$
{2}	(2)	Et
{1}	(3)	$Dt \rightarrow \neg Et$
{1, 2}	(4)	$\neg Dt$
{1, 2}	(5)	$\exists x \neg Dx$

3. The rules UG and E

Universal generalization. According to the rule UG, we may generalize to $\forall xA$ from $A(x/\alpha)$—i.e. infer a line of the form $\langle \Gamma, \forall xA \rangle$ from a line of the form $\langle \Gamma, A(x/\alpha) \rangle$—*provided that the name* α *occurs neither in* $\forall xA$ *nor in any sentence in the premiss set* Γ. This proviso is essential; without it the rule UG is unsound.

An example of UG is provided by the following deduction, which shows that $\forall x(Fx \rightarrow Gx)$ and $\forall x(Gx \rightarrow Hx)$ together imply $\forall x(Fx \rightarrow Hx)$

{1}	(1)	$\forall x(Fx \rightarrow Gx)$	P
{2}	(2)	$\forall x(Gx \rightarrow Hx)$	P
{1}	(3)	$Fa \rightarrow Ga$	UI, 1
{2}	(4)	$Ga \rightarrow Ha$	UI, 2
{1, 2}	(5)	$Fa \rightarrow Ha$	TT, 3, 4
{1, 2}	(6)	$\forall x(Fx \rightarrow Hx)$	UG, 5

In line (6), $\forall x(Fx \rightarrow Hx)$ is inferred from $Fa \rightarrow Ha$, i.e. $(Fx \rightarrow Hx)(x/a)$, on line (5), and the premiss set of line (5) is entered on line (6). Notice that the conditions for an application of UG are satisfied: the name a is new to the sentence inferred, $\forall x(Fx \rightarrow Hx)$, and it is new to all of the sentences in the premiss set {1, 2}, i.e. $\forall x(Fx \rightarrow Gx)$ and $\forall x(Gx \rightarrow Hx)$.

It is useful to compare the steps in the deduction above with ordinary reasoning involving universal generalization. For example, suppose we are given the premisses (1) *Every fiddler is a geologist* and (2) *Every geologist is heretical.* To get to the conclusion (6) *Every fiddler is heretical*, we note that it is enough to show (5) *If Abe is a fiddler then Abe is heretical*—where *Abe* may be thought to represent an arbitrarily chosen individual. We thus use a principle of universal generalization at the outset to replace the original argument by another, simpler argument. Once our reasoning has established *If Abe is a fiddler then Abe is heretical* we are done. By contrast, in our formal deductive system UG is used to "close out" the reasoning by inferring the original, desired conclusion.

Another example of UG is the one below showing that ∀x∀yRxy proves ∀xRxx. It extends a deduction in the subsection above on UI.

{1}	(1)	∀x∀yRxy	P
{1}	(2)	∀yRay	UI, 1
{1}	(3)	Raa	UI, 2
{1}	(4)	∀xRxx	UG, 3

Notice that the inference by UG from line (3) to line (4) is justified: there are no occurrences of the name a in either the sentence inferred (∀xRxx) or the sentence in the premiss set (∀x∀yRxy).

Following is an example of two applications of UG. The deduction shows that ∀x∀yRyx is a consequence of ∀x∀yRxy; it is an extension of one in the earlier subsection on the UI rule.

{1}	(1)	∀x∀yRxy	P
{1}	(2)	∀yRay	UI, 1
{1}	(3)	Rab	UI, 2
{1}	(4)	∀yRyb	UG, 3
{1}	(5)	∀x∀yRyx	UG, 4

In line (4), ∀yRyb is inferred from Rab on line (3). Here the name generalized upon is a; notice that a occurs neither in ∀yRyb nor in the sentence ∀x∀yRxy in the premiss set. Then in line (5), ∀x∀yRyx is inferred from ∀yRyb on line (4). Here the name generalized upon is b, which occurs neither in ∀x∀yRyx nor in the sentence in the premiss set.

For an example of a deduction using UI, EG, and UG, consider the following. It shows that ∀x(Fx → ∃xGx) is implied by ∀x(Fx → Gx).

{1}	(1)	∀x(Fx → Gx)	P
{2}	(2)	Fa	P
{1}	(3)	Fa → Ga	UI, 1

{1, 2}	(4)	Ga	TT, 2, 3
{1, 2}	(5)	∃xGx	EG, 4
{1}	(6)	Fa → ∃xGx	PC, 5
{1}	(7)	∀x(Fx → ∃xGx)	UG, 6

Finally, the following deduction uses UG to show that ∀x(Fx → Fx) is a theorem.

∅	(1)	Fa → Fa	TT
∅	(2)	∀x(Fx → Fx)	UG, 1

Notice that the name generalized upon in the deduction above is new to the sentence inferred, ∀x(Fx → Fx), and to the premiss set, ∅.

The restriction on UG that the name generalized upon not occur in the universal quantification inferred means that the following sequence of lines, for example, is *not* a deduction.

{1}	(1)	∀xRxx	P
{1}	(2)	Raa	UI, 1
{1}	(3)	∀xRax	ERROR

Line (3) above does not follow from line (2) by UG because the name generalized upon, a, still occurs in the inferred sentence ∀xRax. (Observe that ∀xRax is, indeed, not a consequence of ∀xRxx.)

The other half of the restriction on the rule UG—that the name generalized upon not occur in any of the sentences in the premiss set—means that the following sequence, for example, is *not* a deduction.

{1}	(1)	Fa	P
{1}	(2)	∀xFx	ERROR

Line (2) above does not follow from line (1) by UG because the name generalized upon, a, occurs in the sentence Fa in the premiss set. (Observe, too, that ∀xFx is not a consequence of Fa.)

As formulated, however, the rule UG does permit vacuous universal generalization, just as EG permits vacuous existential generalization. Thus, for example, the following sequence *is* a deduction.

{1}	(1)	Fa	P
{1}	(2)	$\forall$xFa	UG, 1

The inference from line (1) to line (2) above is justified by UG since the sentence Fa on line (1) can be regarded as (Fa)(x/b), where b is a name other than a (and hence does not occur in $\forall$xFa or in the sentence Fa in the premiss set). That is to say, one can regard the sentence $\forall$xFa as the result of universally generalizing upon all occurrences of b in Fa, of which there are none. (Notice that $\forall$xFa is equivalent to Fa and so is a consequence of it; hence it is desirable that $\forall$xFa be deducible from Fa.)

The three examples above lead us to this *moral*: Universally generalize on either all or none of the occurrences of a name.

A couple of pitfalls in using the rule UG should be avoided; they are like the ones mentioned above in connection with EG. First, consider this pseudo-deduction:

{1}	(1)	$\forall$x$\exists$yRxy	P
{1}	(2)	$\exists$yRay	UI, 1
{1}	(3)	$\forall$y$\exists$yRyy	ERROR

Here the sentence $\forall$y$\exists$yRyy is erroneously entered on line (3)—erroneously because in attempting to universally generalize on the name a in the sentence $\exists$yRay on line (2), the variable y is introduced in such a way as to become bound by the existential quantifier in $\forall$y$\exists$yRyy. Notice that this sentence is equivalent to $\exists$yRyy, which is not a consequence of $\forall$x$\exists$yRxy. *Moral*: Universally generalize so as to avoid prior bondage of the variable.

Secondly, consider the following pseudo-deduction.

{1}	(1)	$\forall$xRxx	P
{1}	(2)	Raa	UI, 1
{1}	(3)	$\exists$yRay	EG, 2
{1}	(4)	$\exists$y$\forall$xRxy	ERROR

Here line (4) is not justified UG, or by any other rule of inference. (Notice that $\exists$y$\forall$xRxy is not a consequence of $\forall$xRxx.) An application of UG to the sentence $\exists y Ray$ on line (3) would yield a sentence of the form $\forall x \exists y Rxy$. *Moral*: Universally generalize on the outside, never on the inside.

Exemplar. The rule E permits inference of a line of the form $\langle \Gamma \cup \Delta, B \rangle$ from earlier lines of the forms $\langle \Gamma, \exists xA \rangle$ and $\langle \Delta \cup \{A(x/\alpha)\}, B \rangle$ *subject to the proviso that* the name α does not appear in any of: the sentence $\exists xA$, the sentences in Δ, the sentence B. Thus in the context of the deduction α may be thought to represent an "exemplar"—an individual that exemplifies those whose existence is alleged by the existential quantification. (Note that it is okay if α appears in sentences in Γ.) Use of the rule E is illustrated in line (7) of the following deduction.

{1}	(1)	$\forall x(Fx \to Gx)$	P
{2}	(2)	$\exists x(Fx \wedge Hx)$	P
{3}	(3)	$Fa \wedge Ha$	P
{1}	(4)	$Fa \to Ga$	UI, 1
{1, 3}	(5)	$Ga \wedge Ha$	TT, 3, 4
{1, 3}	(6)	$\exists x(Gx \wedge Hx)$	EG, 5
{1, 2}	(7)	$\exists x(Gx \wedge Hx)$	E, 2, 6

In this deduction, line (7) follows from lines (2) and (6) by E. Because on line (6) the sentence $\exists x(Gx \wedge Hx)$ has been shown to follow from $\forall x(Fx \to Gx)$ on line (1) together with the instance $Fa \wedge Ha$ (= $(Fx \wedge Hx)(x/a)$) of $\exists x(Fx \wedge Hx)$ on line (3), the rule E entitles us to infer $\exists x(Gx \wedge Hx)$ again, but this time based on $\forall x(Fx \to Gx)$ (as before) and on the existential premiss $\exists x(Fx \wedge Hx)$ on line (2).

Reasoning by way of an exemplar is used in ordinary discourse to replace an argument having an existential premiss by a simpler argument that does not. To see this, compare the steps in the deduction above with the following reasoning. Suppose our premisses are (1) *Every freebooter is a gangster* and (2) *Some freebooter is a harlequin*, and the conclusion we wish to reach is (7) *Some gangster is a harlequin*. Our first move is to replace (2) by one of its instances—say (3) *Amy is a freebooter and a harlequin*—and note that it is enough, so long as *Amy* may be thought to represent an arbitrarily chosen individual, to argue to the conclusion (6) *Some gangster is a harlequin* from (3) and (1). Thus we relabel (7) as (6), since as in the deduction above, this statement of the conclusion is based on different premisses. The function of the rule E in the formal system is to "close out" the argument by reinstating the existential premiss.

EXERCISES

8.3. Use the rules from sentential logic (including TT!) and UI, EG, UG, and E to justify each of the lines of the following deductions.

(1) (Compare argument (10) in exercise 6.11, page 244.)

{1}	(1)	$\forall x(Sx \to Wx)$
{2}	(2)	$\forall x((Fx \wedge \neg Dx) \to Cx)$
{3}	(3)	$\forall x(Fx \to (Tx \vee \neg Wx))$
{4}	(4)	$\forall x(Fx \to (\neg Sx \to Kx))$
{5}	(5)	$\forall x((Fx \wedge Hx) \to \neg Dx)$
{6}	(6)	$\forall x((Fx \wedge Tx) \to \neg Cx)$
{1}	(7)	$Sa \to Wa$
{2}	(8)	$(Fa \wedge \neg Da) \to Ca$
{3}	(9)	$Fa \to (Ta \vee \neg Wa)$
{4}	(10)	$Fa \to (\neg Sa \to Ka)$
{5}	(11)	$(Fa \wedge Ha) \to \neg Da$
{6}	(12)	$(Fa \wedge Ta) \to \neg Ca$
{1, 2, 3, 4, 5, 6}	(13)	$(Fa \wedge Ha) \to \neg\neg Ka$
{1, 2, 3, 4, 5, 6}	(14)	$\forall x((Fx \wedge Hx) \to \neg\neg Kx)$

(2) (Compare argument (3) in exercise 1.8, page 8.)

{1}	(1)	$\forall x(Ex \to Fx)$
{2}	(2)	$\exists x(Px \wedge \neg Fx)$
{3}	(3)	$Pa \wedge \neg Fa$
{1}	(4)	$Ea \to Fa$
{1, 3}	(5)	$Pa \wedge \neg Ea$
{1, 3}	(6)	$\exists x(Px \wedge \neg Ex)$
{1, 2}	(7)	$\exists x(Px \wedge \neg Ex)$

8.4. Below are three pseudo-deductions each of which contains a line that violates one of the three restrictions on the rule E—that α is new to $\exists xA$, that α is new to B, and that α is new to Δ. (1) Give rules and line numbers to justify the justifiable lines of the pseudo-deductions and identify the ERRORs. (2) Persuade yourself of the need for the restrictions by giving counterexamples that show the invalidity of the arguments represented by the last lines of the deductions.

(1)	{1}	(1)	$\exists x(Fx \land Gx)$
	{2}	(2)	$Fa \land Ga$
	{2}	(3)	Fa
	{1}	(4)	Fa

(2)	{1}	(1)	$\exists xRxa$
	{2}	(2)	Raa
	{2}	(3)	$\exists xRxx$
	{1}	(4)	$\exists xRxx$

(3)	{1}	(1)	$\exists xFx$
	{2}	(2)	$\exists xGx$
	{3}	(3)	Fa
	{4}	(4)	Ga
	{3, 4}	(5)	$Fa \land Ga$
	{3, 4}	(6)	$\exists x(Fx \land Gx)$
	{2, 3}	(7)	$\exists x(Fx \land Gx)$
	{1, 2}	(8)	$\exists x(Fx \land Gx)$

4. The rule QN

The rule QN is an addition to REP. It permits replacement of $\neg \forall xA$ by $\exists x\neg A$ and vice versa, and similarly of $\neg \exists xA$ by $\forall x\neg A$ and vice versa. Examples of QN appear in the following deduction showing that $\exists x(\exists yFy \to Fx)$ is a theorem.

{1}	(1)	$\neg \exists x(\exists yFy \rightarrow Fx)$	P
{1}	(2)	$\forall x \neg (\exists yFy \rightarrow Fx)$	QN, 1
{1}	(3)	$\neg (\exists yFy \rightarrow Fa)$	UI, 2
{1}	(4)	$\exists yFy$	TT, 3
{1}	(5)	$\neg Fa$	TT, 3
{1}	(6)	$\forall y \neg Fy$	UG, 5
{1}	(7)	$\neg \exists yFy$	QN, 6
$\varnothing$	(8)	$\exists x(\exists yFy \rightarrow Fx)$	RAA, 4, 7

(Note that we cannot simply use EG to get the line $\langle \varnothing, \exists x(\exists yFy \rightarrow Fx)\rangle$ from a line of the form $\langle \varnothing, \exists yFy \rightarrow F\alpha\rangle$. Because the inference rules are sound and $\exists yFy \rightarrow F\alpha$ is not valid for any name α, no instance $\exists yFy \rightarrow F\alpha$ is deducible from the empty set.)

EXERCISES

8.5. Use the rules from sentential logic (including TT!) and UI, EG, UG, E, and QN to justify each line of the following deduction. (Compare argument (5) in exercise 1.8, page 8.)

{1}	(1)	$\neg \exists x(Fx \wedge Cx)$
{2}	(2)	$\forall x(Ox \rightarrow Cx)$
{1}	(3)	$\forall x \neg (Fx \wedge Cx)$
{2}	(4)	$Oa \rightarrow Ca$
{1}	(5)	$\neg (Fa \wedge Ca)$
{1, 2}	(6)	$\neg (Oa \wedge Fa)$
{1, 2}	(7)	$\forall x \neg (Ox \wedge Fx)$
{1, 2}	(8)	$\neg \exists x(Ox \wedge Fx)$

5. Strategies

There are several fairly obvious strategies for using the new rules UI, EG, and QN. So we concentrate on strategies for using UG and E.

The UG strategy. This is a strategy for deducing universal quantifications. In order to infer a desired line of the form $\langle \Gamma, \forall xA\rangle$, it is sufficient to deduce an earlier line of the form $\langle \Gamma, A(x/\alpha)\rangle$, where α is a name new to $\forall xA$ and to all the sentences in the premiss set Γ—in short, where α is new to the line $\langle\Gamma, \forall xA\rangle$. For then an application of UG will yield the desired result. The strategy is exemplified above in several deductions. Let us illustrate it here with this deduction:

1	{1}	(1)	$\forall x(Fx \rightarrow \neg Gx)$	P
1	{2}	(2)	$\forall x(Hx \rightarrow Gx)$	P
3	{1}	(3)	$Fa \rightarrow \neg Ga$	UI, 1
4	{2}	(4)	$Ha \rightarrow Ga$	UI, 2
2	{1, 2}	(5)	$Fa \rightarrow \neg Ha$	TT, 3, 4
1	{1, 2}	(6)	$\forall x(Fx \rightarrow \neg Hx)$	UG, 5

The deduction above shows that $\forall x(Fx \rightarrow \neg Gx)$ and $\forall x(Hx \rightarrow Gx)$ together imply $\forall x(Fx \rightarrow \neg Hx)$. The boldface numbering to the left shows the order of development: At step(s) **1** the premisses $\forall x(Fx \rightarrow \neg Gx)$ and $\forall x(Hx \rightarrow Gx)$ and conclusion $\forall x(Fx \rightarrow \neg Hx)$ are installed on lines (1) and (2) and the last line (which turns out to be (6)). At step **2**, because a universal quantification is desired, an appropriate instance is entered on the next-to-last line with the same premiss set as that for the quantification. The instance—using the name a—is appropriate because the name does not occur in the quantification or its premiss set on line (6). Finally, at steps **3** and **4** two applications of UI yield instances of the quantificational premisses, which in turn yield the desired instance by TT.

The UG strategy can be generalized in an obvious way to deduce multiply universal conclusions:

$$\Gamma \quad (\,) \quad \forall x_1 \ldots \forall x_n A$$

To infer such a line, it is sufficient to obtain a line of the form

$$\Gamma \quad (\,) \quad A(x_1/\alpha_1) \ldots (x_n/\alpha_n)$$

—where the names α_i ($i = 1, \ldots, n$) are distinct and do not appear in $\forall x_1 \ldots \forall x_n A$ or in any sentence in the premiss set Γ—and then use the rule UG n times to reach the desired conclusion.

For an example of the UG strategy used twice in a deduction, consider again this variation on the one given earlier showing that ∀x∀yRxy proves ∀x∀yRyx:

1	{1}	(1)	∀x∀yRxy	P
4	{1}	(2)	∀yRby	UI, 1
3	{1}	(3)	Rba	UI, 2
2	{1}	(4)	∀yRya	UG, 3
1	{1}	(5)	∀x∀yRyx	UG, 4

Again, note the order of development. At **1** the premiss and conclusion are registered. Then at **2** an appropriate instance ∀yRya of the desired quantification ∀x∀yRyx is entered with the same premiss set. The instance using the name a is appropriate since no names appear at all in the last line. Now since another quantification—∀yRya—is wanted, at **3** an appropriate instance Rba of this quantification is set down (again with the same premiss set). Here the instance is appropriate since the name b is new to the line featuring the desired ∀yRya. At this point, two uses of UI lead directly to Rba.

The E strategy. Consider once again the earlier deduction showing that ∃x(Gx ∧ Hx) is a consequence of {∀x(Fx → Gx), ∃x(Fx ∧ Hx)}:

1	{1}	(1)	∀x(Fx → Gx)	P
1	{2}	(2)	∃x(Fx ∧ Hx)	P
2	{3}	(3)	Fa ∧ Ha	P
3	{1}	(4)	Fa → Ga	UI, 1
4	{1, 3}	(5)	Ga ∧ Ha	TT, 3, 4
2	{1, 3}	(6)	∃x(Gx ∧ Hx)	EG, 5
1, 5	{1, 2}	(7)	∃x(Gx ∧ Hx)	E, 2, 6

This deduction exhibits a pattern worth noticing, as it demonstrates a method for using existential premisses like ∃x(Fx ∧ Hx) on line (2).

- Set out the premisses (by P, lines (1) and (2)) and the desired conclusion (line (7)), with the premisses making up the premiss set of the conclusion. (Step(s) **1** in the example.)

- Use P to add a fresh instance of the existential quantification as a new hypothesis—Fa ∧ Ha in this case (line (3))—and enter the desired conclusion ∃x(Gx ∧ Hx) again (line (6)), with the same premiss set as the conclusion (line (7)) except that the instance replaces the existential quantification. (Step(s) **2**.)

- Somehow (usually using the new hypothesis) deduce the desired conclusion—∃x(Gx ∧ Hx) (line (6)). (Steps **3** and **4**, in this case.)

- The desired conclusion then follows again, this time with the desired premiss set, by E from this line (6) and the original existential hypothesis (line (2)) (Step **5**.)

This is the E strategy for using an instantiation of an existential quantification. In outline:

	⋮		
Γ	( )	∃xA	[*given somehow*]
	⋮		α *new to* ∃xA, B, *and* Δ
{A(x/α)}	( )	A(x/α)	P
	⋮		
Δ ∪ {A(x/α)}	( )	B	[*deduced somehow*]
	⋮		
Γ ∪ Δ	( )	B	E

It should be emphasized that for the success of this strategy the name α chosen for the instance of the existential quantification ∃xA should be one that does not appear in the quantification itself, in the desired conclusion B, or in any sentence in the premiss set Δ for the desired conclusion B. In short, α should be new to ∃xA, B, and Δ. Otherwise the use of E will be blocked.

The E strategy can be iterated, or "nested", to handle multiply existential sentences,

$$\Gamma \quad (\ ) \quad \exists x_1 \ldots \exists x_n A,$$

by applying the strategy over and over again, one instance at a time. The following deduction illustrates this; it shows that ∃x∃yRxy implies ∃x∃yRyx.

1	{1}	(1)	∃x∃yRxy	P
2	{2}	(2)	∃yRay	P
3	{3}	(3)	Rab	P
4	{3}	(4)	∃yRyb	EG, 3
3	{3}	(5)	∃x∃yRyx	EG, 4
2	{2}	(6)	∃x∃yRyx	E, 2, 5
1	{1}	(7)	∃x∃yRyx	E, 1, 6

What happens when both the UG and the E strategies need to be used? Here is another example, which combines them. It shows that ∀x∃yRyx is implied by ∃x∀yRxy.

1	{1}	(1)	∃x∀yRxy	P
3	{2}	(2)	∀yRby	P
4	{2}	(3)	Rba	UI, 2
3	{2}	(4)	∃yRya	EG, 3
2	{1}	(5)	∃yRya	E, 1, 4
1	{1}	(6)	∀x∃yRyx	UG, 5

Let us follow the steps in the development of this deduction. At step(s) **1** the hypothesis and conclusion are registered with appropriate premiss sets (lines (1) and (6)). Since the desired conclusion is the line ⟨{∃x∀yRxy}, ∀x∃yRyx⟩, we invoke the UG strategy and at step **2** seek to deduce ⟨{∃x∀yRxy}, ∃yRya⟩ on line (5). Notice that the name a is new to the last line. Next we see that the premiss is existential—∃x∀yRxy. So at step(s) **3** we introduce the instance ∀yRby on line (2) as a new hypothesis with an eye to deducing line (4) and then using the E strategy. Notice that b is a name new both to ∃x∀yRxy and to the sentences on line (4). Once all this strategizing is done, it turns out that only line (3) is needed to complete the deduction (step **4**).

EXERCISES

8.6. Explain the strategy in the development of the deduction above that shows that $\exists x \exists y Ryx$ is a consequence of $\exists x \exists y Rxy$.

8.7. Show the validity of each of the following arguments by describing deductions showing the conclusions to be deducible from the premisses. (For (1)–(9), compare exercise 6.11, page 244.)

(1) $\forall x(Mx \rightarrow \neg\neg Sx),\ \forall x(Ex \rightarrow Mx)\ /\ \forall x(\neg Sx \rightarrow \neg Ex)$

(2) $\forall x(Px \rightarrow Sx),\ \forall x(Bx \rightarrow \neg\neg Px)\ /\ \forall x(Bx \rightarrow \neg\neg Sx)$

(3) $\forall x(Bx \rightarrow \neg Lx),\ \forall x(Cx \rightarrow \neg Dx),\ \forall x(\neg Lx \rightarrow Dx)$
$/\ \forall x(Bx \rightarrow \neg Cx)$

(4) $\forall x(Dx \rightarrow \neg Wx),\ \forall x(Ox \rightarrow \neg\neg Wx),\ \forall x(Px \rightarrow Dx)$
$/\ \forall x(Px \rightarrow \neg Ox)$

(5) $\forall x(Sx \rightarrow Lx),\ \forall x(\neg Sx \rightarrow \neg Jx),\ \forall x(Yx \rightarrow \neg Lx)$
$/\ \forall x(Yx \rightarrow \neg Jx)$

(6) $\forall x(Tx \rightarrow \neg Wx),\ \forall x(\neg Wx \rightarrow \neg Cx),\ \forall x(Ux \rightarrow Tx)$
$/\ \forall x(Cx \rightarrow \neg Ux)$

(7) $\forall x(Ex \vee \neg Tx),\ \forall x(Hx \rightarrow \neg Rx),\ \forall x(\neg Rx \rightarrow \neg Ex)$
$/\ \forall x(Hx \rightarrow \neg Tx)$

(8) $\forall x(Bx \rightarrow \neg\neg Sx),\ \forall x(Gx \rightarrow Ix),\ \forall x(Ix \rightarrow \neg Sx)$
$/\ \forall x(Gx \rightarrow \neg Bx)$

(9) $\forall x((Kx \wedge Fx) \rightarrow \neg\neg Tx),\ \forall x((Kx \wedge \neg Lx) \rightarrow \neg Px),$
$\forall x((Kx \wedge Wx) \rightarrow Fx),\ \forall x((Tx \wedge Kx) \rightarrow \neg Gx),$
$\forall x((Kx \rightarrow (Wx \vee \neg Lx))\ /\ \forall x((Kx \wedge Gx) \rightarrow \neg Px)$

(10) $\forall x(Px \rightarrow Bx),\ \forall x(\neg Tx \rightarrow \neg Fx),\ \forall x(Ox \rightarrow \neg Cx),\ \forall x(Bx \rightarrow Fx),$
$\forall x(Px \vee \neg Wx),\ \forall x(\neg Cx \rightarrow \neg Tx)\ /\ \forall x(Ox \rightarrow \neg Wx)$

8.8. Here are arguments (1), (3), (5), (7), and (9) from exercise 6.4 (page 233). Prove their validity by deducing the conclusions from the premisses.

(1) $\forall x(Fx \rightarrow Gx),\ \exists x(Hx \wedge Fx)\ /\ \exists x(Hx \wedge Gx)$

(3) $\forall x(Fx \rightarrow Gx),\ \forall x(Hx \rightarrow \neg Gx)\ /\ \forall x(Fx \rightarrow \neg Hx)$

(5) $\exists x(Fx \wedge Gx),\ \forall x(Gx \rightarrow \neg Hx)\ /\ \exists x(Fx \wedge \neg Hx)$

(7) $\forall x(Fx \rightarrow \neg Gx),\ \forall x(Hx \rightarrow Gx)\ /\ \forall x(Fx \rightarrow \neg Hx)$

(9) $\exists x(Fx \wedge Gx),\ \forall x(Hx \rightarrow \neg Gx)\ /\ \exists x(Fx \wedge \neg Hx)$

8.9. Here again are the arguments, sentences, and pairs of sentences in exercise 7.18 (page 279). Describe deductions showing that the arguments and sentences are valid and that the sentences in each pair are equivalent (recall that in (21)–(30), P is a zero-place predicate). *For* (51) *and* (52) *do not use* QN.

Note that some of the deductions have appeared in the text. Peeking is permitted only as necessary.

(1) $\forall xFx\ /\ Fa$

(2) $Fa\ /\ \exists xFx$

(3) $\forall x\forall yRxy\ /\ Rab$

(4) $\forall x\forall yRxy\ /\ Raa$

(5) $Rab\ /\ \exists x\exists yRxy$

(6) $Raa\ /\ \exists x\exists yRxy$

(7) $\forall x(Fx \wedge Gx),\ \forall xFx \wedge \forall xGx$

(8) $\exists x(Fx \vee Gx),\ \exists xFx \vee \exists xGx$

(9) $\forall xFx \vee \forall xGx\ /\ \forall x(Fx \vee Gx)$

(10) $\exists x(Fx \wedge Gx)\ /\ \exists xFx \wedge \exists xGx$

(11) $\forall x(Fx \rightarrow Gx)\ /\ \forall xFx \rightarrow \forall xGx$

(12) $\forall x(Fx \rightarrow Gx)\ /\ \exists xFx \rightarrow \exists xGx$

(13) $\exists x(Fx \rightarrow Gx),\ \forall xFx \rightarrow \exists xGx$

(14) $\forall xFx \wedge \exists xGx\ /\ \exists x(Fx \wedge Gx)$

(15) $\forall x(Fx \vee Gx)\ /\ \exists xFx \vee \forall xGx$

(16) $\forall x\forall yRxy,\ \forall y\forall xRxy$

(17) $\exists x \exists y Rxy,\ \exists y \exists x Rxy$

(18) $\forall x \forall y Rxy\ /\ \forall x Rxx$

(19) $\exists x Rxx\ /\ \exists x \exists y Rxy$

(20) $\exists x \forall y Rxy\ /\ \forall y \exists x Rxy$

(21) $\forall x P,\ P$

(22) $\exists x P,\ P$

(23) $\forall x(Fx \wedge P),\ \forall x Fx \wedge P$

(24) $\exists x(Fx \wedge P),\ \exists x Fx \wedge P$

(25) $\forall x(Fx \vee P),\ \forall x Fx \vee P$

(26) $\exists x(Fx \vee P),\ \exists x Fx \vee P$

(27) $\forall x(Fx \rightarrow P),\ \exists x Fx \rightarrow P$

(28) $\exists x(Fx \rightarrow P),\ \forall x Fx \rightarrow P$

(29) $\forall x(P \rightarrow Fx),\ P \rightarrow \forall x Fx$

(30) $\exists x(P \rightarrow Fx),\ P \rightarrow \exists x Fx$

(31) $\forall x(Fx \rightarrow Fx)$

(32) $\exists x(Fx \rightarrow Fx)$

(33) $\forall x(Fx \rightarrow \exists x Fx)$

(34) $\exists x(Fx \rightarrow \forall x Fx)$

(35) $\forall x(\forall x Fx \rightarrow Fx)$

(36) $\exists x(\exists x Fx \rightarrow Fx)$

(37) $\exists x(Fx \rightarrow \exists x Fx)$

(38) $\exists x(\forall x Fx \rightarrow Fx)$

(39) $\exists x(\forall x Fx \leftrightarrow Fx)$

(40) $\exists x(\exists x Fx \leftrightarrow Fx)$

(41) $\forall x \exists y(Fx \rightarrow Fy)$

(42) $\forall x \exists y(Fy \rightarrow Fx)$

(43) $\exists x \exists y(Fx \rightarrow Fy)$

(44) $\exists x \forall y(Fx \rightarrow Fy)$

(45) $\exists x \forall y(Fy \rightarrow Fx)$

(46) $\forall x(Rxx \rightarrow Rxx)$

(47) $\forall x \forall y(Rxy \rightarrow Rxy)$

(48) $\forall x \forall y \exists z(Rxy \rightarrow Rxz)$

(49) $\forall x \exists y \forall z(Rxz \rightarrow Ryz)$

(50) $\forall x \exists y \forall z \exists w(Rxz \rightarrow Ryw)$

(51) $\forall xFx, \neg \exists x \neg Fx$

(52) $\exists xFx, \neg \forall x \neg Fx$

8.10. Describe a deduction to show that $\forall x(Fx \leftrightarrow P)$ implies $\exists xFx \leftrightarrow \forall xFx$.

8.11. Describe a deduction to show that $\forall x(Fx \leftrightarrow \exists xFx)$ implies $\exists xFx \leftrightarrow \forall xFx$.

8.12. This is a follow-up to exercises 6.14 and 7.15 (pages 246 and 278). First, show that if there is a barber who shaves all those barbers who do not shave themselves then some barber is a self-shaver; i.e. use a deduction to show that

(1) $\exists x \forall y(\neg Syy \rightarrow Sxy)$ implies $\exists xSxx$.

Now show that if there is a barber who shaves only those barbers who do not shave themselves then not every barber is a self-shaver; i.e. show by means of a deduction that

(2) $\exists x \forall y(Sxy \rightarrow \neg Syy)$ implies $\neg \forall xSxx$.

Is there is a barber who shaves all and only those barbers who do not shave themselves? Use a deduction to show that

(3) $\exists x \forall y(Sxy \leftrightarrow \neg Syy)$ is inconsistent and hence a contradiction.

8.13. Without using UI, describe a deduction showing that $\forall xFx \vdash Fa$. Conclude that the rule UI is derivable given the sentential and other quantificational rules.

8.14. Without using EG, describe a deduction showing that $Fa \vdash \exists xFx$. Conclude that the rule EG is derivable given the sentential and other quantificational rules.

8.15. Without using UG, describe a deduction showing that $\forall x \forall yRxy \vdash \forall xRxx$. Conclude that the rule UG is derivable given the sentential and other quantificational rules.

8.16. Without using E, describe a deduction showing that $\exists xRxx \vdash \exists x \exists yRxy$. Conclude that the rule E is derivable given the sentential and other quantificational rules.

8.17. Without using QN, describe deductions showing that $\neg \forall xFx \sim \exists x \neg Fx$ and that $\neg \exists xFx \sim \forall x \neg Fx$. (*Hint*: See parts (51) and (52) of exercise 8.9.) Conclude that the QN rules are derivable given the sentential and other quantificational rules.

6. Soundness of the deductive system

That the deductive system for predicate logic is sound, which is our main result in this section, follows ultimately from the fact that each rule of inference preserves consequence, and hence for each line $\langle \Gamma, A \rangle$ in a deduction it is the case that $\Gamma \models A$.

We know from the proof of soundness for sentential logic (proposition 4.7, page 170) that the inference rules P, PC, RAA, MP, etc., are sound. So now we need only be assured that the remaining rules—UI, EG, UG, E, and QN—are sound. This is precisely the content of proposition 7.5 (page 286). Therefore:

PROPOSITION 8.3. *Soundness*. The deductive system for predicate logic is sound. That is, if $\Gamma \vdash A$ then $\Gamma \models A$.

As we did in chapter 5, we set out soundness corollaries along with proposition 8.3 itself.

PROPOSITION 8.4. *Soundness and corollaries.*

(1) If $\Gamma \vdash A$ then $\Gamma \models A$.

(2) If $\vdash A$ then $\models A$.

(3) If $A \sim B$ then $A \simeq B$.

(4) If Sat Γ then Con Γ.

Proof. See the corresponding proposition (4.8, page 171) in chapter 4. □

EXERCISES

8.18. Prove the corollaries to soundness in proposition 8.4.

8.19. True or false:

a. There is a formula A such that $\forall xA$ is a theorem and $\exists xA$ is not a theorem.

b. The sentence $\exists xFx \vee \neg \exists xFx$ is deducible from the empty set.

c. The deductive system for predicate logic would be unsound if both the rule UI and the rule EG were dropped.

d. Every negation is deductively equivalent to at least one existential quantification.

e. The set $\{\forall xFx, \neg Fa_0, \neg Fa_1, \neg Fa_2, \ldots\}$ is inconsistent.

f. The deductive system for predicate logic would be incomplete if both the rule UG and the rule E were dropped.

g. A universal quantification $\forall x A$ is a theorem if every instance $A(x/\alpha)$ is a theorem.

h. If nothing but the consequences of a set of sentences are deducible from the set, then a deductive system is complete.

i. There is a formula A such that $\exists xA$ is a theorem and $\forall xA$ is not a theorem.

j. There are at least a trillion universal quantifications that are deductively equivalent to each other.

k. The sentence $\forall x\forall y\forall z\forall w(Rxy \rightarrow Rzw)$ is a theorem.

l. The deductive system for predicate logic would be sound but incomplete if the QN rules were dropped.

m. There is at least one formula A such that $\forall xA$ does not prove $\exists xA$.

n. The sentences $\forall y\exists xRxy$ and $\exists x\forall yRxy$ are deductively equivalent.

o. Every existential quantification is deductively equivalent to at least one conjunction.

p. The set of sentences $\{\neg \forall x Fx, Fa_0, Fa_1, Fa_2, \ldots\}$ is inconsistent.

q. In an unsound deductive system some theorems are not valid.

r. The sentence $\exists x(Fx \rightarrow \forall yFy)$ is not a theorem.

s. In a complete deductive system a sentence is valid unless it is not a theorem.

t. No deduction begins with a line justified by the rule E.

7. Generalizing the rules

In describing deductions we frequently find we wish to instantiate from, generalize to, or otherwise deal with sequences of quantifiers. Following are generalizations of the rules UI, EG, UG, E, and QN.

DEFINITION 8.5. *Generalized rules of inference for predicate logic.* In what follows, A is a formula with at most $x_1, \ldots, x_n$ by way of free variables.

UI*.
$$\frac{\Gamma,\ \forall x_1 \ \ldots \ \forall x_n A}{\Gamma,\ A(x_1/\alpha_1,\ \ldots,\ x_n/\alpha_n)}$$

EG*.
$$\frac{\Gamma,\ A(x_1/\alpha_1,\ \ldots,\ x_n/\alpha_n)}{\Gamma,\ \exists x_1 \ \ldots \ \exists x_n A}$$

UG*.
$$\frac{\Gamma,\ A(x_1/\alpha_1,\ \ldots,\ x_n/\alpha_n)}{\Gamma,\ \forall x_1 \ \ldots \ \forall x_n A}$$

where the names α_i are distinct (per variables x_i) and new to Γ and to A

E*.
$$\frac{\Gamma,\ \exists x_1 \ \ldots \ \exists x_n A;\quad \Delta \cup \{A(x_1/\alpha_1,\ \ldots,\ x_n/\alpha_n)\},\ B}{\Gamma \cup \Delta,\ B}$$

where the names α_i are distinct (per variables x_i) and new to $\exists x_1 \ \ldots \ \exists x_n A$, to B, and to every sentence in Δ

QN*. Move negation signs at will in quantifier sequences, changing $\forall$ to $\exists$ and vice versa.

Despite their fearsome formulations, these generalized rules will considerably simplify and shorten our work. Consider for example the deduction below that shows that $\forall x \forall y Rxy$ implies $\forall x \forall y Ryx$, and compare it with its five-line counterpart in section 3.

{1}	(1)	$\forall x \forall y Rxy$	P
{1}	(2)	Rab	UI*, 1
{1}	(3)	$\forall x \forall y Ryx$	UG*, 2

Similarly, compare the following four-line deduction (showing that $\exists x \exists y Rxy$ implies $\exists x \exists y Ryx$) with its seven-line counterpart in section 3.

{1}	(1)	$\exists x \exists y Rxy$	P
{2}	(2)	Rab	P
{2}	(3)	$\exists x \exists y Ryx$	EG*, 2
{1}	(4)	$\exists x \exists y Ryx$	E*, 1, 3

Of course with the generalized rules come generalized strategies for using them. For example, according to the UG* strategy, to deduce a line of the form

$$\Gamma \quad (\,) \quad \forall x_1 \ldots \forall x_n A,$$

it is sufficient to obtain one of the form

$$\Gamma \quad (\,) \quad A(x_1/\alpha_1, \ldots, x_n/\alpha_n),$$

where each name α_i ($i = 1, \ldots, n$) is correlated uniquely with the corresponding variable x_i. Then the desired line follows by UG*. This is illustrated by the first deduction above.

Note too the E* strategy for dealing with a multiply existential quantification $\exists x_1 \ldots \exists x_n A$:

- By the rule P assume as a new hypothesis an appropriate instance $A(x_1/\alpha_1, \ldots, x_n/\alpha_n)$ of the existential quantification $\exists x_1 \ldots \exists x_n A$.

- Deduce the desired conclusion (usually with the assumed instance $A(x_1/\alpha_1, \ldots, x_n/\alpha_n)$ in the premiss set).

- Use E* to arrive again at the desired conclusion; the premiss set of this line will include the premiss set of the existential quantification $\exists x_1 \ldots \exists x_n A$.

EXERCISES

8.20. Here are some of the items from exercise 8.9. Use the generalized rules to describe deductions showing that the arguments and sentences are valid and that the sentences in each pair are equivalent.

(3) $\forall x \forall y Rxy$ / Rab

(4) $\forall x \forall y Rxy$ / Raa

(5) Rab / $\exists x \exists y Rxy$

(6) Raa / $\exists x \exists y Rxy$

(16) $\forall x \forall y Rxy$, $\forall y \forall x Rxy$

(17) $\exists x \exists y Rxy$, $\exists y \exists x Rxy$

(18) $\forall x \forall y Rxy$ / $\forall x Rxx$

(19) $\exists x Rxx$ / $\exists x \exists y Rxy$

(43) $\exists x \exists y(Fx \to Fy)$

(45) $\forall x \forall y(Rxy \to Rxy)$

(46) $\forall x \forall y \exists z(Rxy \to Rxz)$

8.21. Describe deductions showing that of each of (a)–(f) is a consequence of $\forall x \forall y(Rxy \leftrightarrow P)$.

(a) $\exists x Rxx \leftrightarrow \forall x Rxx$

(b) $\exists x Rxx \leftrightarrow \forall x \forall y Rxy$

(c) $\exists x \exists y Rxy \leftrightarrow \forall x Rxx$

(d) $\exists x \exists y Rxy \leftrightarrow \exists x Rxx$

(e) $\exists x \exists y Rxy \leftrightarrow \forall x \forall y Rxy$

(f) $\forall x Rxx \leftrightarrow \forall x \forall y Rxy$

8.22. Here again from exercises 7.4 and 7.8 (pages 263 and 267) is a list of properties of binary relations:

	Sentence	*Property*
(a)	$\forall x Rxx$	reflexivity
(b)	$\forall x \neg Rxx$	irreflexivity
(c)	$\forall x \forall y Rxy$	universality; totality
(d)	$\forall x \forall y \neg Rxy$	emptiness; nullity
(e)	$\forall x \exists y Rxy$	seriality
(f)	$\forall x \forall y(Rxy \to Ryx)$	symmetry
(g)	$\forall x \forall y(Rxy \to \neg Ryx)$	asymmetry
(h)	$\forall x \forall y(Rxy \vee Ryx)$	strong connectedness; dichotomy
(i)	$\forall x \forall y \forall z((Rxy \wedge Ryz) \to Rxz)$	transitivity
(j)	$\forall x \forall y \forall z((Rxy \wedge Ryz) \to \neg Rxz)$	intransitivity
(k)	$\forall x \forall y \forall z((Rxy \wedge Rxz) \to Ryz)$	euclideanness
(l)	$\forall x \forall y(Rxy \to \exists z(Rxz \wedge Rzy))$	weak density

Use deductions for each of the following problems.

(1) Show that every total relation is symmetric, connected, transitive, and euclidean. That is, show that (c) implies (f), (h), (i), and (k).

(2) Show that every connected relation is reflexive. That is, show that (h) implies (a).

(3) Show that every reflexive relation is serial and weakly dense. That is, show that (a) implies (e) and (l).

(4) Show that the empty relation is symmetric, asymmetric, transitive, intransitive, euclidean, and weakly dense. That is, show that (d) implies (f), (g), (i), (j), (k), and (l).

(5) Show that every asymmetric relation is irreflexive. That is, show that (g) implies (b).

(6) Show that every symmetric transitive relation is euclidean. That is, show that (f), (i) together imply (k).

(7) Show that every reflexive euclidean relation is symmetric. That is, show that (a), (k) together imply (f).

(8) Show that every symmetric euclidean relation is transitive. That is, show that (f), (k) together imply (i).

(9) Show that every serial symmetric euclidean relation is reflexive. That is, show that (e), (f), (k) together imply (a).

(10) Show that every symmetric connected relation is total. That is, show that (f), (h) together imply (c).

(11) Show that every total asymmetric relation is empty. That is, show that (c), (g) together imply (d).

(12) Show that every irreflexive transitive relation is asymmetric. That is, show that (b), (i) together imply (g).

(13) Show that every intransitive relation is irreflexive. That is, show that (j) together imply (b).

(14) Show that every symmetric asymmetric relation is empty. That is, show that (f), (g) together imply (d).

(15) Show that every transitive intransitive relation is asymmetric. That is, show that (i), (j) together imply (g).

(16) Show that symmetry is equivalently expressed:

$$\forall x \forall y(Rxy \leftrightarrow Ryx)$$

That is, show that (f) is equivalent to this sentence.

(17) Show that transitivity is equivalently expressed:

$$\forall x \forall y(\exists z(Rxz \wedge Rzy) \rightarrow Rxy)$$

That is, show that (i) is equivalent to this sentence.

(18) Show that euclideanness is equivalently expressed:

$$\forall x \forall y(\exists z(Rzx \wedge Rzy) \rightarrow Rxy)$$

That is, show that (k) is equivalent to this sentence.

8.23. Describe deductions showing that $\exists y \forall x(Fx \wedge Gy)$ and $\exists x(Fx \wedge Gx)$ are each implied by $\forall x \exists y(Fx \wedge Gy)$.

8.24. Describe a deduction showing that $\exists xRax$ implies $\exists x \neg \neg Rax$. (Note that DN cannot be used to replace Rax by $\neg \neg$ Rax, since REP applies only to sentences and not to open formulas such as Rax.)

8. Postscript on replacement

As it stands, the rule REP of replacement is not as useful as it might be. To see this, consider the following deduction showing that $\exists xRax$ implies $\exists x \neg \neg Rax$ (compare exercise 8.24).

{1}	(1)	$\exists xRax$	P
{2}	(2)	Rab	P
{2}	(3)	$\neg \neg$ Rab	DN, 2
{2}	(4)	$\exists x \neg \neg Rax$	EG, 3
{1}	(5)	$\exists x \neg \neg Rax$	E, 1, 4

This is a rather elaborate way of getting from $\exists xRax$ to $\exists x \neg \neg Rax$, when all that seems to be involved is an inference by double negation (DN). Why not shorten the deduction to

{1}	(1)	$\exists xRax$	P
{1}	(2)	$\exists x \neg \neg Rax$	DN, 1

—where Rax is replaced (using DN) by $\neg \neg$ Rax? The answer is simply that, as it now stands, the rule REP applies only to *sentences* and not to open formulas like Rax and $\neg \neg$ Rax.

The solution is to extend REP to formulas that are open as well as those that are closed, i.e. to all formulas. To do this, we first extend the idea of deductive equivalence to all formulas:

DEFINITION 8.6. *Deductive equivalence of formulas.* Formulas A and B are deductively equivalent—A ~ B—iff any fresh instance of their biconditional A $\leftrightarrow$ B is a theorem. Moreover, A and B are deductively equivalent *in sentential logic* iff any fresh instance of their biconditional is a theorem of sentential logic.

Observe how this definition pronounces such non-sentential pairs as Rax and $\neg\neg$ Rax deductively equivalent: Their biconditional is the formula Rax $\leftrightarrow \neg\neg$ Rax. Clearly, any fresh instance of this, e.g. Rab $\leftrightarrow \neg\neg$ Rab, is a theorem—indeed a theorem of sentential logic, so the formulas Rax and $\neg\neg$ Rax are deductively equivalent in sentential logic. (See section 10 of chapter 6 for the idea of a fresh instance.)

The usefulness of extending deductive equivalence to cover all formulas should be evident in the following generalization of the rule of replacement (REP).

DEFINITION 8.7. *Generalized rule of replacement* (REP*).

REP*. $\dfrac{\Gamma,\ X}{\Gamma,\ X'}$ where X and X′ are alike except that X contains a subformula Y in one or more places where X′ contains a subformula Y′, and Y and Y′ are deductively equivalent in sentential logic or are formulas of the QN form: $\neg\forall$xA and $\exists$x$\neg$ A, or $\neg\exists$xA and $\forall$x$\neg$ A

This generalization is justified by the following proposition.

PROPOSITION 8.8. *Generalized replacement.* Let A and A′ be formulas alike except that A contains a subformula B in one or more places where A′ contains a subformula B′. Then A ~ A′ if B ~ B′.

Proof. The proof (by induction, omitted here) depends on the following lemma, which should be compared with the lemma on page 176 in chapter 4.

> *Lemma.* Suppose formulas A and B are deductively equivalent. Then the following pairs are also deductively equivalent.

		(i)	$\neg A,\ \neg B$
(ii)	$A \wedge C,\ B \wedge C$	(iii)	$C \wedge A,\ C \wedge B$
(iv)	$A \vee C,\ B \vee C$	(v)	$C \vee A,\ C \vee B$
(vi)	$A \rightarrow C,\ B \rightarrow C$	(vii)	$C \rightarrow A,\ C \rightarrow B$
(viii)	$A \leftrightarrow C,\ B \leftrightarrow C$	(ix)	$C \leftrightarrow A,\ C \leftrightarrow B$
		(x)	$\forall xA,\ \forall xB$
		(xi)	$\exists xA,\ \exists xB$

The proof of this, which we leave as an exercise, is similar to that for the lemma on page 176, except that the deductions begin and end with biconditionals. □

EXERCISES

8.25. Prove clauses (i), (ii), (iv), (vi), (viii), and (x) of the lemma above for proposition 8.8.

9. Interderivability of the rules

There is redundancy in our collection of quantificational rules of inference, in the sense that, as we have remarked, each of the rules UI, EG, UG, E, and QN is derivable given the presence of the others (and, of course, the basic rules of sentential logic).

For example, given UI and QN the rule EG can be shown to be implicitly present in the deductive system as follows. Suppose a deduction contains a line

$$\Gamma \quad (\,) \quad A(x/\alpha)$$

from which we wish to infer

$$\Gamma \quad (\,) \quad \exists xA$$

without using EG. In such a circumstance we may introduce these lines:

$\{\neg \exists xA\}$	( )	$\neg \exists xA$	P
$\{\neg \exists xA\}$	( )	$\forall x \neg A$	QN
$\{\neg \exists xA\}$	( )	$\neg A(x/\alpha)$	UI
Γ	( )	$\exists xA$	RAA

In this fashion we reach the desired line (the last one above). (Technically the premiss set of line inferred by RAA should be $\Gamma - \{\neg \exists xA\}$, but we know that $\neg \exists xA$ can always be restored to Γ if necessary.)

We can likewise show that E is derivable given sentential logic and just UG and QN. For suppose a deduction somewhere contains lines corresponding to the two hypotheses of the rule E,

$$\Gamma \quad (\,) \quad \exists xA$$

and

$$\Delta \cup \{A(x/\alpha)\} \quad (\,) \quad B,$$

where we assume that the name α is absent from $\exists xA$, from all the sentences in Δ, and from B. Then the addition of the following lines culminates in a line that we would otherwise wish to infer by the rule E.

$\{\neg B\}$	( )	$\neg B$	P
$\Delta \cup \{\neg B\}$	( )	$\neg A(x/\alpha)$	RAA
$\Delta \cup \{\neg B\}$	( )	$\forall x \neg A$	UG
$\Delta \cup \{\neg B\}$	( )	$\neg \exists xA$	QN
$\Gamma \cup \Delta$	( )	B	RAA

Taken together, the arguments above show that, with sentential logic at hand, we could take UI, UG, and QN as basic rules and dispense with both EG and E. The trio UI, UG, QN thus constitutes a reduced set of quantificational rules of inference.

By means of similar reasoning it is possible to reduce the quantificational rules in several other ways.

PROPOSITION 8.1. *Reduced sets of quantificational rules of inference.* On the basis of sentential logic, for each of the following sets of rules, each rule UI, EG, UG, E, and QN not in the set is derivable using those that are.

(1) UI, UG, QN (2) EG, UG, QN

(3) UI, E, QN (4) EG, E, QN

(5) UI, EG, UG, E

Proof. Part (1) we established above by showing how to derive EG given UI and QN and how to derive E given UG and QN. So for part (2) it is enough to show that UI is derivable given EG and QN, while for part (3) it needs to be shown only that UG can be derived using E and QN. Then the results for (2) and (3) yield (4). Finally, for part (5) it is sufficient to demonstrate the derivability of the QN rules. We leave all these arguments as exercises. □

Each of the reduced sets (1)–(5) in proposition 8.9 is in fact *minimal* (any rule omitted from a set cannot be derived using only those that remain), though we shall not prove this.

EXERCISES

8.26. Prove parts (2)–(5) of proposition 8.9. (The results of exercises 8.13, 8.15, and 8.17 may be helpful.)

9

COMPLETENESS FOR PREDICATE LOGIC

THIS IS the chapter in which we prove that the deductive system in chapter 8 is *complete*, that is, that a sentence A is a consequence of a set of sentences Γ if and only if it is deducible from the set:

$$\text{if } \Gamma \models A \text{ then } \Gamma \vdash A.$$

Recall that *soundness* —

$$\text{if } \Gamma \vdash A \text{ then } \Gamma \models A$$

— was established in chapter 8. So when completeness has been demonstrated we shall have an exact correspondence between consequence and deducibility:

$$\Gamma \models A \text{ iff } \Gamma \vdash A.$$

As in chapter 5 we approach the question of completeness in an equivalent formulation: every consistent set of sentences is satisfiable —

$$\text{if Con } \Gamma \text{ then Sat } \Gamma.$$

Section 1 contains two lemmas, one on *compositionality* and one on *relevance* These properties were required in the proof of soundness in chapter 8, and will be useful again. To prepare for completeness, section 2 introduces the idea of an *EI-closed* set of sentences. In section 3 we use the notion of EI-closure to prove an important *extension of Lindenbaum's lemma*. In section 4 we establish a result about *truth in covered models*. The completeness theorem is proved in section 5. In section 6 we return to the idea of compactness and show that deducibility and consistency are compact in predicate logic. Finally in section 7, we use soundness and completeness to infer a notable corollary, the *Löwenheim–Skolem theorem*.

1. Compositionality and relevance

To show the soundness of our deductive system for predicate logic (proposition 8.3, page 315), we used the property of *compositionality*—a proposition to the effect that if one name is substituted for another in a sentence then the truth value of the sentence will be unaffected so long as the names denote the same thing. Now we state the proposition formally and prove it.

PROPOSITION 9.1. *Compositionality*. $\models_{\mathcal{M}} A(\beta/\alpha)$ iff $\models_{\mathcal{M}^{\beta}_{\mathcal{M}(\alpha)}} A$.

Proof. By induction on the complexity of A. We treat the cases where A is atomic (for the basis) and where A is a universal quantification (in the inductive step); the other cases are left as exercises.

For the basis, we need a simple lemma concerning the compositionality principle as applied to names.

> *Lemma.* For any name γ, $\mathcal{M}(\gamma(\beta/\alpha)) = \mathcal{M}^{\beta}_{\mathcal{M}(\alpha)}(\gamma)$.

To see that this is so, suppose first that β and γ are the same name, i.e. that $\beta = \gamma$. Then $\mathcal{M}(\gamma(\beta/\alpha)) = \mathcal{M}(\beta(\beta/\alpha)) = \mathcal{M}(\alpha) = \mathcal{M}^{\beta}_{\mathcal{M}(\alpha)}(\beta) = \mathcal{M}^{\beta}_{\mathcal{M}(\alpha)}(\gamma)$. If, on the other hand, β and γ are different, then $\mathcal{M}(\gamma(\beta/\alpha)) = \mathcal{M}(\gamma) = \mathcal{M}^{\beta}_{\mathcal{M}(\alpha)}(\gamma)$.

Thus, where A is an atomic sentence, we may argue:

$\models_{\mathcal{M}} (\mathbb{P}\alpha_1 \ldots \alpha_n)(\beta/\alpha)$ iff $\models_{\mathcal{M}} \mathbb{P}\alpha_1(\beta/\alpha) \ldots \alpha_n(\beta/\alpha)$
—definition of replacement;
iff $\langle \mathcal{M}(\alpha_1(\beta/\alpha)), \ldots, \mathcal{M}(\alpha_n(\beta/\alpha)) \rangle \in \mathcal{M}(\mathbb{P})$
—definition 7.4, part (1);
iff $\langle \mathcal{M}^{\beta}_{\mathcal{M}(\alpha)}(\alpha_1), \ldots, \mathcal{M}^{\beta}_{\mathcal{M}(\alpha)}(\alpha_n) \rangle \in \mathcal{M}^{\beta}_{\mathcal{M}(\alpha)}(\mathbb{P})$
—lemma, agreement of $\mathcal{M}$ and $\mathcal{M}^{\beta}_{\mathcal{M}(\alpha)}$ on $\mathbb{P}$;
iff $\models_{\mathcal{M}^{\beta}_{\mathcal{M}(\alpha)}} \mathbb{P}\alpha_1 \ldots \alpha_n$
—definition 7.4, part (1).

Now assume as an inductive hypothesis that the proposition applies to every sentence less complex than A. Where A is a universal quantification $\forall$xB, let γ be a name distinct from both α and β—and hence new to both B and B(β/α). Then we reason as follows.

$\models_{\mathcal{M}} (\forall x B)(\beta/\alpha)$ iff $\models_{\mathcal{M}} \forall x(B(\beta/\alpha))$
— definition of substitution;
iff for every d in $\mathcal{M}$'s domain, $\models_{\mathcal{M}^{\gamma}_{d}} B(\beta/\alpha)(x/\gamma)$
— definition 7.4, part (7);
iff for every d in $\mathcal{M}$'s domain, $\models_{\mathcal{M}^{\gamma}_{d}} B(x/\gamma)(\beta/\alpha)$
— identity of $B(\beta/\alpha)(x/\gamma)$ and $B(x/\gamma)(\beta/\alpha)$;
iff for every d in $\mathcal{M}$'s domain, $\models_{\mathcal{M}^{\gamma,\beta}_{d,\mathcal{M}(\alpha)}} B(x/\gamma)$
— inductive hypothesis;
iff for every d in $\mathcal{M}^{\beta}_{\mathcal{M}(\alpha)}$'s domain, $\models_{\mathcal{M}^{\beta,\gamma}_{\mathcal{M}(\alpha),d}} B(x/\gamma)$
— identity of $\mathcal{M}^{\gamma,\beta}_{d,\mathcal{M}(\alpha)}$ and $\mathcal{M}^{\beta,\gamma}_{\mathcal{M}(\alpha),d}$;
iff $\models_{\mathcal{M}^{\beta}_{\mathcal{M}(\alpha)}} \forall x B$
— definition 7.4, part (7).

This ends the proof. □

The other property we used in paving the way toward soundness was *relevance*.

PROPOSITION 9.2. *Relevance.* The only thing relevant to the truth value of a sentence are the domain of the model and the values of the names and predicates in the sentence. In other words, models agree on the truth value of a sentence A whenever they have the same domain and agree on the values they assign to all the names and predicates in A.

Proof. Consider a sentence A. Let $\mathcal{M}$ and $\mathcal{M}^*$ be models having a common domain, D, and assume that (i) for each name α in A, $\mathcal{M}(\alpha) = \mathcal{M}^*(\alpha)$, and (ii) for each predicate $\mathbb{P}$ (of whatever degree) in A, $\mathcal{M}(\mathbb{P}) = \mathcal{M}^*(\mathbb{P})$. (If we call A's names and predicates its *items*, we can state assumptions (i) and (ii) simply by saying that $\mathcal{M}$ and $\mathcal{M}^*$ agree on the items of A.) Now we wish to show that $\mathcal{M}$ and $\mathcal{M}^*$ agree on the truth value of A:

$$\models_{\mathcal{M}} A \text{ iff } \models_{\mathcal{M}^*} A.$$

The proof is by induction on the complexity of A.

For the basis of the induction, where A is an atomic sentence $\mathbb{P}\alpha_1 \ldots \alpha_n$, we argue as follows.

$\models_{\mathcal{M}} \mathbb{P}\alpha_1 \ldots \alpha_n$ iff $\langle \mathcal{M}(\alpha_1), \ldots, \mathcal{M}(\alpha_n) \rangle \in \mathcal{M}(\mathbb{P})$
—definition 7.4, part (1);
iff $\langle \mathcal{M}^*(\alpha_1), \ldots, \mathcal{M}^*(\alpha_n) \rangle \in \mathcal{M}^*(\mathbb{P})$
—agreement of $\mathcal{M}$ and $\mathcal{M}^*$ on the items of A;
iff $\models_{\mathcal{M}^*} \mathbb{P}\alpha_1 \ldots \alpha_n$
—definition 7.4, part (1).

For the inductive step of the proof, we assume as an inductive hypothesis that the proposition holds for all sentences less complex than A. It is important to note that in this instance this means that if X is any sentence less complex than A then *any* models agree on the truth value of X whenever they have the same domain and agree on the values they assign to all the items of X.

To argue now that the proposition holds for A, first note that the cases in which A is a negation, conjunction, disjunction, conditional, or biconditional are unproblematic, since they were covered in effect in the proof in chapter 5 of the corresponding relevance property (proposition 5.9, page 216). So we are left with the cases in which A is a universal or existential quantification. Let us treat the first of these and leave the other as an exercise.

Suppose, then, that A is a universal quantification, $\forall$xB. The truth conditions for $\forall$xB involve those of a sentence B(x/α) where α is a name new to B, and so new to A. Hence B(x/α), though less complex than A, is not directly covered by the inductive hypothesis. The following lemma and its corollary resolve the problem.

Lemma. Where α is a name new to B (and hence A), and d is any element of the common domain of models $\mathcal{M}$ and $\mathcal{M}^*$,

$$\models_{\mathcal{M}^\alpha_d} \mathrm{B}(\mathrm{x}/\alpha) \text{ iff } \models_{\mathcal{M}^{*\alpha}_d} \mathrm{B}(\mathrm{x}/\alpha).$$

For suppose α and d are as advertised. Then the models $\mathcal{M}^\alpha_d$ and $\mathcal{M}^{*\alpha}_d$ have the same domain and agree on all names and predicates in B(x/α)—because d is the value of α in both α–variants of $\mathcal{M}$, and because $\mathcal{M}$ and $\mathcal{M}^*$ agree on the predicates in B(x/α) and on the rest of the names in the sentence. Since, as noted, $B(x/\alpha)$ is less complex than A, the inductive hypothesis applies with respect to the models $\mathcal{M}_d^\alpha$ and $\mathcal{M}^{*\alpha}_d$—and hence they agree on $B(x/\alpha)$. That is to say, $B(x/\alpha)$ is true in $\mathcal{M}_d^\alpha$ if and only if it is true in $\mathcal{M}^{*\alpha}_d$, as we wished to prove.

Corollary. Where α is a name new to B (and hence to A),

$$\begin{gathered}\text{for every } d \text{ in } \mathcal{M}\text{'s domain, } \models_{\mathcal{M}_d^\alpha} B(x/\alpha)\\ \text{iff}\\ \text{for every } d \text{ in } \mathcal{M}^*\text{'s domain, } \models_{\mathcal{M}^{*\alpha}_d} B(x/\alpha).\end{gathered}$$

This follows at once from the lemma.

Now to show that our proposition holds when A is $\forall xB$, we assume that the name α is new to B, and then reason thus:

$$\begin{aligned}\models_{\mathcal{M}} \forall xB \text{ iff }& \text{for every } d \text{ in } \mathcal{M}\text{'s domain, } \models_{\mathcal{M}_d^\alpha} B(x/\alpha)\\ &\quad\text{—definition 7.4, part (7);}\\ \text{iff }& \text{for every } d \text{ in } \mathcal{M}^*\text{'s domain, } \models_{\mathcal{M}^{*\alpha}_d} B(x/\alpha)\\ &\quad\text{—corollary to the lemma;}\\ \text{iff }& \models_{\mathcal{M}^*} \forall xB\\ &\quad\text{—definition 7.4, part (7).}\end{aligned}$$

With this the proof of proposition 9.2 is complete. □

EXERCISES

9.1. In the proof of compositionality (proposition 9.1), check that in the inductive case the sentences $B(\beta/\alpha)(x/\gamma)$ and $B(x/\gamma)(\beta/\alpha)$ are identical, and that the models $\mathcal{M}^{\gamma,\ \beta}_{d,\mathcal{M}(\alpha)}$ and $\mathcal{M}^{\beta,\ \gamma}_{\mathcal{M}(\alpha),d}$ are identical. Give the proofs of the remaining inductive cases (where A is a negation, conjunction, disjunction, conditional, biconditional, and existential quantification).

9.2. In the proof of relevance (proposition 9.2), check the inductive cases in which A is a negation, conjunction, disjunction, conditional, and biconditional. Then prove the result for the case in which A is an existential quantification. Here it is helpful to note another corollary of the lemma in the proof of 9.2: for some d in $\mathcal{M}$'s domain, $\models_{\mathcal{M}_d^\alpha} B(x/\alpha)$ iff for some d in $\mathcal{M}^*$'s domain, $\models_{\mathcal{M}^{*\alpha}_d} B(x/\alpha)$.

2. Preliminaries to completeness

Unlike the situation in sentential logic, completeness cannot be demonstrated simply by first showing that every consistent set of sentences has a maximal extension and then arguing that every maximal set has a model. Something more is needed—the idea of an EI-closed set of sentences.

A set of sentences is said to be EI-closed just in case whenever it contains an existential quantification it also contains at least one instance of the quantification. Formally:

DEFINITION 9.3. *EI-closed sets.* EI-clsd Γ iff whenever $\exists xA \in \Gamma$, $A(x/\alpha) \in \Gamma$, for some name α.

An EI-closed set of sentences is so to speak "closed under existential instantiation". This syntactic property combines with maximality to powerful effect, some indication of which appears in our next theorem.

PROPOSITION 9.4. Let Γ be a maximal EI-closed set of sentences.

(1) $\forall xA \in \Gamma$ iff for every name α, $A(x/\alpha) \in \Gamma$.

(2) $\exists xA \in \Gamma$ iff for some name α, $A(x/\alpha) \in \Gamma$.

Proof. Assume that Γ is both maximal and EI-closed. The left-to-right of (1) is left as an exercise. For right to left we argue contrapositively. Suppose that $\forall xA$ is not in Γ. Because Γ is maximal, Γ contains $\neg\forall xA$. But $\exists x\neg A$ is deducible from $\neg\forall xA$. So $\exists x\neg A$ is deducible from Γ. By reason of Γ's maximality, then, $\exists x\neg A$ is also a member of Γ. Since Γ is EI-closed, $\neg A(x/\alpha)$ is in Γ, for some name α. But again by maximality, this means that $A(x/\alpha)$ lies outside of Γ. In other words, it is not the case that for every name α, $A(x/\alpha)$ is in Γ—as we wished to prove.

The left-to-right of (2) obtains simply by definition of EI-closure. For right-to-left, suppose that $A(x/\alpha)$ is in Γ, for some name α. Of course $\exists xA$ is deducible from $A(x/\alpha)$, and so it is deducible from Γ. Hence by the maximality of Γ, Γ contains $\exists xA$. □

Let us close this section with one further result that will be useful in the proof of proposition 9.6 in the next section.

PROPOSITION 9.5. Suppose Γ contains the sentence $\exists xA$, and α is a name new to all the sentences in Γ. If Con Γ, then Con $\Gamma \cup \{A(x/\alpha)\}$.

Proof. Note that this proposition does not presuppose that the set Γ is EI-closed. We give the proof contrapositively. Suppose that Γ contains $\exists xA$ and $\Gamma \cup \{A(x/\alpha)\}$ is not consistent, where α is a name not present in any sentence in Γ (and therefore not present in $\exists xA$). Then $\neg A(x/\alpha)$ is deducible from Γ. By UG—since α is new to Γ—it follows that $\forall x \neg A$ is deducible from Γ. By QN, it follows that $\neg \exists xA$ is also deducible from Γ. But then both $\exists xA$ and $\neg \exists xA$ are deducible from Γ, and hence Γ is inconsistent, as we wished to show. □

EXERCISES

9.3. Prove the left-to-right of part (1) of proposition 9.4.

9.4. Let us say that a set of sentences is *UG-closed* ("closed under universal generalization") if and only if it contains a universal quantification whenever it contains all its instances. In other words, Γ is UG-closed just in case whenever $A(x/\alpha) \in \Gamma$, for every name α, then also $\forall xA \in \Gamma$. Show that a maximal set of sentences is UG-closed if and only if it is EI-closed. Note that this means that proposition 9.4 holds as well for maximal UG-closed sets.

9.5. To say that a set of sentences is *ω-complete* ("omega complete") means that a universal quantification is deducible from the set whenever every one of its instances is deducible from it. In other words, Γ is ω-complete if and only if whenever $\Gamma \vdash A(x/\alpha)$, for every name α, then also $\Gamma \vdash \forall xA$. Referring to exercise 9.4, prove that as regards maximal sets of sentences the properties of EI-closure, UG-closure, and ω-completeness all come to the same thing. In other words, show that if a set is maximal then (1) it is ω-complete if and only if it is UG-closed, and, hence, (2) it is ω-complete if and only if it is EI-closed.

9.6. A set of sentences is said to be *ω-consistent* ("omega consistent") just in case a universal quantification is consistent with the set whenever every one of its instances is deducible from it. In other words, Γ is ω-consistent if and only if Con $\Gamma \cup \{\forall xA\}$ whenever $\Gamma \vdash A(x/\alpha)$ for every name α. Prove that every negation complete ω-consistent set of sentences is ω-complete. (Recall negation completeness from exercise 5.4, page 205.)

9.7. Show that whenever Γ is ω-complete so is $\Gamma \cup \{A\}$, for any sentence A.

9.8. In the proof of proposition 9.5, check that (as claimed) $\neg \exists xA$ is deducible from Γ, given that $\neg A(x/\alpha)$ is.

3. Lindenbaum's lemma extended

The following is the analogue for predicate logic of Lindenbaum's lemma (proposition 5.3, page 206); indeed, it contains Lindenbaum's lemma within it. The proposition asserts that every consistent set of sentences lies within some maximal EI-closed set.

PROPOSITION 9.6. *Extension of Lindenbaum's lemma.* If Con Γ, then there exists a set of sentences Δ such that (i) $\Gamma \subseteq \Delta$, (ii) Max Δ, and (iii) EI-clsd Δ.

Proof. Once again we begin with a set of sentences Γ that we assume to be consistent; i.e. we assume that Con Γ.

For the duration of the proof, we again suppose we have a fixed enumeration of the set of sentences,

$$A_1,\ A_2,\ A_3,\ \ldots$$

—but this time with a wrinkle. To wit, we assume also that the sentences in the list are arranged so that whenever an existential quantification $\exists xA$ occurs in it, it is followed immediately by an instance $A(x/\alpha)$—with the proviso that the name α in $A(x/\alpha)$ is new to all the preceding sentences in the list (hence including $\exists xA$) and to all the sentences in the set Γ. That is to say, the enumeration of sentences $A_1, A_2, A_3, \ldots$ is to have the property that

> whenever A_n is of the form $\exists xA$, the sentence A_{n+1} is of the form $A(x/\alpha)$, where α is a name not in any of the sentences up to and including A_n and not in any sentence in Γ.

It is this feature of the enumeration of the sentences that is responsible ultimately for the EI-closure of the set Δ we are about to define.

(There is a small problem, however. How do we know that a sentence of the form $\exists xA$ can always be followed by an instance $A(x/\alpha)$ in which α is new to $\exists xA$ and to all the sentences in Γ? For example, what if our consistent set Γ contains so many sentences that all the names in the language are used up? One solution—ours—is simply to posit the existence of an infinite class of names, all distinct from those ordinarily found in the language, brought in specially for the purposes of the proof. Another solution—discussed in exercise 9.12—is in effect to assume that the names in Γ's sentences have only even-numbered indices, so that the infinitely many oddly-indexed names are left over for the prescribed construction of the instances following the existential quantifications in the enumeration.)

Using Γ and the enumeration of the sentences, we define the sets

$$\Delta_0,\ \Delta_1,\ \Delta_2,\ \ldots$$

just as we did before.

Definition 1.

(i) $\Delta_0 = \Gamma$.

(ii) $$\Delta_n = \begin{cases} \Delta_{n-1} \cup \{A_n\}, \text{ if Con } \Delta_{n-1} \cup \{A_n\} \\ \Delta_{n-1}, \text{ if } \text{C}\not{\text{o}}\text{n } \Delta_{n-1} \cup \{A_n\} \end{cases} \quad n > 0.$$

Compare definition 1 in the proof of Lindenbaum's lemma in chapter 5.

Lemmas 2 and 3 below are counterparts of those in the proof of Lindenbaum's lemma, and counterparts of lemmas 5, 6, 7, 8, and 9 (following definition 4) appear there as well. Because the proofs of the lemmas are unaltered, we will simply register them here.

Lemma 2. Con Δ_n, for each $n \geq 0$.

Lemma 3. $\Delta_i \subseteq \Delta_j$, whenever $i \leq j \geq 0$.

Again as before, we define our desired set Δ as the infinite union of the sets $\Delta_0, \Delta_1, \Delta_2, \ldots$.

Definition 4. $\Delta = \Delta_0 \cup \Delta_1 \cup \Delta_2 \cup \ldots$.

Compare the definition in the proof of Lindenbaum's lemma.

Lemma 5. $\Delta_n \subseteq \Delta$, for each $n \geq 0$.

Lemma 6. $A_k \in \Delta_k$ whenever $A_k \in \Delta$, for each $k > 0$.

Lemma 7. For every finite subset Δ' of Δ, $\Delta' \subseteq \Delta_n$ for some $n \geq 0$.

Lemma 8. $\Gamma \subseteq \Delta$.

Lemma 9. Max Δ.

Thus Δ is a maximal extension of Γ, and it remains to be shown that Δ is also EI-closed.

Lemma 10. EI-clsd Δ.

For lemma 10, we need to show that whenever a sentence of the form $\exists xA$ is in Δ, there is also a sentence of the form $A(x/\alpha)$ in Δ. So suppose that $\exists xA$ is in Δ. Then, for some integer n, $\exists xA$ is the nth sentence in the enumeration $A_1, A_2, A_3, \ldots$, and $A(x/\alpha)$ follows it as the $n + 1$st sentence. In other words:

$$\exists xA = A_n \text{ and } A(x/\alpha) = A_{n+1},$$

where α is a name new to all the sentences in the sequence up to and including $\exists xA$, and new to all the sentences in Γ as well. To say that $\exists xA$ is in Δ is to say, then, that A_n is in Δ. By lemma 6 above, this means that A_n is in the set Δ_n in the sequence $\Delta_0, \Delta_1, \Delta_2, \ldots$. That is to say, $\exists xA$ is in Δ_n.

It is important now to observe that because the name α in $A(x/\alpha)$ (i.e. in A_{n+1}) does not occur in any sentence in Γ and does not occur in any of the sentences ahead of A_n in the enumeration, it does not appear in any sentence in the set Δ_n. Therefore we infer by proposition 9.5 that the set $\Delta_n \cup \{A(x/\alpha)\}$ is consistent. In other words, $\Delta_n \cup \{A_{n+1}\}$ is consistent. But then, by definition 1, $\Delta_{n+1} = \Delta_n \cup \{A_{n+1}\}$—i.e. $\Delta_{n+1} = \Delta_n \cup \{A(x/\alpha)\}$. So $A(x/\alpha)$ is in Δ_{n+1}. But as Δ_{n+1} is a subset of Δ (lemma 5), this means that $A(x/\alpha)$ is in Δ. This ends the proof of lemma 10.

With lemmas 8, 9, and 10 established, the proof of this extension of Lindenbaum's lemma is complete: Every consistent set of sentences is a subset of a maximal EI-closed set of sentences. □

Notice that in the proof just concluded we appealed to various propositions which themselves depend ultimately on the nature of the rules of inference that characterize the deductive system. Thus this extension of Lindenbaum's lemma holds good for any deductive system in which the rules for sentential operators—P, PC, RAA, MP, etc.—appear with the quantificational rules UI, EG, UG, E, and QN. This fact is useful in chapter 10, where we shall want to know that the proof of this proposition can be carried out for a logic stronger than that of the deductive system for predicate logic.

EXERCISES

9.9. Check lemmas 2, 3, 5, 6, 7, 8, and 9 in the proof of proposition 9.6.

9.10. Where A is a sentence, let 2A be the sentence that results when the numerical indices are doubled on all the names that occur in A. Use familiar quantificational rules of inference to explain why A is a theorem if and only if 2A is a theorem.

9.11. Referring to exercise 9.10, explain why A is consistent just in case 2A is.

9.12. Use exercise 9.11 to explain how the problem of "not enough names", mentioned in the proof of proposition 9.6, can be circumvented under the assumption that the names in the sentences in Γ have only even-numbered indices. This boils down to showing that, where 2Γ is the set of "doubles" of sentences in Γ, Γ is consistent if and only if 2Γ is. The compactness property for consistency—proposition 4.28 (page 193)—will play a role in the argument.

4. Truth in covered models

We are nearly to our goal of showing that the deductive system for predicate logic is complete. Our final preliminary has to do with covered models. Recall from chapter 7 that a model is covered when every element in its domain is the value of at least one name (definition 7.2, page 264).

Our interest in covered models stems from the fact, noted earlier, that in them the truth conditions for quantifications can be explained in terms of the truth conditions of their instances, without reference to the notion of variants of a model. This is the content of the following proposition.

PROPOSITION 9.7. *Truth in a covered model.* Let $\mathcal{M}$ be a covered model.

(1) $\models_{\mathcal{M}} \forall xA$ iff for every name α, $\models_{\mathcal{M}} A(x/\alpha)$.

(2) $\models_{\mathcal{M}} \exists xA$ iff for some name α, $\models_{\mathcal{M}} A(x/\alpha)$.

Proof. Let us prove part (2) and leave part (1) as an exercise. Notice that the right-to-left of (2) is simply the lemma for the soundness of the rule EG (see proposition 7.5, page 286). So we may concentrate on left-to-right. Assume, then, that $\exists xA$ is true in $\mathcal{M}$. By the definition of truth (7.4, part (8), page 273), this means that, where β is a name new to A, the sentence $A(x/\beta)$ is true in $\mathcal{M}^{\beta}_{d}$ for some d in the domain of $\mathcal{M}$. Because $\mathcal{M}$ is covered, d is the value in $\mathcal{M}$ of some name α—i.e. there is some name α such that $d = \mathcal{M}(\alpha)$. Let α be such a name. Then $A(x/\beta)$ is true in the model $\mathcal{M}^{\beta}_{\mathcal{M}(\alpha)}$. By compositionality (proposition 9.1), it follows that A(x/β)(β/α) is true in the model $\mathcal{M}$. But since β is new to A, the sentences A(x/β)(β/α) and A(x/α) are the same. Thus A(x/α) is true in $\mathcal{M}$, as we wished to show. □

The usefulness of proposition 9.7 will become apparent in the next section.

EXERCISES

9.13. Prove part (1) of proposition 9.7.

5. Completeness

Proposition 9.6 tells us that every consistent set of sentences is included in a maximal EI-closed set of sentences. Thus if we can show that every maximal EI-closed set is satisfiable, our completeness result will be at hand. For then we shall know that every consistent set of sentences is a subset of a satisfiable set and is thus satisfiable itself. The next proposition asserts the satisfiability of maximal EI-closed sets of sentences.

PROPOSITION 9.8. If Max Γ and EI-clsd Γ, then Sat Γ.

Proof. Let Γ be a maximal EI-closed set of sentences. We define the model $\mathcal{M}$ (the *canonical model* with respect to Γ) as follows.

(i) D (the domain of $\mathcal{M}$) = {α: α is a name}.
(ii) $\mathcal{M}(\alpha) = \alpha$, for each name α.
(iii) $\langle \alpha_1, \ldots, \alpha_n \rangle \in \mathcal{M}(\mathbb{P})$ iff $\mathbb{P}\alpha_1 \ldots \alpha_n \in \Gamma$, for each n-place predicate $\mathbb{P}$ ($n \geq 0$).

That is to say: (i) the domain D of the canonical model $\mathcal{M}$ is the set of all names in the language; (ii) $\mathcal{M}$ assigns each name to itself as value; and (iii) the value under $\mathcal{M}$ of an n-place predicate is the set of all n-tuples of names in the language that occur following the predicate to make atomic sentences that appear in the set Γ. Stated in another, equivalent fashion, clause (iii) says that

$$\mathcal{M}(\mathbb{P}) = \{\langle \alpha_1, \ldots, \alpha_n \rangle : \mathbb{P}\alpha_1 \ldots \alpha_n \in \Gamma\}.$$

Clearly, $\mathcal{M}$ *is* a model: it has a nonempty domain, names are given values in the domain, and the value of each n-place predicate is a set of n-tuples of elements of the domain. Note that the domain of the model $\mathcal{M}$ is covered, in the sense of definition 7.2: every element of D is the value under $\mathcal{M}$ of some name α, namely (!) itself.

Now we prove that just those sentences are true in the model $\mathcal{M}$ as are members of the set Γ. That is, we have the following *truth lemma*:

Lemma. $\models_{\mathcal{M}} \mathrm{A}$ iff $\mathrm{A} \in \Gamma$, for every sentence A.

The proof of this lemma is inductive. We treat only the cases in which A is (a) atomic and (b) a universal quantification. The rest are left as exercises. (Note that the cases in which A is a negation or a conditional are treated in the proof of the corresponding lemma, on page 211, in the proof of proposition 5.4.)

For (a). Suppose A is atomic, $\mathbf{P}\alpha_1 \ldots \alpha_n$. We argue thus:

$$\begin{aligned}
\models_{\mathcal{M}} \mathbf{P}\alpha_1 \ldots \alpha_n \text{ iff } & \langle \mathcal{M}(\alpha_1), \ldots, \mathcal{M}(\alpha_n)\rangle \in \mathcal{M}(\mathbf{P}) \\
& \quad\text{—definition 7.4, part (1);} \\
\text{iff } & \langle \alpha_1, \ldots, \alpha_n\rangle \in \{\langle \alpha_1, \ldots, \alpha_n\rangle : \mathbf{P}\alpha_1 \ldots \alpha_n \in \Gamma\} \\
& \quad\text{—definition of } \mathcal{M}; \\
\text{iff } & \mathbf{P}\alpha_1 \ldots \alpha_n \in \Gamma. \\
& \quad\text{—more simply.}
\end{aligned}$$

So the lemma holds when A is atomic.

For the inductive cases, we assume as an inductive hypothesis that the lemma holds for all sentences less complex than A.

For (b). Suppose A is a universal quantification, $\forall x\mathrm{B}$. Note that for every name α the sentence $\mathrm{B}(x/\alpha)$ is shorter than A. So the inductive hypothesis tells us that

for every name α, $\models_{\mathcal{M}} \mathrm{B}(x/\alpha)$ iff $\mathrm{B}(x/\alpha) \in \Gamma$

—from which it follows that

$$\models_{\mathcal{M}} B(x/\alpha) \text{ for every name } \alpha \text{ iff } B(x/\alpha) \in \Gamma \text{ for every name } \alpha.$$

Therefore:

$$\begin{aligned} \models_{\mathcal{M}} \forall x\, B \text{ iff for every name } \alpha, \models_{\mathcal{M}} B(x/\alpha) \\ \text{—proposition 9.7;} \\ \text{iff for every name } \alpha, B(x/\alpha) \in \Gamma \\ \text{—inductive hypothesis;} \\ \text{iff } \forall x B \in \Gamma \\ \text{—proposition 9.4.} \end{aligned}$$

So the lemma holds when A is a universal quantification.

This ends our proof of the lemma.

We see thus that all the sentences in the maximal EI-closed set Γ are true in the model $\mathcal{M}$. So Γ is satisfiable. This ends the proof of proposition 9.8. □

PROPOSITION 9.9. *Completeness.* The deductive system for predicate logic is complete. That is, if $\Gamma \models A$ then $\Gamma \vdash A$.

Proof. As before, in the case of completeness for sentential logic, we reason for an equivalent statement: every consistent set of sentences is satisfiable. Suppose that Γ is a consistent set of sentences. Then by proposition 9.6, Γ has a maximal EI-closed extension Δ. By proposition 9.8, moreover, the set Δ is satisfiable. Therefore, by diminution for satisfiability, Γ is likewise satisfiable. □

Let us state some corollaries to completeness, along with the proposition itself.

PROPOSITION 9.10. *Completeness and corollaries.*

(1) If $\Gamma \models A$ then $\Gamma \vdash A$.

(2) If $\models A$ then $\vdash A$.

(3) If $A \simeq B$ then $A \sim B$.

(4) If Con Γ then Sat Γ.

Proof. Exercise (compare the proof for proposition 5.6, page 213). □

EXERCISES

9.14. Complete the proof of the lemma in the proof of proposition 9.8 by arguing for the inductive cases in which A is a negation, conjunction, disjunction, conditional, biconditional, and existential quantification.

9.15. Prove proposition 9.10.

9.16. *An alternative canonical model.* Let Γ be a maximal EI-closed set of sentences, and define the model $\mathcal{M}$ as follows.

(i) D (the domain of $\mathcal{M}$) = $\mathbf{N}$.
(ii) $\mathcal{M}(\mathrm{a}_n) = n$, for each $n \in \mathbf{N}$.
(iii) $\langle k_1, \ldots, k_n \rangle \in \mathcal{M}(\mathbb{P})$ iff $\mathbb{P}\alpha_{k_1} \ldots \alpha_{k_n} \in \Gamma$, for each n-place predicate $\mathbb{P}$ ($n \geq 0$) and natural numbers $k_1, \ldots, k_n$.

That is to say: (i) the domain D of the model $\mathcal{M}$ is the set of all natural numbers $\{0, 1, 2 \ldots\}$; (ii) $\mathcal{M}$ assigns each name a_n to its numerical index n as value; and (iii) the value in $\mathcal{M}$ of an n-place predicate is the set of all n-tuples of numbers that index the names that follow the predicate to make atomic sentences that appear in the set Γ. It should be evident that $\mathcal{M}$ is a model: it has a nonempty domain, names take values in the domain, and each n-place predicate's value is a set of n-tuples of elements of the domain. Note that the model $\mathcal{M}$ is covered: every element of D is the value in $\mathcal{M}$ of some (indeed, just one) name.

Prove that just those sentences are true in the model $\mathcal{M}$ as are members of the set Γ. That is, prove that for every sentence A, $\models_{\mathcal{M}} \mathrm{A}$ iff $\mathrm{A} \in \Gamma$. Give an inductive argument for the cases in which A is atomic, a negation, a conditional, and a universal quantification. Thus $\mathcal{M}$ is entitled to be called a canonical model.

This result provides an alternative way of proving that every maximal EI-closed set of sentences has a model (proposition 9.8).

6. Compactness, and another corollary

Of course, compactness holds for deducibility and consistency (propositions 4.12 and 4.28, pages 187 and 193):

$$\text{if } \Gamma \vdash \text{A, then } \Gamma' \vdash \text{A for some finite subset } \Gamma' \text{ of } \Gamma,$$

and

$$\text{Con } \Gamma, \text{ if Con } \Gamma' \text{ for every finite subset } \Gamma' \text{ of } \Gamma.$$

As was the case in sentential logic, the soundness and completeness propositions yield corresponding compactness results for consequence and satisfiability.

PROPOSITION 9.11. *Compactness for consequence and satisfiability.*

(1) If $\Gamma \models$ A, then $\Gamma' \models$ A for some finite subset Γ' of Γ.

(2) Sat Γ, if Sat Γ' for every finite subset Γ' of Γ.

Proof. See the proof for proposition 5.7 (page 214). □

In the light of compactness, we have several equivalent statements of soundness and completeness. The following theorem is identical with 5.8.

PROPOSITION 9.12. In each column below, each statement is equivalent to each of the others.

	Soundness	*Completeness*	*Both*
(1)	$\Gamma \vdash A \Rightarrow \Gamma \models A$	$\Gamma \models A \Rightarrow \Gamma \vdash A$	$\Gamma \models A$ iff $\Gamma \vdash A$
(2)	$\vdash A \Rightarrow \models A$	$\models A \Rightarrow \vdash A$	$\models A$ iff $\vdash A$
(3)	$A \sim B \Rightarrow A \simeq B$	$A \simeq B \Rightarrow A \sim B$	$A \simeq B$ iff $A \sim B$
(4)	Sat $\Gamma \Rightarrow$ Con Γ	Con $\Gamma \Rightarrow$ Sat Γ	Con Γ iff Sat Γ

Proof. Again, we leave it as an exercise to work out the implications and necessary arguments. Compare exercise 5.14 (page 215). □

EXERCISES

9.17. Prove proposition 9.11.

9.18. Prove proposition 9.12.

7. Löwenheim–Skolem

The *size* (technically, the *cardinality*) of a model is the number of elements in the domain of the model. Thus if the domain of a model is the set of provinces in Canada, the model is said to be of size (or cardinality) ten, since there are that many Canadian provinces. Similarly, if the domain is the set $\boldsymbol{N}$ of natural numbers, the model is said to be infinite—more precisely to have the size of the lowest order of infinity, the cardinal number for which is $\aleph_0$ ("aleph nought" or "aleph null"). By contrast, if a model's domain is the set $\boldsymbol{R}$ of all real numbers, its size is that of the continuum, c, an even higher order of infinity.

One of the most interesting propositions about predicate logic is the "Löwenheim–Skolem theorem", according to which if a set of sentences is satisfiable then it has a model with a domain the size of the set of natural numbers.

PROPOSITION 9.13. *Löwenheim–Skolem theorem.* If Sat Γ, then there is a model $\mathcal{M}$ such that (i) the domain of $\mathcal{M}$ is the size of the set $\boldsymbol{N}$, and (ii) every sentence in Γ is true in $\mathcal{M}$.

Proof. Suppose Γ is a satisfiable set. By soundness, Γ is consistent, and hence by proposition 9.6, Γ has a maximal EI-closed extension Δ. By the construction in the proof of proposition 9.8, the domain of the canonical model for Δ is the set $\{a_0, a_1, a_2, \ldots\}$ of names in the language. The size of this set is obviously the same as the set $\{0, 1, 2, \ldots\}$ of natural numbers, and of course every sentence in Γ is true in the model. □

In cahoots with soundness, completeness, and compactness, the Löwenheim–Skolem theorem yields many further results. We set out some of these in the exercises.

EXERCISES

9.19. The Löwenheim–Skolem theorem can also be formulated as follows. If Sat Γ, then there is a model $\mathcal{M}$ such that (i) the domain of $\mathcal{M}$ *is* the set $\mathbf{N}$, and (ii) every sentence in Γ is true in $\mathcal{M}$. Prove this (with reference to exercise 9.16).

9.20. The quantifier $\exists!$ has the meaning *there is exactly one*. In formal terms, the truth conditions for $\exists!$ are:

$$\models_{\mathcal{M}} \exists!\mathrm{xA} \text{ iff for exactly one } d \text{ in } \mathcal{M}\text{'s domain, } \models_{\mathcal{M}_d^\alpha} \mathrm{A(x/\alpha)}$$

—where α is new to A. Using the Löwenheim–Skolem theorem, explain why $\exists!$ is not definable in the language of predicate logic. That is, show that some sentence of the form $\exists!$xA is not equivalent to any sentence B in the language. (*Suggestion*: Consider a sentence of the form $\exists!$xA where A's content is trivial, e.g. $\exists!$x(Fx $\vee \neg$ Fx).) What about the quantifier $\exists?$, meaning *there is at most one*—is it definable in the language?

9.21. Suppose the quantifier Ǝ means *there are finitely many*—i.e.

$$\models_{\mathcal{M}} \text{Ǝ}\mathrm{xA} \text{ iff for finitely many } d\text{s in } \mathcal{M}\text{'s domain, } \models_{\mathcal{M}_d^\alpha} \mathrm{A(x/\alpha)}$$

—where α is new to A. Explain why Ǝ is not definable in the language of predicate logic.

9.22. Suppose the quantifier I means *there are infinitely many*—i.e.

$$\models_{\mathcal{M}} \mathsf{I}x\mathrm{A} \text{ iff for infinitely many } d\text{s in } \mathcal{M}\text{'s domain, } \models_{\mathcal{M}_d^\alpha} \mathrm{A}(x/\alpha)$$

—where α is new to A. Explain why I is not definable in the language.

9.23. Suppose the quantifier ∩ means *there are uncountably, or nondenumerably many*—i.e.

$$\models_{\mathcal{M}} \cap x\mathrm{A} \text{ iff for uncountably many } d\text{s in } \mathcal{M}\text{'s domain, } \models_{\mathcal{M}_d^\alpha} \mathrm{A}(x/\alpha)$$

—where α is new to A. Explain why ∩ is not definable in the language.

9.24. Assume that A implies B, A is satisfiable, and B is invalid. Prove that A and B have at least one item (predicate or name) in common. *Hint*: Use relevance (proposition 9.2) as well as the version of Löwenheim–Skolem in exercise 9.19. (Compare exercise 5.18, page 218.)

9.25. The Löwenheim–Skolem theorem is sometimes called the "*downward* Löwenheim–Skolem theorem", because it guarantees that a set of sentences having a model whose size is any higher order of infinity than that of the natural numbers also has a model whose size is the lowest order of infinity (like the set ***N***). The so-called "*upward* Löwenheim–Skolem theorem" (sometimes referred to as the "Löwenheim–Skolem–Tarski theorem") is the proposition that if a set of sentences has a model of a given size it also has models of every greater size. Prove this in a simple, finitary version: if a sentence has a model of size n it also has a model of size $n + 1$.

Suggestion: Begin with a sentence A true in a model $\mathcal{M}$ with domain D of size n, and let d be some member of D (it does not matter which). Construct a new model $\mathcal{M}^+$ with a domain D^+ consisting of everything in D plus one new thing d^+. Thus $\mathcal{M}^+$ is of size $n + 1$. Next define the values of names in $\mathcal{M}^+$ to agree with their values in $\mathcal{M}$ (hence the new element d^+ is not named in $\mathcal{M}^+$). Then make sure the values of predicates in $\mathcal{M}^+$ are at least what they were in $\mathcal{M}$. Finally, extend the values of predicates in $\mathcal{M}^+$ so that the new element d^+ "behaves" in $\mathcal{M}^+$ just as the old one d does in $\mathcal{M}$. The effect of this construction is to make the new element not only anonymous, but also invisible so far as the language of predicate logic is concerned: because the behavior of d^+ exactly imitates that of d, no sentence in the language distinguishes one of them from the other. Demonstrate this by proving inductively that any sentence is true in $\mathcal{M}^+$ just in case it is true in $\mathcal{M}$. Then conclude, as desired, that A has a model of size $n + 1$.

9.26. Use the "upward" Löwenheim–Skolem theorem to explain why the quantifier $\exists?$ in exercise 9.20 is not definable in the language of predicate logic. Then use the theorem to explain why the quantifier $\exists?_n$, meaning *there are at most n*, is not definable in the language. Finally, use the latter result to show that the quantifier $\exists_n$, meaning *there are at least n*, is not definable in the language.

9.27. Suppose the quantifier $\exists_\omega$ means *there are exactly as many as there are natural numbers*. Explain why $\exists_\omega$ is not definable in the language of predicate logic. (Use the "upward" Löwenheim–Skolem theorem in its most general form.)

PART III

IDENTITY

10

THE LOGIC OF IDENTITY

THE LANGUAGE of predicate logic with *identity* is a modest but powerful extension of that introduced in chapter 6, in which a predicate for identity is introduced and treated as a logical constant on a par with the sentential operators and quantifiers. The logic of the new language extends that of its predecessor particularly by validating inferences involving *substitution*. Thus for example, the argument

> Mark Twain wrote *Roughing it*
> Mark Twain = Sam Clemens
> ———
> Sam Clemens wrote *Roughing it*

can now be shown to be valid. The conclusion follows by substituting the name Sam Clemens for the name Mark Twain in the first premiss, a move licensed by the identity in the premiss Mark Twain = Sam Clemens.

The greater expressive power of the new language is evidenced by its ability to translate sentences that express numerical propositions, i.e. propositions about how many things—or things of this or that kind—there are. As we shall see, can now adequately translate such sentences as There are at least two frogs and The present King of France is bald.

In section 1 we describe the new *language*, and in section 2 we consider its semantics and explain the *truth conditions* of identity sentences in its *models*. Questions of *translation* are taken up in section 3, and *deduction* is the topic of section 4. The *soundness* of the new inference rules is established in section 5, and some *interderivability* is noted in section 6.

The last three sections are devoted to metatheory: *completeness* (section 7); *compactness* (section 8); and a revised *Löwenheim–Skolem theorem* (section 9). These sections rely on earlier such results and may be skipped by those who have not read chapters 5 and 9.

1. The language

The language of predicate logic with identity is obtained by adding to the earlier language a new logical constant—the two-place *identity predicate*:

$$\mathbb{I}$$

Thus in addition to all the familiar formulas, the language now contains new *atomic formulas* called *identity formulas*:

$$\mathbb{I}\tau\sigma$$

Formally, to get a proper definition of the set of *formulas*, we subjoin to clause (1) of definition 6.1 (of formulas of predicate logic) the clause

(1′) $\mathbb{I}\tau\sigma$ is a formula.

Other definitions, e.g. of a *free variable* and of a *sentence*, are extended to cover formulas of the new form, so that, as before, sentences are formulas without free variables. (Hence an identity formula using only names is an *identity sentence*.) Further notions—such as *replacement* and *substitution*—apply here as they did earlier; see chapter 6.

It is convenient to indulge in the informality of writing

$$\tau = \sigma,$$

instead of $\mathbb{I}\tau\sigma$, for identity formulas. Let us also write the negation of an identity formula as

$$\tau \neq \sigma,$$

instead of $\neg\, \tau = \sigma$ (although we occasionally resort to the proper form). Of course we are already using the identity symbol "=" in the metalanguage to express identity, but the context of occurrence should always suffice to resolve ambiguity.

2. Semantics

Consonant with our intention to treat = as a logical constant—i.e. as an expression whose meaning does not change from model to model—we do not specify its meaning in our account of a model. The meaning of = will be fixed in the definition of truth in a model.

Thus a *model* for the language of predicate logic with identity is just what it was before, a nonempty domain together with an assignment of values from the domain to the names and predicates (not including =). So we may use definitions 7.1 and 7.3, of models and variants of models.

To definition 7.4 of truth in a model it is sufficient to add a single clause specifying the truth conditions of identity sentences:

(1′) $\models_{\mathcal{M}} \alpha = \beta$ iff $\mathcal{M}(\alpha) = \mathcal{M}(\beta)$.

Thus an identity sentence is true in a model just in case the names in the sentence have the same value in the model—i.e. just in case, according to the model, the names name the same thing.

The idea of a model of a set of sentences (or sentence) and the idea of a model of n elements are also carried over.

Though the definitions of consequence, validity, equivalence, and satisfiability have not changed in form, their extensions have again increased. Let us consider the following new results, which establish the soundness of the new rules of inference set out in section 4 below.

PROPOSITION 10.1. *Soundness of the rules of inference for identity.*

(1) $\models \alpha = \alpha$.

(2) $\alpha = \beta, A(\gamma/\alpha) \models A(\gamma/\beta)$.

(3) $\beta = \alpha, A(\gamma/\alpha) \models A(\gamma/\beta)$.

Proof. Part (1) expresses instances of the reflexivity of the identity relation—that anything is the same as itself. It provides for the soundness of the new rule of inference ID introduced below in section 4. To see that (1) holds, note that in any model $\mathcal{M}$, for any name α, $\mathcal{M}(\alpha) = \mathcal{M}(\alpha)$.

Parts (2) and (3) are principles of substitution for identity, embodied deductively in the rule SUB in section 4. They affirm that if names α and β name the same thing in a model (i.e. whenever $\alpha = \beta$ or $\beta = \alpha$ is true in the model), then either name can be substituted for the other in any sentence. The proofs of (2) and (3) use compositionality (proposition 9.1). For (2), suppose that both $\alpha = \beta$ and $A(\gamma/\alpha)$ are true in a model $\mathcal{M}$. From the second assumption it follows by proposition 9.1 that A is true in the model $\mathcal{M}^{\gamma}_{\mathcal{M}(\alpha)}$. The first assumption of course means that $\mathcal{M}(\alpha)$ and $\mathcal{M}(\beta)$ are the same individual in $\mathcal{M}$'s domain. Therefore A is true in $\mathcal{M}^{\gamma}_{\mathcal{M}(\beta)}$—which by compositionality means that $A(\gamma/\beta)$ is true in $\mathcal{M}$, as we wished to show. The reasoning for part (3) is similar and is left as an exercise. □

Let us illustrate the principles of substitution in proposition 10.1. Here is a simple instance of (2):

$$a = b,\ Fa \models Fb.$$

According to this, if a and b name the same thing and Fa is true, then Fb, the result of substituting b for a in Fa, is also true. More explicitly, the displayed implication reads:

$$a = b,\ (Fc)(c/a) \models (Fc)(c/b).$$

Of course, (Fc)(c/a) is just Fa, and (Fc)(c/b) is just Fb.

Next consider this implication:

$$a = b,\ Raa \models Rab.$$

Here Rab follows from a = b and Raa by substitution of b for just one occurrence of a in Raa. In this case, the sentence Raa is (Rac)(c/a), and the implication is the following instance of (2).

$$a = b,\ (Rac)(c/a) \models (Rac)(c/b).$$

With the addition of identity as a logical idiom, it is now possible to translate accurately sentences of English affirming the existence of certain numbers of things. We turn to this in the next section.

EXERCISES

10.1. Here is a more complicated example of substitution:

$$a = b,\ \forall x(Rax \to \neg\exists y(Rya \wedge Fa)) \models \forall x(Rbx \to \neg\exists y(Rya \wedge Fb)).$$

Use the replacement notation, (/), to show that this is an instance of part (2) of proposition 10.1 above.

10.2. Confirm that part (2) of proposition 10.1 is equivalent to saying that, for any sentence A and any names α and β, a sentence B is a consequence of $\alpha = \beta$ and A whenever B results from replacing occurrences of the name α in A by the name β.

10.3. Here is a principle of substitution (of the name β for the name α):

$$\alpha = \beta,\ A \models A(\alpha/\beta).$$

Explain why, though correct, this principle is inadequate. *Suggestion*: Find an example of substitution covered by part (2) of proposition 10.1, but not covered by this principle. (Recall section 9 of chapter 6.)

10.4. Prove part (3) of proposition 10.1.

10.5. Explain why $\beta = \alpha \models \alpha = \beta$. Then use this and part (2) of proposition 10.1 to prove part (3).

3. Translation

Let us begin with an example of something that could not be adequately expressed in sentential logic or predicate logic without identity:

$$\exists x \exists y\ x \neq y$$

This sentence clearly affirms that *at least two* (i.e. distinct) things exist. In a similar fashion, the sentence

$$\exists x \exists y \exists z(x \neq y \wedge y \neq z \wedge x \neq z)$$

says that there are at least three things.

To say, on the other hand, that there are *at most two* things, we write:

$$\forall x \forall y \forall z(x = y \vee y = z \vee x = z)$$

That is, of any three (not necessarily distinct) things, at Least two are the same. Note that this is equivalent to the denial that there are at least three things:

$$\neg \exists x \exists y \exists z(x \neq y \wedge y \neq z \wedge x \neq z)$$

The means are obviously at hand, thus, to translate sentences such as

There are exactly two things.

This sentence can be translated simply by conjoining the translations of There are at least two things and There are at most two things:

$$\exists x \exists y\ x \neq y \wedge \forall x \forall y \forall z(x = y \vee y = z \vee x = z)$$

The proposition that there are exactly two things can be rendered more briefly in this way:

$$\exists x \exists y(x \neq y \wedge \forall z(z = x \vee z = y))$$

—i.e. there exist at least two things such that everything is identical with one or the other.

It should be noted that to say that there is *at most one* thing—

$$\forall x \forall y\ x = y$$

(i.e. any two things are identical)—is equivalent to saying that there exists *exactly one* thing:

$$\exists x \forall y\ x = y$$

This is so simply because the domain of a model is always nonempty; it is always the case that something (i.e. at least one thing) exists.

Of course it is also possible now, with identity as a logical constant, to express propositions affirming the existence of a certain number of things having this-or-that property. For example, the sentence

$$\exists x \exists y(Fx \wedge Fy \wedge x \neq y)$$

says—where F translates is a frog—that there are at least two frogs. And the sentence

$$\forall x \forall y \forall z((Fx \wedge Fy \wedge Fz) \rightarrow (x = y \vee y = z \vee x = z))$$

says that there are at most two frogs. So to say that there are exactly two frogs one can write the conjunction of these two sentences; or one can use a shorter form:

$$\exists x \exists y(Fx \wedge Fy \wedge x \neq y \wedge \forall z(Fz \rightarrow (x = z \vee y = z))$$

—i.e. that there are at least two frogs and anything that is a frog is one of the two.

To say that there is exactly one frog, there are a number of natural choices. One can write the conjunction

$$\exists xFx \wedge \forall x \forall y((Fx \wedge Fy) \rightarrow x = y),$$

which says that there is at least one frog and there is at most one frog. Or one can use a shorter form,

$$\exists x(Fx \wedge \forall y(Fy \rightarrow x = y)),$$

which says (equivalently) that there is at least one frog and any frog is it. Or one can use the still shorter form:

$$\exists x \forall y(Fy \leftrightarrow x = y)$$

The equivalence of this last form to the previous two may not be so obvious. It will be easier to understand if we look at an analogous translation of a sentence stating that some named individual is the one and only frog. Using a for Al Jennings, we may translate Al is the one and only frog equivalently by any of these three sentences:

$$Fa \wedge \forall x \forall y((Fx \wedge Fy) \rightarrow x = y)$$

$$Fa \wedge \forall y(Fy \rightarrow a = y)$$

$$\forall y(Fy \leftrightarrow a = y)$$

The first of this trio says that Al is a frog and any two frogs are identical, which clearly leaves Al as the unique frog. The second again says that Al is a frog and adds that every frog is identical with Al. So this version also implies that Al is the unique frog. The third sentence simply says something is a frog just in case it is Al—i.e. that every frog is identical with Al and everything identical with Al is a frog. Now observe that the three sentences set out earlier as translations of There is exactly one frog are simply respective existential generalizations of the three displayed just above.

To say that Al is a frog we may of course write Fa. As may be gathered from the discussion above, to say that Al is a frog we may also write

$$\forall x(x = a \rightarrow Fx)$$

—i.e. anything identical with Al (such as Al!) is a frog. To say, on the other hand, that *only* Al is a frog, we write

$$\forall x(Fx \rightarrow x = a)$$

—i.e. if anything is a frog, it is Al (note that this does not imply that Al is a frog). To say, once again, that Al is the (one and only) frog, we may write the conjunction of the two displayed sentences,

$$\forall x(x = a \rightarrow Fx) \wedge \forall x(Fx \rightarrow x = a),$$

We have already noticed three ways of saying that Al is the one and only frog:

$$Fa \wedge \forall x \forall y((Fx \wedge Fy) \rightarrow x = y)$$

$$Fa \wedge \forall y(Fy \rightarrow a = y)$$

$$\forall y(Fy \leftrightarrow a = y)$$

So now we have four ways.

Later on, in section 4 and its exercises, the equivalence of the various translations will be proved.

Finally, let K mean is a present King of France and B mean is bald. Then the sentence

$$\exists x(Kx \wedge \forall y(Ky \rightarrow x = y) \wedge Bx)$$

means that there is exactly one present King of France, and he is bald. Note that this is often equivalent to

The present King of France is bald.

In 1905 the philosopher Bertrand Russell proposed that sentences of this form—*The so-and-so is* [*a*] *such-and-such*—be analyzed in this way—*There is exactly one so-and-so, and it is* [*a*] *such-and-such*. Russell called expressions of the form *the so-and-so* "definite descriptions" (and they have been so called ever since).

It is important to notice that on this analysis to translate the English sentence

The present King of France is not bald

one may put either

$$\exists x(Kx \wedge \forall y(Ky \rightarrow x = y) \wedge \neg Bx)$$

or

$$\neg \exists x(Kx \wedge \forall y(Ky \rightarrow x = y) \wedge Bx).$$

The first of these says that there is just one present King of France, and he is not bald, while the second says that it is not the case that there is just one present King of France, who is bald. The point is that the first sentence again affirms the unique existence of a present King of France, and denies that he is bald, whereas the second simply denies that there is a unique present King of France, who is bald. Thus the English negation, The present King of France is not bald, is ambiguous, and its ambiguity is revealed by the alternative translations.

In general, the two translations are not equivalent. Indeed, given the current facts of the matter, the first is false and the second is true. We leave it as an exercise to devise translations of sentences similar to these in which more than one entity is affirmed to exist uniquely and have some property, for example, The two team captains loved each other.

The logic of the language of identity is usefully studied in terms of a sound and complete set of rules of inference in a deductive system. So we proceed next to the theory of deduction.

EXERCISES

10.6. Where, as above, K means is a present King of France, B means is bald, and the domain is the set of humans, what do the following sentences mean?

$$\exists x \forall y((Ky \leftrightarrow x = y) \wedge Bx)$$

$$\exists x \forall y(Ky \leftrightarrow x = y) \wedge \forall x(Kx \rightarrow Bx)$$

$$\exists x \forall y(Ky \leftrightarrow x = y) \wedge \exists x(Kx \wedge Bx)$$

$$\exists xKx \wedge \forall x \forall y((Kx \wedge Ky) \rightarrow x = y) \wedge \forall x(Kx \rightarrow Bx)$$

$$\exists xKx \wedge \forall x \forall y((Kx \wedge Ky) \rightarrow x = y) \wedge \exists x(Kx \wedge Bx)$$

10.7. Find some other ways of saying (in the language of predicate logic with identity) that the present King of France is bald.

10.8. Where R means is an author of, a means Asa Mercer, b means *The banditti of the plains* and the domain is the set of humans, the sentence

$$\exists x(Rxb \wedge \forall y(Ryb \rightarrow x = y) \wedge x = a)$$

means that Mercer is the author of *The banditti of the plains*. Find some other ways of saying this.

10.9. Where R means is a father of, a means N.H. (Old Man) Clanton, b means Ike, and the domain is the set of humans, translate Old Man Clanton is the father of Ike.

10.10. Where the domain is the set of humans, let R mean is an author of, a mean Mark Twain, b mean Samuel Langhorne Clemens, and c mean *Tom Sawyer*. Translate the following argument.

Mark Twain is the author of *Tom Sawyer*
Mark Twain is Sam Clemens

Sam Clemens is *an* author of *Tom Sawyer*

10.11. Let F, M, B, and S mean is a father of, is a mother of, is a brother of, and is a sister of, let a and b mean Zerelda Samuel and Jesse James, and let the domain be the set of humans. Translate:

(1) Zerelda is a mother of Jesse

(2) Only Zerelda is a mother of Jesse

(3) Zerelda is the mother of Jesse

(4) Zerelda is a grandmother of Jesse

(5) Zerelda is the paternal grandmother of Jesse

(6) Zerelda is a sister of Jesse

(7) Only Zerelda is a sister of Jesse

(8) Jesse is a maternal uncle of Zerelda

(9) Jesse is a nephew of Zerelda

(10) Zerelda is a paternal great aunt of Jesse

(11) Zerelda has exactly two sisters

(12) Zerelda has exactly two siblings

(13) Jesse is a grandfather of an aunt of one of Jesse's sisters

10.12. Relative to a model in which R means shoots, the name a means Annie Oakley, and the domain is the set of humans, match each sentence (a)–(h) with its translation in (1)–(8).

(a) Annie shoots all those who shoot her

(b) Annie shoots all those who shoot themselves

(c) Annie shoots only those who shoot her

(d) Annie shoots only those who shoot themselves

(e) Only Annie shoots all those who shoot her

(f) Only Annie shoots all those who shoot themselves

(g) Only Annie shoots only those who shoot her

(h) Only Annie shoots only those who shoot themselves

(1) $\forall x(\forall y(Ryx \rightarrow Rxy) \rightarrow x = a)$

(2) $\forall x(x = a \rightarrow \forall y(Rxy \rightarrow Ryx))$

(3) $\exists x(x = a \wedge \forall y(Rxy \rightarrow Ryy))$

(4) $\forall x(\forall y(Ryy \rightarrow Rxy) \rightarrow x = a)$

(5) $\forall x(\forall y(Rxy \rightarrow Ryy) \rightarrow x = a)$

(6) $\exists x(x = a \wedge \forall y(Ryx \rightarrow Rxy))$

(7) $\forall x(x = a \rightarrow \forall y(Ryy \rightarrow Rxy))$

(8) $\forall x(\forall y(Rxy \rightarrow Ryx) \rightarrow x = a)$

10.13. We noted that Al is a frog may be translated by Fa (as usual), and also by $\forall x(x = a \rightarrow Fx)$. Here is a third translation:

$$\exists x(x = a \wedge Fx)$$

This says that there is at least one thing identical with Al that is a frog—i.e. Al is a frog. Argue informally for the equivalence of the three translations of Al is a frog. (Formal proofs are requested in parts (6) and (7) of exercise 10.20.)

10.14. Show the difference between saying that Al is a frog and that only Al is a frog by describing models in which:

(1) Fa is true and $\forall x(Fx \rightarrow x = a)$ is false.

(2) $\forall x(Fx \rightarrow x = a)$ is true and Fa is false.

Observe, moreover, that in each model the truth values of $\forall x(x = a \rightarrow Fx)$ and $\exists x(x = a \wedge Fx)$ are the same as the truth value of Fa.

10.15. Here are some further properties of binary relations:

	Sentence	*Property*
(m)	$\forall x\forall y(Rxy \rightarrow x = y)$	vacuity
(n)	$\forall x\forall y((Rxy \wedge Ryx) \rightarrow x = y)$	antisymmetry
(o)	$\forall x\forall y(Rxy \vee Ryx \vee x = y)$	weak connectedness; trichotomy

(p)	$\forall x \forall y(Rxy \rightarrow$ $\exists z(x \neq z \wedge y \neq z \wedge Rxz \wedge Rzy))$	strong density
(q)	$\forall x \forall y \forall z((Rxy \wedge Rxz) \rightarrow y = z)$	partial functionality
(r)	$\forall x \exists y \forall z(Rxz \leftrightarrow y = z)$	functionality

Using in addition definitions (a)–(l) in exercise 7.4 (page 263):

(1) Give an example of a vacuous relation. That is, show that (m) is satisfiable.

(2) Give an example of an antisymmetric relation. That is, show that (n) is satisfiable.

(3) Give an example of a weakly connected relation. That is, show that (o) is satisfiable.

(4) Give an example of a strongly dense relation. That is, show that (p) is satisfiable.

(5) Give an example of a partially functional relation. That is, show that (q) is satisfiable.

(6) Give an example of a functional relation. That is, show that (r) is satisfiable.

(7) Give an example of an antisymmetric relation that is not also vacuous. That is, show that (n) does not imply (m).

(8) Give an example of a partially functional relation that is not also vacuous. That is, show that (q) does not imply (m).

(9) Give an example of a weakly connected relation that is not also strongly connected. That is, show that (o) does not imply (h).

(10) Give an example of a weakly dense relation that is not also strongly dense. That is, show that (l) does not imply (p).

(11) Give an example of a serial relation that is not also functional. That is, show that (e) does not imply (r).

(12) Give an example of a partially functional relation that is not also functional. That is, show that (q) does not imply (r).

10.16. True or false:

a. The sentences $\exists x \forall y\ x = y$ and $\forall x \forall y\ x = y$ are equivalent.

b. The sentence $\exists x \forall y(Fy \leftrightarrow x = y)$ is a consequence of the sentence $\exists x \forall y\ x = y$.

c. The sentence $\exists x \forall y(Fy \leftrightarrow x = y)$ implies the sentence $\forall x(Fx \rightarrow Gx) \vee \forall x(Fx \rightarrow \neg Gx)$.

d. In predicate logic with identity there are satisfiable sentences that are not true in any infinite model.

e. The sentences $\forall x(Fx \rightarrow x = a)$ and $\exists x(Fx \wedge x = a)$ are equivalent.

10.17. Consider the following two sentences.

$$\exists x \exists y(Rxy \wedge \forall z(Rzx \vee z = x \vee z = y \vee Ryz))$$

$$\forall x \forall y(x = y \vee \exists z((Rxz \wedge Rzy) \vee (Ryz \wedge Rzx)))$$

Describe a model of the first in which the second is false; then describe a model of the second in which the first is false. (*Suggestion*: Let the value of R be < in each model.)

10.18. For each positive integer n the language contains a sentence—call it A(n)—that is true in every model that has at least n elements in its domain and is false in all models of fewer than n elements. For example, let A(2) be $\exists x \exists y\ x \neq y$, let A(3) be $\exists x \exists y \exists z(x \neq y \wedge y \neq z \wedge x \neq z)$, and so on.

Notice that A(1) is valid, since every model has at least one element in its domain. Notice too that whenever $n \geq m$, the sentence A(n) implies the sentence A(m).

Let Ω be the set {A(1), A(2), A(3), . . .}, and let Σ be any set of sentences that includes Ω. Prove that any model of Σ is infinite, i.e. that whenever every sentence in Σ is true in a model $\mathcal{M}$, $\mathcal{M}$ is infinite. (Note that this does not imply that Σ has any models.)

4. Deduction

For a sound and complete deductive system for predicate logic with identity, we first take over the definitions of deduction, deduction of, deducibility, etc. Then it is sufficient to add two rules of inference—ID (identity) and SUB (substitution)—to our stock of sentential and quantificational rules.

DEFINITION 10.2. *Rules of inference for identity.*

ID. $$\frac{}{\varnothing,\ \alpha = \alpha}$$

SUB. $$\frac{\Gamma,\ \alpha = \beta;\quad \Delta,\ A(\gamma/\alpha)}{\Gamma \cup \Delta,\ A(\gamma/\beta)} \qquad\qquad \frac{\Gamma,\ \beta = \alpha;\quad \Delta,\ A(\gamma/\alpha)}{\Gamma \cup \Delta,\ A(\gamma/\beta)}$$

The first rule, ID, permits the introduction anywhere in a deduction of a *self-identity* sentence (with the empty set of premisses). The second rule, SUB, allows for substitution of identicals. In both versions of SUB the name β is substituted for the name α; the difference is that in the first version the substituted name is on the right-hand side of the identity sign, while in the second version it is on the left. In fact, we could get along with either one of the two parts of SUB, since, given ID, each version is derivable from the other (this is proved in section 6). We have included both versions for the sake of convenience in constructing deductions.

Here is a simple deduction using SUB (first version):

{1}	(1)	Fa	P
{2}	(2)	a = b	P
{1, 2}	(3)	Fb	1, 2, SUB

This deduction shows that Fb follows from the premisses Fa and a = b: the identity sentence on line 2 permits the substitution of b for a in Fa on line 1, resulting in Fb on line 3. Thus if F, a, and b translate, respectively, wrote *Rough-*

ing it, Mark Twain, and Sam Clemens, this deduction validates the argument mentioned at the beginning of the chapter. Let us display the same deduction, all over again, with the sentences on lines (1) and (3) described using the replacement notation (/):

{1}	(1)	(Fc)(c/a)	P
{2}	(2)	a = b	P
{1, 2}	(3)	(Fc)(c/b)	1, 2, SUB

Next consider a slightly more complex deduction using SUB, showing that if Mark Twain loved Sam Clemens, then Sam Clemens loved himself:

{1}	(1)	Rab	P
{2}	(2)	a = b	P
{1, 2}	(3)	Rbb	1, 2, SUB

In this deduction, b is again substituted for a in a sentence (Rab) that already contains an occurrence of b. Because it is important to realize that this is indeed licensed by the rule SUB, we set out the deduction once more using replacement notation:

{1}	(1)	(Rcb)(c/a)	P
{2}	(2)	a = b	P
{1, 2}	(3)	(Rcb)(c/b)	1, 2, SUB

The point of course is that the sentence Rab is (Rcb)(c/a), and likewise Rbb is (Rcb)(c/b).

The following deduction uses both identity rules, ID and SUB. In effect, it shows that identity is a symmetric relation.

{1}	(1)	a = b	P
∅	(2)	a = a	ID
{1}	(3)	b = a	1, 2, SUB

This deduction uses SUB (first version) to substitute b for the left-hand occurrence of a in a = a to yield b = a—thus demonstrating that the order of the names in an identity sentence can always be reversed. This deduction provides the key to a proof that each version of SUB is derivable given the other and the rule ID.

The strategies described in connection with sentential and quantificational inference rules continue to apply.

EXERCISES

10.19. The sentences below express some properties of the identity relation. Describe deductions showing that each of the sentences is valid.

(1)	$\forall x \exists y\ x = y$	(seriality)
(2)	$\forall x\ x = x$	(reflexivity)
(3)	$\forall x \forall y (x = y \rightarrow y = x)$	(symmetry)
(4)	$\forall x \forall y \forall z ((x = y \wedge y = z) \rightarrow x = z)$	(transitivity)
(5)	$\forall x \forall y \forall z ((x = y \wedge x = z) \rightarrow y = z)$	(euclideanness)

10.20. Below are a number of principles involving identity. Describe deductions showing that the arguments and sentences are valid and that the paired sentences are equivalent.

(1) $\forall x \forall y (x = y \rightarrow (Fx \leftrightarrow Fy))$

(2) $\forall x \forall y (x = y \leftrightarrow \forall z (x = z \leftrightarrow y = z))$

(3) $\forall x \forall y\ x = y,\ \forall y\ a = y$

(4) $\exists x \forall y\ x = y\ /\ a = b$

(5) $\exists x \forall y\ x = y,\ \forall x \forall y\ x = y$

(6) $Fa,\ \forall x (x = a \rightarrow Fx)$

(7) $Fa,\ \exists x (x = a \wedge Fx)$

(8) $\forall x \forall y\ x = y\ /\ Fa \leftrightarrow Fb$

(9) $\forall x \forall y\ x = y\ /\ \exists x Fx \leftrightarrow \forall x Fx$

(10) $\exists x \forall y\ x = y\ /\ Fa \leftrightarrow Fb$

(11) $\exists x \forall y\ x = y$ / $\exists x Fx \leftrightarrow \forall x Fx$

(12) $\exists x \forall y\ x = y$ / $\forall x Fx \vee \forall x \neg Fx$

(13) $\exists x \forall y (Fy \leftrightarrow x = y)$ / $\exists x (Fx \wedge Gx) \leftrightarrow \forall x (Fx \to Gx)$

(14) $\exists x \forall y (Fy \leftrightarrow x = y)$ / $\forall x (Fx \to Gx) \vee \forall x (Fx \to \neg Gx)$

(15) $\exists x \forall y (Fy \leftrightarrow x = y)$, $\exists x (Fx \wedge \forall y (Fy \to x = y))$

(16) $\exists x \forall y (Fy \leftrightarrow x = y)$, $\exists x Fx \wedge \forall x \forall y ((Fx \wedge Fy) \to x = y)$

(17) $\forall x Fx$, $Fa \wedge \forall x (x \neq a \to Fx)$

(18) $\exists x Fx$, $Fa \vee \exists x (x \neq a \wedge Fx)$

(19) Fa, $\neg Fb$ / $\exists x \exists y\ x \neq y$

(20) $\exists x Fx$, $\exists x \neg Fx$ / $\exists x \exists y\ x \neq y$

(21) Fa, $\neg Fb$ / $\forall x \exists y\ x \neq y$

(22) $\exists x Fx$, $\exists x \neg Fx$ / $\forall x \exists y\ x \neq y$

(23) $\exists x \exists y\ x \neq y$, $\exists y\ a \neq y$

(24) $\forall x \exists y\ x \neq y$, $\exists y\ a \neq y$

(25) $\exists x \exists y\ x \neq y$, $\forall x \exists y\ x \neq y$

(26) $\exists x \exists y (x \neq y \wedge \forall z (x = z \vee y = z))$,
$\exists x \exists y\ x \neq y \wedge \forall x \forall y \forall z (x = y \vee y = z \vee x = z)$

(27) $\forall x (x \neq a \to Fx)$ / $\forall x \forall y (x \neq y \to (Fx \vee Fy))$

(28) $\forall x \forall y \forall z (x = y \vee y = z \vee x = z)$ /
$\forall x Fx \vee \forall x (Fx \to Gx) \vee \forall x (Fx \to \neg Gx)$

10.21. Here are some more principles involving identity. Describe deductions showing that the sentences are valid and that the paired sentences are equivalent.

(1) $\exists x (x \neq a \wedge Fx)$, $\exists x Fx \wedge (Fa \to \exists x \exists y (x \neq y \wedge Fx \wedge Fy))$

(2) $\forall x \forall y (x \neq y \to y \neq x)$

(3) $\forall x \forall y \forall z((x = y \wedge x \neq z) \rightarrow y \neq z)$

(4) $\forall x \forall y \forall z((x = y \wedge y \neq z) \rightarrow x \neq z)$

(5) $\exists x \forall y((Fy \leftrightarrow x = y) \wedge Gx)$, $\exists x(Fx \wedge \forall y(Fy \rightarrow x = y) \wedge Gx)$

10.22. Deduce $\exists x \forall y\ x = y$ from $\forall x(a = x \leftrightarrow P)$.

10.23. Recall from exercise 10.15 the following properties of binary relations.

	Sentence	*Property*
(m)	$\forall x \forall y(Rxy \rightarrow x = y)$	vacuity
(n)	$\forall x \forall y((Rxy \wedge Ryx) \rightarrow x = y)$	antisymmetry
(o)	$\forall x \forall y(Rxy \vee Ryx \vee x = y)$	weak connectedness; trichotomy
(p)	$\forall x \forall y(Rxy \rightarrow$ $\exists z(x \neq z \wedge y \neq z \wedge Rxz \wedge Rzy))$	strong density
(q)	$\forall x \forall y \forall z((Rxy \wedge Rxz) \rightarrow y = z)$	partial functionality
(r)	$\forall x \exists y \forall z(Rxz \leftrightarrow y = z)$	functionality

Using in addition (a) through (l) in exercise 7.4 (page 263), describe deductions for each of the following problems.

(1) Show that the empty relation is vacuous, antisymmetric, strongly dense, and partially functional. That is, show that (d) implies (m), (n), (p), and (q).

(2) Show that every vacuous relation is symmetric, transitive, euclidean, weakly dense, antisymmetric, and partially functional. That is, show that (m) implies (f), (i), (k), (l), (n), and (q).

(3) Show that every serial vacuous relation is reflexive. That is, show that (e), (m) together imply (a).

(4) Show that every asymmetric relation is antisymmetric. That is, show that (g) implies (n).

(5) Show that every strongly connected relation is weakly connected. That is, show that (h) implies (o).

(6) Show that every strongly dense relation is weakly dense. That is, show that (p) implies (l).

(7) Show that every reflexive partially functional relation is vacuous. That is, show that (a), (q) together imply (m).

(8) Show that a relation is functional if and only if it is serial and partially functional. That is, show that (r) is equivalent to the conjunction (e) $\wedge$ (q).

(9) Show that partial functionality is equivalently expressed:

$$\forall x \exists y (\exists z (Rzx \wedge Rzy) \to x = y)$$

That is, show that this sentence is equivalent to (q).

(10) Show that functionality is equivalently expressed:

$$\forall x \exists y (Rxy \wedge \forall z (Rxz \to y = z))$$

That is, show that this sentence is equivalent to (r).

(11) Show that reflexivity is equivalently expressed:

$$\forall x \forall y (x = y \to Rxy)$$

That is, show that this sentence is equivalent to (a).

(12) Show that a reflexive relation in a one-element domain is universal. That is, show that (a), $\exists x \forall y\ x = y$ together imply (c). (Perhaps use the sentence in (11) instead of (a).)

10.24. Show that one version of the rule SUB (the second, say) is derivable from the other, thus demonstrating that theoretically we could make do with just one version (and ID). *Hint*: Recall the proof above that b = a is deducible from a = b, and compare exercise 10.5.

10.25. True or false:

a. The deductive system for predicate logic with identity would be incomplete if both the rule ID the rule SUB were dropped.

b. The sentence $\forall x Fx \vee \forall x \neg Fx$ is deducible from the sentence $\forall x \forall y\ x = y$.

c. The sentence $\exists x(Fx \wedge Gx) \leftrightarrow \forall x(Fx \rightarrow Gx)$ is deducible from the sentence $\exists x \forall y(Fy \leftrightarrow x = y)$.

d. No deduction begins with a line justified by the rule ID.

e. The sentence $\forall x \forall y\ x = y$ is deducible from the sentence $\forall x Fx \vee \forall x \neg Fx$.

f. The sentence $\exists x \forall y(Fy \leftrightarrow x = y)$ is not deducible from the sentence $\exists x \forall y\ x = y$.

g. The sentence $\forall x(x = a \rightarrow \forall y(Ryx \rightarrow Rxy))$ is deductively equivalent to the sentence $\exists x(\forall y(Ryx \rightarrow Rxy) \wedge x = a)$.

h. The sentence $\exists x Fx \vee \exists x \neg Fx$ is deducible from the sentence $\forall x \exists y\ x = y$.

i. The sentences $\exists x \exists y\ x \neq y$ and $\forall x \exists y\ x \neq y$ are deductively equivalent.

j. The deductive system for predicate logic with identity would be unsound if either the rule ID or the rule SUB were dropped.

5. Soundness

Once again, as in chapters 5 and 9, we wish to show that the deductive system at hand is sound, i.e. that

$$\text{if } \Gamma \vdash A \text{ then } \Gamma \models A.$$

Our deductive system is characterized, now, by the basic rules of inference presented in chapters 4 and 8 together with the new rules ID and SUB for identity. Since the soundness of the deductive system itself comes down to the soundness of its rules—and since we have already been convinced that the rules in chapters 4 and 8 preserve consequence—it remains only to be shown that the rules ID and SUB are sound. This of course is the import of proposition 10.1. Therefore:

PROPOSITION 10.3. *Soundness.* The deductive system for predicate logic with identity is sound. That is, if $\Gamma \vdash A$ then $\Gamma \models A$.

PROPOSITION 10.4. *Soundness and corollaries.*

(1) If $\Gamma \vdash A$ then $\Gamma \models A$.

(2) If $\vdash A$ then $\models A$.

(3) If $A \sim B$ then $A \simeq B$.

(4) If Sat Γ then Con Γ.

Proof. Exercise. □

EXERCISES

10.26. Prove proposition 10.4. (Compare proposition 4.8, page 171.)

6. Interderivability

Are the rules ID and SUB minimal? The answer is no: we can show that the two versions of SUB are interderivable given ID, and hence we need not include both in our basic set of rules for predicate logic with identity.

PROPOSITION 10.5. Given the rule ID and either version of SUB, the other version of SUB is derivable.

Proof. Let us show that the second version is derivable given the first and ID. Suppose a deduction contains the lines:

Γ	( )	$\beta = \alpha$
Δ	( )	$A(\gamma/\alpha)$

Then the following lines may be added to this deduction using the rules ID and the first version of SUB.

$\varnothing$	( )	$\beta = \beta$	ID
Γ	( )	$\alpha = \beta$	SUB (first version)
$\Gamma \cup \Delta$	( )	$A(\gamma/\beta)$	SUB (first version)

With regard to the second line in the continuation, $\langle \varnothing, \beta = \beta \rangle$, note that $\beta = \beta$ is the same sentence as $(\gamma = \beta)(\gamma/\beta)$. Thus the full deduction shows that $A(\gamma/\beta)$ is deducible from $\Gamma \cup \Delta$ given both that $\beta = \alpha$ is deducible from Γ and that $A(\gamma/\alpha)$ is deducible from Δ. This means that the second version of SUB is dispensable. □

EXERCISES

10.27. Prove that, given the rule ID and the second version of SUB, the first version of SUB is derivable.

7. Completeness

As in the case of the completeness of predicate logic without identity, the proof of the completeness of predicate logic with identity turns mainly on the facts that every consistent set of sentences has a maximal EI-closed extension, and that every maximal EI-closed set is satisfiable.

Establishing the first of these propositions presents no special difficulty. We remarked at the conclusion of our proof of proposition 9.6 (page 334) that it holds good for any deductive system that includes all the rules of inference of the system set forth in chapter 8 for predicate logic without identity. Because our system of predicate logic *with* identity has in addition to the rules of that system only the rules ID and SUB, it follows that here, too, every consistent set of sentences is a subset of a maximal EI-closed set. So although we appeal to this in our proof of the completeness of predicate logic with identity, there is no need to state it again.

The proof of the second result—that every maximal EI-closed set of sentences is satisfiable—is different in the present case. Having identity as a logical constant accounts for several novelties in the argument. Therefore, we do state this separately:

PROPOSITION 10.6. If Max Γ and EI-clsd Γ, then Sat Γ.

Proof. In structure, the proof is like that for proposition 9.8 (page 339): In terms of a given maximal EI-closed set of sentences we define a canonical model and show that this model verifies precisely the sentences in the set. But we cannot use again the canonical model $\mathcal{M}$ defined in the proof of proposition 9.8, in which the domain is the set of names in the language and each name is assigned to itself as value in $\mathcal{M}$. For if we did, we should be able to argue like this:

$$\begin{aligned} \models_{\mathcal{M}} \alpha = \beta \text{ iff } & \mathcal{M}(\alpha) = \mathcal{M}(\beta) \\ & \quad \text{—definition 7.4, part (1');} \\ \text{iff } & \alpha = \beta \\ & \quad \text{—definition of } \mathcal{M}. \end{aligned}$$

In other words, an identity sentence is true in *that* canonical model if and only if the names in the sentence are the same. But there may very well be sentences of the form $\alpha = \beta$ in the maximal EI-closed set with which we begin, where α and β are different names. After all, $\{a_0 = a_1\}$ is a consistent set and so has a maximal EI-closed extension. Thus using the old canonical model we will not be able to obtain the critical, desired result that

$$\models_{\mathcal{M}} \alpha = \beta \text{ iff } \alpha = \beta \in \Gamma,$$

where Γ is the initially given maximal EI-closed set of sentences. Clearly, in defining $\mathcal{M}$ with respect to Γ we must acknowledge the meaning of an identity sentence in Γ—to wit, that the names in it name the same thing.

For these reasons, we are led to the following argument for proposition 10.6.

Suppose Γ to be a maximal EI-closed set of sentences. We define the relation $\equiv$ in the set of names in the language, in terms of Γ, as follows.

Definition 1. $\alpha \equiv \beta$ iff $\alpha = \beta \in \Gamma$.

(Technically, we should write "$\alpha \equiv_\Gamma \beta$", to indicate the dependence of the relation on the set Γ. But we forbear in the interests of simplicity and perspicuity. The same goes for the notation "$[\alpha]$" defined below.)

Lemma 2. $\equiv$ is an equivalence relation in the set of names.

To prove that $\equiv$ is an equivalence relation, let us show that it is (a) reflexive, (b) symmetric, and (c) transitive. (See the appendix on set theory for more about equivalence relations.)

For (a). By the rule ID, we know that $\alpha = \alpha$ is a theorem, for every name α. Because Γ is maximal, $\alpha = \alpha$ is in Γ (proposition 5.2, page 203). Therefore, by the definition, $\alpha \equiv \alpha$, for every name α. So the relation is reflexive.

For (b). Observe the deduction:

$\{1\}$	(1)	$\alpha = \beta$
$\varnothing$	(2)	$\alpha = \alpha$
$\{1\}$	(3)	$\beta = \alpha$

(Compare the similar deduction in section 4 above.) Thus $\beta = \alpha$ is deducible from $\alpha = \beta$. Now suppose that $\alpha \equiv \beta$, i.e. that $\alpha = \beta$ is a member of Γ. By proposition 5.2, so is $\beta = \alpha$, which means that $\beta \equiv \alpha$. Hence for any names α and β, if $\alpha \equiv \beta$, then $\beta \equiv \alpha$. So the relation is symmetric.

For (c). Observe the deduction:

$\{1\}$	(1)	$\alpha = \beta$
$\{2\}$	(2)	$\beta = \gamma$
$\{1, 2\}$	(3)	$\alpha = \gamma$

(Justifications left as an exercise.) Given this deduction, one can argue that for any names α, β, and γ, if $\alpha \equiv \beta$ and $\beta \equiv \gamma$, then $\alpha \equiv \gamma$. So the relation is transitive. The argument is left as an exercise.

The proof of lemma 2 is thus complete.

Being an equivalence, the relation $\equiv$ partitions the set of names into a collection of disjoint, non-empty equivalence classes of names. In each such class are all and only those names that appear together in identity sentences contained in the set Γ. Thus the relation $\equiv$ so to speak "identifies" those names that form identities in Γ. We record the definition of these equivalence classes:

Definition 3. $[\alpha] = \{\beta: \alpha \equiv \beta\}$, for each name α.

That is to say, the equivalence class $[\alpha]$ is the set of names β for which the sentence $\alpha = \beta$ is a member of Γ. (Note that because the theorem $\alpha = \alpha$ is always in Γ, the set $[\alpha]$ always contains α.) It follows from the definitions that equivalence classes $[\alpha]$ and $[\beta]$ are identical just in case the sentence $\alpha = \beta$ is in Γ. We record this for subsequent reference:

Lemma 4. $[\alpha] = [\beta]$ iff $\alpha = \beta \in \Gamma$, for any names α and β.

We arrive at last at the definition of the canonical model $\mathcal{M}$ with respect to the maximal EI-closed set of sentences Γ. The foregoing constructions make it evident that the equivalence class $[\alpha]$ is a good candidate for the value of the name α under $\mathcal{M}$. So $\mathcal{M}$ is defined as follows.

Definition 5.

(i) D (the domain of $\mathcal{M}$) = $\{[\alpha]: \alpha \text{ is a name}\}$.
(ii) $\mathcal{M}(\alpha) = [\alpha]$, for each name α.
(iii) $\langle [\alpha_1], \ldots, [\alpha_n] \rangle \in \mathcal{M}(\mathbb{P})$ iff $\mathbb{P}\alpha_1 \ldots \alpha_n \in \Gamma$, for each n-place predicate $\mathbb{P}$ ($n \geq 0$).

That is to say: (i) The domain of the canonical model $\mathcal{M}$ is simply the set of equivalence classes of names. That is, each element of the domain is such a class of names. (ii) To each name in the language the model assigns as value the equivalence class that contains that name. Note that this is okay, since each name appears in just one equivalence class (though such a class may contain more than one name and so be named in the model by more than one name). (iii) Lastly, to each n-place predicate the model assigns as value a set of n-tuples of elements of the domain—a set of n-tuples of equivalence classes of names. An n-tuple of equivalence classes $\langle [\alpha_1], \ldots, [\alpha_n] \rangle$ belongs to the value of an n-place predicate $\mathbb{P}$ just in case the atomic sentence formed by following the predicate by the names—i.e. the sentence $\mathbb{P}\alpha_1 \ldots \alpha_n$—occurs in the set Γ. Compare this with the corresponding clause (iii) in the proof of proposition 9.8 in chapter 9.

To see that $\mathcal{M}$, as defined, is indeed a model, observe that the domain is nonempty and that every name takes a value within the domain. We must also check that $\mathcal{M}$ assigns values to predicates unambiguously. That is, we must prove:

Lemma 6. If $[\alpha_1] = [\beta_1], \ldots, [\alpha_n] = [\beta_n]$, then

$$\langle [\alpha_1], \ldots, [\alpha_n] \rangle \in \mathcal{M}(\mathbb{P}) \text{ iff } \langle [\beta_1], \ldots, [\beta_n] \rangle \in \mathcal{M}(\mathbb{P}).$$

To prove this, suppose that $[\alpha_1] = [\beta_1], \ldots, [\alpha_n] = [\beta_n]$. By lemma 4 above, this means that the identity sentences $\alpha_1 = \beta_1, \ldots, \alpha_n = \beta_n$ are all in the set Γ. Consider this deduction schema:

$$\begin{array}{rcl}
\{\alpha_1 = \beta_1\} & (\,) & \alpha_1 = \beta_1 \\
 & \vdots & \\
\{\alpha_n = \beta_n\} & (\,) & \alpha_n = \beta_n \\
\varnothing & (\,) & \mathbb{P}\alpha_1 \ldots \alpha_n \leftrightarrow \mathbb{P}\alpha_1 \ldots \alpha_n \\
\{\alpha_1 = \beta_1\} & (\,) & \mathbb{P}\alpha_1 \ldots \alpha_n \leftrightarrow \mathbb{P}\beta_1 \ldots \alpha_n \\
 & \vdots & \\
\{\alpha_1 = \beta_1, \ldots, \alpha_n = \beta_n\} & (\,) & \mathbb{P}\alpha_1 \ldots \alpha_n \leftrightarrow \mathbb{P}\beta_1 \ldots \beta_n
\end{array}$$

(Justifications are left as an exercise.) Thus we see that

$$\alpha_1 = \beta_1, \ldots, \alpha_n = \beta_n \vdash \mathbb{P}\alpha_1 \ldots \alpha_n \leftrightarrow \mathbb{P}\beta_1 \ldots \beta_n.$$

So by proposition 5.2, the sentence $\mathbb{P}\alpha_1 \ldots \alpha_n \leftrightarrow \mathbb{P}\beta_1 \ldots \beta_n$ is in Γ. Again since Γ is maximal, it follows by 5.2 that $\mathbb{P}\alpha_1 \ldots \alpha_n$ is a member of Γ if and only if $\mathbb{P}\beta_1 \ldots \beta_n$ is. And so by clause (iii) of definition 5 we have the desired result:

$$\langle [\alpha_1], \ldots, [\alpha_n] \rangle \in \mathcal{M}(\mathbb{P}) \text{ iff } \langle [\beta_1], \ldots, [\beta_n] \rangle \in \mathcal{M}(\mathbb{P})$$

In thus concluding the proof of the lemma and the verification that $\mathcal{M}$ is a model, let us note that $\mathcal{M}$ is in fact a covered model: each element $[\alpha]$ of $\mathcal{M}$'s domain is the value of a name in the language.

We come now to the final lemma, the *truth lemma*, which states that a sentence is true in the canonical model $\mathcal{M}$ just in case it is a member of the maximal EI-closed set of sentences Γ.

Lemma 7. $\models_{\mathcal{M}} \mathrm{A}$ iff $\mathrm{A} \in \Gamma$, for every sentence A.

The proof is of course inductive. The cases in which A is molecular or quantificational are the same as they were in the proofs of propositions 5.4 and 9.8 (pages 211 and 339). So we treat here only the cases in which A is atomic, that is, when A is (a) a predication and (b) an identity.

For (a). Suppose A is a predication, $\mathbb{P}\alpha_1 \ldots \alpha_n$. Then:

$$\begin{aligned} \models_{\mathcal{M}} \mathbb{P}\alpha_1 \ldots \alpha_n \text{ iff } & \langle \mathcal{M}(\alpha_1), \ldots, \mathcal{M}(\alpha_n) \rangle \in \mathcal{M}(\mathbb{P}) \\ & \quad \text{—definition 7.4, part (1);} \\ \text{iff } & \langle [\alpha_1], \ldots, [\alpha_n] \rangle \in \{\langle [\alpha_1], \ldots, [\alpha_n] \rangle : \mathbb{P}\alpha_1 \ldots \alpha_n \in \Gamma\} \\ & \quad \text{—definition of } \mathcal{M}; \\ \text{iff } & \mathbb{P}\alpha_1 \ldots \alpha_n \in \Gamma \\ & \quad \text{—more simply.} \end{aligned}$$

So the lemma holds when A is a predication.

For (b). Suppose A is an identity, $\alpha = \beta$. Then:

$$\begin{aligned} \models_{\mathcal{M}} \alpha = \beta \text{ iff } & \mathcal{M}(\alpha) = \mathcal{M}(\beta) \\ & \quad \text{—definition 7.4, part (1');} \\ \text{iff } & [\alpha] = [\beta] \\ & \quad \text{—definition of } \mathcal{M}; \\ \text{iff } & \alpha = \beta \in \Gamma \\ & \quad \text{—lemma 4 above.} \end{aligned}$$

So the lemma holds when A is an identity.

This completes our proof of lemma 7.

Since all the sentences in the maximal EI-closed set Γ are true in the canonical model $\mathcal{M}$, the set is satisfiable. The proof of proposition 10.6 is therefore finished. □

PROPOSITION 10.7. *Completeness.* The deductive system for predicate logic with identity is complete. That is, if $\Gamma \models A$ then $\Gamma \vdash A$.

Proof. As before, we argue for an equivalent statement. Suppose that Γ is a consistent set of sentences. Then by proposition 9.6 (page 334), Γ has a maximal EI-closed extension Δ. By proposition 10.6, moreover, the set Δ is satisfiable. So Γ is satisfiable. Therefore, every consistent set of sentences is satisfiable—which means that the deductive system is complete. □

PROPOSITION 10.8. *Completeness and corollaries.*

(1) If $\Gamma \models A$ then $\Gamma \vdash A$.

(2) If $\models A$ then $\vdash A$.

(3) If $A \simeq B$ then $A \sim B$.

(4) If Con Γ then Sat Γ.

Proof. Compare the proof for proposition 5.6 (page 213). □

EXERCISES

10.28. Justify the lines of the second and third deductions in the proof of lemma 2 in the proof of proposition 10.6. Use the third deduction to prove that the relation $\equiv$ is transitive.

10.29. Justify the lines of the deduction schema in the proof of lemma 6 in the proof of proposition 10.6.

10.30. Prove lemma 4 in the proof of proposition 10.6.

10.31. Complete the proof of lemma 7 in the proof of proposition 10.6 by arguing for the inductive cases in which A is a negation, conjunction, disjunction, conditional, biconditional, universal quantification, and existential quantification. (Before proceeding, compare exercise 9.14, page 342!)

10.32. *An alternative canonical model.* Let Γ be a maximal EI-closed set of sentences. Define the relation $\equiv$, this time *in the set* $\mathbf{N}$, as follows.

$$n \equiv m \text{ iff } \mathrm{a}_n = \mathrm{a}_m \in \Gamma.$$

In other words, natural numbers n and m are $\equiv$-equivalent just in case the sentence $\mathrm{a}_n = \mathrm{a}_m$ is in Γ. The equivalence class $[n]$ is now the set of all numbers that are $\equiv$-equivalent to n; i.e. $[n] = \{m \in \mathbf{N}: n \equiv m\}$.

Define the model $\mathcal{M}$ as follows.

(i) D (the domain of $\mathcal{M}$) $= \{[n]: n \in \mathbf{N}\}$.
(ii) $\mathcal{M}(\mathrm{a}_n) = [n]$, for each $n \in \mathbf{N}$.
(iii) $\langle [k_1], \ldots, [k_n] \rangle \in \mathcal{M}(\mathbb{P})$ iff $\mathbb{P}\mathrm{a}_{k_1} \ldots \mathrm{a}_{k_n} \in \Gamma$, for each n-place predicate $\mathbb{P}$ ($n \geq 0$) and natural numbers $k_1, \ldots, k_n$.

That is to say: (i) the domain D of the canonical model $\mathcal{M}$ is the set of all $\equiv$-equivalence classes of natural numbers; (ii) $\mathcal{M}$ assigns to each name a_n the equivalence class $[n]$ as value; and (iii) the value under $\mathcal{M}$ of an n-place predicate is the set of all n-tuples of equivalence classes of numbers that index the names that follow the predicate to make atomic sentences that appear in the set Γ. It should be evident that $\mathcal{M}$ is a model: it has a nonempty domain, names take values in the domain, and each n-place predicate's value is a set of n-tuples of elements of the domain. Note that the model $\mathcal{M}$ is covered: every element of D is the value under $\mathcal{M}$ of some name.

Prove that just those sentences are true in the model $\mathcal{M}$ as are members of the set Γ (and so $\mathcal{M}$ is a canonical model). That is, prove that for every sentence A, $\models_{\mathcal{M}}$ A iff A $\in \Gamma$. Give an inductive argument for the cases in which A is atomic (i.e. a predication or an identity), a negation, a conditional, and a universal quantification.

This result provides an alternative way of proving proposition 10.6. Explain how. (Compare exercise 9.16, page 342.)

10.33. Modify the definition of $[n]$ in exercise 10.32 so that, instead of being a class of numbers, $[n]$ is the smallest natural number m such that $n \equiv m$. Then in the canonical model defined in exercise 10.32 the domain is simply a set of natural numbers (and not, as there, a set of sets of natural numbers). Check to see that these modifications do not affect the proof of lemma 7 in proposition 10.6.

8. Compactness, and another corollary

Just as in chapters 5 and 9, the soundness and completeness yield compactness results for consequence and satisfiability, but now in predicate logic with identity.

PROPOSITION 10.9. *Compactness for consequence and satisfiability.*

(1) If $\Gamma \models A$, then $\Gamma' \models A$ for some finite subset Γ' of Γ.

(2) Sat Γ, if Sat Γ' for every finite subset Γ' of Γ.

Proof. See the proof for proposition 5.7 (page 214). □

As before (propositions 5.8 and 9.12), compactness yields a number of alternative statements of soundness and completeness.

PROPOSITION 10.10. In each column below, each statement is equivalent to each of the others.

	Soundness	*Completeness*	*Both*
(1)	$\Gamma \vdash A \Rightarrow \Gamma \models A$	$\Gamma \models A \Rightarrow \Gamma \vdash A$	$\Gamma \models A$ iff $\Gamma \vdash A$
(2)	$\vdash A \Rightarrow\ \models A$	$\models A \Rightarrow\ \vdash A$	$\models A$ iff $\vdash A$
(3)	$A \sim B \Rightarrow A \simeq B$	$A \simeq B \Rightarrow A \sim B$	$A \simeq B$ iff $A \sim B$
(4)	Sat $\Gamma \Rightarrow$ Con Γ	Con $\Gamma \Rightarrow$ Sat Γ	Con Γ iff Sat Γ

Proof. Exercise. □

EXERCISES

10.34. Explain why in predicate logic with identity a set of sentences that has models of every finite size also has an infinite model.

Suggestion: Recall from exercise 10.18 the set Ω consisting of all sentences A(1), A(2), A(3), . . .—where A(n) is true in any model of least n elements and is false in all smaller models, for every positive integer n.

Now let Γ be a set of sentences having models of every finite size, and consider $\Gamma \cup \Omega$. As stated in exercise 10.18, any model of $\Gamma \cup \Omega$ is infinite (prove this if you have not already done so). Proceed by showing that every finite subset of $\Gamma \cup \Omega$ has a model. (Let Φ be a finite subset of $\Gamma \cup \Omega$. Argue that Φ has a model of size n, where n is the highest number for which the sentence A(n) is in Φ.) It follows by compactness (proposition 10.9) that $\Gamma \cup \Omega$ has a model, which must be infinite, and from this the desired conclusion concerning Γ follows.

10.35. Recall from exercise 9.21 (page 345) the quantifier Ⅎ meaning *there are finitely many*. Use the result of exercise 10.34 to explain why Ⅎ is not definable in the language of predicate logic with identity.

10.36. Recall from exercise 9.22 (page 346) the quantifier I meaning *there are infinitely many*. Is I definable in predicate logic with identity? Explain.

9. Löwenheim–Skolem

For first order predicate logic without identity, the Löwenheim–Skolem theorem (9.13) says that every satisfiable set of sentences has a model with a domain the size of the set of natural numbers. Now, with identity as a logical constant, only a modified version of the Löwenheim–Skolem theorem holds: every satisfiable set of sentences has a model with a domain the size of *a subset of* the set of natural numbers. In other words, in first order predicate logic with identity, if a set of sentences is satisfiable, then it has a model whose domain is either finite or denumerably infinite in size.

It is easy to see that the original version of Löwenheim–Skolem does not hold when identity is a logical constant. For example, the sentence $\exists x \forall y\ x = y$ is true in a model if and only if the model's domain consists of just one thing. So the set $\{\exists x \forall y\ x = y\}$ is satisfiable, but it does not have any model whose domain is the size of the set of natural numbers. (Similar examples can of course be produced for every number $n > 1$.)

PROPOSITION 10.11. *Löwenheim–Skolem theorem*. If Sat Γ, then there is a model $\mathcal{M}$ such that (i) the domain of $\mathcal{M}$ is the size of a subset of the set $\boldsymbol{N}$, and (ii) every sentence in Γ is true in $\mathcal{M}$.

Proof. Suppose the set Γ is satisfiable. By soundness, Γ is consistent. So by proposition 9.6 (page 334), Γ has a maximal EI-closed extension Δ. Of course every sentence in Γ is true in the canonical model for Γ constructed in the proof of proposition 10.6. In that construction, the domain of the canonical model is the set $\{[\alpha]: \alpha \text{ is a name}\}$. The size of this set is obviously at most the size of the set $\boldsymbol{N}$. In particular, the domain may be finite, i.e. be the size of some finite subset of $\boldsymbol{N}$. (It will be finite, for example, when Γ is the set $\{\exists x \forall y\ x = y\}$ mentioned above.) □

EXERCISES

10.37. The Löwenheim–Skolem theorem for predicate logic with identity can also be formulated as follows. If Sat Γ, then there is a model $\mathcal{M}$ such that (i) the domain of $\mathcal{M}$ is a subset of the set $\boldsymbol{N}$, and (ii) every sentence in Γ is true in $\mathcal{M}$. Use the results of exercises 10.32 and 10.33 to explain how the alternative definition of a canonical model in the latter exercise can be used to establish this formulation of the Löwenheim–Skolem theorem. (Compare exercises 9.16 and 9.19, pages 342 and 345.)

10.38. Recall from exercise 9.20 (page 345) the quantifier $\exists!$ meaning *there is exactly one*. Show how this quantifier may be defined in predicate logic with identity. That is, given any sentence of the form $\exists!xA$, describe an equivalent sentence. What about the quantifier $\exists?$, meaning *there is at most one*—is it definable? If so, show how.

10.39. Recall from exercise 9.23 (page 346) the quantifier $\cap$ meaning *there are uncountably, or nondenumerably many*. Is $\cap$ definable in predicate logic with identity? If so, why? If not, why not?

10.40. In exercise 9.24 (page 346) we found that—for sentences A and B in predicate logic without identity—if A implies B, A is satisfiable, and B is invalid, then A and B have at least one item (predicate or name) in common. Give a counterexample to this for the language of predicate logic with identity.

10.41. The so-called "upward Löwenheim–Skolem theorem" in exercise 9.25 (page 346) no longer holds. Describe a counterexample. An infinitary version of the "upward Löwenheim–Skolem theorem" does hold, however: if a set of sentences has a model of any infinite size it also has models of every greater infinite size. Adapt the argument for the finitary version in exercise 9.25 to prove this.

10.42. Recall from exercise 9.27 (page 347) the quantifier $\exists_\omega$ meaning *there are exactly as many as there are natural numbers*. Explain why $\exists_\omega$ is not definable. (Use the "upward" Löwenheim–Skolem theorem in exercise 10.41.)

APPENDICES

SOME SET-THEORETICAL CONCEPTS

IN PRESENTING various languages and their logics in the text we make use of some notions and notation from elementary set theory. The main ideas are those of *sets*, *ordered pairs* and *n-tuples*, *relations*, and *functions*. In this appendix we briefly canvass the main points concerning these notions. In sections 1–4 we deal with sets, in section 5 with ordered pairs and n-tuples, in sections 6–8 with relations, and in sections 9–11 with functions.

1. Sets

For our purposes, a *set* is any class, collection, or aggregate of things. The things in a set are called its *members* or *elements*. The identity of a set is wholly determined by its membership: sets are identical just in case they have the same members. Notice that the notion of a set is thus "thinner" than the notion of a property or attribute. For example, the property of being a rational animal is distinct from that of being a featherless biped, whereas the set of rational animals is the same as the set of featherless bipeds (plucked chickens do not count; the example is from antiquity).

Among all the possible sets there is one that contains nothing whatsoever. This is the *empty set* (also called the *null* set), which we denote by the symbol

$$\varnothing.$$

Because the identity of a set is determined by its elements, some at first surprising identifications emerge in connection with the empty set. For example, the set of centaurs is identical with the set of unicorns, since each set is empty and therefore identical with $\varnothing$.

To say that a thing x *is a member of*, or *is an element of*, a set A, we use a lowercase greek epsilon and write:

$$x \in A$$

For example, if A is the set of outlaws, then the sentence

$$\text{Billy the Kid} \in A$$

says that Billy the Kid is a member of the set of outlaws (or, more succinctly, that Billy the Kid is an outlaw). We may also indulge in abbreviations like

$$\text{Kootenay Brown, Jerry Potts, John Ware} \in B$$

to express what the following trio of sentences

$$\text{Kootenay Brown} \in B, \text{Jerry Potts} \in B, \text{John Ware} \in B$$

collectively express regarding membership in a set B.

To deny that x is an element of A we can use the slash:

$$x \notin A$$

For example, it is true that for every thing x, $x \notin \varnothing$.

At this point it may be well to elaborate, however feebly, on what sorts of things can be members of sets. Although there are some constraints on candidates for membership, the answer is roughly that anything can be a member of a set—shoes, ships, sealing wax, cabbages, cows, kings, numbers, and sets themselves. Thus, for example, a set A may contain the empty set $\varnothing$, in which case we may write that $\varnothing \in A$. And notice that the set whose sole member is $\varnothing$ is not itself empty, even though $\varnothing$ is.

Membership is a relation between things (which may be sets) and sets. An important relation that holds only between sets and sets is that of *inclusion*. When every member of a set A is also a member of a set B, we say that A *is included in*, or *is a subset of*, B, and we express this by the following symbolism.

$$A \subseteq B$$

For example, let A be the set of Canadian children and let B be the set of Canadians. Then $A \subseteq B$, since every member of A is also a member of B.

Inclusion may be denied by using the slash; for example, $B \not\subseteq A$, since there are Canadians who are not Canadian children.

Inclusion, $\subseteq$, does not mean *proper inclusion*. To express this relation between sets, the underline is omitted. Thus the sentence

$$A \subset B$$

means that the set A *is properly included in*, or *is a proper subset of*, the set B—i.e. that every member of A is a member of B, and B has at least one element that A does not. Where A and B are again the sets of Canadian children and Canadians, we may say that $A \subset B$.

As a corollary to our explanation of the difference between inclusion and proper inclusion, we see that the identity conditions for sets may be characterized in terms of inclusion: $A = B$ if and only if both $A \subseteq B$ and $B \subseteq A$.

Where A and B are sets, *the intersection of A with B* is the set that contains just those things that are members of both A and B. We denote the intersection of sets A and B:

$$A \cap B$$

An example of the intersection of two sets is afforded by the Canadians and the set of human females: the intersection of these two sets is the set of Canadian females, since these are just the members of both sets. For another example, let A be the set of Canadians and let B be the set of presidents of the United States. Then $A \cap B = \varnothing$.

When the intersection of two sets is empty, as in the last example, the sets are said to be *disjoint*. But we shall not introduce a special symbol for this relation.

Intersection is an operation forming sets out of pairs of sets. Union is another such operation. Where A and B are sets, *the union of A with B* is the set that contains just those things that are members of either or both of A and B. The union of sets A and B is denoted:

$$A \cup B$$

For an example of this operation, let A be the set of Canadian males and let B be the set of Canadian females. Then $A \cup B$ is the set of Canadians (barring biological phenomena unfamiliar to the author). Or, if A is the set of Mexicans and B is the set of Mexican males, then $A \cup B = A$.

Finally, there is relative complement. Where A and B are sets, *the (relative) complement of B with respect to A* (more briefly: the *difference* of A and B) is the set that contains just those things that are members of A but not B. We denote the complement of B with respect to A:

$$A - B$$

By way of illustration, let A and B again be the set of Mexicans and the set of Mexican males, respectively. Then $A - B$ is the set of Mexican females (again, barring . . .). Or, if A is the set of prime integers and B is the set of odd integers, then $A - B$ is the set whose sole member is the number two.

In most discussions involving sets, the existence of a *universal set*, V, is assumed. V includes all the sets under discussion, and it is important to realize that the content of V changes from context to context. Relative to a given universal set V there is the operation of (absolute) complement. Where A is a subset of V, *the (absolute) complement of* A is the set that contains just those things that are members of V but not of A. The complement of a set A is denoted:

$$\bar{A}$$

A simple example of complementation is the case in which the universal set is the set of natural numbers and A is the set of odd numbers. Then $\bar{A}$ is the set of even numbers (including zero).

The intersection, union, relative complement, and absolute complement of sets can be represented pictorially by what are known as *Venn diagrams*. In figures 1–4 below, the rectangle represents a universal set, and the labeled circles represent sets within the universal set. The lined areas in the diagrams represent, respectively, $A \cap B$, $A \cup B$, $A - B$, and $\bar{A}$.

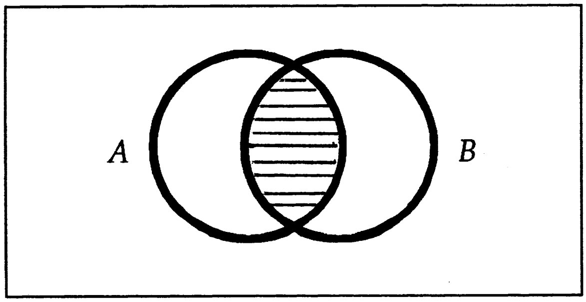

Figure 1. $A \cap B$

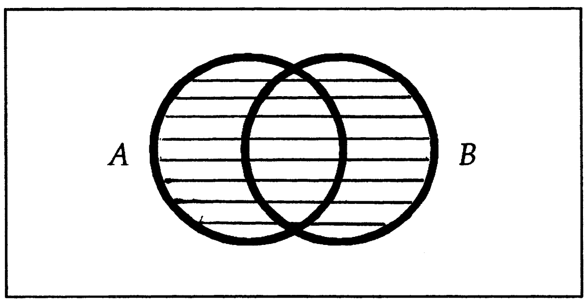

Figure 2. $A \cup B$

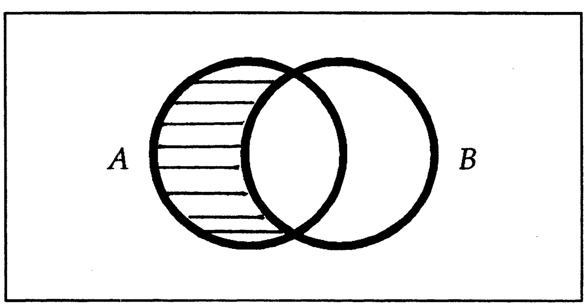

Figure 3. $A - B$

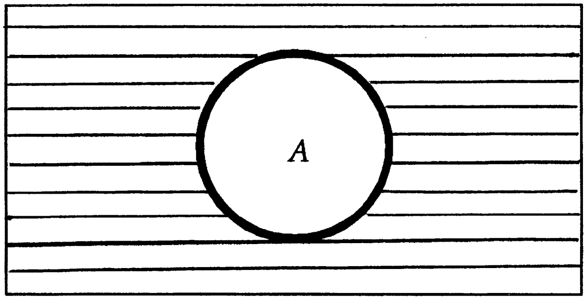

Figure 4. $\bar{A}$

It is possible to "read off" from Venn diagrams such as these a number of general truths about sets and the operations and relations on them. For example, from the diagram in figure 1, we immediately see that the sets $A \cap B$ and $B \cap A$ are the same—i.e. that the operation of intersection is commutative. Likewise for the operation of union: see figure 2.

For truths involving three sets at a time, one can construct more elaborate Venn diagrams. For example, the diagrams in figures 5 and 6 show that intersection is an associative operation on sets, i.e. that for all sets A, B, and C, $A \cap (B \cap C) = (A \cap B) \cap C$.

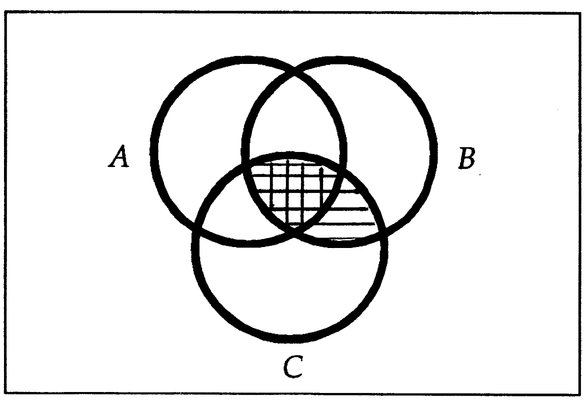

Figure 5. $A \cap (B \cap C)$

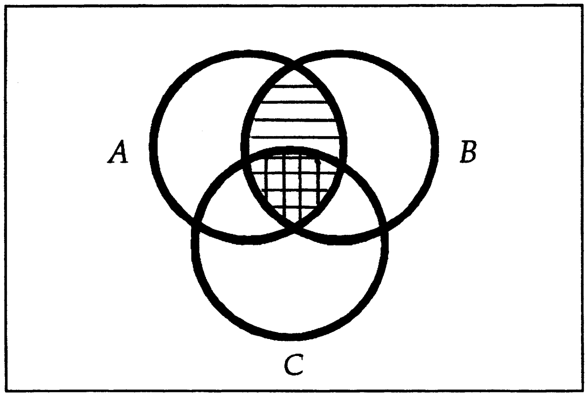

Figure 6. $(A \cap B) \cap C$

In figure 5, the set $B \cap C$ is represented first, by horizontal lining. Then $A \cap (B \cap C)$ is represented by vertical lining. In figure 6, the set $A \cap B$ is represented first, by horizontal lining. Then $(A \cap B) \cap C$ is represented by vertical lining. Because the resulting hatched areas in the two diagrams are the same, $A \cap (B \cap C) = (A \cap B) \cap C$.

In section 4 we list sixty-odd principles about sets, many of which can be shown to be true by means of Venn diagrams.

We have enumerated a number of relations between sets and operations on sets. We turn now to a discussion of the ways in which sets may be described.

2. Describing sets

A set may be described by listing its members between a pair of braces. For example, the set consisting of the leaders of the Seventh Cavalry at the Little Bighorn is denoted by the expression:

{Custer, Reno, Benteen}

The order of the listing is immaterial; both of the following expressions, for example, denote the set just mentioned.

{Reno, Custer, Benteen}

{Benteen, Reno, Custer}

Redundancies in the listing do not matter either. The set denoted by the expression

$$\{\text{Reno, Benteen, Reno, Custer}\}$$

is the same as that denoted by the others. But notice that redundancies can have subtle forms. For example, the set $\{0, 1, 2, 1 + 1\}$ is the same set as $\{0, 1, 2\}$, since $1 + 1 = 2$.

When the members of a finite set are too numerous conveniently to list, ellipsis is permitted. For example, the expression

$$\{0, 1, 2, \ldots, 99\}$$

obviously denotes the set consisting of the first one hundred natural numbers. Similarly, the expression

$$\{1, 3, 5, \ldots, 99\}$$

denotes the set of odd numbers from 1 through 99.

Ellipsis serves essentially in the description of some infinite sets, too. For example, the set of natural numbers may be described by the expression:

$$\{0, 1, 2, \ldots\}$$

The ellipsis indicates how the list continues, just as in the finite case. Similarly, the set of odd numbers greater than 9 is denoted by:

$$\{11, 13, 15, \ldots\}$$

Another way of describing a set is by mentioning a property or condition common to all and only the members of the set. For sets described in this way we often use an expression of the form:

$$\{x: \qquad \}$$

An expression of this form may read "the set of things x such that _______", where in place of the blank a property or condition is mentioned. For example, the expression

$$\{x\text{: } x \text{ is a natural number}\}$$

denotes the set $\{0, 1, 2, \ldots\}$, i.e. the set of things x such that x is a natural numbers. Similarly, the expression

$$\{x\text{: } x \text{ is a leader of the Seventh Cavalry at the Little Bighorn}\}$$

denotes the set {Custer, Reno, Benteen}.

Occasionally, though not frequently, we employ expressions of the form

$$\{f(x)\text{: } \quad\},$$

in which f is some operation or function. Such notation is short for

$$\{y\text{: } y = f(x), \text{ for some } x \text{ such that } \quad\},$$

where the blank is as usual completed by mentioning some property or condition. Expressions of this form are perhaps best understood by means of an example or two. For instance, the expression

$$\{x^2\text{: } x \text{ is a natural number}\}$$

denotes the set of squares of natural numbers, $\{0, 1, 4, 9, 16, \ldots\}$ (compare $\{y\text{: } y = x^2, \text{ for some } x \text{ such that } x \text{ is a natural number}\}$). Or, for another example, the expression

$$\{\text{the father of } x\text{: } x \text{ is human}\}$$

denotes the set of all human fathers (anyway, fathers of humans; various gods may make "human fathers" too narrow).

In contrast to the form just described, we often describe sets by means of expressions of the form

$$\{x \in A: \quad \},$$

in which A is a set. This notation is short for:

$$\{x: x \in A \text{ and } \quad \}$$

For example, let A be the set of Mexicans. Then the expression

$$\{x \in A: x \text{ is male}\}$$

denotes the set of Mexican males. For another example, let A be the set of natural numbers. Then the expression

$$\{x \in A: x \text{ is prime}\}$$

denotes the set $\{2, 3, 5, 7, 11, 13, \ldots\}$ of prime numbers.

Some sets are mentioned so frequently that it is convenient to have (or invent) simple names for them. The empty set, $\varnothing$, is an obvious example. Another is the set $\{0, 1, 2, \ldots\}$ of natural numbers, which we call $\boldsymbol{N}$ for short.

Finally, sets may be described—usually at greater length, but not always—simply by means of English phrases; witness those used throughout the section (and, indeed, the book).

The braces notation for set description can be used to define the empty set ($\varnothing$), and the operations of intersection ($\cap$), union ($\cup$), and relative complement ($-$), as well as that of absolute complement ($\bar{\ }$) with respect to a given universal set V. Where A and B are any subsets of V, the definitions are as follows.

$$\varnothing = \{x \in V: x \notin V\}.$$

$$A \cap B = \{x \in V: x \in A \text{ and } x \in B\}.$$

$$A \cup B = \{x \in V: x \in A \text{ or } x \in B \text{ (or both)}\}.$$

$$A - B = \{x \in V: x \in A \text{ and } x \notin B\}.$$

$$\bar{A} = \{x \in V: x \notin A\}.$$

In case the universal set is tacitly assumed, the condition expressed by "$\in V$" may be omitted throughout—and the definition of the empty set can be altered to read:

$$\varnothing = \{x: x \neq x\}$$

Proofs of some of the general truths about sets and their operations and relations can often be carried out conveniently in terms of these definitions. For example, that $A \cap B = B \cap A$ is easily established simply by noting that for every thing x, $x \in A$ and $x \in B$ if and only if $x \in B$ and $x \in A$—so that $\{x: x \in A \text{ and } x \in B\} = \{x: x \in B \text{ and } x \in A\}$.

3. Sizes of sets

Before moving to a list of principles concerning sets, let us note that sets come in various sizes, or *cardinalities*. The principal distinction in sizes is between those that are *finite* and those that are *infinite*. The set of plug nickels, however large, is clearly finite, as is the empty set, $\varnothing$. The set of natural numbers $\{0, 1, 2, \ldots\}$ is infinite, on the other hand.

Notice that proper inclusion in an infinite set does not guarantee the finitude of the included set. For example, the set $\{1, 3, 5, \ldots\}$ of odd numbers is clearly a proper subset of the set $\{0, 1, 2, \ldots\}$ of natural numbers; but the former set is no less infinite than the latter—indeed, there are the same number of odd numbers as there are natural numbers.

Not all infinite sets are of equal size, however. The set of non-negative real numbers is infinite, but of an order of infinity strictly greater than that of the set of natural numbers. Infinite sets the size of the set of natural numbers are said to be *countably* or *denumerably* or *enumerably infinite*; infinite sets of larger size are said to be *uncountably* or *non-denumerably infinite*.

4. Some principles

We devote this section entirely to the following principles, which hold for all sets A, B, C that are subsets of a universal set V.

(1) $\varnothing \subseteq A$.

(2) $A \subseteq A$.

(3) $A \subseteq V$.

(4) $A \not\subset A$.

(5) $A \subseteq \varnothing$ iff $A = \varnothing$.

(6) $V \subseteq A$ iff $V = A$.

(7) If $A = B$, then $A \subseteq B$.

(8) $A = B$ iff $A \subseteq B$ and $B \subseteq A$.

(9) If $A \subseteq B$ and $B \subseteq C$, then $A \subseteq C$.

(10) If $A \subseteq B$ and $B = C$, then $A \subseteq C$.

(11) If $A = B$ and $B \subseteq C$, then $A \subseteq C$.

(12) $A \subseteq B$ iff $A \cap B = A$.

(13) $A \subseteq B$ iff $A \cup B = B$.

(14) $A \subseteq B$ iff $A - B = \varnothing$.

(15) $A \subset B$ iff $A \subseteq B$ and $A \neq B$.

(16) If $A \subset B$ then $A \subseteq B$.

(17) If $A \subset B$ and $B \subset C$, then $A \subset C$.

(18) If $A \subset B$ and $B \subseteq C$, then $A \subset C$.

(19) If $A \subseteq B$ and $B \subset C$, then $A \subset C$.

(20) If $A \subset B$, then $B \not\subset A$.

(21) $A \cap \varnothing = \varnothing$.

(22) $A \cup \varnothing = A$.

(23) $A - \varnothing = A.$

(24) $\varnothing - A = \varnothing.$

(25) $A \cap V = A.$

(26) $A \cup V = V.$

(27) $A - V = \varnothing.$

(28) $V - A = \bar{A}.$

(29) $A \cap A = A.$

(30) $A \cup A = A.$

(31) $A - A = \varnothing.$

(32) $\bar{\varnothing} = V.$

(33) $\bar{V} = \varnothing.$

(34) $A = \bar{\bar{A}}.$

(35) $A \cap \bar{A} = \varnothing.$

(36) $A \cup \bar{A} = V.$

(37) $A - \bar{A} = A.$

(38) $\bar{A} - A = \bar{A}.$

(39) $\overline{A \cap B} = \bar{A} \cup \bar{B}.$

(40) $\overline{A \cup B} = \bar{A} \cap \bar{B}.$

(41) $A \subseteq B$ iff $\bar{B} \subseteq \bar{A}.$

(42) $A = \bar{B}$ iff $\bar{A} = B.$

(43) $A \cap B = B \cap A.$

(44) $A \cup B = B \cup A.$

(45) $A - B = A \cap \bar{B}.$

(46) $A \cap (B \cap C) = (A \cap B) \cap C.$

(47) $A \cup (B \cup C) = (A \cup B) \cup C.$

(48) $A \cap (A \cup B) = A.$

(49) $A \cup (A \cap B) = A.$

(50) $(A - B) - A = \varnothing.$

(51) $(A - B) - B = A - B.$

(52) $A - (A - B) = A \cap B.$

(53) $A - (B - A) = A.$

(54) $(A \cap B) - B = \varnothing$

(55) $(A \cup B) - B = A - B.$

(56) $(A - B) \cup B = A \cup B.$

(57) $A \cap (B \cup C) = (A \cap B) \cup (A \cap C).$

(58) $A \cup (B \cap C) = (A \cup B) \cap (A \cup C).$

(59) $A - (B \cap C) = (A - B) \cup (A - C).$

(60) $A - (B \cup C) = (A - B) \cap (A - C).$

(61) $A \cap (B - C) = B \cap (A - C).$

(62) $(A \cup B) - C = (A - C) \cup (B - C).$

(63) $(A - B) - C = A - (B \cup C).$

(64) If $A \subseteq B$, then $A \cap C \subseteq B \cap C$.

(65) If $A \subseteq B$, then $A \cup C \subseteq B \cup C$.

(66) If $A \subseteq B$, then $A - C \subseteq B - C$.

(67) If $A \subseteq B$, then $C - B \subseteq C - A$.

5. Ordered pairs and *n*-tuples

An *ordered pair*

$$\langle x, y \rangle$$

is a pair of things x and y in which the order is important: unless $x = y$, the ordered pair $\langle x, y\rangle$ is in general different from the ordered pair $\langle y, x\rangle$. Thus, for example,

$$\langle 0, 1\rangle \neq \langle 1, 0\rangle,$$

since $0 \neq 1$. (Contrast this with sets, where $\{x, y\}$ and $\{y, x\}$ are the same set.)

The parts of an ordered pair are called *coordinates*; x is the first coordinate of the pair $\langle x, y\rangle$, and y is the second.

The coordinates of an ordered pair can be identical:

$$\langle x, x\rangle$$

But this pair is not to be confused or identified with the thing x itself.

In summary, then, an ordered pair is identified solely in terms of its content and order:

$$\langle x, y\rangle = \langle z, w\rangle \text{ iff } x = z \text{ and } y = w.$$

The notion of an ordered pair may be generalized to that of an ordering of n coordinates, for any finite number $n \geq 0$:

$$\langle x_1, \ldots, x_n\rangle$$

—an *n-tuple*. Ordered pairs are thus two-tuples, or couples, and similarly for triples, quadruples, quintuples, etc.

Notice that the generalization goes downward, too, to the cases in which $n = 1$ and $n = 0$. A one-tuple (a single?)

$$\langle x\rangle$$

may be identified with its sole coordinate; i.e. we shall say that

$$\langle x\rangle = x.$$

(This is a good place to emphasize, again, that although $\langle x \rangle = x$, $\langle x, x \rangle \neq x$—since $\langle x, x \rangle$ has two coordinates whereas $\langle x \rangle$ has one.)

The zero-tuple,

$$\langle\,\rangle,$$

need not be otherwise identified. But note that there is just one zero-tuple (just as there is but one empty set).

The identity conditions for ordered pairs can be generalized for n-tuples ($n \geq 0$):

$$\langle x_1, \ldots, x_n \rangle = \langle y_1, \ldots, y_n \rangle$$
$$\text{iff}$$
$$x_1 = y_1, \ldots, x_n = y_n.$$

By now readers have probably come to see that the idea of a *finite sequence* (possibly null) is explicated by the concept of an n-tuple. In some contexts, n-tuples are thought of as *ordered sets*.

Finally—and this is not essential to an understanding of n-tuples—there are various ways in which the notion of an n-tuple can be defined. To take but one example, the ordered pair $\langle x, y \rangle$ can be defined as a certain set, to wit:

$$\langle x, y \rangle = \{\{x\}, \{x, y\}\}.$$

(To see that this is an acceptable definition, simply check that the identity conditions for ordered pairs are satisfied. That is to say, verify that $\{\{x\}, \{x, y\}\} = \{\{z\}, \{z, w\}\}$ if and only if $x = z$ and $y = w$.) From this, one can define a triple $\langle x, y, z \rangle$ as a certain ordered pair; to wit:

$$\langle x, y, z \rangle = \langle \langle x, y \rangle, z \rangle.$$

(To see that this is an acceptable definition, check that the identity conditions for triples are satisfied—i.e. that $\langle\langle x, y \rangle, z \rangle = \langle\langle u, v \rangle, w \rangle$ if and only if $x = u$, $y = v$, and $z = w$.) And in general, for $n > 2$, an n-tuple is defined:

$$\langle x_1, \ldots, x_n \rangle = \langle \langle x_1, \ldots, x_{n-1} \rangle, x_n \rangle.$$

Readers new to these notions can safely ignore the construction of n-tuples in such terms as these. An intuitive understanding is enough for the purposes of the next few sections.

6. Relations

A *relation* is usually thought of as something like a property that holds between or among various things. For example, *love*, or *loving*, is a relation holding between (usually animate) things; and *betweenness* is a relation holding among things three at a time. Love is a two-place or *binary* relation; betweenness is a three-place or *ternary* relation. In general, we may conceive of n-place or *n-ary* relations—relations that hold among things n at a time.

We may construe an n-ary relation in a set as just the set of n-tuples of those things in the set for which the relation holds. Examples: With respect to the set $\{1, 2, 3\}$, the set of ordered pairs

$$\{\langle 1, 2 \rangle, \langle 1, 3 \rangle, \langle 2, 3 \rangle\}$$

is the binary relation of being *less than* (<), since the elements are just the ordered pairs of members of $\{1, 2, 3\}$ whose first coordinates are less than their second coordinates. The set of pairs

$$\{\langle 2, 1 \rangle, \langle 3, 1 \rangle, \langle 3, 2 \rangle\},$$

on the other hand, is the relation of being *greater than* (>) on the set. And the set of pairs

$$\{\langle 1, 1 \rangle, \langle 2, 2 \rangle, \langle 3, 3 \rangle\}$$

is the *identity* relation (=) on the set. With respect to the same set, $\{1, 2, 3\}$, the set of triples

$$\{\langle 3, 2, 1 \rangle, \langle 3, 1, 2 \rangle, \langle 2, 1, 1 \rangle\}$$

is the relation of being *equal to the sum of* (e.g. $3 = 2 + 1$). This, of course, is a ternary or three-place relation in the set.

Notice that by our conventions of the preceding section, simple subsets of $\{1, 2, 3\}$ are *unary* or one-place relations on the set, since they are sets of 1-tuples. For example, the set

$$\{1, 3\},$$

otherwise known as $\{\langle 1\rangle, \langle 3\rangle\}$, must be considered a relation in the set $\{1, 2, 3\}$. Thus simple properties are treated as kinds of relations—unary relations.

Moreover, the set

$$\{\langle\ \rangle\}$$

counts as a relation—as a zero-ary relation, since it consists of 0-tuples. (In a way, $\{\langle\ \rangle\}$ is Truth itself.)

Just as we write

$$\{x: x < 3\}$$

to describe the set $\{1, 2\}$, we can write, for example,

$$\{\langle x, y\rangle: x < y\}$$

to describe the binary relation $\{\langle 1, 2\rangle, \langle 1, 3\rangle, \langle 2, 3\rangle\}$ of being less than, mentioned above. In general we often describe an n-ary relation in a set by the form

$$\{\langle x_1, \ldots, x_n\rangle: \qquad\quad \},$$

where the blank is filled by some condition on $x_1, \ldots, x_n$ (for example, "$x < y$", above).

Now consider these binary relations:

$$\{\langle 1, 2\rangle, \langle 1, 3\rangle, \langle 2, 3\rangle\}$$

$$\{\langle 1, 2\rangle, \langle 1, 3\rangle, \langle 1, 4\rangle, \langle 2, 3\rangle, \langle 2, 4\rangle, \langle 3, 4\rangle\}$$

The first is the less-than relation in the set $\{1, 2, 3\}$; the second is the less-than relation in the set $\{1, 2, 3, 4\}$. Both are less-than relations, however, and this shows that the description

$$\{\langle x, y\rangle : x < y\}$$

employed just above is incomplete: mention must be made of the base set for the relation. For the moment, let us say that $A = \{1, 2, 3\}$ and $B = \{1, 2, 3, 4\}$. Then the following descriptions describe the two less-than relations in question.

$$\{\langle x, y\rangle : x, y \in A \text{ and } x < y\}$$

$$\{\langle x, y\rangle : x, y \in B \text{ and } x < y\}$$

We will shortly find an even shorter way of saying what set a relation is on.

Restricting ourselves for the time being to binary relations, we can recognize three relations that always exist in a set A:

- the *universal* (or *total*) relation in A—the set

  $$\{\langle x, y\rangle : x, y \in A\}$$

  of all possible ordered pairs of elements of A

- the *identity* relation in A—the set

  $$\{\langle x, x\rangle : x \in A\}$$

  of all ordered pairs of identical elements of A, otherwise described $\{\langle x, y\rangle : x, y \in A \text{ and } x = y\}$

- the *empty* (or *null*) relation in A—the set

$$\varnothing$$

of all pairs of elements of A drawn from the empty set, otherwise described $\{\langle x, y\rangle: x, y \in \varnothing\}$

Generalizing from the binary case, we reach the ideas of the *universal, identity,* and *empty* n-ary relations in a set A:

$$\{\langle x_1, \ldots, x_n\rangle: x_1, \ldots, x_n \in A\}$$

$$\{\langle x_1, \ldots, x_n\rangle: x_1 = \ldots = x_n \in A\}$$

$$\{\langle x_1, \ldots, x_n\rangle: x_1, \ldots, x_n \in \varnothing\}$$

Thus we see that just as every binary relation in a set A is a subset of the universal binary relation in A, so likewise every n-ary relation in a set A is a subset of the universal n-ary relation in A. Moreover, the empty set, $\varnothing$, turns out to be an n-ary relation for every $n \geq 0$.

Relations can be regarded as formed by an operation on sets, the operation of forming the *cross* (or *cartesian*) *product*. By the cross product of sets A and B—in symbols,

$$A \times B$$

—we mean the set of all ordered pairs whose first and second coordinates are drawn, respectively, from the sets A and B. That is:

$$A \times B = \{\langle x, y\rangle: x \in A \text{ and } y \in B\}.$$

For example, let $A = \{1, 2\}$ and $B = \{2, 3\}$. Then the sets A and B give rise to the binary relation $A \times B$:

$$\{\langle 1, 2\rangle, \langle 1, 3\rangle, \langle 2, 2\rangle, \langle 2, 3\rangle\}$$

Note that the cross product operation $\times$ is not commutative—i.e. that in general, $A \times B \neq B \times A$. (What is $B \times A$ above?)

Generalizing on the idea of cross product, we have the cross product of n sets, $A_1, \ldots, A_n$ ($n \geq 0$):

$$A_1 \times \ldots \times A_n$$

This is the set of all n-tuples whose coordinates belong to the respective sets—in other words:

$$\{\langle x_1, \ldots, x_n\rangle : x_1 \in A_1 \text{ and } \ldots \text{ and } x_n \in A_n\}$$

In limiting cases, we form the nth cross product of a set A with itself:

$$\underbrace{A \times \ldots \times A}_{\ldots n \text{ times} \ldots}$$

This is the universal n-ary relation in A, and we denote it by A^n. Note that $A^1 = A$, and $A^0 = \{\langle\,\rangle\}$.

Now we see that instead of describing the less-than relation in the set $A = \{1, 2, 3\}$ as $\{\langle x, y\rangle : x, y \in A \text{ and } x < y\}$, we can as easily and more briefly write:

$$\{\langle x, y\rangle \in A^2 : x < y\}$$

In the case of a binary relation we speak of its *domain*, its *counterdomain*, and its *field*. The domain of a binary relation is the set of things occurring as first coordinates of its ordered pairs; the counterdomain of a binary relation is the set of things occurring as second coordinates; and the field is the set of all coordinates of pairs in the relation—i.e. the union of the domain and the counterdomain. These ideas extend to n-ary relations, except that the terminology becomes "first domain", "second domain", "third domain", and so on, corresponding to the appropriate sets of coordinates of the n-tuples in the relation. The expression "counterdomain" is thus special to binary relations.

We introduce now some shorthand for stating that things stand in various kinds of relations to each other. Where R is an n-ary relation in a set A—i.e. where $R \subseteq A^n$—and where $x_1, \ldots, x_n$ are elements of A, we ordinarily write

$$\langle x_1, \ldots, x_n \rangle \in R$$

to mean that $x_1, \ldots, x_n$ are related by R. It is often more convenient to write

$$R(x_1, \ldots, x_n),$$

or even just

$$Rx_1 \ldots x_n,$$

with the same meaning. If R is binary, "$\langle x, y \rangle \in R$" is nicely abbreviated to:

$$xRy$$

Compare "$x < y$", used all along in this section. We follow some of these conventions from now on.

7. Some properties of relations

Binary relations can have interesting properties. We list a few below (note universality, emptiness, and identity again). Let R be a binary relation in a set A. Then:

- R is *reflexive* in A iff for every $x \in A$, xRx.
- R is *irreflexive* in A iff for every $x \in A$, not xRx.
- R is *universal* (or *total*) in A iff for every $x, y \in A$, xRy.
- R is *empty* (or *null*) in A iff for every $x, y \in A$, not xRy.
- R is *serial* in A iff for every $x \in A$ there is at least one $y \in A$ such that xRy.
- R is *vacuous* in A iff for every $x, y \in A$, if xRy, then $x = y$.

- R is *identity* in A iff for every $x, y \in A$, xRy if and only if $x = y$.
- R is *symmetric* in A iff for every $x, y \in A$, if xRy then yRx.
- R is *asymmetric* in A iff for every $x, y \in A$, if xRy then not yRx.
- R is *antisymmetric* in A iff for every $x, y \in A$, if xRy and yRx, then $x = y$.
- R is *strongly connected* (or *satisfies dichotomy*) in A iff for every $x, y \in A$, xRy or yRx.
- R is *weakly connected* (or *satisfies trichotomy*) in A iff for every $x, y \in A$, xRy or yRx or $x = y$.
- R is *transitive* in A iff for every $x, y, z, \in A$, if xRy and yRz, then xRz.
- R is *intransitive* in A iff for every $x, y, z, \in A$, if xRy and yRz, then not xRz.
- R is *euclidean* in A iff for every $x, y, z \in A$, if xRy and xRz, then yRz.
- R is *weakly dense* in A iff for every $x, y \in A$, if xRy, then there is a $z \in A$ such that xRz and zRy.
- R is *strongly dense* in A iff for every $x, y \in A$, if xRy, then there is a $z \in A$ such that $x \neq z$ and $y \neq z$, and xRz and zRy.
- R is *partially functional* (or a *partial function*) in A iff for every $x \in A$ there is at most one $y \in A$ such that xRy—i.e. iff for every $x, y, z \in A$, if xRy and xRz, then $y = z$.
- R is *functional* (or a *function* or *operation*) in A iff for every $x \in A$ there is exactly one $y \in A$ such that xRy—i.e. iff R is serial and partially functional in A.

These properties of binary relations are not all independent. For example, every asymmetric binary relation is irreflexive, and every strongly connected relation is reflexive.

8. Equivalence relations

When a binary relation R in a set A is reflexive, symmetric, and transitive, we say that R is an *equivalence relation* in A. Alternatively, and equivalently (as readers can prove), R is an equivalence relation in A just in case it is both reflexive and euclidean.

Relative to an equivalence relation R, certain elements of the set A are thus equivalent—R-*equivalent*. A little reflection reveals that R divides the set A into a number of mutually exclusive subsets, namely subsets consisting of equivalent elements of A. Each such subset is called an *equivalence class* with respect to R, or an R-*equivalence class*. Where x is any member of an R-equivalence class, the equivalence class itself may be denoted by the expression:

$$[x]_R$$

In other words, we define, for any element x of the set A,

$$[x]_R = \{y \in A: xRy\}.$$

That it matters not which element of an equivalence class $[x]_R$ is chosen to represent the class in the description "$[x]_R$" is the import of the following proposition.

$$\text{For any } x, y \in A\text{: } [x]_R = [y]_R \text{ iff } xRy.$$

(Readers may establish the truth of this for themselves.) Often, when the relation R is known or assumed to be understood, the subscript is omitted in "$[x]_R$".

The collection of R-equivalence classes, for a relation R in a set A, is called the *partition* of A under R. And a partition it is, since each element is an R-equivalence class and these classes exhaustively divide A into mutually exclusive subsets (or *cells*, as they are sometimes called).

Of the three relations always known to exist in a set A, two—the universal relation and the identity relation—are equivalence relations. The universal relation, A^2, on A determines but one A^2-equivalence class, namely the set A itself (since all its members are equivalent with respect to A^2). Thus the partition of A relative to the universal relation is just the set $\{A\}$—the set consisting of the set A only. On the other hand, the identity relation in A—often written I_A (or Δ_A)—divides A into as many I_A-equivalence classes as A has elements—since element of A is equivalent only to itself with respect to I_A. Thus the partition of A relative to the identity relation is just the collection of singleton subsets of A—that is, the set $\{\{x\}: x \in A\}$. The universal relation in a set A is therefore said to provide the *coarsest* partition of A, whereas the identity relation in A is said to provide the *finest* partition of A.

For a simple and familiar example of an equivalence relation, let A = the set of human beings, and $R = \{\langle x, y\rangle \in A^2$: x has the same parents as $y\}$. Clearly R is reflexive, symmetric, and transitive in A; so it is an equivalence relation in A. For each human being x, R collects x together with all his or her siblings in the R-equivalence class $[x]_R$. And we see, for instance, that [Jesse James]$_R$ = [Frank James]$_R$, since ⟨Jesse James, Frank James⟩ $\in R$—i.e. Jesse and Frank are siblings.

For another example, let $A = \{1, 2, 3, \ldots\}$, and $R = \{\langle x, y\rangle \in A^2: x \equiv y \pmod 4\}$. Then the equivalence classes generated by R are these four:

$$\{1, 5, 9, \ldots\}$$

$$\{2, 6, 10, \ldots\}$$

$$\{3, 7, 11, \ldots\}$$

$$\{4, 8, 12, \ldots\}$$

9. Functions

In its simplest form, a *function* is something that delivers a certain output given a certain input. For example, among numbers the addition function + delivers the output 12 given 7 and 5 as input—a fact we usually express by saying that $7 + 5 = 12$.

Functions are also referred to as *injections*.

We refer to the input for a function as its *arguments*, and to the output as its *value*. Where f is a function with arguments $x_1, \ldots, x_n$ and value y, we write:

$$f(x_1, \ldots, x_n) = y$$

Thus in the case of addition we might have written:

$$+(7, 5) = 12$$

Note that not all functions are numerical in nature. There is a function t that takes each day of the week to the next:

$$t(\text{Sunday}) = \text{Monday, etc.}$$

There is a function m that takes (most) human beings to their mothers—

$$m(\text{Jesse James}) = \text{Zerelda Samuel, etc.}$$

Of course functions can be iterated. For example, consider $t(t(t(\text{Sunday})))$. In this case we can calculate:

$$\begin{aligned} t(t(t(\text{Sunday}))) &= t(t(\text{Monday})) \\ &= t(\text{Tuesday}) \\ &= \text{Wednesday} \end{aligned}$$

As the notation "$f(x_1, \ldots x_n) = y$" makes evident, functions come in *degrees*, depending on the number of arguments they take as input. Thus addition is a binary or two-place function, while t and m above are unary (more properly, singulary) or one-place functions. The degree of a function is always finite, but otherwise we recognize no limit to degrees. A not too contrived example of a ternary (three-place) function is s, where

$$s(i, j, k) = j$$

—i.e. s picks out the second of any trio of integers.

There is no harm if we include zero among the degrees of functions. If f is a zero-place (zero-ary?) function then, since it takes no arguments to deliver a value, it is its own value—i.e. we can regard a zero-place function simply as an individual. Or, to put the matter the other way around, each thing can be considered a zero-place function—to wit the zero-place function that delivers itself as output given no input at all!

The example of s above shows that the order of arguments for a function is generally important, for $s(1, 2, 3) = 2$, whereas $s(2, 3, 1) = 3$. When the order of arguments is not important—e.g. $7 + 5 = 5 + 7 = 12$—that fact is sometimes itself important (we say that addition is a *commutative* function).

In dealing with a function, care must be taken to specify the source of its arguments and the location of its values. Thus, above, the function t takes its arguments from the set of days of the week and yields its values in that set as well. By contrast, arguments of the function m are drawn from the set of human beings (perhaps with the exceptions of Adam and Eve), although its values lie wholly within the set of females of the species. And the function s goes from the set of triples of integers to the set of integers.

The set of arguments for a function f is called the *domain* of f, and a set that contains all possible values for f is its *range*. Notice that the notion of range is relative to specification: We said above (in effect) that the range of m is the set of female human beings; but we might also have specified the range as the set of all human beings, even though no males are values (i.e. are mothers).

More generally, where f is an n-ary function, we speak of its first, second, third, . . . , nth domains.

The idea of the domain of a function is actually also somewhat relative. For example we might (zanily) say that the domain of m is the set consisting of all human beings and all positive integers. In this case, the function m is said to be *defined* only at (or with respect to) the human members of the domain, and to be *undefined* at 1, 2, 3, ... A function is *total* if it is defined at every member of its domain; otherwise it is a *partial* function. Unless we say differently, we mean *total function* when we say "function". An example of a function that is defined everywhere in every set is the *identity function*, i, for which the value is the same as the argument:

$$i(x) = x.$$

An arrow notation is used to show that a unary function f has domain X and range Y—that f *maps* X *to* (or *into*) Y:

$$f\colon X \mapsto Y$$

Similarly for an n-ary function f with domains $X_1, \ldots, X_n$ (not necessarily distinct) and range Y:

$$f\colon X_1, \ldots, X_n \mapsto Y$$

When it happens that the domains and range of a function are all the same set, the function is called an *operation*. Thus identity, i above, is an operation in every set. Another example is the successor function, which carries a number to its successor:

$$s(n) = n + 1.$$

The successor function is an operation, e.g., in the set of all integers (positive, negative, and zero).

10. Some properties of functions

Certain properties of functions are noteworthy. Suppose a unary function f maps X to Y and "uses up" its range Y in the sense that each member of Y is the value of f for some argument in its domain X—i.e.

for each y in Y there is at least
one x *in* X such that $f(x) = y$.

Then f is said to be *onto*, or to be an *onto function*, or to *map* X *onto* Y. The function t above, taking each day of the week to the next, is an onto function, since each weekday is output for at least (indeed, exactly) one input. Similarly, in the set of all integers, the successor function s is onto, although it is not onto when its domain is restricted to, e.g., the set of natural numbers (0 is not $s(n)$ for any natural number n). Identity, i, is another example of an onto function. Onto-ness is of course relative to specification of range: m, which takes human beings to their mothers, is onto when the range is the set of human mothers, but it is not onto when the range is the set of human females or the set of all humans.

The property of onto-ness can be had by functions of any number of arguments, so that where f is n-ary, we say that it is onto just in case each y in its range is a value for some combination of arguments $x_1, \ldots, x_n$.

Onto functions are also referred to as *surjections*.

In general, functions are many-one mappings, so that the same value may be determined for more than one argument (e.g. m(Jesse James) = m(Frank James) = Zerelda Samuel, where Jesse James ≠ Frank James). When, however, different arguments for a unary function f always lead to different values—i.e. when

$$f(x) \neq f(y) \text{ whenever } x \neq y,$$

or, equivalently,

$$f(x) = f(y) \text{ only if } x = y$$

—the function is said to be *one-one* (or *one-to-one*). The next-day function t and the successor function s are one-one, as is, of course, identity. One-one functions need not be onto, as is shown by the example of successor function s with natural numbers as domain.

A one-one onto function is also called a *bijection*.

11. Functions as relations, and as sets

To each function there corresponds a relation, a relation that relates arguments to values. For example, to the unary next-day function T there corresponds the binary next-day relation T. The function *maps* each day to the next; the relation *relates* each day to the next. The connection between t and T is succinctly stated:

$$T(x, y) \text{ iff } t(x) = y.$$

The point generally is that to each function of degree n there corresponds a relation of degree $n + 1$. That is, where f is n-ary there is an $(n + 1)$-ary relation R such that

$$R(x_1, \ldots, x_n, y) \text{ iff } f(x_1, \ldots, x_n) = y.$$

Looking at the matter the other way around, a relation R can be regarded as a function if it satisfies the condition that for every $x_1, \ldots, x_n$ there is exactly one y such that $R(x_1, \ldots, x_n, y)$. We can split this into conditions of *existence* and *uniqueness*, or what might be called "end-point seriality" and "end-point partial functionality":

> R is *end-point serial* iff for every $x_1, \ldots, x_n$ there is at least one y such that $R(x_1, \ldots, x_n, y)$.
>
> R is *end-point partially functional* iff for every $x_1, \ldots, x_n$ there is at most one y such that $R(x_1, \ldots, x_n, y)$.

Compare the conditions of seriality and partial functionality for binary relations in section 7 above.

Thus functions may be identified, often usefully, with relations of a certain kind. Since, as we saw earlier, relations are construable as sets (of n-tuples), it follows that functions too can be construed as sets.

SELECTED SOLUTIONS TO EXERCISES

FOLLOWING are selected solutions to exercises in chapters 1–4, 6–8, and 10. Be aware that solutions may sometimes also be found in the text and in other exercises. Note, too, that in describing deductions we have omitted the braces ({ , }) for premiss sets and left a blank to indicate the empty set (∅).

Exercises in chapter 1

1.3 Every bear has hair, Custer is a bear / Custer has hair

1.8 See exercises 1.11 and 1.12. (What about the others?)

1.12

Expression	*Interpretation*
these bonbons	dogs
chocolate cream	animal
delicious	canine

1.15

Expression	*Interpretation*
European	male human
French	male
Spanish	non-male
Libyan	human

1.18 Show that the displayed proposition follows from proposition 1.4: Assume proposition 1.4—that whenever a set implies a sentence so does any superset of the set. Now suppose that Γ implies A. Since $\Gamma \cup \Delta$ is a superset of Γ (for any set Δ), $\Gamma \cup \Delta$ also implies A.

1.33 For part (1): Suppose both that Γ is satisfiable and that Γ implies A. This means that there is a model of Γ (i.e. a model in which every sentence in Γ is true) and that A is true in every model of Γ. So A has a model too. This means that A is satisfiable. For part (2): Given part (1), it follows that if Γ implies an unsatisfiable sentence A, then Γ is likewise unsatisfiable.

1.35 For proposition 1.11: Suppose that A and B are valid. This means that A is true in every model and so is B. Hence A and B agree in truth value in every model. This means that A and B are equivalent.

1.37 Only c, e, and h are true; the rest are false. (*Why?*)

Exercises in chapter 2

2.2 (5) If Arapaho do not rule, neither Blackfoot nor Cheyenne do
(8) If either Arapaho or Cheyenne do not rule, then both Blackfoot and Dakota do not
(14) If Arapaho rule only if Blackfoot don't, then either Cheyenne or Dakota do

2.3 See exercise 2.7.

2.4 (5), (8), and (14) Compare exercise 2.2, parts (5), (8), and (14).

2.5 (3) $(F \wedge G) \leftrightarrow ((H \vee I) \wedge \neg (H \wedge I))$ (10) $\neg (\neg E \vee F)$

2.10 (1) Sn (5) Sbsn(Γ) (7) Atm(Γ) (9) ∅

Exercises in chapter 3

3.7 See exercise 3.25.

3.16 (1), (4)–(6), (11)–(12)—and others?

3.17 (1)–(3), (8)–(16)—and others?

3.18 (1)–(9), (12), (15)–(20)—and others?

3.19 (1) and (3)—and others?

3.20 (1)–(2), (4)–(5)—and others?

3.21 All

3.22 (1)–(4), (7)–(10)—and others?

3.23 (3)–(9), (13)–(22)—and others?

3.24 (2) and (4)—and others?

3.25 (3)–(7)—and others?

3.26 (1)–(2), (4)–(9), (15)–(16)—and others?

3.45 A and B are equivalent if and only if the conditionals $A \rightarrow B$ and $B \rightarrow A$ are both valid.

3.50 (3) <u>For right-to-left</u>: Assume that A implies its negation $\neg$ A. This means that there is no model in which A is true and $\neg$ A is false. Then there is no model in which A is true and A is true. In short, there is no model in which A is true. This means that A is unsatisfiable.

3.57 *True* b, e, h, q, s *False* a, d, l (*Why*—and what about the others?)

3.70 See exercise 3.82.

3.78 According to part (3) of exercise 3.13, any sentence is equivalent to its self-disjunction.

3.82 <u>For</u> $\{\Downarrow, \rightarrow\}$: The sentence $A \rightarrow \Downarrow$ expresses the truth function for negation. Since the set $\{\neg, \rightarrow\}$ is truth-functionally complete, so is $\{\Downarrow, \rightarrow\}$.

3.83 <u>For</u> $\clubsuit$: (1) $(\neg A \wedge B \wedge \neg C) \vee (\neg A \wedge \neg B \wedge \neg C)$ *or* $\neg A \wedge \neg C$
(2) $(A \wedge B) \vee (A \wedge \neg B) \vee (\neg A \wedge B)$ *or* $A \vee B$

3.84 <u>For</u> ☻: $(A \wedge B \wedge C \wedge D \wedge E \wedge F) \vee (\neg A \wedge \neg B \wedge \neg C \wedge \neg D \wedge \neg E \wedge \neg F)$

3.89 If a two-place operator is truth-functionally complete by itself, its truth function it begins with $\perp$ and ends with $\top$ (otherwise sentences constructed using the operator will always express a truth function that begins with $\top$ or ends with $\perp$, in which case the truth function for negation cannot be expressed). There are four possible ways of filling in the interior two places in the truth function: $\top\,\top$, $\top\,\perp$, $\perp\,\top$, $\perp\,\perp$. Two of these yield negation (which alone is not truth-functionally complete); the other ways yield the truth functions for the Sheffer stroke and the dagger.

Exercises in chapter 4

4.4 (4) (1)–(3) P (4) RAA, 2, 3 (5) PC, 4 (6) P (7) PC, 6 (8) DIL, 1, 5, 7
(6) (10) RAA, 7, 9 (13) RAA, 8, 12

4.6 *True* (*Why?*)

4.10 (15)

1	(1)	$\neg B \rightarrow \neg A$
2	(2)	A
3	(3)	$\neg B$
1, 3	(4)	$\neg A$
1, 2	(5)	B
1	(6)	$A \rightarrow B$
	(8)	$(\neg B \rightarrow \neg A) \rightarrow (A \rightarrow B)$

4.10 (17)
Peirce's Law

1	(1)	$(A \rightarrow B) \rightarrow A$
2	(2)	$\neg A$
3	(3)	A
2, 3	(4)	B
2	(5)	$A \rightarrow B$
1, 2	(6)	A

		1	(7)	A
			(8)	$((A \to B) \to A) \to A$
4.15	(12)	1	(1)	$\neg ((A \to B) \vee (B \to C))$
		2	(2)	B
		2	(3)	$A \to B$
		2	(4)	$(A \to B) \vee (B \to C)$
		1, 2	(5)	C
		1	(6)	$B \to C$
		1	(7)	$(A \to B) \vee (B \to C)$
			(8)	$(A \to B) \vee (B \to C)$
4.16	(3)	1	(1)	$A \vee A$
right-to-left		2	(2)	A
			(3)	$A \to A$
		1	(4)	A

4.22 Suppose that A and B are deductively equivalent. This means that each is deducible from the other. By part (1), each is a consequence of the other. So A and B are equivalent (by proposition 1.9).

4.25	(8)	1	(1)	$\neg (A \to B)$
		2	(2)	B
		2	(3)	$A \to B$
		1	(4)	$\neg B$
			(5)	$\neg (A \to B) \to \neg B$
4.25	(11)	1	(1)	$\neg ((A \to B) \vee (B \to A))$
		2	(2)	A
		2	(3)	$B \to A$
		2	(4)	$(A \to B) \vee (B \to A)$
		1	(5)	$\neg A$
		1	(6)	$A \to B$
		1	(7)	$(A \to B) \vee (B \to A)$

		(8)	$(A \to B) \vee (B \to A)$
4.26 (12)	1	(1)	$A \wedge (B \vee C)$
left-to-right	1	(2)	A
	1	(3)	$B \vee C$
	4	(4)	B
	1, 4	(5)	$A \wedge B$
	1	(6)	$B \to (A \wedge B)$
	7	(7)	C
	1, 7	(8)	$A \wedge C$
	1	(9)	$C \to (A \wedge C)$
	1	(10)	$(A \wedge B) \vee (A \wedge C)$
4.29 (5)	1	(1)	$A \vee (B \wedge \neg B)$
right-to-left	2	(2)	B
		(3)	$B \to B$
		(4)	$\neg (B \wedge \neg B)$
	1	(8)	A
4.29 (10)	1	(1)	$B \leftrightarrow (A \leftrightarrow B)$
right-to-left	1	(2)	$B \leftrightarrow (B \leftrightarrow A)$
	1	(3)	$(B \leftrightarrow B) \leftrightarrow A$
	1	(4)	$(B \leftrightarrow B) \to A$
	5	(5)	B
		(6)	$B \to B$
		(7)	$B \leftrightarrow B$
	1	(8)	A
4.29 (27)	1	(1)	$A \vee (B \to C)$
left-to-right	2	(2)	B
	3	(3)	A
		(4)	$A \to A$
	5	(5)	$B \to C$

2, 5	(6)	C
2	(7)	$(B \to C) \to C$
1, 2	(8)	$A \vee C$
1	(9)	$B \to (A \vee C)$

Note: No REP rules are used above. But compare the solution for right-to-left.

4.29 (27)	1	(1)	$B \to (A \vee C)$
right-to-left	1	(2)	$\neg B \vee (A \vee C)$
	1	(3)	$(\neg B \vee A) \vee C$
	1	(4)	$(A \vee \neg B) \vee C$
	1	(5)	$A \vee (\neg B \vee C)$
	1	(6)	$A \vee (B \to C)$

4.38 A and B are deductively equivalent if and only if each is deducible from the other; if and only if (by proposition 4.18) $A \to B$ and $B \to A$ are theorems; if and only if $A \leftrightarrow B$ is a theorem (via lemma 1 and deductions using BI and BE).

4.51 *True* a, f, j, o, v *False* g, i, n, t, y (*Why*—and what about the others?)

Exercises in chapter 6

6.2 (3) and (5) If cows moo then cows moo (Huh? *Why?*)

6.3 (4) $\exists x \neg Fx \to \forall x \neg Fx$ *or* $\exists x \neg Fx \to \neg \exists x Fx$

6.4 (5) Some oysters are silent, Nothing that is silent is amusing / Some oysters are not amusing

6.7 (3) $\forall x(Px \to \neg Sx)$ *or* $\neg \exists x(Px \wedge Sx)$

(11) $\forall x((Px \wedge \neg Fx) \to \neg Lx)$ *or* $\forall x(Px \to (\neg Lx \vee Fx))$ *or* $\forall x(Px \to (Lx \to Fx))$ *or* $\neg \exists x(Px \wedge (Lx \wedge \neg Fx))$

(14) $\forall x \forall y((Px \wedge (Ry \wedge Fxy)) \to Mx)$ *or* $\forall x((Px \wedge \exists y(Ry \wedge Fxy)) \to Mx)$ *or* $\forall x(Px \to \forall y((Ry \wedge Fxy) \to Mx))$ *or* $\neg \exists x((Px \wedge \exists y(Ry \wedge Fxy)) \wedge \neg Mx)$ *or* . . .

6.9 (9) Everyone is loved by everyone

(15) There is at least one person whom everyone loves

(23) No one loves anyone

6.10 (5) No one loves anyone

6.11 (7)

$\forall x(\neg Tx \lor Ex)$	*or*	$\forall x(\neg Ex \to \neg Tx)$	*or*	...
$\forall x(Hx \to \neg Rx)$	*or*	$\neg \exists x(Hx \land Rx)$	*or*	...
$\underline{\forall x(\neg Rx \to \neg Ex)}$	*or*	$\underline{\neg \exists x(Ex \land \neg Rx)}$	*or*	...
$\forall x(Hx \to \neg Tx)$	*or*	$\forall x(Tx \to \neg Hx)$	*or*	...

6.12 (a) 3 (b) 2 (c) 5 (d) 4 (e) 1

6.14 (3) See exercise 7.15, part (3).

6.16 $(R\underline{x}a \lor \exists xRx\underline{y}) \leftrightarrow \neg \forall y(R\underline{x}y \to Rya)$ Underlining indicates free occurrences; others are bound. (Note that a is not a variable.)

6.18 (Rbb)(b/a) (Rba)(b/a) (Rab)(b/a) (Raa)(b/a)

Exercises in chapter 7

7.1 (10) domain = [your choice] $\mathcal{M}(F) = \varnothing$ $\mathcal{M}(G)$ = domain $\mathcal{M}(H)$ = domain

7.3 (8) in exercise 6.9 domain = $\mathbf{N} = \{0, 1, 2, \ldots\}$ $\mathcal{M}(R) = \leq = \{\langle m, n\rangle : m \leq n\}$

7.4 (g) domain = $\mathbf{N}$ $\mathcal{M}(R) = < = \{\langle m, n\rangle : m < n\}$
(l) domain = $\mathbf{N}$ $\mathcal{M}(R) = \leq$

7.6 *First question* yes (*Why?*) *Second question* no (*Why?*)

7.7 (3) domain = $\mathbf{N}$ $\mathcal{M}(F) = \{0, 2, 4, \ldots\}$ $\mathcal{M}(G) = \{1, 3, 5, \ldots\}$
(12) domain = $\mathbf{N}$ $\mathcal{M}(a) = 0$ $\mathcal{M}(F) = \{1, 3, 5, \ldots\}$
(20) domain = $\mathbf{N}$ $\mathcal{M}(R) = <$

7.8 (6) domain = [your choice] $\mathcal{M}(R) = \varnothing$ (See also exercise 8.22, part (14).)
(12) domain = $\mathbf{N}$ $\mathcal{M}(R) = <$

7.9 (4) domain = $\{0, 1\}$ $\mathcal{M}(a_n) = 0$ for each name a_n $\mathcal{M}(F) = \{0\}$

7.10 *Second question* infinitely many (*Why?*)

7.23 *True* c, d, j *False* b, f, m (*Why*—and what about the others?)

Exercises in chapter 8

8.2 (2) (1)–(2) P (3) UI, 1 (4) TT, 2, 3 (5) EG, 4

8.3 (2) (1)–(3) P (4) UI, 1 (5) TT, 3, 4 (6) EG, 5 (7) E, 2, 6

8.7	(7)	1	(1)	$\forall x(Ex \vee \neg Tx)$
		2	(2)	$\forall x(Hx \rightarrow \neg Rx)$
		3	(3)	$\forall x(\neg Rx \rightarrow \neg Ex)$
		1	(4)	$Ea \vee \neg Ta$
		2	(5)	$Ha \rightarrow \neg Ra$
		3	(6)	$\neg Ra \rightarrow \neg Ea$
		1, 2, 3	(7)	$Ha \rightarrow \neg Ta$
		1, 2, 3	(8)	$\forall x(Hx \rightarrow \neg Tx)$

8.8	(5)	1	(1)	$\exists x(Fx \wedge Gx)$
		2	(2)	$\forall x(Gx \rightarrow \neg Hx)$
		3	(3)	$Fa \wedge Ga$
		2	(4)	$Ga \rightarrow \neg Ha$
		2, 3	(5)	$Fa \wedge \neg Ha$
		2, 3	(6)	$\exists x(Fx \wedge \neg Hx)$
		1, 2	(7)	$\exists x(Fx \wedge \neg Hx)$

8.9	(17)	1	(1)	$\exists x \exists y Rxy$
left-to-right		2	(2)	$\exists y Ray$
		3	(3)	Rab
		3	(4)	$\exists x Rxb$
		3	(5)	$\exists y \exists x Rxy$
		2	(6)	$\exists y \exists x Rxy$
		1	(7)	$\exists y \exists x Rxy$

Note: Lines (2), (4), and (6) can be omitted if *-rules are used.

8.9	(33)	1	(1)	Fa
		1	(2)	$\exists x Fx$
			(3)	$Fa \rightarrow \exists x Fx$
			(4)	$\forall x(Fx \rightarrow \exists x Fx)$

8.9	(36)	1	(1)	$\neg \exists x(\exists x Fx \rightarrow Fx)$
		1	(2)	$\forall x \neg (\exists x Fx \rightarrow Fx)$

		1	(3)	$\neg(\exists xFx \to Fa)$
		1	(4)	$\exists xFx$
		1	(5)	$\neg Fa$
		1	(6)	$\forall x \neg Fx$
		1	(7)	$\neg \exists xFx$
			(8)	$\exists x(\exists xFx \to Fx)$

Note: This theorem cannot be deduced from the empty set by EG.

8.9	(42)		(1)	$Fa \to Fa$
			(2)	$\exists y(Fy \to Fa)$
			(3)	$\forall x\exists y(Fy \to Fx)$

8.9	(52)	1	(1)	$\neg \forall x \neg Fx$
right-to-left		2	(2)	$\neg \exists xFx$
		3	(3)	Fa
		3	(4)	$\exists xFx$
		2	(5)	$\neg Fa$
		2	(6)	$\forall x \neg Fx$
		1	(7)	$\exists xFx$

8.10		1	(1)	$\forall x(Fx \leftrightarrow P)$
		2	(2)	$\exists xFx$
		3	(3)	Fb
		1	(4)	$Fa \leftrightarrow P$
		1	(5)	$Fb \leftrightarrow P$
		1, 3	(6)	Fa
		1, 2	(7)	Fa
		1, 2	(8)	$\forall xFx$
		1	(9)	$\exists xFx \to \forall xFx$
		10	(10)	$\forall xFx$
		10	(11)	Fa
		10	(12)	$\exists xFx$
			(13)	$\forall xFx \to \exists xFx$

		1	(14)	$\exists x Fx \leftrightarrow \forall x Fx$

8.13		1	(1)	$\forall x Fx$
		2	(2)	$\neg Fa$
		2	(3)	$\exists x \neg Fx$
		2	(4)	$\neg \forall x Fx$
		1	(5)	Fa

8.19 *True* d, f, i *False* l, m, s (*Why*—and what about the others?)

8.22	(7)	1	(1)	$\forall x Rxx$
		2	(2)	$\forall x \forall y \forall z((Rxy \wedge Rxz) \rightarrow Ryz)$
		1	(3)	Raa
		2	(4)	$(Rab \wedge Raa) \rightarrow Rba$
		1, 2	(5)	$Rab \rightarrow Rba$
		1, 2	(6)	$\forall x \forall y(Rxy \rightarrow Ryx)$

8.22	(9)	1	(1)	$\forall x \exists y Rxy$
		2	(2)	$\forall x \forall y(Rxy \rightarrow Ryx)$
		3	(3)	$\forall x \forall y \forall z((Rxy \wedge Rxz) \rightarrow Ryz)$
		1	(4)	$\exists y Ray$
		5	(5)	Rab
		2	(6)	$Rab \rightarrow Rba$
		3	(7)	$(Rba \wedge Rba) \rightarrow Raa$
		2, 3, 5	(8)	Raa
		1, 2, 3	(9)	Raa
		1, 2, 3	(10)	$\forall x Rxx$

Exercises in chapter 10

10.6 *First three sentences* The present King of France is bald

10.11 (1) Mab *or* $\forall x(x = a \rightarrow \text{Mxb})$ *or* $\exists x(x = a \wedge \text{Mxb})$
(3) $\forall x(\text{Mxb} \leftrightarrow x = a)$ *or* $\text{Mab} \wedge \forall x(\text{Mxb} \rightarrow x = a)$ *or* . . .
(4) $\exists x(\text{Max} \wedge (\text{Fxb} \vee \text{Mxb}))$
(11) $\exists x \exists y(x \neq y \wedge \text{Sxa} \wedge \text{Sya} \wedge \forall z(\text{Sza} \rightarrow (x = z \vee y = z)))$

10.12 (b) 7 (c) 2 (e) 1 (h) 5

10.15 (7) and (10) $\leq$ in the set of natural numbers
(4), (7), and (9) $<$ in the set of real numbers
(7) and (12) $= 1 +$ in the set of natural numbers

10.20 (4)

left-to-right

1	(1)	$\exists x \forall y\ x = y$
2	(2)	$\forall y\ c = y$
2	(3)	$c = a$
2	(4)	$c = b$
2	(5)	$a = b$
1	(6)	$a = b$

10.20 (7)

right-to-left

1	(1)	$\exists x(x = a \wedge Fx)$
2	(2)	$b = a \wedge Fb$
2	(3)	$b = a$
2	(4)	Fb
2	(5)	Fa
1	(6)	Fa

10.20 (13)

1	(1)	$\exists x \forall y(Fy \leftrightarrow x = y)$
2	(2)	$\forall y(Fy \leftrightarrow a = y)$
3	(3)	$\exists x(Fx \wedge Gx)$
4	(4)	$Fb \wedge Gb$
5	(5)	Fc
2	(6)	$Fb \leftrightarrow a = b$
2	(7)	$Fc \leftrightarrow a = c$
2, 4	(8)	$a = b$
2, 5	(9)	$a = c$

	2, 4, 5	(10)	$b = c$
	4	(11)	Gb
	2, 4, 5	(12)	Gc
	2, 4	(13)	$Fc \rightarrow Gc$
	2, 4	(14)	$\forall x(Fx \rightarrow Gx)$
	2, 3	(15)	$\forall x(Fx \rightarrow Gx)$
	2	(16)	$\exists x(Fx \wedge Gx) \rightarrow \forall x(Fx \rightarrow Gx)$
	17	(17)	$\forall x(Fx \rightarrow Gx)$
	17	(18)	$Fa \rightarrow Ga$
	2	(19)	$Fa \leftrightarrow a = a$
		(20)	$a = a$
	2, 17	(21)	$Fa \wedge Ga$
	2, 17	(22)	$\exists x(Fx \wedge Gx)$
	2	(23)	$\forall x(Fx \rightarrow Gx) \rightarrow \exists x(Fx \wedge Gx)$
	2	(24)	$\exists x(Fx \wedge Gx) \leftrightarrow \forall x(Fx \rightarrow Gx)$
	1	(25)	$\exists x(Fx \wedge Gx) \leftrightarrow \forall x(Fx \rightarrow Gx)$

10.20 (17)	1	(1)	$Fa \wedge \forall x(x \neq a \rightarrow Fx)$
right-to-left	2	(2)	$\neg Fb$
	1	(3)	$\forall x(x \neq a \rightarrow Fx)$
	1	(4)	$b \neq a \rightarrow Fb$
	1, 2	(5)	$b = a$
	1	(6)	Fa
	1, 2	(7)	Fb
	1	(8)	Fb
	1	(9)	$\forall xFx$

10.23 (2)	1	(1)	$\forall x\forall y(Rxy \rightarrow x = y)$
for (f)	2	(2)	Rab
	1	(3)	$Rab \rightarrow a = b$
	1, 2	(4)	$a = b$
	1, 2	(5)	Raa

	1, 2	(6)	Rba
	1	(7)	$Rab \rightarrow Rba$
	1	(8)	$\forall x \forall y(Rxy \rightarrow Ryx)$

10.23 (8)	1	(1)	$\forall x \exists y Rxy$
(e), (q) $\models$ (r)	2	(2)	$\forall x \forall y \forall z((Rxy \wedge Rxz) \rightarrow y = z)$
	1	(3)	$\exists y Ray$
	4	(4)	Rab
	2	(5)	$(Rab \wedge Rac) \rightarrow b = c$
	2, 4	(6)	$Rac \rightarrow b = c$
	7	(7)	$b = c$
	4, 7	(8)	Rac
	4	(9)	$b = c \rightarrow Rac$
	2, 4	(10)	$Rac \leftrightarrow b = c$
	2, 4	(11)	$\forall z(Raz \leftrightarrow b = z)$
	2, 4	(12)	$\exists y \forall z(Raz \leftrightarrow y = z)$
	1, 2	(13)	$\exists y \forall z(Raz \leftrightarrow y = z)$
	1, 2	(14)	$\forall x \exists y \forall z(Rxz \leftrightarrow y = z)$

10.25 *True* g, h *False* e (*Why*—and what about the others?)

INDEXES

INDEX OF SYMBOLS

INDEX OF RULES OF INFERENCE

INDEX OF SUBJECTS

UNIVERSITY OF CALGARY

BOOKSTORE

Custom Course Materials

For any questions regarding the production of this coursepack please call our bookstore (403) 220-5937 or email bkstore@ucalgary.ca

University of Calgary Bookstore
2500 University Drive NW
Calgary, Alta
T2N 1N4

2014-11-18

Prepared by

Brian Chellas

$39.00

1018 6268