forallχ

Leeds PHIL1250 2016-17

P.D. Magnus
University at Albany, State University of New York

Modified for the University of Leeds PHIL1250 module by:
E.J.R. Elliott
University of Leeds

P. D. Magnus would like to thank the people who made this project possible. Notable among these are Cristyn Magnus, who read many early drafts; Aaron Schiller, who was an early adopter and provided considerable, helpful feedback; and Bin Kang, Craig Erb, Nathan Carter, Wes McMichael, Selva Samuel, Dave Krueger, Brandon Lee, Toan Tran, and the students of Introduction to Logic, who detected various errors in previous versions of the book.

E. J. R. Elliott would like to thank P.D. Magnus for his generosity in making `forall`χ available to everyone. He would also like to thank Jessica Isserow for discussion and helpful comments on drafts.

The style for typesetting proofs is based on fitch.sty (v0.4) by Peter Selinger, University of Ottawa.

How to Read This Book

Chapters marked '[$\mathcal{R}eq.$]' are required reading for the module; each one corresponds to a lecture. All other chapters are supplementary—you may read them if you like, and if you want to deepen you understanding of logic, but you will not be required to do so. Nor will you be examined on anything contained within the supplementary chapters.

Important concepts are **bolded** when they're first introduced. Furthermore, key ideas are often summarised in boxes, like so:

A KEY IDEA is summarised in here. It is strongly advised that you pay attention to what's written inside boxes like this one.

Finally, each chapter ends with a series of practice problems. The answers to some of the problems are provided at the end of the book in appendix A; those specific problems that are solved in the appendix are marked with a '$\star$'. In some cases, more than one solution to the same problem is possible; where this is the case, the solutions in the appendix only represent one possible way of proceeding.

Contents

Part I

Propositional Logic

Chapter 1

[$\mathcal{Req}$.] What is Logic?

Central to the study of logic is the evaluation of arguments. In everyday language, we sometimes use the word 'argument' to refer to heated disagreements between two or more people. If you and a friend have an argument in this sense, things are not going well between the two of you. However, when studying logic, we are not interested in the teeth-gnashing, hair-pulling kind of argument. Rather, we are interested in logical arguments, which are intended to provide audiences with reasons for accepting some conclusion.

This chapter starts with a brief review of some of the basic concepts involved in the evaluation of arguments. In particular, it is important to begin the study of formal logic with a clear understanding of what arguments are, and of what it means for an argument to be sound and valid.

Later, we will start to translate arguments from English into a simple formal language. In doing so, we will ultimately want to characterize a notion of formal validity. An argument is formally valid if it has a valid form—that is, if it has a logical form which guarantees its validity.

1.1 Arguments, Sentences, and Propositions

Here is an example of a very simple argument:

P1 No snakes have fur.
P2 All pythons are snakes.

C No pythons have fur.

P1 and P2 are the **premises** of the argument, and C is the **conclusion**. The line between the premises and the conclusion is called an **inference bar**; it indicates that the premises which come before it are supposed to support (give reasons for accepting) the conclusion which follows. If the argument is a good one, and you accept the premises as true, then the argument should give you reasons to accept the conclusion as well.

Very generally, we can define an **argument** as a sequence of two or more propositions, exactly one of which is the conclusion and the rest of which are premises. The final proposition in the sequence is always the conclusion. (I'll explain in a moment what a proposition is.)

> An ARGUMENT is a sequence of two or more propositions, exactly one of which—the final one in the sequence—is the conclusion, and the rest of which are premises.

On this definition, an argument only ever has one conclusion. However, the conclusion of one argument can also be a premise in another argument; hence it is possible to chain together arguments into an *extended argument*. In this text, though, we will restrict our attention to arguments with only one conclusion.

What's a **proposition**? That's a matter of much contemporary debate amongst philosophers. For our purposes, it's enough if we say that a proposition is the factual content of a declarative sentence, as uttered on a given occasion; i.e., it is the *meaning* of an uttered declarative sentence.

A **declarative sentence** is the kind of sentence which can be either true or false. In the jargon, we can say that declarative sentences are **truth-apt**. For a declarative sentence, it should make sense to ask: *is this sentence true, or false?* (It would be just as good to say that declarative sentences are falsity-apt, but that wouldn't flow off the tongue quite as nicely.)

For example, 'The moon is round' is a declarative sentence. It expresses the proposition *The moon is round*. The sentence is true because things really are as the sentence says they are; it would be false otherwise. Other declarative sentences include:

- ▷ 'The moon is made of cheese.'
- ▷ 'An asteroid wiped out the dinosaurs.'
- ▷ 'Dogs and cats are not usually purple.'
- ▷ 'All humans are either mortal or immortal.'
- ▷ 'If I am well-fed, then I am happy.'

Note that one and the same proposition can be expressed by different declarative sentences, especially if those sentences are in different languages. For example, the English sentence 'Snow is white' expresses the same proposition as the French sentence 'La neige est blanche': that *Snow is white*. Furthermore, some special sentences (involving words like 'I', 'you', 'here', 'now', 'that', 'there') can express different propositions depending on the context in which they are uttered. For example, if Alice were to say 'I am happy', she would express a different proposition—viz., that *Alice is happy*—than would be expressed if Bob were to say 'I am happy'—viz., that *Bob is happy*. So, propositions and declarative sentences are very closely connected to one another, but they are not the same thing.

Because individual propositions are expressed by declarative sentences, arguments are expressed by sequences of declarative sentences. All of the ordinary language arguments you'll find in this textbook is expressed using a sequence of sentences in English. Consequently, we will spend a lot of time talking about the meanings of English sentences in the chapters that follow.

Like the declarative sentences which express them, propositions can be either true or false. A declarative sentence is true just in case the proposition it expresses is true, and false if the proposition it expresses is false. We will say that 'true' and 'false' are two possible **truth-values**.

In the kinds of logic that we will be considering, we will assume that every proposition must be either true or false, and not both. That is, we will assume that there are exactly two truth-values, and every proposition has one of them, and no more than one. Some 'non-classical' logics exist which do not make these assumptions, but it is helpful to begin with just two mutually exclusive truth-values.

> A PROPOSITION is the factual content of a declarative (truth-apt) sentence, as uttered on a given occasion. Propositions can be either true or false, and not both or neither.

There are many types of sentence in English which are not declarative:

Questions: Questions like 'Are you sleepy yet?' can be called *interrogative sentences*. It does not make sense to say that a question is either true or false; they are not truth-apt.

Imperatives: Commands are often phrased as imperatives like 'Wake up!', 'Sit up straight', and so on. In a grammar class, these would count as *imperative sentences*. Although it might be good for you to sit up straight or it might not be, the command itself is neither true nor false.

Exclamations: 'Ouch!' and 'Boo!' are sometimes called *exclamatory sentences*. They are neither true nor false. It would not make sense to response 'Yes, that's true' to someone who just yelled out 'Ouch!'

These kinds of sentences do not express propositions, and so they are not fit to express the premises and/or conclusions in any logical argument. Hence, the simple formal languages that we will develop in this book will not deal with them.

Nevertheless, our definition of an argument is very, very general. Consider this:

P1 There is coffee in the coffee pot.
P2 There is a dragon playing bassoon in the attic.

C Salvador Dali was a poker player.

It may seem odd to call this an argument. Certainly, you wouldn't find anyone trying to assert that P1 and P2 give reasons to believe the conclusion. But that is only because it's an especially terrible argument. The two premises have nothing at all to do with the conclusion. However, given our definition, it counts as an argument. So does the following:

P1 There are dogs.
P2 If there are dogs then there are animals.
P3 If there are animals then there are things.

C There are things.

This argument is much better. The goal for the rest of this chapter is to characterize what makes the former argument bad, and the latter argument good.

1.2 Validity and Soundness

Consider the following argument:

P1 It is raining heavily.
P2 If you do not take an umbrella, you will get soaked.

C You should take an umbrella.

If premise P1 is false—if it is sunny outside—then the argument gives you no reason to carry an umbrella. And, even if it is raining outside, you might not need an umbrella. You might wear a rain pancho or keep to covered walkways to keep yourself dry. In these cases, premise P2 would be false: you can forego the umbrella without getting soaked.

Suppose now that both of the premises are true, and that you do not own a rain pancho or any other wet-weather gear. You need to go places where there are no covered walkways, so you will very likely get soaked. Now does the argument show you that you should take an umbrella? Not necessarily. Perhaps you enjoy walking in the rain, and you would like to get soaked. In that case, even if the premises were both true, the conclusion would *still* be false.

For any argument, there are two main ways that it could be bad that we will be interested in. First, one or more of the premises might be false. An argument gives you a reason to accept its conclusion only if you accept its premises, and if you reject its premises then you may not find the argument very convincing at all. Second, the premises may simply fail to support the conclusion. Regardless of whether you accept the premises or not, an argument which is flawed in the second way won't give you very good reasons to accept its conclusion. The umbrella argument is, in the cases we've described, flawed in both of these ways.

Consider another example:

P1 Jane is reading a logic textbook.
P2 *Almost all* people who read logic textbooks are logic students.

C Jane a logic student.

This is *not* a terrible argument. The two premises do seem to provide strong reasons for believing the conclusion. However, the premises do not *guarantee* that the conclusion is true—they merely make the conclusion very likely, absent any evidence to the contrary. We can make the argument even stronger by adding another premise: you have no evidence that Jane is not a logic student. But even then, the premises wouldn't guarantee the conclusion. Perhaps Jane is just one of the very few non-students who enjoy reading logic textbooks.

Now compare that argument with the following argument, where this time the two premises *do* guarantee the truth of the conclusion:

P1 Jane is reading a logic textbook.
P2 *All* people who read logic textbooks are logic students.

C Jane a logic student.

The difference between the two arguments is that the latter is **deductively valid** (or just 'valid'). That is, if both of its premises were true, then the conclusion would have to be true as a matter of necessity. To say the same thing in a different way: an argument is valid just in case, necessarily, if the premises are true then the conclusion is true.

Notice that an argument can have true premises and a true conclusion, yet still be invalid. For instance:

P1 London is in England.
P2 Beijing is in China.

C Dogs exist.

This argument is invalid because it is *possible* for P1 and P2 to both be true while C is false. The important point about a valid argument is that *if* we accept each of its premises, then we *cannot* rationally deny its conclusion.

Furthermore, a valid argument may have a false conclusion, if (and only if) at least one of its premises are false. For example, the following argument is valid:

P1 Dogs do not exist.
P2 Either dogs exist or Aristotle was a famous guitarist.

C Aristotle was a famous guitarist.

Accepting the premises of a valid argument commits us to the argument's conclusion, but it's often possible to deny an argument's premises. In this case, P1 is clearly false, as a matter of fact. We'll say that a valid argument is **sound** just in case all of its premises are actually true. Sound arguments always have true conclusions.

> An argument is VALID whenever it is impossible for the premises to be true and the conclusion false. An argument is SOUND just in case it is valid, *and* all of its premises are true.

When studying logic, we are primarily concerned with the property of validity rather than soundness. By focusing on validity, we abstract away from particular matters of fact. This lets us study the special sense in which some sets of premises can guarantee (or *entail*) a specific conclusion.

1.3 Formal languages

Here is a famous valid argument:

P1 All humans are mortal.
P2 All Greeks are human.

C All Greeks are mortal.

Logically, this is an iron-clad argument. The only way you could challenge the conclusion is by denying one of the premises—structurally, the argument is impeccable. What about this next argument?

P1 All hippos are morally inept.
P2 All galaxies are hippos.

C All galaxies are morally inept.

This argument is in one respect much worse than the first argument, because its second premise is obviously false (and its first premise quite questionable). Yet the latter argument is also valid. To see this, notice that both arguments have same basic **logical form**. The medieval logicians referred to the form of such arguments as *Barbara* (I'll explain why in a moment); they go like this:

P1 All H are M.
P2 All G are H.

C All G are M.

In first argument, 'H' stands for *human*, 'M' stands for *mortal*, and 'G' stands for *Greek*. In the second argument, 'H' stands for *hippos*, 'M' stands for *morally inept*, and 'G' stands for *galaxies*. Both arguments have the same basic logical form, and since the first argument is valid, so too is the second. Indeed, *every* argument which has this form is valid. These arguments are **formally valid**: you can tell that they are valid merely by considering their logical form, without even checking what 'H', 'M', and 'G' stand for.

We can now say that **formal logic** is the study of arguments with valid logical form; i.e., the study of which arguments have valid logical form and which arguments lack it, and what the difference between these two kinds of arguments might amount to.

Not every valid argument has a valid form. For example, from the singular premise 'Andrew is a *bachelor*', you can validly infer that 'Andrew is a *male*', because bachelors are (by definition) male. However, the basic form of this argument is:

P1 A is a B.

C A is an M.

There are many invalid arguments of this form. For example, if we suppose that 'Andrew is a *boa constrictor*', we cannot infer that 'Andrew is a *mammal*'. So there are valid arguments which are not formally valid. On the other hand, every argument that has a valid form is a valid argument. So, helpfully, if you know that an argument has valid form, you know that it's valid *simpliciter*.

In order to characterize formal validity, we need a way to characterize the **logical form** of an argument. This, in turn, requires a way of characterising the logical form of the sentences that make up the argument. We do this by translating our arguments into a **formal language**, which uses symbols to represent sentences and the parts thereof.

There are formal languages that work like the symbolization we gave for the two arguments we just looked at earlier in this section. A logic like this was developed by Aristotle in the 4th century BC. Aristotle's logic, with some revisions, was the dominant logic in the western world for more than two millennia. In Aristotelean logic, categories (like *human* and *vegetable*) are replaced with capital letters. Every line of an argument is then represented as having one of four forms, which medieval logicians labelled in this way:

A All As are Bs
E No As are Bs
I Some A is a B
O Some A is not a B

It is then possible to describe valid *syllogisms*, which are simple three-line arguments similar to the two we considered just above. Medieval logicians gave mnemonic names to all of the valid argument forms. The form of our two arguments, for instance, was called *Barbara*—the three As in the name 'Barbara' represent the fact that the two premises and the conclusion are all **A** form sentences.

There are many limitations to Aristotelean logic. One is that it cannot capture a very large range of valid inferences we can make. For instance, the following argument is valid (and sound), but is not an Aristotelean syllogism:

P1 Either $1 + 1 = 2$, or $1 + 1 = 3$.
P2 If $1 + 1 = 3$, then $3 - 1 = 1$.
P3 It's not the case that $3 - 1 = 1$.

C $1 + 1 = 2$.

More generally, there are many sentences—and valid inferences involving them—which do not have a logical form captured by **A**, **E**, **I**, and **O**. This makes it impossible to apply Aristotelian logic to arguments involving such sentences. Luckily for us, Aristotelean logic has been superceded.

The remainder of Part I will develop a simple formal language, which we will label $\mathcal{L}_S$. The subscript 'S' stands for 'sentential,' as $\mathcal{L}_S$ is the kind of language used for **sentential logic**. (Sentential logic is sometimes also called *propositional logic*, *statement logic*, and *truth-functional logic*.) In $\mathcal{L}_S$, we will use letters to represent atomic declarative sentences, and other special symbols to represent words like 'and' and 'or', which are used to link atomic sentences together to create more complicated sentences. This language will also be useful when we come to mathematical probabilities in Part III.

In Part II, we will begin to develop a more complicated language, $\mathcal{L}_P$. The subscript 'P' stands for 'predicate,' and $\mathcal{L}_P$ is the language used when studying **predicate logic** (sometimes also called *quantificational logic*, *lower predicate calculus*, or *first order logic*). In $\mathcal{L}_P$, we formalise the *internal* logical structure of a wide range of sentences, which will let us represent the valid form of a very large range of different arguments.

1.4 Atomic and Non-Atomic Sentences

In $\mathcal{L}_S$, italicised capital letters (e.g., A, B, C, …, Z) are used to represent whole declarative sentences. We will call these **sentence letters**.

Considered purely as a symbol, the sentence letter A could stand for any declarative sentence whatsoever. So, when translating from English into $\mathcal{L}_S$ (and *vice versa*), it is important to provide a **symbolization key**. The key provides the intended reference for each sentence letter used in the symbolization.

For example, consider this argument:

P1 Jack went for a swim.
P2 If Jack went for a swim, then Jack went to the river.

C Jack went to the river.

This is clearly a valid argument in English. In symbolizing it, we want to preserve what it is about the *structure* of the argument that makes it valid. What happens if we replace each step in the argument with a letter? In that case, our symbolization key might look like this:

A : Jack went for a swim.
B : If Jack went for a swim, then Jack went to the river.
C : Jack went to the river.

We would then symbolize the argument in this way:

P1 A
P2 B

C C

This isn't what we want: our symbolisation of the argument has not faithfully captured what it is about the *form* of the argument which makes it valid.

Now, the key thing to note about the argument is that the second premise is not merely *any* sentence, logically independent of the other steps of the argument. Rather, the second premise contains the first premise and the conclusion *as parts*. So our symbolization key for the argument only needs to include meanings for A and C, and we can construct the second premise out of those two parts. So we could instead symbolize the argument this way:

P1 A
P2 If A, then C

C C

This is a much better representation of the structure of the argument that makes it valid. It still makes use of the English expression 'If... then...', which is a kind of connective expression: it can be used to connect different sentences together to form a new sentence. Although we will ultimately want to replace *all* of the English expressions in P2 with logical notation, this partial symbolization is a good start for highlighting what it is about the structure of the argument that makes it valid.

In general, then, it is helpful to distinguish between simple (or **atomic**) sentences and complex (or **non-atomic**) sentences:

> An ATOMIC SENTENCE is a sentence which does not contain any other sentences as proper parts. Any non-atomic sentence contains at least one atomic sentence as a proper part.

Atomic sentences are so-called because they are the basic building blocks of our language $\mathcal{L}_S$. For example, the sentence 'Roses are red' is atomic, whereas 'Roses are red and violets are blue' is not—the latter has two atomic sentences ('Roses are red' and 'Violets are blue') as parts, which have been connected together using the word 'and.' The sentence 'It is not the case that roses are red' is also non-atomic: it contains 'Roses are red' as a proper part.

Connectives are used to build non-atomic sentences out of other (atomic and non-atomic) sentences. For example, suppose we have three atomic sentences, symbolized using the sentence letters 'A', 'B', and 'C'. We can then build the non-atomic sentences like 'A and B', which we can use to construct even more complex sentences like 'If C, then A and B'. (We can keep going and make sentences as complex as we like.) There are five connectives that are symbolised in $\mathcal{L}_S$. The table that follows summarizes them. Chapter 2 focuses on the first three connectives, $\neg$, $\wedge$ and $\vee$, while Chapter 3 focuses on the conditionals, $\rightarrow$ and $\leftrightarrow$.

symbol	what it is called	what it means
$\neg$	negation	'It is not the case that…'
$\wedge$	conjunction	'Both… and …'
$\vee$	disjunction	'Either… or …'
$\rightarrow$	conditional	'If … then …'
$\leftrightarrow$	biconditional	'… if and only if …'

To best represent the logical form of the premises and conclusions in an argument, we generally do well to let sentence letters stand for atomic sentences. We then want to represent non-atomic sentences as being composed out of atomic sentences, using connectives.

There are only twenty-six letters of the alphabet, but there is no limit to the number of atomic sentences we might want to consider. So there are many more atomic sentences than we have letters for. To deal with this, we can use the same letter to symbolize different atomic sentences by adding subscripts. For example, we could have a symbolization key that looks like this:

A_1 : The apple is under the armoire.
A_2 : Arguments in $\mathcal{L}_S$ always contain atomic sentences.
A_3 : Adam Ant is taking an airplane from Anchorage to Albany.
$\vdots$
A_{294} : Alliteration angers otherwise affable astronauts.

Keep in mind that each of these is a different 'sentence letter'. When there are subscripts in the symbolization key, it is important to keep track of them.

Practice Exercises

⋆ **Part A**
Which of the following sentences are declarative?

1. England is larger than China.
2. Greenland is south of Jerusalem.
3. Is New Jersey east of Wisconsin?
4. The atomic number of helium is 2.
5. The atomic number of helium is $3\pi^2$.
6. I hate seafood.
7. Blech—seafood!
8. Please hurry up.
9. Jack is over there.
10. Go over there.

⋆ **Part B**
Which of the following are true?

1. A valid argument can have one false premise and one true premise.
2. A valid argument can have a false conclusion.
3. A valid argument cannot have a false conclusion and all true premises.
4. Every valid argument is sound.
5. Every valid argument with true premises has a true conclusion.
6. Every unsound argument has a false conclusion.
7. Every argument with a false conclusion and only true premises is invalid.
8. There are sound arguments that are invalid.
9. There are arguments with a valid form that are invalid.
10. There are arguments with an invalid form that are valid.

Chapter 2

[$\mathcal{Req.}$] Negation, Conjunction, and Disjunction

This chapter introduces another fragment of the formal language $\mathcal{L}_S$. Three logical connectives and their symbolizations are introduced: negation ($\neg$), conjunction ($\wedge$), and disjunction ($\vee$).

2.1 Negation

Consider how we might symbolize the following sentences of English:

1. Mary is in Barcelona.
2. Mary is not in Barcelona.
3. Mary is somewhere other than Barcelona.

In order to symbolize sentence 1, we will need just one sentence letter. This is because 1 is an atomic sentence. So we can start with the following symbolization key:

B : Mary is in Barcelona.

(Note that here we are giving 'B' a different interpretation than we did in the previous chapter. The symbolization key only specifies what 'B' means *in a* *specific context*. It is vital that we continue to use this meaning of 'B' so long as we are talking about Mary and Barcelona. Later, when we are symbolizing different arguments and sentences, we can write a new symbolization key and let 'B' mean something else.)

Sentence 2, on the other hand, is not atomic: it contains 'B' as a proper part. Indeed, 2 just says the same thing as '*It is not the case that* Mary is in Barcelona'—that is, '*It is not the case that* B.' In order to fully symbolize 2, we will need a symbol for logical **negation**. We will use $\neg$. Now we can symbolize 2 as '$\neg B$.'

Sentence 3 does not contain the word 'not' anywhere. However, it clearly *means* the same thing as 2. That is, both 2 and 3 express the very same proposition. As such, because we symbolized 2 as '$\neg B$,' we should also symbolize 3 as '$\neg B$.' Generally speaking:

> A sentence can be symbolized as '$\neg\mathcal{A}$' if it means the same thing in English as 'It is not the case that $\mathcal{A}$.'

(Here, the fancy symbol $\mathcal{A}$ is *not* a sentence letter. Rather, it is a *variable* which can be used to refer to any atomic or non-atomic sentence symbolized in $\mathcal{L}_S$. So, $\mathcal{A}$ might refer to an atomic sentence 'A,' 'B,' ..., or to non-atomic sentences like '$\neg B$,' or any of the other more complex symbolizations that we will be considering below.)

Now consider these further examples:

4. The widget can be replaced.
5. The widget is irreplaceable.
6. The widget is not irreplaceable.

We can start with this symbolization key:

R : The widget is replaceable.

In this case, sentence 4 can be symbolized as just 'R,' as 'can be replaced' and 'is replaceable' are just two different ways of saying the same thing.

What about sentence 5? Saying the widget is irreplaceable means that it is not the case that the widget is replaceable. So, sentence 5 expresses the negation of 4, and since we've already symbolized 4 using 'R,' we should symbolize 5 using the negation symbol to get '$\neg R$.' To do otherwise would be to neglect the important logical relationship that holds between 4 and 5.

Now, sentence 6 is a little more tricky. It can be paraphrased as 'It is not the case that the widget is irreplaceable.' Since we have symbolized sentence 5 as '$\neg R$', to symbolize 6 we should use negation twice and represent it as '$\neg\neg R$'. In other words, the sentence means the same thing as '*It is not the case that* [*it is not the case that* [the widget is replaceable]].' Where it's helpful, here and in what follows I will sometimes use square brackets to help indicate the underlying structure of an English sentence.

That's enough for the symbolization of negation. Let's now look at an interesting property that negation has. In particular, notice that for any sentence '$\mathcal{A}$', if '$\mathcal{A}$' is true, then '$\neg\mathcal{A}$' is false. And similarly, if '$\mathcal{A}$' is false, then '$\neg\mathcal{A}$' is true. Very simply, putting a $\neg$ out the front of any sentence '$\mathcal{A}$' gives us a new sentence with the opposite truth-value. This means that the truth-value of '$\neg\mathcal{A}$', whatever it may be, depends *entirely* on the truth-value of '$\mathcal{A}$': the former is a *function of* the latter.

Using T for *true* and F for *false*, we can summarize these truth-functional properties in what we'll call the **characteristic truth table** for negation:

$\mathcal{A}$	$\neg\,\mathcal{A}$
T	F
F	T

(For reasons that we will get back to in Chapter 4, we will put the Ts and Fs underneath the $\neg$ symbol.) The truth table for negation lays out in an easy to read manner exactly how the truth-value of '$\neg\mathcal{A}$' always depends entirely on the truth-value of '$\mathcal{A}$'.

A brief point of clarification. Unlike the other connectives that we will discuss in this book (like 'and' and 'or'), negation is not really used to *connect* two different sentences together to form a bigger sentence. Nevertheless, we will call it a one-place sentential connective, because it shares interesting properties with the two-place connectives that we will discuss below—namely, it allows us to create non-atomic sentences out of atomic sentences.

2.2 Conjunction

Consider the following three claims:

7. Adam is athletic.
8. Barbara is athletic.
9. Adam is athletic, and Barbara is also athletic.

We will need separate sentence letters for 7 and 8, so we define this symbolization key:

A : Adam is athletic.
B : Barbara is athletic.

Sentence 7 can then be straightforwardly symbolized as 'A', and sentence 8 can be symbolized as 'B'. On the other hand, sentence 9 can be paraphrased as simply 'A and B.' So, in order to fully symbolize 9, we need another symbol for 'and.' We will use $\wedge$, and we will symbolize 'A and B' as '$A \wedge B$'. The logical connective $\wedge$ is called **conjunction**, and 'A' and 'B' are its **conjuncts**.

Notice that we have made no attempt to symbolize the word 'also' in 9. Words like 'both' and 'also' function to draw our attention to the fact that two things are being conjoined. However, they are not doing any further logical work: 9 means exactly the same thing as 'Adam is athletic, and Barbara is athletic.' So, we do not need to represent words like 'both' and 'also' in $\mathcal{L}_S$.

> A declarative sentence can be symbolized as '$\mathcal{A} \wedge \mathcal{B}$' if it means the same thing in English as 'Both $\mathcal{A}$, and $\mathcal{B}$.'

Some more examples:

10. Barbara is athletic and energetic.
11. Barbara and Adam are both athletic.
12. Although Barbara is energetic, she is not athletic.
13. Barbara is athletic, but Adam is more athletic than she is.

Sentence 10 is obviously a conjunction. It says two things about Barbara, so in English it is permissible to refer to Barbara only once. It might be tempting to try this when symbolizing 10: since B means 'Barbara is athletic', one might paraphrase the sentence as 'B and energetic.' But this would be a mistake: the conjuncts of a conjunction must be whole sentences, and 'energetic' by itself is not a sentence. We should instead introduce a new sentence letter C to stand for 'Barbara is energetic', and symbolize sentence 10 as '$B \wedge C$.'

Sentence 11 says one thing about two different subjects. It says of both Barbara and Adam that they are athletic, and in English we use the word 'athletic' only once. When symbolizing in $\mathcal{L}_S$, it is important to realize that the sentence means the same thing as 'Barbara is athletic, and Adam is athletic.' Hence we can symbolize 11 as '$B \wedge A$.'

Sentence 12 is a bit more complicated. The word 'although' sets up a contrast between the first part of the sentence and the second part. Nevertheless, it says both that Barbara is energetic and that she is not athletic. In order to make each of the conjuncts an atomic sentence, we need to replace 'she' with 'Barbara.' So we can paraphrase 12 as '*Both* Barbara is energetic, *and* Barbara is not athletic.' The second conjunct contains a negation, so we paraphrase further: '*Both* [Barbara is energetic] *and* [*it is not the case that* [Barbara is athletic]].' So the overall symbolization of 12 is '$C \wedge \neg B$.'

Sentence 13 contains a similar contrastive structure: the word 'but' is usually used to indicate a contrast between two clauses in a complex sentence. Nevertheless, this contrast is irrelevant for the purposes of translating into $\mathcal{L}_S$, so we can paraphrase 13 as '*Both* Barbara is athletic, *and* Adam is more athletic than Barbara.' (Notice that we once again replace the pronoun 'she' with 'Barbara.') How should we symbolize the second conjunct? We already have the sentence letter 'A' which is about Adam's being athletic, and 'B' which is about Barbara's being athletic, but neither is about one of them being *more* athletic than the other. So we need a new sentence letter. Let D be our symbol for the atomic sentence 'Adam is more athletic than Barbara.' Now 13 gets symbolized as '$B \wedge D$.'

> Declarative sentences that can be paraphrased as '$\mathcal{A}$, but $\mathcal{B}$' or 'Although $\mathcal{A}$, $\mathcal{B}$' can be symbolized as '$\mathcal{A} \wedge \mathcal{B}$.'

One last example:

14. Barbara and Adam lifted the jukebox together.

How should we symbolize sentence 14? You might be tempted to think that it is a conjunction, because of the presence of the word 'and.' However, note that the sentence *doesn't* mean the same thing as 'Barbara lifted the jukebox and Adam lifted the jukebox', which would imply that Barbara and Adam each lifted the jukebox *individually*. Sentence 14 is actually an atomic sentence, rather than a conjunction. It's important to keep this in mind: a declarative sentence only expresses a conjunction *if* it can be paraphrased as 'Both $\mathcal{A}$ and $\mathcal{B}$.' So, sometimes a sentence can use the word 'and' but not be a conjunction.

Like negation, conjunction also has special truth-functional properties. To see this, note that for any two sentences '$\mathcal{A}$' and '$\mathcal{B}$,' '$\mathcal{A} \wedge \mathcal{B}$' is true just in case both '$\mathcal{A}$' and '$\mathcal{B}$' are true. This means that $\wedge$ is a **truth-functional connective**: the truth-value of '$\mathcal{A} \wedge \mathcal{B}$' depends entirely on the truth-values of its two parts, '$\mathcal{A}$' and '$\mathcal{B}$.' So if you know the truth-values of the parts, you can work out the truth-value of any conjunction containing those parts.

We can summarize conjunction's truth-functional character using the characteristic truth table for conjunction:

$\mathscr{A}$	$\mathscr{B}$	$\mathscr{A} \wedge \mathscr{B}$
T	T	T
T	F	F
F	T	F
F	F	F

Note that conjunction is **symmetrical**, in that we can swap the order of the conjuncts without any changes in the truth-value of the conjunction overall. That is, regardless of what '$\mathscr{A}$' and '$\mathscr{B}$' are, '$\mathscr{A} \wedge \mathscr{B}$' is true just in case '$\mathscr{B} \wedge \mathscr{A}$' is true. The order in which the conjuncts appear in the conjunction makes no difference—only the truth-values of the conjuncts makes a difference. This is an important property of conjunction (and several of the other connectives that we will consider), which we will make use of later on in the book.

2.3 Disjunction

Two new sentences to consider:

15. Either Denison will play golf with me, or he will watch movies.
16. Either Denison or Ellery will play golf with me.

We will use this symbolization key:

D : Denison will play golf with me.
E : Ellery will play golf with me.
M : Denison will watch movies.

Sentence 15 has the basic form 'Either D or M.' To fully symbolize this we will want a new symbol for 'or'. For this we'll use $\vee$. Thus sentence 15 becomes '$D \vee M$.' The connective $\vee$is called **disjunction**, and 'D' and 'M' are (in this case) its **disjuncts**.

Sentence 16 is only slightly more complicated. This time, there are two subjects, but the English sentence only gives the verb once. In symbolizing, we can paraphrase it as 'Either Denison will play golf with me, or Ellery will play golf with me.' Hence, it symbolizes as '$D \vee E$.'

A declarative sentence can be symbolized as' $\mathscr{A} \vee \mathscr{B}$' if it means the same thing in English as 'Either $\mathscr{A}$, or $\mathscr{B}$.'

Sometimes in English, the word 'or' can be plausibly understood as excluding the possibility that both disjuncts are true. This is sometimes called an *exclusive or*. For instance, an exclusive or is clearly intended when it says, on a restaurant menu, 'Entrees come with either soup or salad.' You may have soup, or you may have salad—but if you want *both*, then you have to pay extra.

At other times, the word 'or' allows for the possibility that both disjuncts might be true. This is probably the case with 16, above. I might play with Denison, with Ellery, or with both Denison and Ellery. Sentence 16 merely says that I will play with *at least* one of them. This is called an **inclusive or**.

The symbol $\vee$ represents an inclusive or. So '$D \vee E$' is true if 'D' is true, if 'E' is true, or if *both* 'D' and 'E' are true. This means that it is false *only* if both 'D' and 'E' are false.

Disjunction is, therefore, another kind of truth-function: for any two sentences '$\mathcal{A}$' and '$\mathcal{B}$,' if you know the truth-values of '$\mathcal{A}$' and '$\mathcal{B}$' then you can work out the truth value of '$\mathcal{A} \vee \mathcal{B}$.' We can summarize this with the characteristic truth table for disjunction:

$\mathcal{A}$	$\mathcal{B}$	$\mathcal{A} \vee \mathcal{B}$
T	T	T
T	F	T
F	T	T
F	F	F

Like conjunction, disjunction is symmetrical. '$\mathcal{A} \vee \mathcal{B}$' is just another way of saying '$\mathcal{B} \vee \mathcal{A}$,' and they always have the same truth-value.

The following examples are somewhat more complicated:

17. Either you will not have soup, or you will not have salad.
18. You will have neither soup nor salad.
19. You will have soup or salad, and a sweet surprise.
20. You get either soup or salad, but not both.

We will use the following key:

S_1 : You will have soup.
S_2 : You will have salad.
S_3 : You will have a sweet surprise.

Then, sentence 17 can be paraphrased in this way: '*Either* [it is not the case that you get soup], *or* [it is not the case that you get salad].' Translating this requires both disjunction and negation. It becomes '$\neg S_1 \vee \neg S_2$.'

Now, sentence 18 also requires negation. It can be paraphrased as, '*It is not the case that* [either you get soup or you get salad].' To symbolize this we will need some way of indicating that the negation does not only negate the right or the left disjunct, but instead negates the entire disjunction. That is, the **scope** of the $\neg$ is the entire disjunction. In order to faithfully represent the logical form of 18, we can put parentheses around the disjunction: 'It is not the case that $(S_1 \vee S_2)$.' So the proper symbolization of 18 becomes simply '$\neg(S_1 \vee S_2)$.'

Notice that the parentheses are doing very important work here. If we tried to symbolize 18 as just '$\neg S_1 \vee S_2$,' then it could be taken to mean '*Either* [it is not the case that you will have soup], *or* [you will have salad].' In this case we would say that the scope of the negation is limited to 'S_1,' and this is not what we want. The correct placement of parentheses is therefore critically important for converting complex sentences into symbolic form.

The parentheses are also very important for symbolizing sentence 19. Suppose we just wrote it down as '$S_1 \vee S_2 \wedge S_3$.' This would be ambiguous, with the two possible readings being:

(i) [You will have soup or salad], *and* [a sweet surprise].

(ii) *Either* [you will have soup], *or* [you will have salad and a sweet surprise].

If we use parentheses, however, we can disambiguate these two readings. The former becomes '$(S_1 \vee S_2) \wedge S_3$,' while the latter becomes '$S_1 \vee (S_2 \wedge S_3)$.'

Finally, despite its relatively simple expression in English, sentence 20 is surprisingly complex when symbolized in $\mathcal{L}_S$. We can break what it says down into two parts. The first part says that you get one or the other. This is a simple disjunction which we can symbolize as '$S_1 \vee S_2$.' The second part says you don't get both. We can paraphrase this as '*It is not the case that* [you get soup and salad].' Note that the negation is here taking scope over the entire conjunction, so it should be symbolized as '$\neg(S_1 \wedge S_2)$' rather than '$\neg S_1 \wedge S_2$.' (The latter would say that you don't get soup, but you do get salad.)

Now we just need to put the two parts, '$S_1 \vee S_2$' and '$\neg(S_1 \wedge S_2)$,' together. As we saw above, 'but' can usually be symbolized as a conjunction, so we'll use $\wedge$ to join them. Since '$S_1 \vee S_2$' is one whole conjunct in the conjunction, we'll need to put parentheses around it to avoid ambiguity. So, finally, sentence 20 can thus be symbolized as '$(S_1 \vee S_2) \wedge \neg(S_1 \wedge S_2)$.'

Practice Exercises

⋆ Part A
Using the symbolization key given, symbolize each of the following English-language sentences into $\mathcal{L}_S$.

M : Those creatures are men in suits.
C : Those creatures are chimpanzees.
G : Those creatures are gorillas.

1. Those creatures are not men in suits.
2. Those creatures are gorillas, not chimpanzees.
3. Those creatures are men in suits, or they are not.
4. Those creatures are either gorillas or chimpanzees.
5. Those creatures are neither gorillas nor chimpanzees.
6. Although those creatures are not men in suits, they are not gorillas either.
7. Those creatures are gorillas or chimpanzees, but not both.

Part B
Using the symbolization key given, symbolize each English-language sentence into $\mathcal{L}_S$.

E_a : Ava is an electrician.
E_h : Harrison is an electrician.
F_a : Ava is a firefighter.
F_h : Harrison is a firefighter.
S_a : Ava is satisfied with her career.
S_h : Harrison is satisfied with his career.

1. Ava and Harrison are both electricians.
2. Harrison is an unsatisfied electrician.
3. Neither Ava nor Harrison is an electrician.
4. Both Ava and Harrison are electricians, but neither find it satisfying.
5. Ava is satisfied with her career but Harrison is not satisfied with his.
6. It cannot be that Harrison is both an electrician and a firefighter.
7. Harrison and Ava are both firefighters, and neither is an electrician.
8. Although Ava is a firefighter, she is not satisfied with her career.

⋆ Part C
Give a symbolization key and symbolize the following into $\mathcal{L}_S$.

1. Alice and Bob each lifted the couch.
2. Either Alice or Bob lifted the couch, but not both.
3. Alice and Bob together lifted the couch.
4. Although Alice didn't lift the couch, Bob did.
5. Neither Alice nor Bob lifted the couch.

Chapter 3

[$\mathcal{Req}$.] Conditionals

This chapter two more truth-functional connectives to our simple formal language $\mathcal{L}_S$, the material conditional and the biconditional.

At the end of the chapter, we will also give a complete characterization of what it is to be a *sentence* in the language of $\mathcal{L}_S$.

3.1 Material Conditional

For the following two examples, we will use this symbolization key:

R : You cut the red wire.
B : The bomb will explode.

1. If you cut the red wire, then the bomb will explode.
2. The bomb will explode only if you cut the red wire.

Sentence 1 can be partially symbolized as 'If R, then B.' Sentences of this form are called **conditionals**. In $\mathcal{L}_S$, we make use of a special kind of conditional that we will call the **material conditional**, symbolised using an arrow $\to$. Then we can symbolize 1 as '$R \to B$.' The sentence on the left-hand side of the conditional ('R,' in this example) is called the **antecedent**. The sentence on the right-hand side ('B') is called the **consequent**.

Sentence 2 is also a conditional. Since the word 'if' appears in the second half of the sentence, it might be tempting to symbolize this in the same way as that we have symbolized 1. That would be a mistake. Actually, 1 and 2 say quite different things.

The conditional '$R \rightarrow B$' says '*If R, then B.*' In other words, 'R' being true is a **sufficient condition** for the truth of 'B.' Any time 'R' is true, 'B' must be true too—or, the truth of 'R' is "enough for" the truth of 'B'. But '$R \rightarrow B$' does *not* say that your cutting the red wire is the *only* way to make it true that the bomb explodes. For example, even if you don't cut the wire, someone else might cut it, or the bomb might be on a timer and explode anyway regardless of whether the wire is cut. So 1 does not say anything about what to expect if 'R' is *false.*

On the other hand, sentence 2 says that the *only* condition under which the bomb will explode involves your having cut the red wire. That is, *if* the bomb explodes, *then* you must have cut the wire. Any time 'B' is true, 'R' must also be true. So the truth of 'R,' in this case, is a **necessary condition** for the truth of 'B'—that is, the truth of 'R' is required for the truth of 'B.' Given this, sentence 2 should really be symbolized as '$B \rightarrow R$.'

> For any material conditional '$\mathcal{A} \rightarrow \mathcal{B}$': the truth of the antecedent is a SUFFICIENT CONDITION for the truth of the consequent; and the truth of the consequent is a NECESSARY CONDITION for the truth of the antecedent.

It is important to remember that the connective $\rightarrow$says only that, if the antecedent is true, then the consequent is true. It says nothing about the *causal* connection between the two events. Translating 2 as '$B \rightarrow R$' does not mean that the bomb exploding would somehow have *caused* your cutting the wire. Both of 1 and 2 can reasonably be taken to *suggest* that, if you cut the red wire, your cutting the red wire would be the cause of the bomb exploding. What they differ on is the *logical* connection between 'B' and 'R' that they describe. If 2 is true, then an explosion would imply—for those of us safely away from the bomb—that you must have cut the red wire. Without an explosion, 2 tells us nothing. If 1 is true, then an explosion *may* mean that you cut the wire, but not necessarily.

> A declarative sentence can be symbolized as '$\mathcal{A} \rightarrow \mathcal{B}$' if it means the same thing in English as 'If $\mathcal{A}$, then $\mathcal{B}$' (or '$\mathcal{A}$ only if $\mathcal{B}$').

'If $\mathcal{A}$ then $\mathcal{B}$' means that if '$\mathcal{A}$' is true then so is '$\mathcal{B}$.' So we know that if the antecedent '$\mathcal{A}$' is true but the consequent '$\mathcal{B}$' is false, then the whole conditional 'If $\mathcal{A}$ then $\mathcal{B}$' must be false.

What is the truth value of 'If $\mathcal{A}$ then $\mathcal{B}$' under other circumstances? Suppose, for instance, that the *antecedent* happened to be false. In English, the truth of conditionals often depends on what *would* be the case if the antecedent *were* true—even if, as a matter of fact, the antecedent is false. This poses a problem for translating conditionals into $\mathcal{L}_S$. Considered as sentences of $\mathcal{L}_S$, 'R' and 'B' in the above examples have nothing intrinsic to do with each other. In order to consider what the world would be like if 'R' were true, we would need to analyze what 'R' says about the world. Since 'R' is an atomic symbol of $\mathcal{L}_S$, however, there is no further structure to be analyzed. When we use sentence letters, we consider it merely as some sentence that might be true or false.

In order to translate conditionals into $\mathcal{L}_S$, we will not try to capture all the subtleties of the English language 'If… then….' Instead, the symbol $\rightarrow$will be a *material conditional*. This means that when '$\mathcal{A}$' is false, the conditional '$\mathcal{A} \rightarrow \mathcal{B}$' is automatically true, regardless of the truth value of '$\mathcal{B}$.' And if both '$\mathcal{A}$' and '$\mathcal{B}$' are true, then the conditional '$\mathcal{A} \rightarrow \mathcal{B}$' is true.

In short, '$\mathcal{A} \rightarrow \mathcal{B}$' is false if and only if '$\mathcal{A}$' is true and '$\mathcal{B}$' is false. We can summarize this with a characteristic truth table for the conditional.

$\mathcal{A}$	$\mathcal{B}$	$\mathcal{A}\rightarrow\mathcal{B}$
T	T	T
T	F	F
F	T	T
F	F	T

Notice that the material conditional is **asymmetrical**. You cannot in general swap the antecedent and consequent without changing the meaning of the sentence, because '$\mathcal{A} \rightarrow \mathcal{B}$' and '$\mathcal{B} \rightarrow \mathcal{A}$' do not mean the same thing. Indeed, as the truth table makes clear, if '$\mathcal{A} \rightarrow \mathcal{B}$' is false, then '$\mathcal{B} \rightarrow \mathcal{A}$' must be true.

Consider next these examples of that use the English connective 'unless':

3. Unless you wear a jacket, you will catch cold.
4. You will catch cold unless you wear a jacket.

Let J mean 'You will wear a jacket' and let D mean 'You will catch a cold.'

We can paraphrase sentence 3 as 'Unless J, D.' This means that if you do not wear a jacket, then you will catch cold; with this in mind, we might translate it as '$\neg J \rightarrow D$.' It also means that if you do not catch a cold, then you must have worn a jacket; with this in mind, we might translate it as '$\neg D \rightarrow J$.'

Which of these is the correct translation of 3? Both translations are correct, because the two translations are logically equivalent in $\mathcal{L}_S$.

Sentence 4, in English, is logically equivalent to sentence 3. It can be translated as '$\neg J \rightarrow D$,' or as '$\neg D \rightarrow J$.'

When symbolizing things like 3 and 4, it is easy to get turned around. Since the conditional is not symmetric, it would be wrong to translate either as '$J \rightarrow \neg D$.' Fortunately, there are other logically equivalent expressions. Both 3 and 4 mean that you will wear a jacket or—if you do not wear a jacket—then you will catch a cold. So we can translate them as '$J \vee D$.' (You might worry that the 'or' here should be an *exclusive or*. However, 3 and 4 do not exclude the possibility that you might *both* wear a jacket *and* catch a cold; jackets do not protect you from all the possible ways that you might catch a cold.) Thus:

> A declarative sentence of the form 'Unless $\mathcal{A}$, $\mathcal{B}$' can be symbolized as '$\mathcal{A} \vee \mathcal{B}$,' or as '$\neg\mathcal{A} \rightarrow B$.'

3.2 Biconditional

Consider these:

5. That figure is a triangle only if it has exactly three sides.
6. That figure is a triangle if it has exactly three sides.
7. That figure is a triangle if and only if it has exactly three sides.

We'll use this symbolization key for the atomic sentences:

T: That figure is a triangle.
S: That figure has exactly three sides.

Sentence 5, for the reasons that we've already discussed, can be symbolized as '$T \rightarrow S$.' On the other hand, 6 is importantly different. It can be paraphrased as, 'If the figure has three sides, then it is a triangle.' So it should be symbolized as '$S \rightarrow T$.'

Now, 7 says that 'T' is true *if and only if* 'S' is true. So we can infer 'S' from 'T,' and we can infer 'T' from 'S.' This is called a **biconditional**, because it entails the two conditionals '$S \rightarrow T$' and '$T \rightarrow S$.' We will use $\leftrightarrow$to represent the biconditional, so 7 can be symbolized as '$S \leftrightarrow T$.' Unlike the other two-place connectives, the different parts of a biconditional don't usually get fancy-sounding names; instead, they're usually just referred to as 'the left-hand side' and 'the right-hand side.'

As we said in the last section, whenever '$S \rightarrow T$' is true, then the truth of 'S' is a sufficient condition for the truth of 'T,' and the truth of 'T' is a necessary condition for the truth of 'S.' That means that, whenever '$S \leftrightarrow T$' is true, the truth of 'S' must be *both* a sufficient *and* necessary condition for the truth of 'T'. Likewise, the truth of 'T' must be necessary and sufficient for the truth of 'S.' In other words, '$S \leftrightarrow T$' says that 'S' and 'T' are necessary and sufficient for each other. You can't have either one being true without the other one being true, and if one of them is false then they both have to be false.

Because of this, $\leftrightarrow$ is also a truth-functional connective. Simply, '$\mathcal{A} \leftrightarrow \mathcal{B}$' is true just in case '$\mathcal{A}$' and '$\mathcal{B}$' are both are true, or both are false. Another way to put that is that '$\mathcal{A} \leftrightarrow \mathcal{B}$' is true just in case '$\mathcal{A}$' and '$\mathcal{B}$' *always have the same truth-value*. Hence, here is the characteristic truth table for the biconditional:

$\mathcal{A}$	$\mathcal{B}$	$\mathcal{A} \leftrightarrow \mathcal{B}$
T	T	T
T	F	F
F	T	F
F	F	T

Of course, if we wanted to, we could abide without any new symbols for the biconditional. Since 7 just means '[If that figure is a triangle then it has exactly three sides] *and* [if that figure has exactly three sides, then it is a triangle]', we could always symbolize it as '$(T \to S) \wedge (S \to T)$.' (Note again the need for parentheses to indicate that '$T \to S$ and $S \to T$' are separate conjuncts; without them, the expression '$T \to S \wedge S \to T$' would be highly ambiguous.)

So we *could* just write '$(\mathcal{A} \to \mathcal{B}) \wedge (\mathcal{B} \to \mathcal{A})$' instead of '$\mathcal{A} \leftrightarrow \mathcal{B}$,' and we do not strictly speaking *need* to introduce a new symbol for the biconditional. Nevertheless, logical languages usually have such a symbol for simplicity—and as we'll see in the next chapter, the biconditional is very important. $\mathcal{L}_S$ will have therefore have a symbol for the biconditional, which will also make it easier for us to symbolize phrases involving 'if and only if.'

3.3 Sentences of $\mathcal{L}_S$

The sentence 'It's not true that either apples are blue, or berries are pink' is a sentence of English, and the sentence '$\neg(A \vee B)$' is a sentence of $\mathcal{L}_S$. Although we can identify sentences of English when we encounter them, we do not have a formal definition of 'sentence of English'. But in this section, we shall offer a complete definition of what counts as a sentence of $\mathcal{L}_S$. This is one respect in which a formal language like $\mathcal{L}_S$ is more precise than a natural language like English.

We have seen that there are three basic kinds of symbols in $\mathcal{L}_S$:

Atomic sentences	$A, B, C, \ldots, Z$
with subscripts, as needed	$A_1, B_1, Z_1, A_2, A_{25}, J_{375}, \ldots$
Connectives	$\neg, \wedge, \vee, \rightarrow, \leftrightarrow$
Brackets	(,)

We can define an **expression of $\mathcal{L}_S$** as *any* string of symbols of $\mathcal{L}_S$. Take any of the symbols of $\mathcal{L}_S$ and write them down, in any order, and you have an expression of $\mathcal{L}_S$. Of course, many expressions of $\mathcal{L}_S$ will be total gibberish: for example, '$()\wedge A)(\rightarrow)B \rightarrow\rightarrow\rightarrow (AB))$' would not symbolize anything meaningful. We want to know when an expression of $\mathcal{L}_S$ amounts to a **sentence of $\mathcal{L}_S$**.

Obviously, our symbols for individual atomic sentences, like 'A' and 'G', should count as sentences of $\mathcal{L}_S$. We can create further sentences out of these by using the various connectives. Using negation, we can get '$\neg A$' and $\neg G$. Using conjunction, we can get '$(A\wedge G)$,' '$(G\wedge A)$,' '$(A\wedge A)$,' and '$(G\wedge G)$.' We could also apply negation repeatedly to get sentences like '$\neg\neg A$,' '$\neg\neg\neg A$,' '$\neg\neg\neg\neg A$'. Or we could apply negation along with conjunction to get sentences like '$\neg(A\wedge G)$' and '$\neg(G \wedge \neg G)$.'

The possible combinations are endless, even starting with just these two sentence letters, and there are infinitely many sentence letters. So there's no point in trying to list all the possible sentences one by one.

Instead, we will describe the process by which sentences can be systematically *constructed*. Consider negation: Given any sentence '$\mathcal{A}$ of $\mathcal{L}_S$, '$\neg\mathcal{A}$' will also be a sentence of $\mathcal{L}_S$. That is a very simple rule that we can apply to anything we already know is a sentence, in order to create a new sentence. We can say similar things for each of the other connectives. For instance, if '$\mathcal{A}$' and '$\mathcal{B}$' are sentences of $\mathcal{L}_S$, then '$(\mathcal{A} \wedge \mathcal{B})$' is a sentence of $\mathcal{L}_S$. Providing clauses like this for all of the connectives, we arrive at the following formal definition for a **sentence of $\mathcal{L}_S$**:

Definition of a SENTENCE OF $\mathcal{L}_S$:

1. Every atomic sentence is a sentence.
2. If '$\mathcal{A}$' is a sentence, then '$\neg\mathcal{A}$' is a sentence.
3. If '$\mathcal{A}$' and '$\mathcal{B}$' are sentences, then '$(\mathcal{A} \wedge \mathcal{B})$' is a sentence.
4. If '$\mathcal{A}$' and '$\mathcal{B}$' are sentences, then '$(\mathcal{A} \vee \mathcal{B})$' is a sentence.
5. If '$\mathcal{A}$' and '$\mathcal{B}$' are sentences, then '$(\mathcal{A} \rightarrow \mathcal{B})$' is a sentence.
6. If '$\mathcal{A}$' and '$\mathcal{B}$' are sentences, then '$(\mathcal{A} \leftrightarrow \mathcal{B})$' is a sentence.
7. Nothing else is a sentence.

(You may notice that in this definition, I've put brackets around '$(\mathcal{A} \wedge \mathcal{B})$,' '$(\mathcal{A} \vee \mathcal{B})$,' and so on. I will explain why I've done this shortly, and why it differs from how we've been doing it over the past couple of chapters.)

Definitions like this are called *recursive*. Recursive definitions begin with some specifiable base elements, and then present ways to generate indefinitely many more elements by compounding together previously established ones. To give you a better idea of what a recursive definition is, we can give a recursive definition of the idea of *an ancestor of mine*. We specify a base clause:

▷ My parents are ancestors of mine.

and then offer further clauses like:

▷ If x is an ancestor of mine, then x's parents are ancestors of mine.

▷ Nothing else is an ancestor of mine.

Using this definition, we can easily check to see whether someone is my ancestor: just check whether she is the parent of the parent of. . . one of my parents. And the same is true for our recursive definition of sentences of $\mathcal{L}_S$. Just as the recursive definition allows complex sentences to be built up from simpler parts, the definition allows us to decompose sentences into their simpler parts. And if we get down to atomic sentences, then we are ok.

Let's consider some examples.

Suppose we want to know whether or not '$\neg\neg\neg D$' is a sentence of $\mathcal{L}_S$. Looking at the second clause of the definition, we know that '$\neg\neg\neg D$' is a sentence *if* '$\neg\neg D$' is a sentence. So now we need to ask whether or not '$\neg\neg D$' is a sentence. Again looking at the second clause of the definition, '$\neg\neg D$' is a sentence *if* '$\neg D$' is. Again, '$\neg D$' is a sentence *if* 'D' is a sentence. Now 'D' is an atomic sentence in $\mathcal{L}_S$, so we know that 'D' is a sentence by the first clause of the definition. So for a compound sentence like '$\neg\neg\neg D$,' we must apply the definition repeatedly. Eventually we arrive at the atomic sentences from which the sentence is built up.

Next, consider the example '$\neg(P \wedge \neg(\neg Q \vee R))$.' Looking at the second clause of the definition, this is a sentence if '$(P \wedge \neg(\neg Q \vee R))$' is. And this is a sentence if *both* 'P' *and* '$\neg(\neg Q \vee R)$' are sentences. The former is an atomic sentence, and the latter is a sentence if '$(\neg Q \vee R)$' is a sentence. It is. Looking at the fourth clause of the definition, this is a sentence if both '$\neg Q$' and 'R' are sentences. And both are!

Ultimately, every sentence of $\mathcal{L}_S$ either is an atomic sentence, or it is constructed out of atomic sentences and connectives. The recursive structure of sentences in

$\mathcal{L}_S$ will be important in the next chapter, when we consider the circumstances under which a particular sentence would be true or false. The sentence '$\neg\neg\neg D$' is true if and only if the sentence '$\neg\neg D$' is false, and so on through the structure of the sentence, until we arrive at the atomic components.

Bracketing conventions

Strictly speaking, the brackets in '$(Q \wedge R)$' are an indispensable part of the sentence. Part of the reason for this is because we might want to use '$(Q \wedge R)$' as a sub-sentence in a more complicated sentence. For example, we might want to negate '$(Q \wedge R)$', obtaining '$\neg(Q \wedge R)$'. If we just had '$Q \wedge R$' without the brackets and put a negation in front of it, we would have '$\neg Q \wedge R$', which means something quite different than '$\neg(Q \wedge R)$'. It is most natural to read '$\neg Q \wedge R$' as meaning the same thing as '$(\neg Q \wedge R)$'. Strictly speaking, then, '$Q \wedge R$' is *not* a sentence. It is a mere *expression*.

With that said, when working with $\mathcal{L}_S$, it will make our lives much easier if we are sometimes a little less than strict when it comes to writing down sentences. So, we allow ourselves to omit the *outermost* brackets of a sentence. Thus, for example, we will allow ourselves to write '$Q \wedge R$' instead of the sentence '$(Q \wedge R)$', wherever it is convenient to do so.

However, when we want to turn '$Q \wedge R$' into a sub-sentence for a larger, more complex sentence, we need to put the brackets back around it.

Practice Exercises

Part A
Using the symbolization key given, translate each of the following into $\mathcal{L}_S$ as close as possible.

A : Mister Ace was murdered.
B : The butler did it.
C : The cook did it.
D : The Duchess is lying.
E : Mister Edge was murdered.
F : The murder weapon was a frying pan.

1. Either Mister Ace or Mister Edge was murdered.
2. If Mister Ace was murdered, then the cook did it.
3. If Mister Edge was murdered, then the cook did not do it.
4. Either the butler did it, or the Duchess is lying.
5. The cook did it only if the Duchess is lying.
6. If the murder weapon was a frying pan, then the culprit must have been the cook.
7. If the murder weapon was not a frying pan, then the culprit was either the cook or the butler.
8. Mister Ace was murdered if and only if Mister Edge was not murdered.
9. The Duchess is lying, unless it was Mister Edge who was murdered.
10. If Mister Ace was murdered, he was done in with a frying pan.
11. Since the cook did it, the butler did not.
12. Of course the Duchess is lying!

⋆ **Part B**
Using the symbolization key given, translate the following into $\mathcal{L}_S$.

E_1 : Adam is an entertainer.
E_2 : Heather is an entertainer.
F_1 : Adam is a philosopher.
F_2 : Heather is a philosopher.
S_1 : Adam is satisfied with his career.
S_2 : Heather is satisfied with her career.

1. Adam and Heather are both entertainers.
2. If Adam is a philosopher, then he is satisfied with her career.
3. Adam is a philosopher, unless he is an entertainer.
4. Heather is an unsatisfied entertainer.
5. Neither Adam nor Heather is an entertainer.

6. Both Adam and Heather are entertainers, but neither of them find it satisfying.
7. Heather is satisfied only if she is a philosopher.
8. If Adam is not an entertainer, then neither is Heather, but if he is, then she is too.
9. Adam is satisfied with his career if and only if Heather is not satisfied with hers.
10. If Heather is both an entertainer and a philosopher, then she must be satisfied with her career.
11. It cannot be that Heather is both an entertainer and a philosopher.
12. Heather and Adam are both philosophers if and only if neither of them is an entertainer.

$\star$ Part C

Give a symbolization key and symbolize the following in $\mathcal{L}_S$.

1. Alice and Bob are both spies.
2. If either Alice or Bob is a spy, then the code has been broken.
3. If neither Alice nor Bob is a spy, then the code remains unbroken.
4. The German embassy will be in an uproar, unless someone has broken the code.
5. Either the code has been broken or it has not, but the German embassy will be in an uproar regardless.
6. Either Alice or Bob is a spy, but not both.

Part D

For each argument, write a symbolization key and translate the argument as well as possible into $\mathcal{L}_S$.

1. If Dorothy plays the piano in the morning, then Roger wakes up cranky. Dorothy plays piano in the morning unless she is distracted. So if Roger does not wake up cranky, then Dorothy must be distracted.
2. It will either rain or snow on Tuesday. If it rains, Neville will be sad. If it snows, Neville will be cold. Therefore, Neville will either be sad or cold on Tuesday.
3. If Zoog remembered to do his chores, then things are clean but not neat. If he forgot, then things are neat but not clean. Therefore, things are either neat or clean—but not both.

Chapter 4

[$\mathcal{Req}$.] Truth Tables

This chapter introduces a way of evaluating the logical status of complex sentences and arguments formalised in the language of $\mathcal{L}_S$. Although it can be laborious, the truth table method is a purely mechanical procedure that requires no intuition or special insight. That may sound like a bad thing, but in fact it's incredibly useful: so long as you follow the method correctly, you're guaranteed to get the correct result.

4.1 Complete truth tables

For reference, the following tables summarise the characteristic truth tables for the five truth-functional connectives outlined in the previous two chapters.

$\mathcal{A}$	$\neg\mathcal{A}$
T	F
F	T

$\mathcal{A}$	$\mathcal{B}$	$\mathcal{A}\wedge\mathcal{B}$	$\mathcal{A}\vee\mathcal{B}$	$\mathcal{A}\rightarrow\mathcal{B}$	$\mathcal{A}\leftrightarrow\mathcal{B}$
T	T	T	T	T	T
T	F	F	T	F	F
F	T	F	T	T	F
F	F	F	F	T	T

Keep in mind that '$\mathcal{A}$' and '$\mathcal{B}$' can represent *any* sentences in $\mathcal{L}_S$ whatsoever, and not only atomic sentences. For example, in §2.1 we said that 'The widget is irreplaceable' should be symbolized as '$\neg R$' and 'The widget is not irreplaceable' as '$\neg\neg R$.' The characteristic truth table for negation tells us how the truth-value for '$\neg R$' depends on the truth-value for 'R':

R	$\neg R$
T	F
F	T

But the table *also* tells us how the truth-value of '$\neg\neg R$' depends on the value of '$\neg R$.' To see this, suppose that we use the special symbol $\mathcal{R}$ to represent the complex sentence '$\neg R$'; so, writing '$\mathcal{R}$' is just another way of writing '$\neg R$'. Then, '$\neg\neg R$' just means the same thing as '$\neg\mathcal{R}$.' We can therefore apply the characteristic truth table for negation to '$\neg\mathcal{R}$':

$\mathcal{R}$	$\neg\mathcal{R}$
T	F
F	T

In other words, '$\neg\neg R$' is related to '$\neg R$' in just the same way that '$\neg R$' is related to 'R.' So, '$\neg\neg R$' is true whenever '$\neg R$' is false, and false whenever '$\neg R$' is true. It's easy to see how we could extend this reasoning for indefinitely many iterations of the negation operation. So, for example, '$\neg\neg\neg R$' is just the negation of '$\neg\neg R$,' and therefore has the opposite truth-value of '$\neg\neg R$.' And '$\neg\neg\neg\neg R$' is just the negation of '$\neg\neg\neg R$' ... and so on, *ad infinitum.*

This means that we can use truth tables to work out how the truth-values of even very complex non-atomic sentences ultimately depend on the truth-values of their atomic parts. To get the feel for how we can do this, let's look at a slightly more complicated example.

Complete truth tables

Suppose we have a non-atomic sentence '$(H \wedge I) \rightarrow H$.' What we want to do is work out the conditions (if any) under which '$(H \wedge I) \rightarrow H$' will be true, and when it will be false.

The first step is to mark down all possible combinations of truth and falsity for 'H' and 'I.' As there are only two atomic sentences here, there are *four* possible combinations and therefore four rows on our truth table, like so:

H	I	$(H \wedge I) \rightarrow H$
T	T	
T	F	
F	T	
F	F	

We then copy the truth-values for 'H' and 'I' over and write them underneath the relevant letters on the right-hand side of the table:

H	I	(H	$\wedge$	I)	$\rightarrow$	H
T	T	**T**		**T**		**T**
T	F	**T**		**F**		**T**
F	T	**F**		**T**		**F**
F	F	**F**		**F**		**F**

Now, the sentence '$(H \wedge I) \rightarrow H$' *as a whole* is a conditional; i.e., it has the form '$\mathcal{A} \rightarrow \mathcal{B}$,' where $\mathcal{A}$ = '$H \wedge I$' and the $\mathcal{B}$ = 'H.' Since the antecedent is itself a complex sentence with the form of a conjunction, we can break it down into its parts as well. So we will begin by filling in the truth table just for the conjunctive antecedent.

'H' and 'I' are both true on the first row. The first row represents the possible situation where 'H' and 'I' are both true. Since the characteristic truth table for conjunction tells us that a conjunction is true when both conjuncts are true, we write a 'T' underneath the conjunction symbol for the first row. The second row represents the possible situation where 'H' is true and 'I' is false. So, since a conjunction is false whenever one of its conjuncts is false, we write an 'F' underneath the conjunction symbol for the second row. We continue this process for final two rows and get this:

H	I	(H	$\wedge$	I)	$\rightarrow$	H
		$\mathcal{A}$	$\wedge$	$\mathcal{B}$		
T	T	T	**T**	T		T
T	F	T	**F**	F		T
F	T	F	**F**	T		F
F	F	F	**F**	F		F

Almost done! As we noted just above, the entire sentence is a conditional: it has the form '$\mathcal{A} \rightarrow \mathcal{B}$,' with $\mathcal{A}$ = '$H \wedge I$' and $\mathcal{B}$ = 'H.' And, as the characteristic truth table for the material conditional then tells us, '$(H \wedge I) \rightarrow H$' is false just in case '$(H \wedge I)$' is true and 'H' is false. It is true otherwise. So now we're going to use what we've just worked out about the truth-value of the antecedent '$(H \wedge I)$' to determine the truth-value of the whole sentence '$(H \wedge I) \rightarrow H$.'

For the first row, we've just worked that '$(H \wedge I)$' is true. 'H' is also true on that row. Therefore, the whole conditional must be true on that row, and we put down a T for the first row underneath the $\rightarrow$symbol. This tells us that '$(H \wedge I) \rightarrow H$' is true if '$H$' and '$I$' are both true. Then, on the second row, '$(H \wedge I)$' is false, and 'H' is true. Since a conditional is true whenever the antecedent is false, we can again write a T in the second row underneath the conditional symbol. We continue for the final two rows and get this:

H	I	$(H \wedge I) \to H$
		$\mathcal{A} \quad \to \mathcal{B}$
T	T	T **T** T
T	F	F **T** T
F	T	F **T** F
F	F	F **T** F

Now the column of Ts underneath the conditional tells us that '$(H \wedge I) \to I$' is always true, regardless of the truth-values of 'H' and 'I.' This should be intuitive: if 'H' *and* 'I' is true, then 'H' by itself is true. We can know that this must be the case purely by our logical reasoning; the truth table confirms it.

In this example, we have not repeated all of the entries in every successive table. When actually writing truth tables on paper, however, it is impractical to erase whole columns or rewrite the whole table for every step. Although it is more crowded, the complete truth table can be written in this way:

H	I	$(H \wedge I) \to H$
T	T	T T T **T** T
T	F	T F F **T** T
F	T	F F T **T** F
F	F	F F F **T** F

Each row represents a different possibility; e.g., the possibility where 'H' and 'I' are both true, where one is true but not the other, and where both are false. Since $\wedge$ and $\to$are truth-functions, the truth-value of '$(H \wedge I) \to I$' depends ultimately on the truth values of 'H' and 'I.' It's therefore important that we consider each row; otherwise, we won't have considered all of the possibilities. The truth-value of the sentence on each row is just the column underneath the **main logical connective** of the sentence; in this case, the column underneath the conditional.

The main logical connective

Working out which connective is the main connective in a complex sentence can sometimes be quite difficult. The main connective is defined as the connective with the **widest scope**. We will begin with an example, after which I will lay out a very simple rule for deciding which is the main connective.

We'll work through the connectives in the following rather complex example of a sentence:

'$((A \leftrightarrow B) \to C) \wedge \neg(\neg D \to C)$'

In this case, the main connective is the $\wedge$. This is because it has the *widest scope* of all the connectives in the sentence. That is, its scope covers '$((A \leftrightarrow B) \rightarrow C)$' on its left side, and '$\neg(D \rightarrow C)$' on its right side. None of the other connectives in the sentence have anything as big as these in their scope. For example, the scope of the first $\rightarrow$ is limited to '$(A \leftrightarrow B)$' on its left side, and 'C' on its right side; the scope of the second $\rightarrow$ is limited to '$\neg D$' on its left side, and 'C' on its right side.

Consider then the first conjunct:

'$((A \leftrightarrow B) \rightarrow C)$'

In this case, $\rightarrow$ has a wider scope than $\leftrightarrow$: on its left side it has, '$(A \leftrightarrow B)$,' and on its right side, 'C.' The scope of the $\leftrightarrow$ is narrower, covering only 'A' and 'B' on its left and right sides respectively. Since $\leftrightarrow$ has only simple sentence letters in its scope, it has the narrowest possible scope.

We can also take a look at the second conjunct:

'$\neg(\neg D \rightarrow C)$'

Here, the first $\neg$ (the one outside the parentheses) has the widest scope: it covers all of '$(\neg D \rightarrow C)$,' whereas the $\rightarrow$ only covers 'D' and 'C.'

And, finally:

'$(\neg D \rightarrow C)$'

In this case, the scope of $\neg$ is limited only to the 'D' on the left, so $\rightarrow$ has the wider scope of the two. Again: as $\neg$ only has a simple sentence letter in its scope, it has the narrowest possible scope. (It's possible for more than one part of a complex sentence to have maximally narrow scope; in this case, $\leftrightarrow$ and the second $\neg$ have maximally narrow scope.)

When we're filling in truth tables for a complex sentence, we always begin with the connectives which have the narrowest scope, and work out way up to the main logical connective. The connectives with the *narrowest* scope always only connect simple sentence letters—there should be no complex sentences within their scope. In the end, there should only ever be *one* main logical connective.

The main connective tells you what *kind* of sentence you're dealing with. So, since the main connective of '$((A \leftrightarrow B) \rightarrow C) \wedge \neg(\neg D \rightarrow C)$' is a $\wedge$, we're dealing with a conjunction; and since the main connective of '$(H \wedge I) \rightarrow I$' earlier was $\rightarrow$, it was a conditional sentence.

> The MAIN LOGICAL CONNECTIVE of any complex sentence is the connective with the widest scope. There is only ever one main logical connective. The column of the truth table that is underneath the main logical connective indicates the conditions under which the sentence is true.

Fortunately, if you want to work out the main connective for any sentence symbolized in $\mathcal{L}_S$, you can always use the following steps. First, we apply our notational convention and remove the outer-most brackets (3.3). In this case, the main logical connective will always be *outside* of any brackets which may still be left. This will often be sufficient to work out which is the main connective, as there will often only be one connective which sits outside of all the brackets.

However, in some cases there may be more than one connective which is not contained within some pair of brackets. Then, there are two possibilities: first, all of the remaining connectives are negations; and second, only some of the remaining connectives are negations. In the first case, the *left-most* negation is the main connective. In the second case, *none* of the negations is the main connective—so whatever other connective (whether $\wedge$, $\vee$, $\rightarrow$, or $\leftrightarrow$) sits outside of all of the brackets must be the main logical connective.

Truth table rows

A complete truth table for any sentence should have a row for all the possible combinations of T and F, for all of the sentence letters that can be found in the symbolization. The size of the complete truth table therefore depends on the number of different sentence letters in the table. This is true even if the same letter is repeated many times, as in '$((C \leftrightarrow C) \rightarrow C) \wedge \neg(\neg C \rightarrow C)$.' The complete truth table for this requires only two lines because there are only two possibilities: 'C' can be true, or it can be false. The truth table for this sentence looks like this:

C	((	C	$\leftrightarrow$	C	)	$\rightarrow$	C	)	$\wedge$	$\neg$	(	$\neg$	C	$\rightarrow$	C	)
T		T	T	T		T	T		**F**	F		F	T	T	T	
F		F	T	F		F	F		**F**	T		T	F	F	F	

Note that a single sentence letter can never be marked both T and F *on the same row*. Each row represents a possible truth-value for 'C,' and it can never be both true and false at once.

A complex sentence that contains two atomic sentences as parts requires *four* lines for a complete truth table, as in the characteristic truth tables and the table for' $(H \wedge I) \rightarrow I$.' Here there are four possible combinations: TT, TF, FT, and FF. A complex sentence that contains three atomic sentences as parts requires *eight* lines. For example:

M	N	P	$M \wedge (N \vee P)$
T	T	T	T **T** T T T
T	T	F	T **T** T T F
T	F	T	T **T** F T T
T	F	F	T **F** F F F
F	T	T	F **F** T T T
F	T	F	F **F** T T F
F	F	T	F **F** F T T
F	F	F	F **F** F F F

From this table, we know that '$M \wedge (N \vee P)$' might be true or false, depending on the truth-values of 'M,' 'N,' and 'P.' In particular, the truth table tells us that '$M \wedge (N \vee P)$' is true in exactly three cases:

1. When 'M,' 'N,' and 'P' are all true.
2. When 'M' and 'N' are true, and 'P' false.
3. When 'M' and 'P' are true, and 'N' false.

This holds *regardless* of what the sentence letters 'M,' 'N,' and 'P' refer to. This is because the connectives $\wedge$ and $\vee$ are truth-functional: the truth of conjunction/disjunction depends only on the truth of the conjuncts/disjuncts. Thus we can break down complex sentences into their smaller parts, work out the truth-values of those parts, and work back up to the truth-value of the whole.

A complete truth table for a complex sentence that contains four atomic sentences as parts requires 16 lines. Five atomic sentences as parts, 32 lines. Six atomic sentences as parts, 64 lines. And so on. To be perfectly general: if a complex sentence has n-many atomic sentences as parts, then its complete truth table must have 2^n rows. So if a complex sentence has 10 atomic sentences as its parts, its truth table must have $2^{10} = 1024$ rows.

In order to fill in the columns of a complete truth table, begin with the right-most sentence letter and alternate Ts and Fs. In the next column to the left, write two Ts, write two Fs, and repeat. For the third sentence letter, write four Ts followed by four Fs. This yields an eight line truth table like the one above. For a 16 line truth table, the next column of sentence letters should have eight Ts followed by eight Fs. And so on.

4.2 Tautologies and Contradictions

Using truth tables, we can quickly come to see that there may be some sentences that are *always* and *necessarily* true just by virtue of their logical form. That is, there may be some complex sentences which are always true regardless of the truth-values that their atomic parts take.

We have already seen some examples of this in the previous section, but here is a rather simple example:

1. Either dinosaurs lived on Earth, or they didn't.

Let's symbolise this and then fill in its truth table. We'll use the following symbolization key:

D : Dinosaurs lived on Earth.

With this key, we should symbolize sentence 1 as '$D \vee \neg D$.' Now when we fill in its truth table, we'll see that it must be true regardless of whether 'D' is true or false:

D	$D \vee \neg D$
T	T **T** F T
F	F **T** T F

This should be intuitive: regardless of what happened in the past, either dinosaurs lived on Earth, or they didn't. One of these two disjuncts must be true. We can know this without having to go out and check; merely doing some logical reasoning is enough.

Indeed, the same can be said of any sentence with the same logical form. For instance:

2. The dinosaurs were wiped out by an asteroid, or they weren't.
3. I am human, or non-human.
4. Whales are mammals, or they are not.

The specific content of these sentences does not matter so much for working out their truth-value—what matters is that they have a logical form which forces them to be true. Their truth tables reflect this fact, in that they have all Ts underneath their main connective. We will call any sentence which, once it has been symbolized in $\mathcal{L}_S$, has all Ts underneath its main connective, a **tautology in** $\mathcal{L}_S$.

For simplicity, in the discussion that follows we'll usually leave out the 'in $\mathcal{L}_S$' clause, but it's important to keep it in the back of your mind. Whether a sentence of English has a tautologous logical form depends in part upon how *we* formalize it—that is, it depends on the formal language we use when symbolizing it. As we will see in Chapter 6, there are some sentences which seem to be true purely as a matter of logic, but which *aren't* tautologies with respect to $\mathcal{L}_S$. However, they will be tautologies with respect to a more powerful formal language, $\mathcal{L}_P$. The good news is that any English sentence which can be symbolized as a tautology in $\mathcal{L}_S$ will *also* end up being a tautology in $\mathcal{L}_P$. Therefore, anything which is a tautology in $\mathcal{L}_S$ can be known to be true merely by logical reasoning—but there may also be some sentences which are *not* tautologies when symbolized in $\mathcal{L}_S$ but which are still knowable using (a more complicated form of) logical reasoning alone.

Another simple tautology can be seen from the following truth table:

D	$D \rightarrow D$
T	T **T** T
F	F **T** F

Again, this should be very intuitive: *if* dinosaurs lived on Earth, *then* dinosaurs lived on Earth. Propositions with this form are also true as a matter of logic, they cannot be false. One more example before we move on:

D	$\neg\,(D \wedge \neg\, D)$
T	**T** T F F T
F	**T** F F T F

In this case we have '*It is not the case that* dinosaurs lived on Earth and they didn't live on Earth'. There are in fact infinitely many tautologies in $\mathcal{L}_S$, and many of them are very, very complex.

> A sentence, when symbolized in $\mathcal{L}_S$, is TAUTOLOGY IN $\mathcal{L}_S$ just in case its truth table shows all Ts underneath its main logical connective.

Notice that the truth-values underneath the $\wedge$ in '$\neg(D \wedge \neg D)$' are always false. So if we were to fill in the truth table for '$D \wedge \neg D$,' we'd find that there are only Fs underneath its main connective. We will say that any sentence which has this property is a **contradiction in $\mathcal{L}_S$**.

As with tautologies, every contradiction in $\mathcal{L}_S$ will also be a contradiction in $\mathcal{L}_P$. So, if something is a contradiction in $\mathcal{L}_S$, we can know that it must be false merely using logic. But there will also be some contradictions in the more powerful logical language $\mathcal{L}_P$ which are *not* contradictions in $\mathcal{L}_S$.

Contradictions are the opposite of tautologies: they have such a form that they *cannot* be true, regardless of how the world turns out to be. '$\neg(D \wedge \neg D)$' is a tautology precisely because it says of some contradiction—viz., '$D \wedge \neg D$'—that it is not true. Since contradictions are always false, the negation of a contradiction is always true. Likewise, the negation of every tautology is a contradiction.

A sentence, when symbolized in $\mathcal{L}_S$, is a CONTRADICTION IN $\mathcal{L}_S$ just in case its truth table shows all Fs underneath its main logical connective.

We can say that a sentence (when symbolized in $\mathcal{L}_S$) is **contingent in** $\mathcal{L}_S$ if its truth table shows some Ts and some Fs underneath its main connective. If we are dealing with a contingent sentence, then for all the truth table method tells us, the sentence may be true or it may be false.

Finally, we can say that a sentence is **consistent in** $\mathcal{L}_S$ just in case its truth table has at least one T underneath its main connective. Thus, a consistent sentence can be either a tautology or contigent, but not a contradiction.

A sentence, when symbolized in $\mathcal{L}_S$, is CONTINGENT IN $\mathcal{L}_S$ just in case its truth table shows at least one T and at least one F underneath its main logical connective. It is CONSISTENT IN $\mathcal{L}_S$ if its truth table shows at least one T under its main logical connective.

4.3 Logical Equivalence

We can also ask about the logical relations *between* collections of sentences. We will begin with what it means for two (or more) sentences to be **logically equivalent** when symbolized in $\mathcal{L}_S$. For example:

J_1 : John went to the store and to the park.
J_2 : John went to the park and to the store.

'J_1' and 'J_2' are both contingent, since John might not have gone anywhere at all. Yet they must have the same truth-value. If either is true, then they both are. And if either is false, then they both are. When two sentences have the same truth values under all the same conditions, we say that they are logically equivalent.

Consider the sentences '$\neg(A \vee B)$' and '$\neg A \wedge \neg B$.' Are they logically equivalent? We can check by constructing a **combined truth table** for both sentences at once:

A	B	$\neg\,(A \vee B)$	$\neg A \wedge \neg B$
T	T	**F** T T T	F T **F** F T
T	F	**F** T T F	F T **F** T F
F	T	**F** F T T	T F **F** F T
F	F	**T** F F F	T F **T** T F

Look at the columns for the main connectives; negation for the first sentence, and conjunction for the second. On the first three rows, both are F. On the final row, both are T. Since they match on every row, the two sentences are logically equivalent. In other words, each of '$\neg(A \vee B)$' and '$\neg A \wedge \neg B$' are true if and only if 'A' is false and 'B' is false; they are true under all and only the same conditions.

Notice that '$\neg(A \vee B)$' and would *not* be logically equivalent to '$\neg(C \vee D)$,' even though they have the same *form* as one another. The truth table for '$\neg(A \vee B)$' says that it is true whenever 'A' and 'B' are both false; the truth table for '$\neg(C \vee D)$' says that it is true whenever 'C' and 'D' are both false. If 'A' and 'B' are different sentences than 'C' and 'D,' then these represent different possibilities.

> Two sentences, when symbolized in $\mathcal{L}_S$, are LOGICALLY EQUIVALENT IN $\mathcal{L}_S$ just in case they have the same truth table.

As it turns out, there are many interesting logical equivalencies. For example, there is also the equivalence of '$\neg(A \wedge B)$' and '$\neg A \vee \neg B$.' Another example is that '$A \rightarrow B$' is equivalent to '$\neg(A \wedge \neg B)$,' which is equivalent to '$\neg A \vee B$,' which is equivalent to '$\neg B \rightarrow \neg A$.'

As we said in §3.2, a biconditional is true just in case the left-hand side of the biconditional always has the same truth-value as the right-hand side. Therefore, if we have two logically equivalent sentences—for example, '$\neg(A \vee B)$' and '$\neg A \wedge \neg B$'—we can put them on either side of a biconditional to create a tautology:

A	B	$\neg\,(A \vee B) \leftrightarrow (\neg A \wedge \neg B)$
T	T	F T T T **T** F T F F T
T	F	F T T F **T** F T F T F
F	T	F F T T **T** T F F F T
F	F	T F F F **T** T F T T F

So, every pair of logically equivalent sentences corresponds to a tautologous biconditional (and *vice versa*). In other words, if a biconditional sentence is a tautology, then the left-hand side and the right-hand side of the biconditional are logically equivalent. And, in the other direction: if two sentences are logically equivalent, then the biconditional between them is a tautology.

4.4 Joint Consistency

Next consider these two sentences:

B_1 : Bob went for a swim today.
B_2 : Bob did not go for a swim today.

Logic alone cannot tell us which, if either, of these two sentences is true. Yet we can say that *if* the first ('B_1') is true, *then* the second ('B_2') must be false. And if the second is true, then the first must be false. It cannot be the case that both of these sentences are true at the same time.

If a set of sentences cannot all be true at the same time, then the *set* is said to be **inconsistent**. We can also say that the sentences within the set are **jointly inconsistent**; this is just another way of saying that the set is inconsistent. If every sentence within the set can all be true at the same time, we say that the set is consistent, and its members are jointly consistent. A set of sentences might be *jointly* inconsistent even if each of the sentences within it is *individually* consistent.

We can ask about the consistency of any set of sentences. For example, consider the following list:

G_1 : There are at least four giraffes at the wild animal park.
G_2 : There are exactly seven gorillas at the wild animal park.
G_3 : There are not more than two martians at the wild animal park.
G_4 : Every giraffe at the wild animal park is a martian.

'G_1' and 'G_4' together imply that there are at least four martian giraffes at the park. This conflicts with 'G_3', which implies that there are no more than two martian giraffes at the park. So the collection of sentences, 'G_1' to 'G_4', is jointly inconsistent. Notice that the inconsistency of the set $\{G_1, G_2, G_3, G_4\}$ has nothing at all to do with 'G_2'—it just happens to be a member of an inconsistent set of sentences, but its membership does not contribute to that inconsistency.

We can use truth tables to test for joint consistency of any set of sentences. Suppose we have three sentences, '$P \wedge Q$,' '$P \rightarrow R$,' and '$R \rightarrow \neg Q$.' Individually, the truth tables for these three are:

P	Q	$P \wedge Q$
T	T	T **T** T
T	F	T **F** F
F	T	F **F** T
F	F	F **F** F

P	R	$P \rightarrow R$
T	T	T **T** T
T	F	T **F** F
F	T	F **T** T
F	F	F **T** F

R	Q	$R \rightarrow \neg Q$
T	T	T **F** F T
T	F	T **T** T F
F	T	F **T** F T
F	F	F **T** T F

As you can see, each sentence in the set is individually consistent. To work out whether they're jointly consistent we'll need to put all three into a single combined truth table:

P	Q	R	$P \wedge Q$	$P \to R$	$R \to \neg Q$
T	T	T	T **T** T	T **T** T	T **F** F T
T	T	F	T **T** T	T **F** F	F **T** F T
T	F	T	T **F** F	T **T** T	T **T** T F
T	F	F	T **F** F	T **F** F	F **T** T F
F	T	T	F **F** T	F **T** T	T **F** F T
F	T	F	F **F** T	F **T** F	F **T** F T
F	F	T	F **F** F	F **T** T	T **T** T F
F	F	F	F **F** F	F **T** F	F **T** T F

Now, if we go through each row, we'll see that there is no row where all three of the sentences are true. That means that there is no possible condition—no combination of truth-values for 'P,' 'Q,' and 'R'—under which all three of '$P \wedge Q$,' '$P \to R$,' and '$R \to \neg Q$' can be true at the same time. So they are jointly inconsistent.

> A set of sentences, when symbolized in $\mathcal{L}_S$, are JOINTLY INCONSISTENT IN $\mathcal{L}_S$ whenever there is no row in their combined truth table under which each is true.

Interestingly, if some collection of sentences are jointly inconsistent, then the *conjunction* of all of those sentences together is a contradiction. We can see this with the following truth table for the conjunction of '$P \wedge Q$,' '$P \to R$,' and '$R \to \neg Q$.' (As we noted in §2.2, conjunction is symmetrical, so it does not matter what order they are conjoined together.)

P	Q	R	$(P \wedge Q) \wedge ((P \to R) \wedge (R \to \neg Q))$
T	T	T	T T T **F** T T T F T F F T
T	T	F	T T T **F** T F F F F T F T
T	F	T	T F F **F** T T T F T F F T
T	F	F	T F F **F** T F F F F T F T
F	T	T	F F T **F** F T T T T T T F
F	T	F	F F T **F** F T F T F T T F
F	F	T	F F F **F** F T T T T T T F
F	F	F	F F F **F** F T F T F T T F

So, every jointly inconsistent set of sentences corresponds to a contradictory conjunction (and *vice versa*). This makes sense: a conjunction is only true when each of the conjuncts are true, and a set of sentences is only consistent when each member of the set *can* be true at the same time.

4.5 Truth Tables and Validity

As we've defined it in Chapter 1, an argument is valid just in case it's not possible its premises to be true and its conclusion false. And, just as we've used truth tables to test for consistency, equivalence, and so on, we can also use truth tables to test for formal validity.

Consider this argument:

P1 $\neg L \to (J \vee L)$
P2 $\neg L$

C J

Is it valid? There are two ways to test this using truth tables. The easiest is to construct a combined truth table that contains each of the premises as well as the conclusion. If the conclusion is true whenever the premises are all true, then the argument is valid.

J	L	$\neg L \to (J \vee L)$	$\neg L$	J
T	T	F T **T** T T T	**F** T	**T**
T	F	T F **T** T T F	**T** F	**T**
F	T	F T **T** F T T	**F** T	**F**
F	F	T F **F** F F F	**T** F	**F**

Yes, the argument is valid. The only row on which both the premises are T is the second row, and on that row the conclusion is also T.

> An argument is FORMALLY VALID IN $\mathcal{L}_S$ whenever there is no row in its combined truth table such that the premises are all true and the conclusion false.

Note that we *don't* have to worry about the rows where one or more of the premises is false. This means that if there is *no* row where all of the premises are true, then the argument is automatically valid. You might find this strange, but it fits with the definition of validity: that it is not possible for the premises to all be true and the conclusion false. If it's not possible for the premises to all be true, then the definition is trivially satisfied, regardless of the truth-value of the conclusion. Hence, arguments with jointly inconsistent premises are always valid.

Another consequence of the definition is that any argument with a tautologous conclusion is automatically valid—regardless of what its premises are! Again, this fits with the definition of validity: if the conclusion is necessarily true, then it cannot be false; therefore, it cannot be the case that the premises are true and the conclusion false.

We can also test for validity by turning the argument into a material conditional. First, we conjoin each of the premises. This conjunction will form the antecedent. Then, the conclusion will be our consequent. If the conditional is a tautology, then the argument is valid. So, for the current example, we begin by conjoining our two premises to get '$(\neg L \rightarrow (J \vee L)) \wedge \neg L$.' This will be the antecedent. The conclusion is just 'J.' Hence, we have '$((\neg L \rightarrow (J \vee L)) \wedge \neg L) \rightarrow J$,' the truth table of which is:

J	L	$((\neg$	L	$\rightarrow$	$(J$	$\vee$	$L))$	$\wedge$	$\neg$	$L)$	$\rightarrow$	J
T	T	F	T	T	T	T	T	F	F	T	**T**	T
T	F	T	F	T	T	T	F	T	T	F	**T**	T
F	T	F	T	T	F	T	T	F	F	T	**T**	F
F	F	T	F	F	F	F	F	F	T	F	**T**	F

In general, every argument that is valid when symbolized in $\mathcal{L}_S$ corresponds to material conditional that is a tautology in $\mathcal{L}_S$ (and *vice versa*). This makes sense. The material conditional says that the antecedent is sufficient for the truth of the consequent, and a material conditional is a tautology just in case its antecedent is sufficient for its consequent by virtue of their respective logical forms. On the other hand, a valid argument is one where the premises are jointly sufficient for the truth of the conclusion, by virtue of logical form. Hence, the *conjunction* of the premises is sufficient for the truth of the conclusion, by virtue of logical form.

Practice Exercises

If you want additional practice, you can construct truth tables for any of the sentences and arguments in the exercises for the previous chapter.

⋆ **Part A**
Determine whether each sentence is a tautology, a contradiction, or a contingent sentence (in $\mathcal{L}_S$). Justify your answer with a truth table where appropriate.

1. $A \rightarrow A$
2. $\neg B \wedge B$
3. $C \rightarrow \neg C$
4. $\neg D \vee D$
5. $(A \leftrightarrow B) \leftrightarrow \neg(A \leftrightarrow \neg B)$
6. $(A \wedge B) \vee (B \wedge A)$
7. $(A \rightarrow B) \vee (B \rightarrow A)$
8. $\neg(A \rightarrow (B \rightarrow A))$
9. $(A \wedge B) \rightarrow (B \vee A)$
10. $A \leftrightarrow (A \rightarrow (B \wedge \neg B))$
11. $\neg((A \vee B) \leftrightarrow (\neg A \wedge \neg B))$
12. $\neg((A \wedge B) \leftrightarrow A)$
13. $((A \wedge B) \wedge \neg(A \wedge B)) \wedge C$
14. $A \rightarrow (B \vee C)$
15. $((A \wedge B) \wedge C) \rightarrow B$
16. $(A \wedge \neg A) \rightarrow (B \vee C)$
17. $\neg\big((C \vee A) \vee B$
18. $(B \wedge D) \leftrightarrow (A \leftrightarrow (A \vee C))$

⋆ **Part B**
Determine whether each pair of sentences below are logically equivalent. Justify your answer with a complete or partial truth table where appropriate.

1. A
   $\neg A$

2. A
   $A \vee A$

3. $A \rightarrow A$
   $A \leftrightarrow A$

4. $A \vee \neg B$
   $A \rightarrow B$

5. $A \wedge \neg A$
   $\neg B \leftrightarrow B$

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

7. $\neg(A \rightarrow B)$
   $\neg A \rightarrow \neg B$

8. $A \rightarrow B$
   $\neg B \rightarrow \neg A$

9. $(A \vee B) \vee C$
   $A \vee (B \vee C)$

10. $(A \vee B) \wedge C$
    $A \vee (B \wedge C)$

⋆ Part C
Determine whether each set of sentences is consistent or inconsistent. Justify your answer with a truth table where appropriate.

1. $A \rightarrow A$
   $\neg A \rightarrow \neg A$
   $A \wedge A$
   $A \vee A$

2. $A \wedge B$
   $C \rightarrow \neg B$
   C

3. $A \vee B$
   $A \rightarrow C$
   $B \rightarrow C$

4. $A \rightarrow B$
   $B \rightarrow C$
   A
   $\neg C$

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

6. $A \vee B$
   $B \vee C$
   $C \rightarrow \neg A$

7. $A \leftrightarrow (B \vee C)$
   $C \rightarrow \neg A$
   $A \rightarrow \neg B$

8. A
   B
   C
   $\neg D$
   $\neg E$
   F

⋆ Part D
Determine whether each argument is valid or invalid. Justify your answer with a truth table where appropriate.

1. $A \rightarrow A$
   therefore A

2. $A \vee \big(A \rightarrow (A \leftrightarrow A)\big)$
   therefore A

3. $A \rightarrow (A \wedge \neg A)$
   therefore $\neg A$

4. $A \leftrightarrow \neg(B \leftrightarrow A)$
   therefore A

5. $A \vee (B \rightarrow A)$
   therefore $\neg A \rightarrow \neg B$

6. $A \rightarrow B$
   B
   therefore A

7. $A \vee B$
   $B \vee C$
   $\neg A$
   therefore $B \wedge C$

8. $A \vee B$
   $B \vee C$

$\neg B$
therefore $A \wedge C$

9. $(B \wedge A) \to C$
   $(C \wedge A) \to B$
   therefore $(C \wedge B) \to A$

10. $A \leftrightarrow B$
    $B \leftrightarrow C$
    therefore $A \leftrightarrow C$

⋆ Part E
Answer each of the questions below and justify your answer.

1. Suppose that '$\mathcal{A}$' and '$\mathcal{B}$' are logically equivalent. What is the status of '$\mathcal{A} \leftrightarrow \mathcal{B}$'? (i.e., is it a tautology, contradiction, contingent?)
2. Suppose that '$(\mathcal{A} \wedge \mathcal{B}) \to \mathcal{C}$' is contingent. What can you say about the argument "$\mathcal{A}$, $\mathcal{B}$, therefore $\mathcal{C}$"?
3. Suppose that the set of sentences $\{\mathcal{A}, \mathcal{B}, \mathcal{C}\}$ is inconsistent. What can you say about the status of '$(\mathcal{A} \wedge \mathcal{B}) \wedge \mathcal{C}$'? What about the status of '$\mathcal{B} \wedge (\mathcal{A} \wedge \mathcal{C})$'?
4. Suppose that '$\mathcal{A}$' is a contradiction. What can you say about the argument "$\mathcal{A}$, $\mathcal{B}$, therefore $\mathcal{C}$"?
5. Suppose that '$\mathcal{C}$' is a tautology. What can you say about the argument "$\mathcal{A}$, $\mathcal{B}$, therefore $\mathcal{C}$"?
6. Suppose that '$\mathcal{A}$' and '$\mathcal{B}$' are logically equivalent. What can you say about the status of '$\mathcal{A} \vee \mathcal{B}$'?
7. Suppose that '$\mathcal{A}$' and '$\mathcal{B}$' are *not* logically equivalent. What can you say about the status of '$\mathcal{A} \vee \mathcal{B}$'?

Part F
We could leave the symbol for the biconditional $\leftrightarrow$ out of the language if we wanted to. Instead of writing '$A \leftrightarrow B$' we can just write '$(A \to B) \wedge (B \to A)$.' The resulting language would be formally equivalent to $\mathcal{L}_S$, since '$A \leftrightarrow B$' and '$(A \to B) \wedge (B \to A)$' are logically equivalent. The fact that they are logically equivalent means that we can *define* the biconditional $\leftrightarrow$ in terms of $\to$ and $\wedge$ without any loss in our ability to express things. The only loss would be in the simplicity of our symbolization.

Once we start defining some logical connectives in terms of others, we find out that we actually don't need very many logical connectives at all. Indeed, in many formal languages there are only two connectives.

It would be enough to have only negation and conjunction. For example, we can define the material conditional in terms of $\neg$ and $\wedge$ as follows:

$\mathcal{A}$	$\mathcal{B}$	$\mathcal{A} \to \mathcal{B}$	$\neg(\mathcal{A} \wedge \neg \mathcal{B})$
T	T	T **T** T	**T** T F F T
T	F	T **F** F	**F** T T T F
F	T	F **T** T	**T** F F F T
F	F	F **T** F	**T** F F T F

We can similarly "define away" the biconditional and the disjunction without needing anything more than negation and conjunction. Show that this is true by creating sentences equivalent to each of the following without using anything other than negation and conjunction:

⋆ 1. $A \vee B$
⋆ 2. $A \leftrightarrow B$

We could have a language that is equivalent to $\mathcal{L}_S$ with only negation and disjunction as connectives. Show this: using negation, and disjunction, work out what sentences are logically equivalent to each of the following:

⋆ 3. $A \to B$
⋆ 4. $A \wedge B$
⋆ 5. $A \leftrightarrow B$

Part G
The *Sheffer stroke*, $|$, is a two-place logical connective with the following characteristic truthtable:

$\mathcal{A}$	$\mathcal{B}$	$(\mathcal{A} \mid \mathcal{B})$
T	T	**F**
T	F	**T**
F	T	**T**
F	F	**T**

6. Write a sentence using any of two of the connectives of $\mathcal{L}_S$ that is logically equivalent to '$(A|B)$.'

Interestingly, every sentence written using a connective of $\mathcal{L}_S$ can be rewritten as a logically equivalent sentence using one or more Sheffer strokes. This means that, strictly speaking, we only need *one* logical connective in our language without losing any expressive capabilities. (The resulting symbolizations do tend to be very complex, however.)

7. Using only the Sheffer stroke, work out what sentences are logically equivalent to '$\neg A$,' '$A \wedge B$,' '$A \vee B$,' '$A \to B$,' and '$A \leftrightarrow B$.'

Chapter 5

Proofs in Propositional Logic

In this supplementary chapter, we will look at the standard **proof theory** that's applicable when we're using $\mathcal{L}_S$. You will not be tested on anything in this chapter, but it's helpful stuff to know—especially if you want to continue your studies in logic!

Consider two arguments in $\mathcal{L}_S$:

Argument **A**		Argument **B**	
P1	$(P \vee Q)$	**P1**	$(P \to Q)$
P2	$\neg P$	**P2**	P
C	Q	**C**	Q

Clearly, these are valid arguments. You can confirm that they are valid by constructing four-line truth tables. Argument **A** makes use of an inference form that is always valid: given a disjunction and the negation of one of the disjuncts, the other disjunct follows as a valid consequence. This rule is called **disjunctive syllogism**.

Argument **B** makes use of a different valid form: given a conditional and its antecedent, the consequent follows as a valid consequence. This is called **modus ponens**.

When we construct truth tables, we do not need to give names to different inference forms. There is no reason to distinguish modus ponens from a disjunctive syllogism. For this same reason, however, the method of truth tables does not clearly show *why* an argument is valid. If you were to do a 1024-line truth table for an argument that contains ten sentence letters, then you could check to see if there were any lines on which the premises were all true and the conclusion were false. If you did not see such a line and provided you made no mistakes in constructing the table, then you would know that the argument was valid. Yet you would not be able to say anything further about why this particular argument was a valid argument form.

The aim of a **proof system** is to show that particular arguments are valid in a way that allows us to understand the reasoning involved in the argument. We begin with basic argument forms, like disjunctive syllogism and modus ponens. These forms can then be combined to make more complicated arguments, like this one:

P1 $(\neg L \rightarrow (J \vee L))$

P2 $\neg L$

C J

By modus ponens, P1 and P2 entail '$(J \vee L)$.' This is an **intermediate conclusion**. It follows logically from the premises, but it is not the conclusion we want. Now '$(J \vee L)$' and P2 entail the conclusion, 'J,' by disjunctive syllogism. We do not need a new rule for this argument. The proof of the argument shows that it is really just a combination of rules we have already introduced.

Formally, a **proof** is a sequence of sentences. The first sentences of the sequence are often called **assumptions**; these are the premises of a valid argument. Every sentence later in the sequence follows from earlier sentences by one of the rules of proof. The final sentence of the sequence is the conclusion of a valid argument.

Very generally, a **rule of proof** is a general rule that says: if you have already written down a line or lines with such-and-such a form in your proof, then you are permitted to write down a new line with such-and-such a form as well. Some rules of proof take you from a single line in a proof to a new line, while others require two or more lines before they can be applied. Every rule of proof outlined in this chapter corresponds to a valid inference. The idea is that we should eventually be able to chain together sequences of valid inferences using these rules of proof to be able to prove the validity of any valid argument.

This chapter develops a proof system for $\mathcal{L}_S$, which is then extended to cover $\mathcal{L}_P$ (and $\mathcal{L}_P$ with identity) in Chapter 9.

5.1 Basic rules for $\mathcal{L}_S$

In designing a proof system, we could just start with disjunctive syllogism and modus ponens. Whenever we discovered a valid argument which could not be proven with rules we already had, we could introduce new rules. Proceeding in this way, we would have an unsystematic grab bag of rules. We might accidentally add some strange rules, and we would surely end up with more rules than we need.

Instead, we will develop what is called a **natural deduction** system. In a natural deduction system, there will be two rules for each logical operator: an **introduction** rule that allows us to prove a sentence that has it as the main logical operator; and an **elimination** rule that allows us to prove something given a sentence that has it as the main logical operator.

In addition to the rules for each logical operator, we will also have a reiteration rule. If you already have shown something in the course of a proof, the reiteration rule allows you to repeat it on a new line. For instance:

1	$\mathcal{A}$	
2	$\mathcal{A}$	R 1

This says: if you have any sentence '$\mathcal{A}$' written down as a line in your proof, you're allowed to re-write it again on a separate (later) line. When we add a line to a proof, we write the rule that justifies that line. We also write the numbers of the lines to which the rule was applied. The reiteration rule above is justified by one line, the line that you are reiterating. So the 'R 1' on line 2 of the proof means that the line is justified by the reiteration rule (which we'll label 'R'), applied to line 1.

Obviously, the reiteration rule will never allow us to prove anything *new*. For that, we will need more rules. The remainder of this chapter will give introduction and elimination rules for each of the sentential connectives introduced in Chapter 2 and Chapter 3. This will give us a complete proof system for $\mathcal{L}_S$. All of the rules introduced in this chapter are summarized starting on p. 159.

Conjunction

Think for a moment: what would you need to show in order to prove $(E \wedge F)$?

Of course, you could show '$(E \wedge F)$' by proving 'E' and separately proving 'F.' This holds even if the two conjuncts are not atomic sentences. For instance, if you can prove '$((A \vee J) \rightarrow V)$' and '$((V \rightarrow L) \leftrightarrow (F \vee N))$,' then you have effectively proven the following as well: '

$$(((A \vee J) \rightarrow V) \wedge ((V \rightarrow L) \leftrightarrow (F \vee N)))$$

' So this will be our conjunction introduction rule, which we abbreviate $\wedge$I:

$$\begin{array}{r|ll} m & \mathcal{A} & \\ n & \mathcal{B} & \\ & (\mathcal{A} \wedge \mathcal{B}) & \wedge\mathrm{I}\ m,\ n \end{array}$$

This says: if you have two lines in your proof, one of which says '$\mathcal{A}$' and the other of which says '$\mathcal{B}$,' then you're allowed to write down on a new line of proof '$(\mathcal{A} \wedge \mathcal{B})$.' A line of proof must be justified by some rule, and for the new line we justify it using '$\wedge$I m, n.' This means: conjunction introduction applied to line m and line n. Here, we're using m and n as variables, not real line numbers; m is some line and n is some other line. In an actual proof, the lines are numbered $1, 2, 3, \ldots$ and rules must be applied to specific line numbers. When we define the rule, however, we use variables to underscore the point that the rule may be applied to any two lines that are already in the proof. If you have 'K' on line 8 and 'L' on line 15, you can prove '$(K \wedge L)$' at some later point in the proof with the justification '$\wedge$I 8, 15.'

Now, consider the elimination rule for conjunction. What are you entitled to conclude from a sentence like '$(E \wedge F)$'? Surely, you are entitled to conclude 'E'; if '$(E \wedge F)$' were true, then 'E' would be true. Similarly, you are entitled to conclude 'F.' This will be our conjunction elimination rule, which we abbreviate $\wedge$E:

$$\begin{array}{r|ll} m & (\mathcal{A} \wedge \mathcal{B}) & \\ & \mathcal{A} & \wedge\mathrm{E}\ m \\ & \mathcal{B} & \wedge\mathrm{E}\ m \end{array}$$

This says: if you have a line in your proof of the form '$(\mathcal{A} \wedge \mathcal{B})$,' then you're allowed to write down either '$\mathcal{A}$' or '$\mathcal{B}$' on a new line of proof below it. That is, when you have a conjunction on some line of a proof, you can use $\wedge$E to derive either of the conjuncts. The $\wedge$E rule requires only one line of proof to be applied, so we only ever write one line number as the justification.

Even with just these two rules, we can provide some proofs. Consider:

$$\begin{array}{rl} \textbf{P1} & ((A \vee B) \rightarrow (C \vee D)) \wedge ((E \vee F) \rightarrow (G \vee H)) \\ \hline \textbf{C} & ((E \vee F) \rightarrow (G \vee H)) \wedge ((A \vee B) \rightarrow (C \vee D)) \end{array}$$

The main logical operator in both the premise and conclusion is conjunction. Since conjunction is symmetric, the argument is obviously valid. In order to provide a proof, we begin by writing down the premises—in this case, there's only one. After the premises, we draw a horizontal line—everything below this line must be justified by a rule of proof. So the beginning of the proof is:

1	$((A \vee B) \to (C \vee D)) \wedge ((E \vee F) \to (G \vee H))$	

From the premise, we can get each of the conjuncts by $\wedge$E:

1	$((A \vee B) \to (C \vee D)) \wedge ((E \vee F) \to (G \vee H))$	
2	$((A \vee B) \to (C \vee D))$	$\wedge$E 1
3	$((E \vee F) \to (G \vee H))$	$\wedge$E 1

The rule $\wedge$I requires that we have each of the conjuncts available somewhere in the proof. They can be separated from one another, and they can appear in any order. So by applying the $\wedge$I rule to lines 3 and 2, we arrive at the desired conclusion. The finished proof looks like this:

1	$((A \vee B) \to (C \vee D)) \wedge ((E \vee F) \to (G \vee H))$	
2	$((A \vee B) \to (C \vee D))$	$\wedge$E 1
3	$((E \vee F) \to (G \vee H))$	$\wedge$E 1
4	$((E \vee F) \to (G \vee H)) \wedge ((A \vee B) \to (C \vee D))$	$\wedge$I 3, 2

This proof is trivial, but it shows how we can use rules of proof together to demonstrate the validity of an argument form. Also: using a truth table to show that this argument is valid would have required a staggering 256 lines, since there are eight sentence letters in the argument.

Disjunction

If 'M' were true, then '$(M \vee N)$' would also be true. This can be easily checked by $\vee$'s truth table: a disjunction is true whenever at least one of its disjuncts is true. So the disjunction introduction rule ($\vee$I) allows us to derive a disjunction if we have assumed the truth of either one of the two disjuncts:

m	$\mathcal{A}$	
	$(\mathcal{A} \vee \mathcal{B})$	$\vee$I m
	$(\mathcal{B} \vee \mathcal{A})$	$\vee$I m

This says: if you have a sentence '$\mathcal{A}$' written down, you are allowed to write on a new line another sentence of either the form '$(\mathcal{A} \vee \mathcal{B})$,' or '$(\mathcal{B} \vee \mathcal{A})$.' In this case, '$\mathcal{B}$' can be any sentence whatsoever. So the following is a legitimate proof:

1	M	
2	$(M \vee (((A \leftrightarrow B) \rightarrow (C \wedge D)) \leftrightarrow (E \wedge F)))$	$\vee$I 1

It may seem odd that just by knowing 'M' we can derive a conclusion that includes things like 'A,' 'B,' and the rest—sentences that intuitively have nothing to do with 'M.' Yet the conclusion follows immediately by $\vee$I. This is as it should be: the truth conditions for the disjunction mean that, if '$\mathcal{A}$ is true, then '$(\mathcal{A} \vee \mathcal{B})$' is true regardless of what '$\mathcal{B}$' is. So the conclusion could not be false if the premise were true; the argument is valid.

Now consider the disjunction elimination rule. What can you conclude from '$(M \vee N)$'? You cannot conclude 'M.' It might be 'M' truth that makes '$(M \vee N)$' true, as in the example above, but it might not. From '$(M \vee N)$' alone, you cannot conclude anything about either 'M' or 'N' specifically. If you also knew that 'N' was false, however, then you would be able to conclude 'M.'

This is just disjunctive syllogism, it will be the disjunction elimination rule ($\vee$E). Strictly speaking, there are two rules here, depending on what disjunct is being negated:

m	$(\mathcal{A} \vee \mathcal{B})$	
n	$\neg\mathcal{B}$	
	$\mathcal{A}$	$\vee$E m, n

m	$(\mathcal{A} \vee \mathcal{B})$	
n	$\neg\mathcal{A}$	
	$\mathcal{B}$	$\vee$E m, n

Each rule says: if you have a line of proof of the form '$(\mathcal{A} \vee \mathcal{B})$,' and another line of proof of the form '$\neg\mathcal{A}$' (or '$\neg\mathcal{B}$'), you're allowed to write '$\mathcal{B}$' (or '$\mathcal{A}$') down on a new line.

Conditional

Consider this argument:

P1 $(R \vee F)$

C $(\neg R \rightarrow F)$

The argument is certainly a valid one. What should the conditional introduction rule be, such that we can draw this conclusion?

We begin the proof by writing down the premise of the argument and drawing a horizontal line, like this:

$$\begin{array}{r|l}
1 & (R \lor F) \\ \cline{2-2}
\end{array}$$

If we had '$\neg R$' as a further premise, we could derive 'F' by the $\lor$E rule. We do not have '$\neg R$' as a premise of this argument, nor can we derive it directly from the premise we do have—so we cannot simply prove 'F.' What we will do instead is start a **subproof**, a proof within the main proof. When we start a subproof, we draw another vertical line to indicate that we are no longer in the main proof. Then we write in an assumption for the subproof. This can be anything we want. Here, it will be helpful to assume '$\neg R$.' (Over time, you'll get the hang of what you should assume for subproofs.) Our proof now looks like this:

$$\begin{array}{r|ll}
1 & (R \lor F) & \\ \cline{2-3}
2 & & \multicolumn{1}{|l}{\neg R} \\ \cline{3-3}
\end{array}$$

It is important to notice that we are not claiming to have proven '$\neg R$.' We do not need to write in any justification for the assumption line of a subproof. You can think of the subproof as posing the question: what could we show *if* we assumed that '$\neg R$' were true? For one thing, we can derive 'F.' So we do:

$$\begin{array}{r|lll}
1 & (R \lor F) & & \\ \cline{2-3}
2 & & \multicolumn{1}{|l}{\neg R} & \\ \cline{3-3}
3 & & \multicolumn{1}{|l}{F} & \lor\text{E } 1, 2 \\
\end{array}$$

This has shown that *if* we had '$\neg R$' as a premise, *then* we could prove 'F.' In effect, we have proven '$(\neg R \to F)$.' So the conditional introduction rule ($\to$I) will allow us to close the subproof and derive '$(\neg R \to F)$' in the main proof. Our final proof looks like this:

$$\begin{array}{r|lll}
1 & (R \lor F) & & \\ \cline{2-3}
2 & & \multicolumn{1}{|l}{\neg R} & \\ \cline{3-3}
3 & & \multicolumn{1}{|l}{F} & \lor\text{E } 1, 2 \\
4 & (\neg \mathrm{R} \to \mathrm{F}) & & \to\text{I } 2\text{–}3 \\
\end{array}$$

Notice that the justification for applying the $\rightarrow$I rule is the entire subproof. Usually that will be more than just two lines.

It may seem as if the ability to assume anything at all in a subproof would lead to chaos: Does it allow you to prove any conclusion from any premises? The answer is no, it does not. Consider this proof:

$$\begin{array}{r l l}
1 & \mid \underline{\mathcal{A}} & \\
2 & \mid \quad \mid \underline{\mathcal{B}} & \\
3 & \mid \quad \mid \mathcal{B} & \text{R } 2
\end{array}$$

It may seem as if this is a proof that you can derive any conclusions '$\mathcal{B}$' from any premise '$\mathcal{A}$.' When the vertical line for the subproof ends, the subproof is *closed*. In order to complete a proof, you must close all of the subproofs. And you cannot close the subproof and use the R rule again on line 4 to derive '$\mathcal{B}$' in the main proof. Once you close a subproof, you cannot refer back to individual lines inside it.

Closing a subproof is called *discharging* the assumptions of that subproof. So we can put the point this way: You cannot complete a proof until you have discharged all of the assumptions besides the original premises of the argument.

Of course, it is legitimate to do this:

$$\begin{array}{r l l}
1 & \mid \underline{\mathcal{A}} & \\
2 & \mid \quad \mid \underline{\mathcal{B}} & \\
3 & \mid \quad \mid \mathcal{B} & \text{R } 2 \\
4 & \mid (\mathcal{B} \rightarrow \mathcal{B}) & \rightarrow\text{I } 2\text{–}3
\end{array}$$

This should not seem so strange, though. Since '$(\mathcal{B} \rightarrow \mathcal{B})$' is a tautology, no particular premises should be required to validly derive it. (Indeed, as we will see, a tautology follows from any premises.)

Put in a general form, the $\rightarrow$I rule looks like this:

$$\begin{array}{r l l}
m & \mid \quad \mid \underline{\mathcal{A}} & \text{want } \mathcal{B} \\
n & \mid \quad \mid \mathcal{B} & \\
 & \mid (\mathcal{A} \rightarrow \mathcal{B}) & \rightarrow\text{I } m\text{–}n
\end{array}$$

This says that if you are able to derive a sentence '$\mathcal{B}$' from a subproof assumption

'$\mathcal{A}$,' then you are allowed to write down on a later line of proof '$(\mathcal{A} \to \mathcal{B})$.' When we introduce a subproof, we typically write what we want to derive in the column. This is just so that we do not forget why we started the subproof if it goes on for five or ten lines. There is no 'want' rule. It is a note to ourselves and not formally part of the proof.

Although it is always permissible to open a subproof with any assumption you please, there is some strategy involved in picking a useful assumption. Starting a subproof with an arbitrary, wacky assumption would just waste lines of the proof. In order to derive a conditional by the →I, for instance, you must assume the antecedent of the conditional in a subproof. In general, if you want to prove a conditional, it's usually a good strategy to start by assuming the antecedent of that conditional.

The →I rule also requires that the consequent of the conditional be the last line of the subproof. It is always permissible to close a subproof and discharge its assumptions, but it will not be helpful to do so until you get what you want.

Now consider the conditional elimination rule. Nothing follows from '$(M \to N)$' alone, but if we have both '$(M \to N)$' *and* 'M,' then we can conclude 'N.' This rule, *modus ponens*, will be the conditional elimination rule (→E).

m	$(\mathcal{A} \to \mathcal{B})$	
n	$\mathcal{A}$	
	$\mathcal{B}$	→E m, n

That is, if you have two lines of proof, one of the form '$(\mathcal{A} \to \mathcal{B})$' and the other of the form '$\mathcal{A}$' (i.e., the same as the antecedent in '$(\mathcal{A} \to \mathcal{B})$'), then you're allowed to write down '$\mathcal{B}$' in a later line of proof.

Now that we have rules for the conditional, consider this argument:

P1 $(P \to Q)$

P2 $(Q \to R)$

———

C $(P \to R)$

We begin the proof by writing the two premises as assumptions. Since the main logical operator in the conclusion is a conditional, we can expect to use the →I rule. For that, we need a subproof—so we write in the antecedent of the conditional as assumption of a subproof:

1	$(P \to Q)$	
2	$(Q \to R)$	
3	$\quad P$	

We made 'P' available by assuming it in a subproof, allowing us to use $\to$E on the first premise. This gives us 'Q,' which allows us to use $\to$E on the second premise. Having derived 'R,' we close the subproof. By assuming 'P' we were able to prove 'R,' so we apply the $\to$I rule and finish the proof.

1	$(P \to Q)$	
2	$(Q \to R)$	
3	$\quad P$	want R
4	$\quad Q$	$\to$E 1, 3
5	$\quad R$	$\to$E 2, 4
6	$(P \to R)$	$\to$I 3–5

Biconditional

The rules for the biconditional will be like double-barreled versions of the rules for the conditional.

In order to derive '$(W \leftrightarrow X)$,' for instance, you must be able to prove 'X' by assuming 'W' *and* prove 'W' by assuming 'X.' The biconditional introduction rule ($\leftrightarrow$I) requires two subproofs. The subproofs can come in any order, and the second subproof does not need to come immediately after the first—but schematically, the rule works like this:

m	$\quad \mathcal{A}$	want $\mathcal{B}$
n	$\quad \mathcal{B}$	
p	$\quad \mathcal{B}$	want $\mathcal{A}$
q	$\quad \mathcal{A}$	
	$(\mathcal{A} \leftrightarrow \mathcal{B})$	$\leftrightarrow$I m–n, p–q

The biconditional elimination rule ($\leftrightarrow$E) lets you do a bit more than the conditional rule. If you have the left-hand side of the biconditional, you can derive the right-hand side. If you have the right-hand side, you can derive the left-hand side. This is the rule:

m	$(\mathscr{A} \leftrightarrow \mathscr{B})$	
n	$\mathscr{A}$	
	$\mathscr{B}$	$\leftrightarrow$E m, n

m	$(\mathscr{A} \leftrightarrow \mathscr{B})$	
n	$\mathscr{B}$	
	$\mathscr{A}$	$\leftrightarrow$E m, n

Negation

Here is a simple mathematical argument in English:

P1 Assume there is some greatest natural number. Call it α.
P2 That number plus one is also a natural number.
P3 Obviously, $a + 1 > a$.
P4 So there is a natural number greater than α.
P5 This is impossible, as α is assumed to be the greatest natural number.

C There is no greatest natural number.

This argument form is traditionally called a *reductio*. Its full Latin name is *reductio ad absurdum*, which means 'reduction to absurdity.' In a reductio, we assume something for the sake of argument—for example, that there is a greatest natural number. Then we show that the assumption leads to two contradictory sentences—for example, that α is the greatest natural number and that it is not. In this way, we show that the original assumption must have been false.

The basic rules for negation will allow for arguments like this. If we assume something and show that it leads to contradictory sentences, then we have proven the negation of the assumption. This is the negation introduction ($\neg$I) rule:

m	$\quad\mid \mathscr{A}$	for reductio
n	$\quad\mid \mathscr{B}$	
$n+1$	$\quad\mid \neg\mathscr{B}$	
$n+2$	$\neg\mathscr{A}$	$\neg$I m–$n+1$

That is, if you can derive a contradiction '($\mathscr{B}$ and $\neg\mathscr{B}$)' from a subproof assumption '$\mathscr{A}$,' then you can write down on a later line of proof '$\neg\mathscr{A}$.' For the rule to apply, the last two lines of the subproof must be an **explicit contradiction**: some sentence followed on the next line by its negation. We write 'for reductio' as a note to ourselves, a reminder of why we started the subproof. It is not formally part of the proof, and you can leave it out if you find it distracting.

To see how the rule works, suppose we want to prove the law of non-contradiction: '$\neg(G \wedge \neg G)$.' We can prove this without any premises by immediately starting a subproof. We want to apply $\neg$I to the subproof, so we assume '$(G \wedge \neg G)$.' We then get an explicit contradiction by $\wedge$E. The proof looks like this:

1	$(G \wedge \neg G)$	for reductio
2	G	$\wedge$E 1
3	$\neg G$	$\wedge$E 1
4	$\neg(G \wedge \neg G)$	$\neg$I 1–3

The $\neg$E rule will work in much the same way. If we assume '$\neg\mathcal{A}$' and show that it leads to a contradiction, we have effectively proven '$\mathcal{A}$.' So the rule looks like this:

m	$\neg\mathcal{A}$	for reductio
n	$\mathcal{B}$	
$n+1$	$\neg\mathcal{B}$	
$n+2$	$\mathcal{A}$	$\neg$E m–$n+1$

5.2 Derived rules

The rules of the natural deduction system are meant to be systematic. There is an introduction and an elimination rule for each logical operator, but why these basic rules rather than some others? Many natural deduction systems have a disjunction elimination rule that works like this:

m	$(\mathcal{A} \vee \mathcal{B})$	
n	$(\mathcal{A} \rightarrow \mathcal{C})$	
o	$(\mathcal{B} \rightarrow \mathcal{C})$	
	$\mathcal{C}$	DIL m, n, o

Let's call this rule *Dilemma* (DIL) It might seem as if there will be some proofs that we cannot do with our proof system, because we do not have this as a basic rule. Yet this is not the case. Any proof that you can do using the Dilemma rule can be done with basic rules of our natural deduction system. Consider this proof:

1	$(\mathcal{A} \vee \mathcal{B})$	
2	$(\mathcal{A} \to \mathcal{C})$	
3	$(\mathcal{B} \to \mathcal{C})$	want $\mathcal{C}$
4	$\quad \neg\mathcal{C}$	for reductio
5	$\quad\quad \mathcal{A}$	for reductio
6	$\quad\quad \mathcal{C}$	$\to$E 2, 5
7	$\quad\quad \neg\mathcal{C}$	R 4
8	$\quad \neg\mathcal{A}$	$\neg$I 5–7
9	$\quad\quad \mathcal{B}$	for reductio
10	$\quad\quad \mathcal{C}$	$\to$E 3, 9
11	$\quad\quad \neg\mathcal{C}$	R 4
12	$\quad \mathcal{B}$	$\vee$E 1, 8
13	$\quad \neg\mathcal{B}$	$\neg$I 9–11
14	$\mathcal{C}$	$\neg$E 4–13

'$\mathcal{A}$,' '$\mathcal{B}$,' and '$\mathcal{C}$' are meta-variables. They are not symbols of $\mathcal{L}_S$, but stand-ins for arbitrary sentences symbolized in $\mathcal{L}_S$. So this is not, strictly speaking, a proof in $\mathcal{L}_S$. It is more like a recipe. It provides a pattern that can prove anything that the Dilemma rule can prove, using only the basic rules of $\mathcal{L}_S$. This means that the Dilemma rule is not really necessary. Adding it to the list of basic rules would not allow us to derive anything that we could not derive without it.

Nevertheless, the Dilemma rule would be convenient. It would allow us to do in one line what requires eleven lines and several nested subproofs with the basic rules. So we will add it to the proof system as a **derived rule**. A derived rule is a rule of proof that does not make any new proofs possible. Anything that can be proven with a derived rule can be proven without it. You can think of a short proof using a derived rule as shorthand for a longer proof that uses only the basic rules. Any time you use the Dilemma rule, you could always take ten extra lines and prove the same thing without it.

For the sake of convenience, we will add several other derived rules. One is *modus tollens* (MT).

m	$(\mathcal{A} \to \mathcal{B})$	
n	$\neg\mathcal{B}$	
	$\neg\mathcal{A}$	MT m, n

We leave the proof of this rule as an exercise. Note that if we had already proven the MT rule, then the proof of the DIL rule could have been done in only five lines.

We also add *hypothetical syllogism* (HS) as a derived rule. We have already given a proof of it on p. 62.

m	$(\mathcal{A} \to \mathcal{B})$	
n	$(\mathcal{B} \to \mathcal{C})$	
	$(\mathcal{A} \to \mathcal{C})$	HS m, n

5.3 Rules of replacement

Consider how you would prove this argument:

P1 $(F \to (G \land H))$

C $(F \to G)$

Perhaps it is tempting to write down the premise and apply the $\land$E rule to the conjunction '$(G \land H)$.' This is impermissible, however, because the basic rules of proof can only be applied to whole sentences. We need to get '$(G \land H)$' on a line by itself. We can prove the argument in this way:

1	$(F \to (G \land H))$	
2	$\quad F$	want G
3	$\quad (G \land H)$	$\to$E 1, 2
4	$\quad G$	$\land$E 3
5	$(F \to G)$	$\to$I 2–4

We will now introduce some derived rules that may be applied to a *part* of a sentence. These are called **rules of replacement**, because they can be used to replace part of a sentence with a logically equivalent expression. One simple rule of replacement is *commutivity* (abbreviated Comm), which says that we can swap the order of conjuncts in a conjunction or the order of disjuncts in a disjunction. We define the rule this way:

$$\begin{aligned}(\mathcal{A} \wedge \mathcal{B}) &\Longleftrightarrow (\mathcal{B} \wedge \mathcal{A})\\ (\mathcal{A} \vee \mathcal{B}) &\Longleftrightarrow (\mathcal{B} \vee \mathcal{A})\\ (\mathcal{A} \leftrightarrow \mathcal{B}) &\Longleftrightarrow (\mathcal{B} \leftrightarrow \mathcal{A}) \quad \text{Comm}\end{aligned}$$

The bold arrow means that you can take a subformula on one side of the arrow and replace it with the subformula on the other side. The arrow is double-headed because rules of replacement work in both directions.

Consider this argument:

$$\begin{array}{rl} \textbf{P1} & ((M \vee P) \to (P \wedge M)) \\ \hline \textbf{C} & ((P \vee M) \to (M \wedge P)) \end{array}$$

It is possible to give a proof of this using only the basic rules, but it will be long and inconvenient. With the Comm rule, we can provide a proof easily:

$$\begin{array}{r|ll} 1 & ((M \vee P) \to (P \wedge M)) & \\ \hline 2 & ((P \vee M) \to (P \wedge M)) & \text{Comm } 1 \\ 3 & ((P \vee M) \to (M \wedge P)) & \text{Comm } 2 \end{array}$$

Another useful rule of replacement is *double negation* (DN). With the DN rule, you can remove or insert a pair of negations anywhere in a symbolized sentence. This is the rule:

$$\neg\neg\mathcal{A} \Longleftrightarrow \mathcal{A} \quad \text{DN}$$

Two more replacement rules are called *De Morgan's Laws*, named for the 19th-century British logician August De Morgan. (Although De Morgan did discover these laws, he was not the first to do so.) The rules capture useful relations between negation, conjunction, and disjunction. Here are the rules, which we abbreviate DeM:

$$\begin{aligned}\neg(\mathcal{A} \vee \mathcal{B}) &\Longleftrightarrow (\neg\mathcal{A} \wedge \neg\mathcal{B})\\ \neg(\mathcal{A} \wedge \mathcal{B}) &\Longleftrightarrow (\neg\mathcal{A} \vee \neg\mathcal{B}) \quad \text{DeM}\end{aligned}$$

Because '$(\mathcal{A} \to \mathcal{B})$' is a *material conditional*, it is equivalent to '$(\neg\mathcal{A} \vee \mathcal{B})$.' A further replacement rule captures this equivalence. We abbreviate the rule MC, for 'material conditional.' It takes two forms:

$$(\mathcal{A} \to \mathcal{B}) \Longleftrightarrow (\neg\mathcal{A} \lor \mathcal{B})$$
$$(\mathcal{A} \lor \mathcal{B}) \Longleftrightarrow (\neg\mathcal{A} \to \mathcal{B}) \quad \text{MC}$$

Now consider this argument:

P1 $\neg(P \to Q)$

C $(P \land \neg Q)$

As always, we could prove this argument using only the basic rules. With rules of replacement, though, the proof is much simpler:

1	$\neg(P \to Q)$	
2	$\neg(\neg P \lor Q)$	MC 1
3	$(\neg\neg P \land \neg Q)$	DeM 2
4	$(P \land \neg Q)$	DN 3

A final replacement rule captures the relation between conditionals and biconditionals. We will call this rule *biconditional exchange* and abbreviate it $\leftrightarrow$ex.

$$((\mathcal{A} \to \mathcal{B}) \land (\mathcal{B} \to \mathcal{A})) \Longleftrightarrow (\mathcal{A} \leftrightarrow \mathcal{B}) \quad \leftrightarrow\text{ex}$$

5.4 Proof strategy

There is no simple recipe for proofs, and there is no substitute for practice. Here, though, are some rules of thumb and strategies to keep in mind.

Work backwards from what you want. The ultimate goal is to derive the conclusion. Look at the conclusion and ask what the introduction rule is for its main logical operator. This gives you an idea of what should happen *just before* the last line of the proof. Then you can treat this line as if it were your goal. Ask what you could do to derive this new goal.

For example: If your conclusion is a conditional '$(\mathcal{A} \to \mathcal{B})$,' plan to use the $\to$I rule. This requires starting a subproof in which you assume '$\mathcal{A}$.' In the subproof, you want to derive '$\mathcal{B}$.'

Work forwards from what you have. When you are starting a proof, look at the premises; later, look at the things that you have derived so far. Think about the elimination rules for the main logical connectives. These will tell you what your options are.

For a short proof, you might be able to eliminate the premises and introduce the conclusion. A long proof is formally just a number of short proofs linked together, so you can fill the gap by alternately working back from the conclusion and forward from the premises.

Change what you are looking at. Replacement rules can often make your life easier. If a proof seems impossible, try out some different substitutions. For example: It is often difficult to prove a disjunction using the basic rules. If you want to show '$(\mathcal{A} \vee \mathcal{B})$,' it is often easier to show '$(\neg\mathcal{A} \rightarrow \mathcal{B})$' and use the MC rule.

Some replacement rules should become second nature. If you see a negated disjunction, for instance, you should immediately think of DeMorgan's rule.

Do not forget indirect proof. If you cannot find a way to show something directly, try assuming its negation. Remember that most proofs can be done either indirectly or directly. One way might be easier—or perhaps one sparks your imagination more than the other—but either one is formally legitimate.

Repeat as necessary. Once you have decided how you might be able to get to the conclusion, ask what you might be able to do with the premises. Then consider the target sentences again and ask how you might reach them.

Persist. Try different things. If one approach fails, then try something else.

Practice Exercises

⋆ Part A

Provide a justification (rule and line numbers) for each line of proof that requires one.

$$\begin{array}{r|l}
1 & (W \to \neg B)\\
2 & (A \wedge W)\\
3 & (B \vee (J \wedge K))\\ \hline
4 & W\\
5 & \neg B\\
6 & (J \wedge K)\\
7 & K
\end{array}$$

$$\begin{array}{r|l}
1 & (L \leftrightarrow \neg O)\\
2 & (L \vee \neg O)\\ \hline
3 & \quad \vert\ \neg L\\
4 & \quad \vert\ \neg O\\
5 & \quad \vert\ L\\
6 & \quad \vert\ \neg L\\
7 & L
\end{array}$$

$$\begin{array}{r|l}
1 & (Z \to (C \wedge \neg N))\\
2 & (\neg Z \to (N \wedge \neg C))\\ \hline
3 & \quad \vert\ \neg(N \vee C)\\
4 & \quad \vert\ (\neg N \wedge \neg C)\\
5 & \quad \vert\ \quad \vert\ Z\\
6 & \quad \vert\ \quad \vert\ (C \wedge \neg N)\\
7 & \quad \vert\ \quad \vert\ C\\
8 & \quad \vert\ \quad \vert\ \neg C\\
9 & \quad \vert\ \neg Z\\
10 & \quad \vert\ (N \wedge \neg C)\\
11 & \quad \vert\ N\\
12 & \quad \vert\ \neg N\\
13 & (N \vee C)
\end{array}$$

⋆ Part B

Give a proof for each argument in $\mathcal{L}_S$.

1. $(K \wedge L)$
   therefore $(K \leftrightarrow L)$

2. $(A \to (B \to C))$
   therefore $((A \wedge B) \to C)$

3. $(P \wedge (Q \vee R))$
   $(P \to \neg R)$
   therefore $(Q \vee E)$

4. $((C \wedge D) \vee E)$
   therefore $(E \vee D)$

5. $(\neg F \rightarrow G)$
   $(F \rightarrow H)$
   therefore $(G \vee H)$

6. $((X \wedge Y) \vee (X \wedge Z))$
   $\neg(X \wedge D)$
   $(D \vee M)$ therefore M

Part C
Give a proof for each argument in $\mathcal{L}_S$.

1. $(Q \rightarrow (Q \wedge \neg Q))$
   therefore $\neg Q$

2. $(J \rightarrow \neg J)$
   therefore $\neg J$

3. $(E \vee F)$
   $(F \vee G)$
   $\neg F$
   therefore $(E \wedge G)$

4. $(A \leftrightarrow B)$
   $(B \leftrightarrow C)$
   therefore $(A \leftrightarrow C)$

5. $(M \vee (N \rightarrow M))$
   therefore $(\neg M \rightarrow \neg N)$

6. $(S \leftrightarrow T)$
   therefore $(S \leftrightarrow (T \vee S))$

7. $((M \vee N) \wedge (O \vee P))$, $(N \rightarrow P)$
   $\neg P$, therefore $(M \wedge O)$

8. $((Z \wedge K) \vee (K \wedge M))$, $(K \rightarrow D)$
   therefore D

Part D
Show that each of the following pairs are provably equivalent in $\mathcal{L}_S$.

1. $\neg\neg\neg\neg G$
   G

2. $(T \to S)$
   $(\neg S \to \neg T)$

3. $(R \leftrightarrow E)$
   $(E \leftrightarrow R)$

4. $(\neg G \leftrightarrow H)$
   $\neg(G \leftrightarrow H)$

5. $(U \to I)$
   $\neg(U \wedge \neg I)$

Part E
Provide proofs to show each of the following.

1. $(M \wedge (\neg N \to \neg M))$ implies $((N \wedge M) \vee \neg M)$
2. $(C \to (E \wedge G))$ and $(\neg C \to G)$ imply G
3. $((Z \wedge K) \leftrightarrow (Y \wedge M))$ and $(D \wedge (D \to M))$ imply $(Y \to Z)$
4. $((W \vee X) \vee (Y \vee Z))$, $(X \to Y)$ and $\neg Z$ imply $(W \vee Y)$

Part F
For the following, provide proofs using only the basic rules. The proofs will be longer than proofs of the same claims would be using the derived rules.

1. Show that MT is a legitimate derived rule. Using only the basic rules, prove the following: $(\mathcal{A} \to \mathcal{B})$, $\neg\mathcal{B}$, therefore $\neg\mathcal{A}$
2. Show that Comm is a legitimate rule for the biconditional. Using only the basic rules, prove that '$(\mathcal{A} \leftrightarrow \mathcal{B})$' and '$(\mathcal{B} \leftrightarrow \mathcal{A})$' are equivalent.
3. Using only the basic rules, prove the following instance of DeMorgan's Laws: $(\neg A \wedge \neg B)$, therefore $\neg(A \vee B)$
4. Show that $\leftrightarrow$ex is a legitimate derived rule. Using only the basic rules, prove that '$(D \leftrightarrow E)$' and '$((D \to E) \wedge (E \to D))$' are equivalent.

Part II

Quantificational Logic

Chapter 6

[$\mathcal{Req}$.] Names, Predicates, and Quantifiers

Consider the following argument:

P1 Willard is a logician.
P2 All logicians are pedantic.

C Willard is pedantic.

There is no possible way that both of the premises can be true and the conclusion false, so the argument is clearly valid. However, suppose we try to symbolize it in $\mathcal{L}_S$, using the following symbolization key:

L : Willard is a logician.
A : All logicians are pedantic.
P : Willard is pedantic.

The choice of the particular letters 'L', 'A' and 'P' is immaterial—what matters is that each of the three sentences does not contain an atomic sentence as a proper part, and so cannot be symbolized in $\mathcal{L}_S$ as complex sentences. The argument would then be represented as follows:

P1 L
P2 A

C P

It's easy to see that this argument doesn't have a valid form when symbolized in $\mathcal{L}_S$.

This situation is unfortunate. To be sure, we noted in §1.3 that there are some valid *arguments* with non-valid *argument forms*, so we should sometimes expect that our logic won't capture everything there is to validity. However, it would be nice if there were a way to capture what's clearly valid about the present argument—especially since, in this case, the argument's validity *does* seem to have something to do with its logical form.

For instance, the argument seems to have the same structure as the following valid argument:

P1 Froggy is a frog.
P2 All frogs are green.

C Froggy is green.

The first premise of the 1^{st} argument says that some individual (Willard) is a member of a kind (logician). The second premise says that all members of that kind (logicians) have some property (they are pedantic). The conclusion says that the individual, therefore, also has that property (Willard is pedantic). Likewise, the first premise of the 2^{nd} argument says that an individual (Froggy) is a member of a kind (frogs). The second premise says that all members of that kind (frogs) have some property (greenness). The conclusion says that the individual, therefore, also has that property (Froggy is green).

What this tells us is that the method of symbolization in $\mathcal{L}_S$ leaves out some important logical structure. In particular, the symbolization does not allow us to represent the **internal logical structure** of the atomic sentences which make up the premises and the conclusions of these two arguments. The *smallest unit of analysis* in $\mathcal{L}_S$ is the atomic sentence. $\mathcal{L}_S$ is therefore very good for symbolizing the way in which atomic sentences can be joined together to form non-atomic sentences using truth-functional connectives. Sometimes, however, we need to consider the internal structure of the atomic sentences themselves in order to see what makes an argument have a valid form.

It is therefore useful to have a way of representing the internal structure of atomic sentences. For this we need a more sophisticated formal language, which we'll label $\mathcal{L}_P$. The 'P' stands for *predicate*, which is a concept that we'll introduce shortly.

Luckily, while we're developing our new language $\mathcal{L}_P$, we can keep what we've learnt so far while developing $\mathcal{L}_S$. Coming up with a new formal language doesn't mean throwing out all our hard work with propositional logic and replacing it with something entirely different. Instead, we're going to *supplement* the old language with some new tools and techniques. In the end, we'll want a symbol system which includes *names*, *predicates*, *quantifiers* and *variables.*

6.1 Names and Individuals

For the purposes of logic, an **individual** is any *specific* person, place, or thing. So, for example, you and I are individuals, as are Frank Jackson, Mr. T, and Hilary Clinton. Each is an individual person, and hence an individual. London and Sydney are also individuals: they are individual cities, and (by virtue of this) they are individuals. This book is another kind of individual. And in the arguments above, Willard and Froggy were individuals.

In English, we have many different ways of referring to individuals. For example, we can use the *proper name* 'John Cleese' to refer to the individual John Cleese. Or, we could use a description which uniquely characterises John Cleese, such as 'the tallest member of Monty Python'. Descriptions which uniquely pick out an individual are called *definite descriptions*, and they usually have the form 'the so-and-so'. If he were around, we could refer to him by pointing and using a *demonstrative* like 'him' or 'that person there.' Or, if we were talking to him, we could refer to him using the simple pronoun 'you'.

In $\mathcal{L}_P$, we will use **proper names** to refer directly to individuals. We will use the italicised lower-case letters a through w for our names. We can add subscripts if needed. So $a, b, c, \ldots w, a_1, f_{32}, j_{390}$, and m_{12} are all proper names in $\mathcal{L}_P$, and each must refer to exactly one individual. In the jargon, these letters are often called **constants** because they pick out just one individual. This is to contrast them with x, y, and z, which are *not* constants in $\mathcal{L}_P$. We reserve those letters for **variables**, which we will describe later when we introduce quantifiers.

> An INDIVIDUAL is a specific person, place, or thing. In $\mathcal{L}_P$, we use NAMES (also called CONSTANTS) to refer to specific individuals.

The symbolization of definite descriptions into a formal language like $\mathcal{L}_P$ can be a contentious matter. On some philosophical theories, we should treat definite descriptions just like we've treated proper names. That is, they should be symbolized as constants. However, since Bertrand Russell's seminal theory of definite descriptions in 1905, most logicians and philosophers have treated definite descriptions in a very different way. For now, we will avoid any examples involving definite descriptions, and we will come back to their symbolization in the §7.6.

An closely related issue concerns *non-referring names*. In English, we have lots of proper names that do not refer to any existing thing. For example (and, I'm sorry to break it to you) Santa Claus does not exist—so the name 'Santa Claus' is non-referring. Dealing with non-referring names is a tricky matter that philosophers still debate about. For our purposes, though, we don't have to worry about the problem of non-referring names. As we said above, every constant in our language $\mathcal{L}_P$ must refer to exactly one individual—no more, and *no less*. In other words, we will assume that there are no non-referring names in our logical language.

6.2 Predicates, Properties, and Relations

Properties are attributes or qualities had by individuals. For instance, you and I have the property of *being human*, and neither of us have the property of *being green*. (I hope!) We also have the properties of *being living things*, *being mammals*, *being Earth-dwelling things*, and so on. A dog shares many of these properties with us, but not all of them—clearly, a dog doesn't have the property of *being human*, and humans don't have the property of *being a dog*.

In English, we use **one-place predicates** refer to properties. So, for example, the phrase 'is red' is a one-place predicate that picks out the property of *being red*. (The reason for the 'one-place' will become clearer below, when we discuss relations.) In $\mathcal{L}_P$, we symbolize predicates using italicised capital letters; A, B, C, etc., with or without subscripts. Because of this, logics which use a formal language like $\mathcal{L}_P$ are sometimes called *predicate logic*.

Now, we also used italicised capital letters in $\mathcal{L}_S$ to refer to whole sentences. However, there is no danger of confusion in also using these letters here for predicates, because in $\mathcal{L}_P$ we will not have any symbols for atomic sentences. Instead, every whole sentence will be symbolized in terms of its parts.

So we now have symbols for properties and individuals. We can use these to symbolize some simple atomic sentences.

1. Abbie is angry.
2. Bob is red.
3. Bob is angry.

Let's use 'a' as our logical name for Abbie, and 'b' as our name for Bob; and we'll use 'A' for the property of *being angry* and 'R' for the property of *being red*. By convention, we always put the predicate out the front. Hence, we should translate the first sentence 1 into 'Aa.' Likewise, 2 gets symbolized as 'Rb,' and 3 gets symbolized as 'Ab.'

We will continue to use the logical connectives that we developed for $\mathcal{L}_S$ in

Chapter 2 and Chapter 3. This means that we can also start symbolizing some non-atomic sentences:

4. Abbie is angry and Bob is red.
5. If Abbie is angry then Abbie is red.

Hence, sentence 4 can be symolized as '$Aa \wedge Rb$,' and sentence 5 as '$Aa \rightarrow Ra$.'

Besides properties, we also need a way of talking about **relations**. Relations are ways in which individuals can be connected to one another. A **two-place** (or **dyadic**, or **binary**) relation connects two individuals. For example, if we say that 'Frank is married to Jane', we are saying that Frank is connected to Jane via the *is married to* relation. And if we say 'Jane is to the left of Frank', we are saying that Jane is connected to Frank via the *is to the left of* relation. Note that individuals can also be related to themselves in various ways. For example, 'Frank is identical to Frank' just means that Frank is connected to himself via the *identity* relation.

We can also have **three-place** (or **triadic**, or **ternary**) relations. For example, to say that 'Frank sits between Bob and Jane' is to say that Frank is related to Bob and Jane via the *sits between* relation. Similarly, '2 is the sum of 1 and 1'; in this case, we're saying that the number 2 is related to 1 and 1—the same individual, twice over—via the *is the sum of* relation.

For any natural number n, we can have ***n*-place** (or ***n*-adic**, or ***n*-ary**) relations—though natural examples of these in English become progressively more difficult to find. A four-place relation which sometimes gets used in mathematics is *the difference between … and … is greater than the difference between … and …*.

In English, n-place predicates are used to refer to n-place relations. (If you want to, you can think of a property as a one-place relation; hence it is referred to by a one-place predicate. Alternatively, you can think of an n-place relation as a n-place property.) As such, in $\mathcal{L}_P$, we use exactly the same kind of symbols to refer to both properties and relations. The difference is that an n-place predicate will be followed by n-many lower-case letters. In the logical jargon, this is sometimes described by saying that an n-place predicate has n-many **arguments**. In his context, an "argument" can be thought of as a 'slot' for the placement of a name.

PROPERTIES are attributes or qualities had by individuals, and RELATIONS are ways in which individuals can be connected to one another. An n-place relation connects n-many (not necessarily distinct) individuals. In $\mathcal{L}_P$, we use PREDICATES to refer both to properties and to relations.

We can now symbolize the following sentences in $\mathcal{L}_P$:

6. Abbie is to the left of Bob.
7. Abbie is identical to Abbie.
8. Abbie is identical to Caroline.
9. Abbie is between Bob and Caroline.
10. Abbie is to the left of Bob, and between Bob and Caroline.

We will continue to use 'a' and 'b' for Abbie and Bob respectively, and we'll introduce 'c' to refer to Caroline. Furthermore, we'll use 'L' for the relation *is to the left of*, 'B' for the 3-place *betweenness* relation, and 'I' for the identity relation.

We then symbolize sentence 6 as 'Lab.' Notice that the order of the constants matters: because we put the 'a' first, 'Lab' says that Abbie is to the left of Bob. If we had instead put down 'Lba,' we'd have that Bob is to the left of Abbie.

In the case of sentence 7, we should have the symbolization 'Iaa.' Here, the order of the constants does not matter—but only because we have used the same name twice. In general, for a two-place relation R, the left-most constant represents the individual which is R-related to the right-most constant.

In the case of sentence 8, we should translate it as 'Iac.' Note that in this case, we're using two distinct names ('a' and 'c') to refer to the same individual. This is the same as it is in ordinary English: every proper name refers to one and only one individual, but a single individual can have many different proper names.

Betweenness is a three-place relation, so we can symbolize 9 as '$Babc$.' In this case, Abbie (a) is the one who is B-related to Bob (b) and Caroline (c), so the 'a' should be the constant placed furthest to the left.

Finally, we can symbolize sentence 10 as '$Lab \wedge Babc$.' Even though the name 'Abbie' does not appear after the word 'and', the meaning of 10 is obviously the same as '[Abbie is to the left of Bob] *and* [Abbie is between Bob and Caroline].'

6.3 Quantifiers and Variables

We are now ready to introduce quantifiers. **Quantifiers** are words like 'every' and 'some'. There are many quantifiers besides these (e.g., 'most'), but in $\mathcal{L}_P$ we only symbolize 'every' and 'some'. Later on in the book, we will see how we can define up many other quantifiers in terms of 'every' and 'some. Indeed, we'll see how we can define 'every' in terms of 'some', and 'some' in terms of 'every'.

Consider these sentences:

11. Everything is red.
12. Everything is to the left of Bob.
13. Something is red.

We'll continue to use 'R' and 'L' to refer to the same properties as we did before.

How could we translate sentence 11? If we had at least one name for every single individual of whom we might be speaking, we could try to translate it as '$(Ra_1 \wedge Ra_2 \wedge Ra_3 \wedge \ldots \wedge Rb_1 \wedge Rb_2 \wedge Rb_3 \wedge \ldots)$.' (I am ignoring internal parentheses for simplicity: because the order of the conjuncts does not matter, there is no danger of confusion.) This isn't very convenient. Indeed, if there were infinitely many people, we'd need infinitely many constants and infinitely many conjunctions. In fact, in many logical systems infinitely long sentences are not even allowed, and we might not have names for every individual!

It would be nice to have a simpler way of symbolizing sentence 11. In order to do this, we introduce the '$\forall$' symbol. This is called the **universal quantifier**. There are a number of slightly different notational conventions for dealing with quantifiers, but we will use the following: a quantifier must always be followed by a single **variable**, and then a formula that includes that variable (which must be enclosed in brackets). A **formula** is a string of symbols which can be used to represent a proper sentence once all of the variables within it have been replaced by constants. An example will help.

We can translate the sentence 11 as '$\forall x(Rx)$.' We can read this as saying 'For all things x: x is red.' The universal quantifier $\forall$ is followed immediately by the variable 'x,' and then a formula 'Rx' enclosed in brackets. 'Rx' is a formula because if we replace the 'x' within it with a constant (like 'a'), the result ('Ra') symbolizes a complete sentence—in this case, that *Abbie is red*.

The bracketing around the formula 'Rx' describes the *scope* of the quantifier. We will give a formal definition of scope later, but intuitively it is the part of the sentence that the quantifier quantifies over. In '$\forall x(Rx)$,' the scope of the universal quantifier is 'Rx.' The variable 'x' in 'Rx' then acts as a "stand in" for the (possibly infinitely many) constants that name specific individuals. So, '$\forall x(Rx)$' essentially says that for individual a, 'Ra' is true, and for individual b, 'Rb' is true, and ... so on, for every single individual. In other words, assuming we have names for every individual, it means the same thing as the (possibly infinite) conjunction '$(Ra_1 \wedge Ra_2 \wedge Ra_3 \wedge \ldots \wedge Rb_1 \wedge Rb_2 \wedge Rb_3 \wedge \ldots)$.'

Sentence 12 can be paraphrased as, 'For all things x: x is to the left of Bob.' This translates as '$\forall x(Lxb)$.' So, if we have the name 'a' in our language, '$\forall x(Lxb)$' says that 'Lab' is true; and if we have the name 'b' in our language, then 'Lbb' is true; and ... so on, for every single name in our logical language.

In sum, the expression '$\forall x$' means that you can pick any particular constant that names an individual and substitute them in place of the 'x' in the following formula to produce a true sentence. Note that there is no special reason to use 'x' rather than some other variable. For example, '$\forall x(Rx)$' means exactly the same thing as '$\forall y(Ry)$,' which means the same thing as '$\forall z(Rz)$.' (Recall that we always use the lower-case letters 'x,' 'y,' and 'z,' with or without subscripts, to symbolize variables.) Usually, when more than one variable is needed in a given symbolization, it's easiest to use 'x' for the first variable, 'y' for the second, and 'z' for the third. But this is just for bookkeeping purposes.

> The symbol '$\forall$' is the UNIVERSAL QUANTIFIER. A universal quantifier is always followed by a variable, and then a formula including that variable (the latter enclosed in brackets). If $\mathcal{F}$ is any formula, then '$\forall x(\mathcal{F})$' can be read as saying 'For all things x: $\mathcal{F}$.'

To symbolize sentence 13, we can introduce another new symbol: the **existential quantifier**, $\exists$. Like the universal quantifier, the existential quantifier must always be followed by a variable and then a formula that includes that variable. Thus, 13 can be translated as '$\exists x(Rx)$.' This can be read as '*There exists an x such that*: x is red.' Saying this should not be taken to imply that there is only *one* red thing. For all '$\exists x(Rx)$' says, there could be many red things; indeed, it is compatible with saying that *everything* is red. Instead, it says that there's *at least one* red thing. You could therefore also read '$\exists x(Rx)$' as saying '*There is at least one thing x such that*: x is red.'

Once again, the variable 'x' acts as a kind of place-holder. '$\exists x(Rx)$' says that there's at least one constant which names an individual which can be substituted for 'x' in 'Rx' to produce a true sentence. Equivalently, if you have a name for every individual, you could take it to mean the same thing as the (possibly infinite) disjunction '$(Ra_1 \lor Ra_2 \lor Ra_3 \lor \ldots \lor Rb_1 \lor Rb_2 \lor Rb_3 \lor \ldots)$.' After all, the latter is true just in case *at least one* of the disjuncts is true, and '$\exists x(Rx)$' says that at least one of these disjuncts is true.

> The symbol '$\exists$' represents the EXISTENTIAL QUANTIFIER. If $\mathcal{F}$ is any formula, then '$\exists x(\mathcal{F})$' can be read as saying 'There exists an x such that: $\mathcal{F}$'.

Now let's consider these further sentences:

14. Nothing is red.
15. There is something which is not red.
16. Not everything is red.

Sentence 14 can be paraphrased as, '*It is not the case that* [something is red].' In other words, it is the negation of sentence 13, and can therefore be translated as '$\neg\exists x(Rx)$.'

Interestingly, 14 could *also* be paraphrased as, 'Everything is not red.' In that case, we might want to symbolize it as '$\forall x(\neg Rx)$.' Is this a problem? Not at all! In fact, both are perfectly acceptable symbolizations of the same sentence, because they are logically equivalent. In general, where $\mathcal{F}$ is any formula containing the appropriate variable (in this case, 'x'), each of following logical equivalences hold:

$\forall x(\neg\mathcal{F})$ if and only if $\neg\exists x(\mathcal{F})$

$\exists x(\neg\mathcal{F})$ if and only if $\neg\forall x(\mathcal{F})$

$\forall x(\mathcal{F})$ if and only if $\neg\exists x(\neg\mathcal{F})$

$\exists x(\mathcal{F})$ if and only if $\neg\forall x(\neg\mathcal{F})$

This means that any declarative sentence in English which can be symbolized with a universal quantifier can be symbolized with an existential quantifier, and *vice versa*. One particular means of symbolizing might seem more natural than the other on a given occasion, but there is no logical difference in translating with one quantifier rather than the other. For some sentences, it will simply be a matter of taste. Strictly speaking, our formal language really only needs only one of either '$\forall$' or '$\exists$'. However, we make things somewhat easier for ourselves if we have two quantifiers, even if one of them is redundant given the existence of the other.

Sentence 15 is most naturally paraphrased as, 'There is some x such that: x is not red.' This becomes '$\exists x(\neg Rx)$.' Equivalently, we could write it as '$\neg\forall x(Rx)$'; i.e., '*It's not the case that* [everything is red].' That's just what sentence 16 also says, so sentence 15 and 16 are logically equivalent.

6.4 The Universe of Discourse

Given the symbolization key we have been using, '$\forall x(Rx)$' means 'Everything is red.' What if we wanted to say that 'Everyone is red,' where 'everyone' refers only to—let's say—the people in the room? There are two main options here.

The first option is to introduce a symbol for the predicate *is a person in this room*, and paraphrase 'Everyone is red' as 'Every person in this room is red'.

P : ... is a person in this room

R : . . . is red

With this symbolization key, we can translate 'Every person in this room is red' as '$\forall x(Px \to Rx)$.' This says 'For all things x, if x is a person in this room, then x is red.' In this case, the antecedent tells us that of all the individuals that exist, we are only interested in people currently in the room; the consequent then says that each of *those* individuals are red.

It might be tempting to translate 'Every person in this room is red' using a conjunction. However, '$\forall x(Px \wedge Rx)$' would mean that everything that exists is both a person in this room and also red. This would be a crazy thing to say, and it means something very different than '$\forall x(Px \to Rx)$.'

In general, we can translate any sentence which means something of the form 'Every thing which is $\mathcal{P}$ is $\mathcal{Q}$' as '$\forall x(\mathcal{P}x \to \mathcal{Q}x)$.' Similarly, we can translate any sentence which means something of the form 'All things which are $\mathcal{P}$ are $\mathcal{Q}$' as '$\forall x(\mathcal{P}x \to \mathcal{Q}x)$.' In English we recognise an 'is'/'are' distinction which is not as important when we translate into $\mathcal{L}_P$.

In this case, the '$\forall x$' means *everything in the universe*; i.e., every individual that exists, anywhere. That is a huge number of individuals—indeed, infinitely many. If we wanted to, however, we could make '$\forall x$' mean something more restricted. For example, we could just make it mean 'every person in this room.' In that case, when we write '$\forall x(Rx)$,' we would be saying 'For every person x in this room: x is red.'

To alter the meaning of '$\forall x$' in this way, we need to specify what we'll call the **universe of discourse**, abbreviated UoD. The UoD contains every individual that we might potentially be talking about *in the present context.* It's up to us to specify which UoD we intend, and it can be almost anything we like. The only restriction in $\mathcal{L}_P$ is that the UoD must be *non-empty*; that is, it must contain at least one individual. But besides that rule, we're free to choose our own universe of discourse.

The universe of discourse could be taken to be everything *simpliciter*. In that case, we would need to translate 'Everyone is red' as $\forall x(Px \to Rx)$, as above. There are many things this UoD which aren't people, so it would be wrong in this case to symbolize the sentence as '$\forall x(Rx)$.' However, we could also stipulate that, for present purposes, the UoD is just the set of all people in the room. In this case, '$\forall x(Rx)$' is perfectly appropriate as a translation of 'Everyone is red.' Logicians would say that in each case, the variable 'x' *ranges over* a different universe of discourse.

In general, then, '$\forall x$' should really be interpreted as 'For every individual x *in the universe of discourse*: . . .'. Similarly, '$\exists x$' should be read as 'There is at least one individual x *in the universe of discourse*.' Usually, the intended restrictions on the universe of discourse are left implicit, and can easily be discovered from context. For example, when we use words like 'everybody', 'everyone', 'anyone', 'somebody', 'someone', and 'no one', we are often restricting the universe of discourse to some salient set of people. When we say 'everywhere', 'somewhere', and 'nowhere', we are restricting the universe of discourse to a set of locations. And, quite often when we say 'Everything' and 'nothing', we don't mean *everything in the universe*, just *everything of a certain kind around here.*

Restricted quantification occurs frequently in natural language, and we are very good at interpreting it when it happens. When we're doing logic, it's often helpful to avoid the possibility of ambiguity. So, whenever there is some potential for confusion, you should remember to specify the intended UoD.

6.5 Symbolization

We now have all of the pieces of $\mathcal{L}_P$. Translating more complicated sentences will only be a matter of knowing the right way to combine predicates, constants, quantifiers, variables, and connectives. In this section, we'll look at sentences which only make use of a single quantifier. Sentences with multiple quantifiers are usually much harder to deal with, and are covered in the next chapter.

Consider these sentences:

17. Every coin in my pocket is a quarter.
18. Some coin on the table is a dime.
19. Not all the coins on the table are dimes.
20. None of the coins in my pocket are dimes.

In providing a symbolization key, we need to specify a UoD. Since we are talking about coins in my pocket and on the table, the UoD must *at least* contain all of those coins. Since we are not talking about anything besides coins, we let the UoD be all coins. Since we are not talking about any specific coins, we do not need to define any constants. So we define this key:

UoD : All coins
P : . . . is in my pocket.
T : . . . is on the table.
Q : . . . is a quarter.
D : . . . is a dime.

Sentence 17 is most naturally translated with a universal quantifier. The universal quantifier says something about everything in the UoD, not just about the coins in my pocket. Sentence 17 means that (for any coin), *if* that coin is in my pocket, *then* it is a quarter. So we can translate it as '$\forall x(Px \rightarrow Qx)$.'

Sentence 18 is most naturally translated with an existential quantifier. It says that there is some coin which is both on the table and which is a dime. So we can translate it as '$\exists x(Tx \wedge Dx)$.'

Notice that we needed to use a conditional with the universal quantifier, but we used a conjunction with the existential quantifier. What would it mean to write '$\exists x(Tx \rightarrow Dx)$'? Probably not what you think. It means that there is some member of the UoD which would satisfy the formula '$(Tx \rightarrow Dx)$'; roughly speaking, there is some individual a such that '$(Ta \rightarrow Da)$' is true. In our earlier language $\mathcal{L}_S$, '$(\mathcal{A} \rightarrow \mathcal{B})$' was logically equivalent to '$(\neg\mathcal{A} \vee \mathcal{B})$,' and this also holds in $\mathcal{L}_P$. So '$\exists x(Tx \rightarrow Dx)$' is true if there is some a such that '$(\neg Ta \vee Da)$' holds; i.e., it is true if some coin is *either* not on the table *or* is a dime. Of course there is a coin that is not on the table—there are coins in lots of other places. So '$\exists x(Tx \rightarrow Dx)$' is trivially true.

A conditional will usually be the natural connective to use with a universal quantifier, but a conditional within the scope of an existential quantifier can do very strange things. As a general rule, do not put conditionals in the scope of existential quantifiers unless you are certain that you need one to say what you want.

Sentence 19 can be paraphrased as, 'It is not the case that every coin on the table is a dime.' So we can translate it as '$\neg\forall x(Tx \rightarrow Dx)$.' You might look at sentence 19 and paraphrase it instead as, 'Some coin on the table is not a dime.' You would then symbolize it as '$\exists x(Tx \wedge \neg Dx)$.' Although it is probably not obvious, these two translations are logically equivalent. This is due to the logical equivalence between '$\neg\forall x(\mathcal{F})$' and '$\exists x(\neg\mathcal{F})$,' along with the equivalence between '$\neg(\mathcal{A} \rightarrow \mathcal{B})$' and '$(\mathcal{A} \wedge \neg\mathcal{B})$.'

Sentence 20 can be paraphrased as, 'It is not the case that there is some dime in my pocket.' This can be symbolized as '$\neg\exists x(Px \wedge Dx)$.' It might also be paraphrased as, 'Everything in my pocket is not a dime,' and then could be symbolized as '$\forall x(Px \rightarrow \neg Dx)$.' Again the two translations are logically equivalent. Both are perfectly adequate ways of symbolizing 20.

We can now translate the arguments from p. 74, the ones that motivated the need for $\mathcal{L}_P$ in the first place:

P1	Willard is a logician.	**P1**	Froggy is a Frog.
P2	All logicians are pedantic.	**P2**	All frogs are green.
C	Willard is pedantic.	**C**	Froggy is green.

We'll assume that the UoD for both arguments is *everything*, and use the following symbolization keys:

L :	... is a logician	F :	... is a frog
P :	... is pedantic	G :	... is green
w :	Willard	f :	Froggy

P1 Lw
P2 $\forall x(Lx \rightarrow Px)$
C Pw

P1 Ff
P2 $\forall x(Fx \rightarrow Gx)$
C Gf

As you can see, we've managed to capture the structure that the two arguments share, in a way that wasn't possible to do using $\mathcal{L}_S$. The only thing that differs between the arguments is the particular predicates and names involved, but the basic *internal structure* of their premises and conclusions are the same. Both of these are formally valid arguments in $\mathcal{L}_P$.

Quantifiers and scope

In the sentence '$\exists x(Gx) \rightarrow Gl$,' the scope of the existential quantifier was the expression Gx. Would it matter if the scope of the quantifier were the whole sentence? That is, does the sentence '$\exists x(Gx \rightarrow Gl)$' mean something different?

With the key given above, '$\exists x(Gx) \rightarrow Gl$' means that if there is some guitarist, then Lemmy is a guitarist. '$\exists x(Gx \rightarrow Gl)$' means that there is some person such that if that person is a guitarist, then Lemmy is a guitarist. Recall that the conditional here is a material conditional; the conditional is true if the antecedent is false. Let the constant 'p' denote the author of this book, someone who is certainly not a guitarist. '$Gp \rightarrow Gl$' is true because 'Gp' is false. Since someone (namely p) satisfies the sentence, then '$\exists x(Gx \rightarrow Gl)$' is true. The sentence is true because there is a non-guitarist, regardless of Lemmy's skill.

Something strange happened when we changed the scope of the quantifier, because the conditional in $\mathcal{L}_P$ is a material conditional. In order to keep the meaning the same, we would have to change the quantifier: '$\exists x(Gx) \rightarrow Gl$' means the same thing as '$\forall x(Gx \rightarrow Gl)$,' and '$\exists x(Gx \rightarrow Gl)$' means the same thing as '$\forall x(Gx \rightarrow Gl)$.'

This oddity does not arise with other connectives, or if the variable is in the consequent of the conditional. For example, '$\exists x(Gx) \wedge Gl$' means just the same thing as '$\exists x(Gx \wedge Gl)$.' Similarly, '$Gl \rightarrow \exists x(Gx)$' means the same thing as '$\exists x(Gl \rightarrow Gx)$.'

6.6 Empty Predicates and a fallacy

Let's close with a brief discussion about predicates, and a famous fallacy committed within Aristotelian logics.

Nothing we've said so far guarantees that a given predicate will always apply to something in the UoD. A predicate that applies to nothing in the UoD is called an **empty predicate**. For example, suppose we want to symbolize these two sentences:

21. Every monkey knows sign language.
22. Some monkey knows sign language.

It is possible to write the symbolization key for these sentences in this way:

M : ... is a monkey
S : ... knows sign language

Sentence 21 can now be translated as '$\forall x(Mx \rightarrow Sx)$,' and 22 becomes '$\exists x(Mx \wedge Sx)$.'

It is tempting to think that sentence 21 entails sentence 22. That is: if every monkey knows sign language, then it must be that some monkey knows sign language. Indeed, Aristotle thought that the argument had a valid form, which in his system was characterised along the following lines:

P1 All Ms are S.

C Some M is S.

However, the entailment does *not* hold in $\mathcal{L}_P$. It is possible for the sentence '$\forall x(Mx \rightarrow Sx)$' to be true even though the sentence '$\exists x(Mx \wedge Sx)$' is false. How can this be? The answer comes from considering whether the premise and the conclusion would be true *if there were no monkeys.*

We have characterised $\forall$ and $\exists$ in such a way that '$\forall x(\mathcal{F})$' is equivalent to '$\neg\exists x(\neg\mathcal{F})$.' As such, the universal quantifier doesn't imply the existence of anything—it only implies the non-existence of certain things. If 21 is true, then there are *no* monkeys who don't know sign language. Now, if there were no monkeys, then '$\forall x(Mx \rightarrow Sx)$' would be true. Hence, '$\exists x(Mx \wedge Sx)$' would be false. So 21 does not imply 22. If there was at least one monkey, the entailment would go through—but logic alone does not tell us whether any monkeys exist.

It is important to allow empty predicates into our language, because we want to be able to say things like, 'I do not know if there are any monkeys, but any monkeys that there are must know sign language.' That is, we want to be able to speak about properties that do not (or might not) belong to anything.

Alternatively, consider the property of *being a unicorn*. We might use 'U' to refer to this property. We can then translate 'There are no unicorns' as '$\neg\exists x(Ux)$'; i.e., '*It is not the case that* [there is at least one x such that: x is a unicorn]'. We would find this very difficult to say if we didn't allow empty predicates into our formal language. Indeed, our ability to formulate such sentences is a huge advantage of $\mathcal{L}_P$; philosophers of the past worried a great deal about how we can make sense of sentences like 'There are no unicorns.'

Once we have empty predicates in our language, some interesting things happen. For example, consider the sentence 'Every unicorn knows sign language.' Translated, this becomes '$\forall x(Ux \rightarrow Sx)$,' which is *true*. This is counterintuitive, since we do not want to say that there are a whole bunch of unicorns out there that know sign language. It is important to remember, though, that '$\forall x(Ux \rightarrow Sx)$' just means that any member of the UoD which is a unicorn is also something that knows sign language. Since there are no unicorns in the UoD, the sentence is trivially true.

Practice Exercises

⋆ **Part A** Using the symbolization key given, symbolize each of the following in $\mathcal{L}_P$.

UoD :	The set of all animals
A :	... is an alligator
M :	... is a monkey
R :	... is a reptile
Z :	... lives at the zoo
L :	... loves ...
a :	Amos
b :	Bouncer
c :	Cleo

1. Amos, Bouncer, and Cleo all live at the zoo.
2. Bouncer is a reptile, but not an alligator.
3. If Cleo loves Bouncer, then Bouncer is a monkey.
4. If both Bouncer and Cleo are alligators, then Amos loves them both.
5. Some reptile lives at the zoo.
6. Every alligator is a reptile.
7. There are reptiles which are not alligators.
8. Cleo loves a reptile.
9. Bouncer loves all the monkeys that live at the zoo.

⋆ **Part B** Using the symbolization key given, translate each English-language sentence into $\mathcal{L}_P$.

UoD :	The set of all candies
C :	... has chocolate in it
M :	... has marzipan in it
S :	... has sugar in it
T :	Boris has tried ...
B :	... is better than ...

1. Boris has never tried any candy.
2. Marzipan is always made with sugar.
3. Some candy is sugar-free.
4. Every sugar-free candy has marzipan in it.
5. No candy is better than itself.
6. Boris has never tried sugar-free chocolate.

Part C These are syllogistic figures identified by Aristotle and his successors, along with their medieval names. Translate each argument into $\mathcal{L}_P$.

Barbara: All Bs are Cs. All As are Bs. Therefore, all As are Cs.

Baroco: All Cs are Bs. Some A is not B. Therefore, some A is not C.

Bocardo: Some B is not C. All As are Bs. Therefore, some A is not C.

Celantes: No Bs are Cs. All As are Bs. Therefore, no Cs are As.

Celarent: No Bs are Cs. All As are Bs. Therefore, no As are Cs.

Cemestres: No Cs are Bs. No As are Bs. Therefore, no As are Cs.

Cesare: No Cs are Bs. All As are Bs. Therefore, no As are Cs.

Dabitis: All Bs are Cs. Some A is B. Therefore, some C is A.

Darii: All Bs are Cs. Some A is B. Therefore, some A is C.

Datisi: All Bs are Cs. Some A is B. Therefore, some A is C.

Disamis: Some B is C. All As are Bs. Therefore, some A is C.

Ferison: No Bs are Cs. Some A is B. Therefore, some A is not C.

Ferio: No Bs are Cs. Some A is B. Therefore, some A is not C.

Festino: No Cs are Bs. Some A is B. Therefore, some A is not C.

Baralipton: All Bs are Cs. All As are Bs. Therefore, some C is A.

Frisesomorum: Some B is C. No As are Bs. Therefore, some C is not A.

Part D Using the symbolization key given, translate each English-language sentence into $\mathcal{L}_P$.

UoD :	The set of all animals
D :	. . . is a dog
S :	. . . likes samurai movies
L :	. . . is larger than . . .
b :	Bertie
e :	Emerson
f :	Fergis

1. Bertie is a dog who likes samurai movies.
2. Bertie, Emerson, and Fergis are all dogs.
3. Emerson is larger than Bertie, and Fergis is larger than Emerson.
4. All dogs like samurai movies.
5. Only dogs like samurai movies.
6. There is a dog that is larger than Emerson.
7. If there is a dog larger than Fergis, then there is a dog larger than Emerson.

8. No animal that likes samurai movies is larger than Emerson.
9. No dog is larger than Fergis.
10. Any animal that dislikes samurai movies is larger than Bertie.
11. There is an animal that is between Bertie and Emerson in size.
12. There is no dog that is between Bertie and Emerson in size.
13. No dog is larger than itself.

Chapter 7

Advanced Symbolization

In this supplementary chapter, we will look at a few tricky issues that arise when we're translating English language sentences into $\mathcal{L}_P$. Key ideas introduced are how to deal with multiple quantifiers within a single sentence, and a useful way to symbolize the relation of identity. As with the other supplementary chapters in this textbook, you will not be tested on anything herein.

7.1 Translating pronouns

When translating into $\mathcal{L}_P$, it is important to understand the structure of the sentences you want to translate. What matters is the final translation in $\mathcal{L}_P$, and sometimes you will be able to move from an English language sentence directly to a sentence of $\mathcal{L}_P$. Other times, it helps to paraphrase the sentence one or more times. Each successive paraphrase should move from the original sentence closer to something that you can translate directly into $\mathcal{L}_P$.

For the next several examples, we will use this symbolization key:

UoD : The set of all people
G : ... can play guitar
R : ... is a rock star
l : Lemmy

Now consider these sentences:

1. If Lemmy can play guitar, then he is a rock star.

2. If a person can play guitar, then he is a rock star.

Sentences 1 and 2 have the same consequent ('he is a rock star'), but they cannot be translated in the same way.

It helps to paraphrase the original sentences, replacing pronouns with explicit references. Sentence 1 can be paraphrased as, 'If Lemmy can play guitar, then *Lemmy* is a rockstar.' This can obviously be translated as '$(Gl \rightarrow Rl)$.' We do not need any quantifiers for this. On the other hand, 2 must be paraphrased differently: 'If a person can play guitar, then *that person* is a rock star.' This sentence is not about any particular person, so we need a variable. We can re-paraphrase the sentence as, 'For any person x: if x can play guitar, then x is a rockstar.' Now this can be translated as '$\forall x(Gx \rightarrow Rx)$.' This says just the same thing as 'Every person who can play guitar is a rock star.'

Consider these further sentences:

3. If anyone can play guitar, then Lemmy can.
4. If anyone can play guitar, then he or she is a rock star.

These two sentences have the same antecedent ('If anyone can play guitar'), but they have different logical structures.

Sentence 3 can be paraphrased, 'If someone can play guitar, then Lemmy can play guitar.' The antecedent and consequent are separate sentences, so it can be symbolized with a conditional as the main logical operator: '$\exists x(Gx) \rightarrow Gl$.' That is, '*If* [there exists an x such that: x can play guitar], *then* [Lemmy can play the guitar].'

On the other hand, 4 can be paraphrased, 'For any particular person, if that person can play guitar, then he or she is a rock star.' It would be a mistake to symbolize this with an existential quantifier, because it is talking about everybody. The sentence is equivalent to 'All guitar players are rock stars.' It is best translated as '$\forall x(Gx \rightarrow Rx)$.'

The English words 'any' and 'anyone' should typically be translated using quantifiers. As these two examples show, they sometimes call for an existential quantifier (as in 3) and sometimes for a universal quantifier (as in 4). If you have a hard time determining which is required, paraphrase the sentence with an English language sentence that uses words besides 'any' or 'anyone.'

7.2 Ambiguous predicates

Suppose we just want to translate this sentence:

5. Adina is a skilled surgeon.

Let the UoD be the set of all people, let 'K' mean '... is a skilled surgeon', and let 'a' refer to Adina. Sentence 5 is then simply 'Ka.'

Suppose instead that we want to symbolize this argument:

P1 Billy is a surgeon, but is not skilled.
P2 All surgeons are greedy.
P3 The hospital will only hire a skilled surgeon.

C Billy is greedy, but the hospital will not hire him.

We need to distinguish the property of *being a skilled surgeon* from merely *being a surgeon*. So we define this symbolization key:

UoD :	The set of all people
G :	... is greedy.
H :	The hospital will hire ....
R :	... is a surgeon.
K :	... is skilled.
b :	Billy

Now the argument can be fully symbolized in this way:

P1 $Rb \land \neg Kb$
P2 $\forall x(Rx \to Gx)$
P3 $\forall x\big(\neg(Rx \land Kx) \to \neg Hx\big)$

C $Gb \land \neg Hb$

Next, suppose that we want to translate this argument:

P1 Carol is a skilled surgeon and a tennis player.

C Therefore, Carol is a skilled tennis player.

If we start with the symbolization key we used for the previous argument, we could add a predicate (let 'T' mean '... is a tennis player') and a constant (let 'c' mean Carol). Then the argument becomes:

P1 $Rc \wedge (Kc \wedge Tc)$

C $Tc \wedge Kc$

This translation is a disaster! It takes what in English is a terrible argument and translates it as a valid argument in $\mathcal{L}_P$. The problem is that there is a difference between being *skilled as a surgeon* and *skilled as a tennis player*. Translating this argument correctly requires two separate predicates, one for each type of skill. If we let 'K_1' mean '... is skilled as a surgeon' and 'K_2' mean '... is skilled as a tennis player,' then we can symbolize the argument in this way:

P1 $(Rc \wedge K_1c) \wedge Tc$

C $Tc \wedge K_2c$

Like the English language argument it symbolizes, this is invalid.

The moral of these examples is that you need to be careful when symbolizing predicates, and ensure you don't do so in an ambiguous way. Similar problems can arise with predicates like *good*, *bad*, *big*, and *small*. Just as skilled surgeons and skilled tennis players have different skills, big dogs, big mice, and big problems are big in different ways.

Is it enough to have a predicate that means '... is a skilled surgeon', rather than two predicates '... is skilled' and '... is a surgeon'? Sometimes. As sentence 5 shows, sometimes we do not need to distinguish between skilled surgeons and other surgeons.

Must we always distinguish between different ways of being skilled, good, bad, or big? No. As the argument about Billy shows, sometimes we only need to talk about one kind of skill. If you are translating an argument that is just about dogs, it is fine to define a predicate that means '... is big.' If the UoD includes dogs and mice, however, it is probably best to make the predicate mean '... is big for a dog.'

7.3 Multiple Quantifiers

Consider this following symbolization key and the sentences that follow it:

F: ... is a friend of ...
U: ... is unhappy with ...
g: Gerald

6. Gerald is the friend of someone.
7. Someone is Gerald's friend.
8. Someone has a friend.
9. All of Gerald's friends are unhappy with everyone.
10. Someone is friends with everyone.
11. Everyone is friends with someone.

Sentence 6 can be paraphrased as 'There is an x such that: [Gerald is x's friend].' Thus, we can translate it as '$\exists x(Fgx)$.' This is not to be confused with the correct translation of 7, which is '$\exists x(Fxg)$.'

Sentence 8 can be paraphrased as 'There exists an x such that: [x is the friend of someone].' Note that the subsentence, 'x is a friend of someone,' is just the same as 6 but with 'Gerald' replaced by the variable 'x'. Alternatively, sentence 8 could also be paraphrased as 'There exists a y such that: [someone is y's friend].' In this case, the subsentence 'someone is y's friend' is just the same as 7, but with 'Gerald' replaced by the variable 'y'.

So on both ways of paraphrasing the sentence, the subsentences that follow the quantifier contain yet another quantifier. In order to symbolize 8, then, we will need to make use of multiple quantifiers. Thus, we can write it as '$\exists x \exists y(Fxy)$.' This can be read as 'There exists an x and there exists a y such that: [x is the friend of y].'

Sentence 9 can be paraphrased as 'For all x: [if x is a friend of Gerald, then x is unhappy with everyone].' It could be equally well paraphrased as 'For all y: [all of Gerald's friends are unhappy with y].' So, again, we're going to need to make use of multiple quantifiers. We can write it as '$\forall x \forall y(Fxg \to Uxy)$.' In semi-formal English, this can be read as 'For all x and all y: [*if* [x is friends with Gerald], *then* [x is unhappy with y]].'

What about sentence 10? Here, we're going to need to start mixing quantifiers. We can paraphrase 10 as 'There is an x such that: [x is friends with everyone].' The subsentence 'x is friends with everyone' involves a universal quantifier, so we can alternatively paraphrase 10 as 'There is an x such that: [for all y: [x is friends with y]].' Thus, we can symbolize 10 as '$\exists x \forall y(Fxy)$.'

Note that the order in which the quantifiers appear in our symbolization of sentence 10. When we're mixing quantifiers, the order matters. To see this, compare 10 with sentence 11. Here, the paraphrase would be: 'For all x: [there exists a y such that: [x is friends with y]].' In other words, we should translate 11 as '$\forall x \exists y(Fxy)$.'

Let's now consider this symbolization key and the sentences that follow it:

UoD : The set of all people

L : ... likes ...
i : Imre
k : Karl

12. Imre likes everyone that Karl likes.
13. There is someone who likes everyone who likes everyone that he likes.

Sentence 12 can be partially translated as '$\forall x$(Karl likes $x \to$ Imre likes x).' This becomes '$\forall x(Lkx \to Lix)$.' On the other hand, sentence 13 is almost a tongue-twister. There is little hope of writing down the whole translation immediately, but we can proceed by small steps. An initial, partial translation might look like this: '

$$\exists x \text{ (everyone who likes everyone that } x \text{ likes is liked by } x\text{)}$$

'

The part that remains in English is a universal sentence, so we translate further: '

$$\exists x \forall y (\text{if } y \text{ likes everyone that } x \text{ likes, then } x \text{ likes } y)$$

' Now the consequent of the conditional is straightforward to symbolize: 'Lxy.' Notice also that the antecedent is structurally just like sentence 12, but with 'y' and 'x' in place of 'i' and 'k.' So, sentence 13 can be completely translated in this way '

$$\exists x \forall y \big(\forall z (Lxz \to Lyz) \to Lxy\big)$$

'

When symbolizing sentences with multiple quantifiers, it is best to proceed by small steps. In the first step, paraphrase until you come up with an English sentence so that the logical structure is readily symbolized in $\mathcal{L}_P$. Then, translate piecemeal, replacing the daunting task of translating a long sentence with the simpler task of translating shorter formulae.

7.4 Identity

Consider this sentence:

14. Pavel owes money to everyone else.

Let the UoD be the set of all people; this will allow us to translate 'everyone' as a universal quantifier. Let 'O' mean '... owes money to ...', and let 'p' refer to Pavel. Now we can symbolize sentence 14 as '$\forall x(Opx)$.' Unfortunately, this translation has some odd consequences. It says that Pavel owes money to every member of the UoD, including Pavel himself. However, sentence 14 does not say that Pavel owes money to himself; he owes money to everyone *else*. This is a problem, because '$\forall x(Opx)$' is the best translation we can give of this sentence into $\mathcal{L}_P$.

In Chapter 6 we used 'I' to pick out the identity relation. However, as we will see in what follows, it can be exceedingly useful to treat it as a distinctive logical symbol all of its own.

Since it has a special logical meaning, we use $=$ a bit differently than we've been using other predicates: instead of placing its two arguments after it, we will place the arguments on either side of it. So, we should read '$a = b$' as meaning 'a is identical to b.' This does not mean merely that a and b are indistinguishable or that all of the same predicates are true of them. Rather, it means that a and b are *the very same thing*.

When we write '$a \neq b$,' we mean that 'a and b are not identical.' There is no reason to introduce this as an additional predicate. Instead, '$a \neq b$' is simply an abbreviation of '$\neg(a = b)$.'

Now suppose we want to symbolize this sentence:

15. Pavel is Mister Checkov.

Let the constant 'c' mean Mister Checkov. Then, 15 can be symbolized as '$p = c$.' This means that the constants 'p' and 'c' both refer to the same person.

This is all well and good, but how does it help with sentence 14? That sentence can be paraphrased as, 'Everyone who is not Pavel is owed money by Pavel.' This is a sentence structure we already know how to symbolize: 'For all x: if x is not Pavel, then x is owed money by Pavel.' In $\mathcal{L}_P$ with identity, this becomes '$\forall x(x \neq p \rightarrow Opx)$.'

In addition to sentences that use the word 'else', identity will be helpful when symbolizing some sentences that contain the words 'besides' and 'only.' Consider these examples:

16. No one besides Pavel owes money to Hikaru.
17. Only Pavel owes Hikaru money.

We add the constant 'h', which means Hikaru.

Sentence 16 can be paraphrased as, 'No one who is not Pavel owes money to Hikaru.' This can be translated as '$\neg\exists x(x \neq p \wedge Oxh)$.'

Sentence 17 can be paraphrased as, 'Pavel owes Hikaru *and* no one besides Pavel owes Hikaru money.' We have already translated one of the conjuncts, and the other is straightforward. Sentence 17 becomes '$(Oph \wedge \neg\exists x(x \neq p \wedge Oxh))$.'

7.5 Expressions of quantity

We can also use identity to say how many things there are of a particular kind. For example, consider these sentences:

18. There is at least one apple on the table.
19. There are at least two apples on the table.
20. There are at least three apples on the table.

Let the UoD be *things on the table*, and let A pick out the property of *being on the table.*

Sentence 18 does not require the use of the identity symbol. It can be translated adequately as '$\exists x(Ax)$.' That is, there is some apple on the table—perhaps many, but at least one.

It might be tempting to also translate sentence 19 without identity. Yet consider the sentence '$\exists x\exists y(Ax \wedge Ay)$.' It means that there is some apple x in the UoD and some apple y in the UoD. Since nothing precludes x and y from picking out the same member of the UoD, this would be true even if there were only one apple. In order to make sure that there are two *different* apples, we need an identity predicate. Sentence 19 needs to say that the two apples that exist are not identical, so it can be translated as '$\exists x\exists y(Ax \wedge Ay \wedge x \neq y)$.'

Sentence 20 requires talking about three different apples. It can be translated as '$\exists x\exists y\exists z(Ax \wedge Ay \wedge Az \wedge x \neq y \wedge y \neq z \wedge x \neq z)$.' Continuing in this way, we could translate 'There are at least n apples on the table.' There is a summary of how to symbolize sentences like these on p. 154.

Now consider these sentences:

21. There is at most one apple on the table.
22. There are at most two apples on the table.

Sentence 21 can be paraphrased as, 'It is not the case that there are at least *two* apples on the table.' This is just the negation of sentence 19: '

$$\neg\exists x\exists y(Ax \wedge Ay \wedge x \neq y)$$

' Sentence 21 can also be approached in another way. It means that any apples that there are on the table must be the selfsame apple, so it can be translated as '$\forall x \forall y\big((Ax \wedge Ay) \rightarrow x = y\big)$.' The two translations are logically equivalent, so both are correct.

In a similar way, sentence 22 can be translated in two equivalent ways. It can be paraphrased as, 'It is not the case that there are *three* or more distinct apples', so it can be translated as the negation of sentence 20. Using universal quantifiers, it can also be translated as: '

$$\forall x \forall y \forall z\big((Ax \wedge Ay \wedge Az) \rightarrow (x = y \vee x = z \vee y = z)\big).$$

'

The examples above are sentences about apples, but the logical structure of the sentences translates mathematical inequalities like '$a \geq 3$,' '$a \leq 2$,' and so on. We also want to be able to translate statements of equality which say exactly how many things there are. For example:

23. There is exactly one apple on the table.
24. There are exactly two apples on the table.

Sentence 23 can be paraphrased as, 'There is *at least* one apple on the table, and there is *at most* one apple on the table.' This is just the conjunction of 18 and 21 from above: '$\big(\exists x(Ax) \wedge \forall x \forall y\big((Ax \wedge Ay) \rightarrow x = y\big)\big)$.' This is a somewhat complicated way of going about it. It is perhaps more straightforward to paraphrase sentence 23 as, 'There is a thing which is the only apple on the table.' Thought of in this way, the sentence can be translated '$\exists x\big(Ax \wedge \neg\exists y(Ay \wedge x \neq y)\big)$.'

Similarly, sentence 24 may be paraphrased as, 'There are two different apples on the table, and these are the only apples on the table.' This can be translated as '$\exists x \exists y\big(Ax \wedge Ay \wedge x \neq y \wedge \neg\exists z(Az \wedge x \neq z \wedge y \neq z)\big)$.'

Finally, consider this sentence:

25. There are at most two things on the table.

It might be tempting to add a predicate so that 'T' would mean '... is a thing on the table.' However, this is unnecessary. Since the UoD is the set of things on the table, all members of the UoD are on the table. If we want to talk about a *thing on the table*, we need only use a quantifier. Sentence 25 can be symbolized like sentence 22 (which said that there were at most two apples), but leaving out the predicate entirely. That is, sentence 25 can be translated as '$\forall x \forall y \forall z(x = y \vee x = z \vee y = z)$.'

Techniques for symbolizing expressions of quantity ('at most', 'at least', and 'exactly') are summarized on p. 154.

7.6 Definite descriptions

Recall that a constant of $\mathcal{L}_P$ must refer to some member of the UoD. This constraint allows us to avoid the problem of non-referring terms (§6.1). Given a UoD that included only actually existing creatures but a constant 'c' that meant 'chimera' (a mythical creature), sentences containing 'c' would become impossible to evaluate.

The most widely influential solution to this problem was introduced by Bertrand Russell in 1905. Russell asked how we should understand this sentence:

26. The present king of France is bald.

The sub-phrase 'the present king of France' is supposed to pick out an individual by means of a definite description. However, there was no king of France in 1905 and there is none now. Since the description is a non-referring term, we cannot just define a constant 'f' to mean 'the present king of France' and translate the sentence as Kf.

Russell's idea was that sentences that contain definite descriptions have a different logical structure than sentences that contain proper names, even though they share the same grammatical form. What do we mean when we use an unproblematic, referring description, like 'the highest peak in Washington state'? We mean that there is such a peak, because we could not talk about it otherwise. We also mean that it is the only such peak. If there was another peak in Washington state of exactly the same height as Mount Rainier, then Mount Rainier would not be *the* highest peak.

According to this analysis, sentence 26 is saying three things. First, it makes an *existence* claim: there is some present king of France. Second, it makes a *uniqueness* claim: this guy is the only present king of France. Third, it makes a claim of *predication*: this guy is bald.

In order to symbolize definite descriptions in this way, we need the identity predicate. Without it, we could not translate the uniqueness claim which (according to Russell) is implicit in the definite description.

Let the UoD be the set of *people actually living*, let 'F' mean '... is the present king of France', and let 'B' mean '... is bald.' Sentence 26 can then be translated as '$\exists x\big(Fx \wedge \neg\exists y(Fy \wedge x \neq y) \wedge Bx\big)$.' This says that *there is* someone who is the present king of France, he is the *only* present king of France, and he is bald. Understood in this way, sentence 26 is meaningful but false. It says that a certain kind of person exists, but when no such person does.

The problem of non-referring terms is most vexing when we try to translate negations. So consider this sentence:

27. The present king of France is not bald.

According to Russell, this sentence is ambiguous in English. It could mean either of two things:

27a. It is not the case that the present king of France is bald.
27b. The present king of France is non-bald.

Both possible meanings negate sentence 26, but they put the negation in different places.

Sentence 27a is called a **wide-scope negation**, because it negates the entire sentence. It can be translated as '$\neg\exists x\big(Fx \wedge \neg\exists y(Fy \wedge x \neq y) \wedge Bx\big)$.' This does not say anything about the present king of France, but rather says that some sentence about the present king of France is false. Since sentence 26 if false, sentence 27a is true.

Sentence 27b says something about the present king of France. It says that he lacks the property of baldness. Like sentence 26, it makes an existence claim and a uniqueness claim; it just denies the claim of predication. This is called **narrow-scope negation**. It can be translated as '$\exists x\big(Fx \wedge \neg\exists y(Fy \wedge x \neq y) \wedge \neg Bx\big)$.' Since there is no present king of France, this sentence is false.

Russell's theory of definite descriptions resolves the problem of non-referring terms and also explains why it seemed so paradoxical. Before we distinguished between the wide-scope and narrow-scope negations, it seemed that sentences like 27 should be both true and false. By showing that such sentences are ambiguous, Russell showed that they are true understood one way but false understood another way.

Practice Exercises

Part A Choosing your own UoD and symbolization key, translate each English-language sentence into $\mathcal{L}_P$.

1. All the food is on the table.
2. If the food has not run out, then it is on the table.
3. Everyone likes some of the food.
4. If anyone likes the food, then Eli does.
5. Francesca only likes the dishes that have run out.
6. Francesca likes no one, and no one likes Francesca.
7. Eli likes anyone who likes some of the food.
8. Eli likes anyone who likes the people that he likes.
9. If there is a person on the table already, then all of the food must have run out.

$\star$ **Part B** Using the symbolization key given, translate each English-language sentence into $\mathcal{L}_P$.

UoD :	The set of all people
D :	. . . dances ballet
F :	. . . is female
M :	. . . is male
C :	. . . is a child of . . .
S :	. . . is a sibling of . . .
e :	Elmer
j :	Jane
p :	Patrick

1. All of Patrick's children are ballet dancers.
2. Jane is Patrick's daughter.
3. Patrick has a daughter.
4. Jane is an only child.
5. All of Patrick's daughters dance ballet.
6. Patrick has no sons.
7. Jane is Elmer's niece.
8. Patrick is Elmer's brother.
9. Patrick's brothers have no children.
10. Jane is an aunt.
11. Everyone who dances ballet has a sister who also dances ballet.
12. Every man who dances ballet is the child of someone who dances ballet.

Part C Using the symbolization key given, translate each English-language sentence into $\mathcal{L}_P$ with identity. The last sentence is ambiguous and can be translated two ways; you should provide both translations. (Hint: Identity is only required for the last four sentences.)

UoD : The set of all people
K : ... knows the combination to the safe
S : ... is a spy
V : ... is a vegetarian
T : ... trusts ...
h : Hofthor
i : Ingmar

1. Hofthor is a spy, but no vegetarian is a spy.
2. No one knows the combination to the safe unless Ingmar does.
3. No spy knows the combination to the safe.
4. Neither Hofthor nor Ingmar is a vegetarian.
5. Hofthor trusts a vegetarian.
6. Everyone who trusts Ingmar trusts a vegetarian.
7. Everyone who trusts Ingmar trusts someone who trusts a vegetarian.
8. Only Ingmar knows the combination to the safe.
9. Ingmar trusts Hofthor, but no one else.
10. The person who knows the combination to the safe is a vegetarian.
11. The person who knows the combination to the safe is not a spy.

⋆ **Part D** Using the symbolization key given, translate each English-language sentence into $\mathcal{L}_P$ with identity. The last two sentences are ambiguous and can be translated two ways; you should provide both translations for each.

UoD : The set of cards in a standard deck
B : ... is black
C : ... is a club
D : ... is a deuce
J : ... is a jack
M : ... is a man with an axe
O : ... is one-eyed
W : ... is wild

1. All clubs are black cards.
2. There are no wild cards.
3. There are at least two clubs.
4. There is more than one one-eyed jack.
5. There are at most two one-eyed jacks.
6. There are two black jacks.

7. There are four deuces.
8. The deuce of clubs is a black card.
9. One-eyed jacks and the man with the axe are wild.
10. If the deuce of clubs is wild, then there is exactly one wild card.
11. The man with the axe is not a jack.
12. The deuce of clubs is not the man with the axe.

Part E Using the symbolization key given, translate each English-language sentence into $\mathcal{L}_P$ with identity. The last two sentences are ambiguous and can be translated two ways; you should provide both translations for each.

UoD :	The set of animals in the world
B :	. . . is in Farmer Brown's field.
H :	. . . is a horse.
P :	. . . is a Pegasus.
W :	. . . has wings.

1. There are at least three horses in the world.
2. There are at least three animals in the world.
3. There is more than one horse in Farmer Brown's field.
4. There are three horses in Farmer Brown's field.
5. There is a single winged creature in Farmer Brown's field; any other creatures in the field must be wingless.
6. The Pegasus is a winged horse.
7. The animal in Farmer Brown's field is not a horse.
8. The horse in Farmer Brown's field does not have wings.

Chapter 8

Proofs in Quantificational Logic

In this supplementary chapter, we will develop the classical proof theory that's applicable when we're dealing with sentences formalized in $\mathcal{L}_P$.

For proofs in $\mathcal{L}_P$, we use all of the basic rules of $\mathcal{L}_S$ plus four new basic rules: these will be the introduction and elimination rules for each of the quantifiers. Since all of the derived rules of $\mathcal{L}_S$ are derived from the basic rules, they will also hold in $\mathcal{L}_P$. We will add another derived rule, a replacement rule called *quantifier negation*. Remember, all of the proof rules for $\mathcal{L}_S$ and $\mathcal{L}_P$ are summarised in appendix D.

8.1 Substitution instances

In order to concisely state the rules for the quantifiers, we need a way to mark the relation between quantified sentences and their instances. For example, the sentence 'Pa' is a particular instance of the general claim '$\forall x(Px)$.'

For a well-formed formula $\mathscr{A}$ written in $\mathcal{L}_P$, a constant 'c,' and a variable 'x,' define '$\mathscr{A}\boxed{\mathcal{X}\Rightarrow c}$' to mean the formula that we get by replacing every occurrence of 'x' in '$\mathscr{A}$' with 'c.' '$\mathscr{A}\boxed{\mathcal{X}\Rightarrow c}$' is called a **substitution instance** of '$\forall x\mathscr{A}$' and '$\exists x\mathscr{A}$,' and 'c' is called the **instantiating constant**. For example:

- '$Aa \to Ba$,' '$Af \to Bf$,' and '$Ak \to Bk$' are all substitution instances of '$\forall x(Ax \to Bx)$'; the instantiating constants are 'a,' 'f,' and 'k,' respectively.

- ‘Raj,’ ‘Rdj,’ and ‘Rjj’ are substitution instances of ‘$\exists z(Rzj)$’; the instantiating constants are ‘a,’ ‘d,’ and ‘j,’ respectively.

8.2 Universal elimination

If you have ‘$\forall x(Ax)$,’ it is legitimate to infer that anything is an A. You can infer ‘Aa,’ ‘Ab,’ ‘Az,’ ‘Ad_3.’ This is, you can infer any substitution instance—in short, you can infer ‘$A\mathcal{c}$’ for any constant ‘$\mathcal{c}$.’ This is the general form of the universal elimination rule ($\forall$E):

m	$\forall\chi(\mathcal{A})$	
	$\mathcal{A}\boxed{\chi\Rightarrow\mathcal{c}}$	$\forall$E m

Remember that the box mark for a substitution instance is not a symbol of $\mathcal{L}_P$, so you cannot write it directly in a proof. Instead, you write the substituted sentence with the constant ‘$\mathcal{c}$’ replacing all occurrences of the variable ‘χ’ in ‘$\mathcal{A}$.’ For example:

1	$\forall x(Mx \rightarrow Rxd)$	
2	$Ma \rightarrow Rad$	$\forall$E 1
3	$Md \rightarrow Rdd$	$\forall$E 1

8.3 Existential introduction

When is it legitimate to infer ‘$\exists x(Ax)$’? If you know that something is an A—for instance, if you have ‘Aa’ available in the proof.

This is the existential introduction rule ($\exists$I):

m	$\mathcal{A}$	
	$\exists\chi(\mathcal{A}\boxed{\boxed{\chi\Rightarrow\mathcal{c}}})$	$\exists$I m

It is important to notice that ‘$\mathcal{A}\boxed{\boxed{\chi\Rightarrow\mathcal{c}}}$’ is not necessarily a substitution instance. We write it with double boxes to show that the variable ‘χ’ does not need to replace all occurrences of the constant ‘$\mathcal{c}$.’ You can decide which occurrences to replace and which to leave in place. For example:

1	$Ma \to Rad$	
2	$\exists x(Ma \to Rax)$	$\exists$I 1
3	$\exists x(Mx \to Rxd)$	$\exists$I 1
4	$\exists x(Mx \to Rad)$	$\exists$I 1
5	$\exists y \exists x(Mx \to Ryd)$	$\exists$I 4
6	$\exists z \exists y \exists x(Mx \to Ryz)$	$\exists$I 5

8.4 Universal introduction

A universal claim like '$\forall x(Px)$' would be proven if every substitution instance of it had been proven, if every line 'Pa,' 'Pb,' ... were available in a proof. Alas, there is no hope of proving *every* substitution instance. That would require proving 'Pa,' 'Pb,' ..., 'Pj_2,' ..., 'Ps_7,' ..., and so on to infinity. There are infinitely many constants in $\mathcal{L}_P$, and so this process would never come to an end.

Consider a simple argument: $\forall x(Mx)$, therefore $\forall y(My)$

It makes no difference to the meaning whether we use the variable 'x' or the variable 'y,' so this argument is obviously valid. Suppose we begin in this way:

1	$\forall x(Mx)$	want $\forall y(My)$
2	Ma	$\forall$E 1

We have derived 'Ma.' Nothing stops us from using the same justification to derive 'Mb,' ..., 'Mj_2,' ..., 'Ms_7,' ..., and so on until we run out of space or patience. We have effectively shown the way to prove '$M\mathcal{c}$' for any constant $\mathcal{c}$. From this, '$\forall y(My)$' follows.

1	$\forall x(Mx)$	
2	Ma	$\forall$E 1
3	$\forall y(My)$	$\forall$I 2

It is important here that 'a' was just some arbitrary constant. We had not made any special assumptions about it. If 'Ma' were a premise of the argument, then this would not show anything about *all* y. For example:

$$
\begin{array}{r|ll}
1 & \forall x(Rxa) & \\
\hline
2 & Raa & \forall\text{E } 1 \\
3 & \forall y(Ryy) & \text{not allowed!}
\end{array}
$$

This is the schematic form of the universal introduction rule ($\forall$I):

$$
\begin{array}{r|ll}
m & \mathcal{A} & \\
 & \forall\chi(\mathcal{A}\boxed{c^*\Rightarrow\chi}) & \forall\text{I } m
\end{array}
$$

* The constant c must not occur in any undischarged assumption.

Note that we can do this for any constant that does not occur in an undischarged assumption and for any variable.

Note also that the constant may not occur in any **undischarged** assumption, but it may occur as the assumption of a subproof that we have already closed. For example, we can prove '$\forall z(Dz \to Dz)$' without any premises.

$$
\begin{array}{r|l|ll}
1 & & Df & \text{want } Df \\
\cline{3-3}
2 & & Df & \text{R } 1 \\
3 & \multicolumn{2}{l}{Df \to Df} & \to\text{I } 1\text{–}2 \\
4 & \multicolumn{2}{l}{\forall z(Dz \to Dz)} & \forall\text{I } 3
\end{array}
$$

8.5 Existential elimination

A sentence with an existential quantifier tells us that there is *some* member of the UoD that satisfies a formula. For example, '$\exists x(Sx)$' tells us (roughly) that there is at least one S. It does not tell us *which* member of the UoD satisfies S, however. We cannot immediately conclude 'Sa,' 'Sf_{23},' or any other substitution instance. What can we do?

Suppose that we knew both '$\exists x(Sx)$' and '$\forall x(Sx \to Tx)$.' We could reason in this way:

> Since '$\exists x(Sx)$,' there is something that is an S. We do not know which constants refer to this thing, if any do, so call this thing 'Ishmael'. From '$\forall x(Sx \to Tx)$,' it follows that if Ishmael an S, then it is a T. Therefore, Ishmael is a T. Because Ishmael is a T, we know '$\exists x(Tx)$.'

In this paragraph, we introduced a name for the thing that is an S. We gave it an arbitrary name ('Ishmael') so that we could reason about it and derive some consequences from there being an S. Since 'Ishmael' is just a bogus name introduced for the purpose of the proof and not a genuine constant, we could not mention it in the conclusion. Yet we could derive something that does not mention Ishmael; namely, '$\exists x(Tx)$.' This does follow from the two premises.

We want the existential elimination rule to work in a similar way. Yet since English language worlds like 'Ishmael' are not symbols of $\mathcal{L}_P$, we cannot use them in formal proofs. Instead, we will use constants of $\mathcal{L}_P$ which do not otherwise appear in the proof.

A constant that is used to stand in for whatever it is that satisfies an existential claim is called a **proxy**. Reasoning with the proxy must all occur inside a subproof, and the proxy cannot be a constant that is doing work elsewhere in the proof.

This is the schematic form of the existential elimination rule ($\exists$E):

$$\begin{array}{l|l l}
m & \exists \chi(\mathcal{A}) & \\
n & \quad \begin{array}{|l} \mathcal{A} \; \boxed{c^* \Rightarrow \chi} \\ \hline \end{array} & \\
p & \quad \begin{array}{|l} \mathcal{B} \end{array} & \\
 & \mathcal{B} & \exists\text{E } m,\ n\text{–}p
\end{array}$$

* The constant c must not appear outside the subproof.

Remember that the proxy constant cannot appear in $\mathcal{B}$, the sentence you prove using $\exists$E.

It would be enough to require that the proxy constant not appear in '$\exists\chi(\mathcal{A})$,' in '$\mathcal{B}$,' or in any undischarged assumption. In recognition of the fact that it is just a place holder that we use inside the subproof, though, we require an entirely new constant which does not appear anywhere else in the proof.

With this rule, we can give a formal proof that '$\exists x(Sx)$' and '$\forall x(Sx \rightarrow Tx)$' together entail '$\exists x(Tx)$.'

1	$\exists x(Sx)$	
2	$\forall x(Sx \to Tx)$	want $\exists x(Tx)$
3	$\quad Sa$	
4	$\quad Sa \to Ta$	$\forall$E 2
5	$\quad Ta$	$\to$E 3, 4
6	$\quad \exists x(Tx)$	$\exists$I 5
7	$\exists x(Tx)$	$\exists$E 1, 3–6

Notice that this has effectively the same structure as the English-language argument with which we began, except that the subproof uses the proxy constant 'a' rather than the bogus name 'Ishmael'.

8.6 Quantifier negation

When translating from English to $\mathcal{L}_P$, we noted that '$\neg\exists x(\neg\mathcal{A})$' is logically equivalent to '$\forall x(\mathcal{A})$.' In $\mathcal{L}_P$, they are provably equivalent. We can prove one half of the equivalence with a rather gruesome proof:

1	$\forall x(Ax)$	want $\neg\exists x(\neg Ax)$
2	$\quad \exists x(\neg Ax)$	for reductio
3	$\quad\quad \neg Ac$	for $\exists$E
4	$\quad\quad\quad \forall x(Ax)$	for reductio
5	$\quad\quad\quad Ac$	$\forall$E 1
6	$\quad\quad\quad \neg Ac$	R 3
7	$\quad\quad \neg\forall x(Ax)$	$\neg$I 4–6
8	$\quad \forall x(Ax)$	R 1
9	$\quad \neg\forall x(Ax)$	$\exists$E 3–7
10	$\neg\exists x(\neg Ax)$	$\neg$I 2–9

In order to show that the two sentences are genuinely equivalent, we need a second proof that assumes '$\neg\exists x(\neg\mathcal{A})$' and derives '$\forall x(\mathcal{A})$.' We leave that proof as an exercise for the reader.

It will often be useful to translate between quantifiers by adding or subtracting negations in this way, so we add two derived rules for this purpose. These rules are called *quantifier negation* (QN):

$$\neg\forall\chi(\mathcal{A}) \iff \exists\chi(\neg\mathcal{A})$$
$$\neg\exists\chi(\mathcal{A}) \iff \forall\chi(\neg\mathcal{A}) \quad \text{QN}$$

Since QN is a replacement rule, it can be used on whole sentence or on subformulae.

8.7 Rules for identity

The identity predicate is not part of $\mathcal{L}_P$, but we add it when we need to symbolize certain sentences. For proofs involving identity, we add two rules of proof.

Suppose you know that many things that are true of a are also true of b. For example: '$(Aa \wedge Ab)$,' '$(Ba \wedge Bb)$,' '$(\neg Ca \wedge \neg Cb)$,' '$(Da \wedge Db)$,' '$(\neg Ea \wedge \neg Eb)$,' and so on. This would not be enough to justify the conclusion '$a = b$.' In general, there are no sentences in $\mathcal{L}_P$ that do not already contain the identity predicate that could justify the conclusion '$a = b$.' This means that the identity introduction rule will not justify '$a = b$' or any other identity claim containing two different constants.

However, it is always true that '$a = a$.' In general, no premises are required in order to conclude that something is identical to itself. So this will be the identity introduction rule, abbreviated =I:

$$\mathcal{c} = \mathcal{c} \qquad =\text{I}$$

Notice that the =I rule does not require referring to any prior lines of the proof. For any constant $\mathcal{c}$, you can write '$\mathcal{c} = \mathcal{c}$' on any point with only the =I rule as justification.

If you have shown that '$a = b$,' then anything that is true of a must also be true of b. For any symbolized sentence with 'a' in it, you can replace some or all of the occurrences of 'a' with 'b' and produce an equivalent sentence. For example, if you already know 'Raa,' then you are justified in concluding 'Rab,' 'Rba,' 'Rbb.' Recall that '$\mathcal{A}\boxed{\boxed{\mathcal{a}\Rightarrow\mathcal{b}}}$' is the sentence produced by replacing '$\mathcal{a}$' in '$\mathcal{A}$' with '$\mathcal{b}$.' This is not the same as a substitution instance, because '$\mathcal{b}$' may replace some or all occurrences of '$\mathcal{a}$'. The identity elimination rule (=E) justifies replacing terms with other terms that are identical to it:

m	$\mathcal{a} = \mathcal{b}$	
n	$\mathcal{A}$	
	$\mathcal{A}\boxed{\mathcal{a}\Rightarrow\mathcal{b}}$	=E m, n
	$\mathcal{A}\boxed{\mathcal{b}\Rightarrow\mathcal{a}}$	=E m, n

To see the rules in action, consider this proof:

1	$\forall x \forall y\ (x = y)$	
2	$\exists x(Bx)$	
3	$\forall x(Bx \to \neg Cx)$	want $\neg\exists x(Cx)$
4	$\quad Be$	
5	$\quad \forall y\ (e = y)$	$\forall$E 1
6	$\quad e = f$	$\forall$E 5
7	$\quad Bf$	=E 6, 4
8	$\quad Bf \to \neg Cf$	$\forall$E 3
9	$\quad \neg Cf$	$\to$E 8, 7
10	$\neg Cf$	$\exists$E 2, 4–9
11	$\forall x(\neg Cx)$	$\forall$I 10
12	$\neg\exists x(Cx)$	QN 11

8.8 Proof-theoretic concepts

We will use the symbol '$\vdash$' to indicate that a proof is possible. This symbol is called the **turnstile**. When we write '$\{\mathcal{A}_1, \mathcal{A}_2, \ldots\} \vdash \mathcal{B}$,' this means that it is possible to give a proof of $\mathcal{B}$ with '$\mathcal{A}_1$,' '$\mathcal{A}_2$,' … as premises. With just one premise, we leave out the curly braces, so '$\mathcal{A} \vdash \mathcal{B}$' means that there is a proof of '$\mathcal{B}$' with '$\mathcal{A}$' as a premise. Naturally, '$\vdash \mathcal{B}$' means that there is a proof of '$\mathcal{B}$' that has no premises.

Often, logical proofs are called **derivations**. So '$\mathcal{A} \vdash \mathcal{B}$' can be read as '$\mathcal{B}$ is derivable from $\mathcal{A}$.' A **theorem** is a sentence that is derivable without any premises; i.e., $\mathcal{T}$ is a theorem if and only if '$\vdash \mathcal{T}$.'

It is not too hard to show that something is a theorem—you just have to give a proof of it. How could you show that something is *not* a theorem? If its negation is a theorem, then you could provide a proof. For example, it is easy to prove '$\neg(Pa \wedge \neg Pa)$,' which shows that '$Pa \wedge \neg Pa$' cannot be a theorem. For a sentence that is neither a theorem nor the negation of a theorem, however, there is no easy way to show this. You would have to demonstrate not just that certain proof strategies fail, but that no proof is possible. Even if you fail in trying to prove a sentence in a thousand different ways, perhaps the proof is just too long and complex for you to make out.

Two sentences '$\mathcal{A}$' and '$\mathcal{B}$' are **provably equivalent** if and only if each can be derived from the other; i.e., '$\mathcal{A} \vdash \mathcal{B}$' and '$\mathcal{B} \vdash \mathcal{A}$.' It is relatively easy to show that two sentences are provably equivalent—it just requires a pair of proofs. Showing that sentences are *not* provably equivalent would be much harder. It would be just as hard as showing that a sentence is not a theorem. (In fact, these problems are interchangeable. Can you think of a sentence that would be a theorem if and only if '$\mathcal{A}$' and '$\mathcal{B}$' were provably equivalent?)

The set of sentences $\{\mathcal{A}_1, \mathcal{A}_2, \ldots\}$ is **provably inconsistent** if and only if a contradiction is derivable from it; i.e., for some sentence $\mathcal{B}$, $\{\mathcal{A}_1, \mathcal{A}_2, \ldots\} \vdash \mathcal{B}$ and $\{\mathcal{A}_1, \mathcal{A}_2, \ldots\} \vdash \neg\mathcal{B}$. It is easy to show that a set is provably inconsistent: you just need to assume the sentences in the set and prove a contradiction. Showing that a set is *not* provably inconsistent will be much harder. It would require more than just providing a proof or two; it would require showing that proofs of a certain kind are *impossible.*

Practice Exercises

⋆ Part A

Provide a justification (rule and line numbers) for each line of proof that requires one.

$$\begin{array}{r|l}
1 & \forall x\exists y(Rxy \lor Ryx) \\
2 & \underline{\forall x(\neg Rmx)} \\
3 & \exists y(Rmy \lor Rym) \\
4 & \quad\vert\ \underline{(Rma \lor Ram)} \\
5 & \quad\vert\ \neg Rma \\
6 & \quad\vert\ Ram \\
7 & \quad\vert\ \exists x Rxm \\
8 & \exists x(Rxm)
\end{array}$$

$$\begin{array}{r|l}
1 & \forall x(Jx \rightarrow Kx) \\
2 & \exists x\forall y(Lxy) \\
3 & \underline{\forall x(Jx)} \\
4 & \quad\vert\ \underline{\forall y(Lay)} \\
5 & \quad\vert\ Ja \\
6 & \quad\vert\ (Ja \rightarrow Ka) \\
7 & \quad\vert\ Ka \\
8 & \quad\vert\ Laa \\
9 & \quad\vert\ (Ka \land Laa) \\
10 & \quad\vert\ \exists x(Kx \land Lxx) \\
11 & \exists x(Kx \land Lxx)
\end{array}$$

$$\begin{array}{r|l}
1 & \forall x(\exists y(Lxy) \rightarrow \forall z(Lzx)) \\
2 & \underline{Lab} \\
3 & \exists y(Lay) \rightarrow \forall z(Lza) \\
4 & \exists y(Lay) \\
5 & \forall z(Lza) \\
6 & Lca \\
7 & \exists y(Lcy) \rightarrow \forall z(Lzc) \\
8 & \exists y(Lcy) \\
9 & \forall z(Lzc) \\
10 & Lcc \\
11 & \forall x(Lxx)
\end{array}$$

$$\begin{array}{r|l}
1 & \quad\vert\ \underline{\neg(\exists x(Mx) \lor \forall x(\neg Mx))} \\
2 & \quad\vert\ \neg\exists x(Mx) \land \neg\forall x(\neg Mx) \\
3 & \quad\vert\ \neg\exists x(Mx) \\
4 & \quad\vert\ \forall x(\neg Mx) \\
5 & \quad\vert\ \neg\forall x(\neg Mx) \\
6 & \exists x(Mx) \lor \forall x(\neg Mx)
\end{array}$$

⋆ Part B

Provide a proof of each claim.

1. $\vdash \forall x(Fx) \lor \neg\forall x(Fx)$
2. $\{\forall x(Mx \leftrightarrow Nx), Ma \land \exists x(Rxa)\} \vdash \exists x(Nx)$
3. $\{\forall x(\neg Mx \lor Ljx), \forall x(Bx \rightarrow Ljx), \forall x(Mx \lor Bx)\} \vdash \forall x(Ljx)$

4. $\forall x(Cx \wedge Dt) \vdash \forall x(Cx) \wedge Dt$
5. $\exists x(Cx \vee Dt) \vdash \exists x(Cx) \vee Dt$

Part C
Provide a proof of the argument about Billy on p. 94.

Part D
Look back at Part C on p. 89. Provide proofs to show that each of the argument forms is valid in $\mathcal{L}_P$.

Part E
Aristotle and his successors identified other syllogistic forms. Symbolize each of the following argument forms in $\mathcal{L}_P$ and add the additional assumptions 'There is an A' and 'There is a B.' Then prove that the supplemented arguments forms are valid in $\mathcal{L}_P$.

Darapti: All As are Bs. All As are Cs. Therefore, some B is C.

Felapton: No Bs are Cs. All As are Bs. Therefore, some A is not C.

Barbari: All Bs are Cs. All As are Bs. Therefore, some A is C.

Camestros: All Cs are Bs. No As are Bs. Therefore, some A is not C.

Celaront: No Bs are Cs. All As are Bs. Therefore, some A is not C.

Cesaro: No Cs are Bs. All As are Bs. Therefore, some A is not C.

Fapesmo: All Bs are Cs. No As are Bs. Therefore, some C is not A.

Part F
Provide a proof of each claim.

1. $\forall x \forall y(Gxy) \vdash \exists x(Gxx)$
2. $\forall x \forall y(Gxy \rightarrow Gyx) \vdash \forall x \forall y(Gxy \leftrightarrow Gyx)$
3. $\{\forall x(Ax \rightarrow Bx), \exists x(Ax)\} \vdash \exists x(Bx)$
4. $\{Na \rightarrow \forall x(Mx \leftrightarrow Ma), Ma, \neg Mb\} \vdash \neg Na$
5. $\vdash \forall z(Pz \vee \neg Pz)$
6. $\vdash \forall x(Rxx) \rightarrow \exists x \exists y(Rxy)$
7. $\vdash \forall y \exists x(Qy \rightarrow Qx)$

Part G
Show that each pair is provably equivalent.

1. $\forall x(Ax \rightarrow \neg Bx)$, $\neg \exists x(Ax \wedge Bx)$

2. $\forall x(\neg Ax \to Bd)$, $\forall x(Ax) \vee Bd$
3. $\exists x(Px) \to Qc$, $\forall x(Px \to Qc)$

Part H
Show that each of the following is provably inconsistent.

1. $\{(Sa \to Tm),\ (Tm \to Sa),\ (Tm \wedge \neg Sa)\}$
2. $\{\neg\exists x(Rxa),\ \forall x\forall y(Ryx)\}$
3. $\{\neg\exists x\exists y(Lxy),\ Laa\}$
4. $\{\forall x(Px \to Qx),\ \forall z(Pz \to Rz),\ \forall y(Py),\ \neg(Qa \wedge \neg Rb)\}$

⋆ **Part I**
Write a symbolization key for the following argument, translate it, and prove it:

> There is someone who likes everyone who likes everyone that he likes. Therefore, there is someone who likes himself.

Part J
Provide a proof of each claim.

1. $\{(Pa \vee Qb), (Qb \to b = c), \neg Pa\} \vdash Qc$
2. $\{(m = n \vee n = o), An\} \vdash (Am \vee Ao)$
3. $\{\forall x(x = m), Rma\} \vdash \exists x(Rxx)$
4. $\neg\exists x(x \neq m) \vdash \forall x\forall y(Px \to Py)$
5. $\forall x\forall y(Rxy \to x = y) \vdash (Rab \to Rba)$
6. $\{\exists x(Jx), \exists x(\neg Jx)\} \vdash \exists x\exists y(x \neq y)$
7. $\{\forall x(x = n \leftrightarrow Mx), \forall x(Ox \vee \neg Mx)\} \vdash On$
8. $\{\exists x(Dx), \forall x(x = p \leftrightarrow Dx)\} \vdash Dp$
9. $\{\exists x\big(Kx \wedge \forall y(Ky \to x = y) \wedge Bx\big), Kd\} \vdash Bd$
10. $\vdash Pa \to \forall x(Px \vee x \neq a)$

⋆ **Part K**
For each of the following pairs of sentences: if they are logically equivalent in $\mathcal{L}_P$, give proofs to show this.

1. $\forall x(Px) \to Qc$, $\forall x(Px \to Qc)$
2. $\forall x(Px) \wedge Qc$, $\forall x(Px \wedge Qc)$
3. $Qc \vee \exists x(Qx)$, $\exists x(Qc \vee Qx)$
4. $\forall x\forall y\forall z(Bxyz)$, $\forall x(Bxxx)$
5. $\forall x\forall y(Dxy)$, $\forall y\forall x(Dxy)$
6. $\exists x\forall y(Dxy)$, $\forall y\exists x(Dxy)$

⋆ Part L

For each of the following arguments: if it is valid in $\mathcal{L}_P$, give a proof.

1. $\forall x \exists y(Rxy)$, therefore $\exists y \forall x(Rxy)$
2. $\exists y \forall x(Rxy)$, therefore $\forall x \exists y(Rxy)$
3. $\exists x(Px \wedge \neg Qx)$, therefore $\forall x(Px \rightarrow \neg Qx)$
4. $\forall x(Sx \rightarrow Ta)$, Sd, therefore Ta
5. $\forall x(Ax \rightarrow Bx)$, $\forall x(Bx \rightarrow Cx)$, therefore $\forall x(Ax \rightarrow Cx)$
6. $\exists x(Dx \vee Ex)$, $\forall x(Dx \rightarrow Fx)$, therefore $\exists x(Dx \wedge Fx)$
7. $\forall x \forall y(Rxy \vee Ryx)$, therefore Rjj
8. $\exists x \exists y(Rxy \vee Ryx)$, therefore Rjj
9. $\forall x(Px) \rightarrow \forall x(Qx)$, $\exists x(\neg Px)$, therefore $\exists x(\neg Qx)$
10. $\exists x(Mx) \rightarrow \exists x(Nx)$, $\neg \exists x(Nx)$, therefore $\forall x(\neg Mx)$

Part M

1. If you know that $\mathcal{A} \vdash \mathcal{B}$, what can you say about $(\mathcal{A} \wedge \mathcal{C}) \vdash \mathcal{B}$? Explain your answer.
2. If you know that $\mathcal{A} \vdash \mathcal{B}$, what can you say about $(\mathcal{A} \vee \mathcal{C}) \vdash \mathcal{B}$? Explain your answer.

Part III

Probability Theory

Chapter 9

[$\mathcal{R}eq.$] Introduction to Probabilities

In this final part of the book, we're going to take a brief look at mathematical probabilities. Probabilities are important for a very large number of reasons, not least because we cannot do all our reasoning with deductive logic alone. Deductive logic works best when we're dealing with *certainties*, propositions which we know (and know we know). But most of the time, you will be uncertain about a great many things—indeed you're probably at least a *little bit* doubtful of every thought you consider. Are you *absolutely* certain you're not a brain in a vat? Are you sure you shut the door to the fridge this morning?

For reasoning with uncertain premises, we need a grasp of how probabilities work. Probability theory forms the backbone of many areas in the sciences, including physics, biology, and the social sciences.

In this chapter, we will start with the basic theory of probabilities. The exposition will be relatively informal, and we will not cover probabilities in cases involving infinities. (For those kinds of cases, a slightly more sophisticated theory is needed than will be developed here.)

9.1 Background Concepts

Though there are some disagreements here and there, the *formal theory* of probabilities is something that is very widely agreed upon. On the other hand, there are a great many opposing views on what it *means* to say that something is probable or improbable.

We will not survey each of the many theories that have been put forward as to what probability statements mean. (See Bowell & Kemp, pp. 105-106 for a very brief overview of some of the main theories.) Instead, we're going to work with one very common theory—that probabilities reflect **rational degrees of confidence**.

Say that you are certain that some proposition P is false just in case you have 0% confidence in P, and that you are certain that P is true if you're 100% confident in P. The different degrees of confidence that you might have towards P can then be described by all the values between 0% and 100%. So, for example, having 99% confidence in P means that you are very confident in it indeed, but not *quite* certain that it's true—you reserve just a tiny bit of doubt about the truth of P. Having 50% confidence towards P means you're just as confident in P's truth as in its falsity. Having 5% confidence towards P means you're very confident it's false, but not yet certain.

If you're a good, rational thinker, your degrees of confidence should themselves be rational. That is, the amount of confidence you have towards a proposition should be supported by your evidence for and/or against that proposition. For example, suppose a very highly regarded meteorologist tells you that the probability that it will rain tomorrow is 90%. You have every reason to trust what the meteorologist says, and no reasons to think that she might be mistaken. How confident should you be that it will rain tomorrow? Presumably, you should be 90% confident. Any other level of confidence would seem unjustified in the present situation.

Likewise, suppose a very large body of evidence suggests that taking the dietary supplement 'Super Creatine Mix 2000' causes unstoppable hiccups in around 5% of the population. You're wondering whether you want to take the supplement, but worried about the hiccups. You have no reasons to believe that you have some special immunity to hiccups, nor that you're particularly susceptible to them either. In this case we'd usually say that the probability that you would come down with unstoppable hiccups if you took the supplement is around 5%. But how confident should you be that you'll get the hiccups if you decide to go with the supplement? Again, the right answer looks to be: around 5%.

These examples suggest that there is a very close relationship between probability statements—i.e., statements like 'It's 30% likely that it will rain tomorrow', or 'The probability that dinosaurs had pink feathers is 98%'—and the degrees of confidence you *should* have if you're rational.

The second example also suggests that probabilities are also very closely related to **frequencies**. We said that Super Creatine Mix 2000 causes unstoppable hiccups in "around 5% of the population." This is a statement about the frequency with which the supplement causes hiccups in the wider population. Frequencies and rational degrees of confidence are closely connected, because frequencies are an important source of evidence. A lot of the time, we can work out what our rational degree of confidence should be just by looking at the relevant frequencies. For instance, if you're unsure how confident you should be that a coin will land heads, you can simply check the frequency with which it has landed heads when tossed before in similar kinds of circumstances.

In summary, the probability of a proposition reflects how confident you should be that the proposition is true, given the evidence you have. This isn't the only way to understand what probabilities are, or what probability statements might mean, but it will be a useful way to think about them for present purposes.

Representing probabilities

Since probabilities attach to whole propositions, we can go back to using the simple formal language $\mathcal{L}_S$ that we developed in Part I. So, for example, the letters S and W might represent 'Snow is white' and 'Water is wet' respectively; in which case '$S \vee W$' will represent the disjunction 'Snow is white *or* water is wet' and '$S \wedge W$' will represent the conjunction 'Snow is white *and* water is wet.'

In the mathematical theory of probability, probabilities are usually represented as numbers somewhere between 0 and 1, inclusive. (The use of 1 and 0 rather than 100 and 0 isn't just to make things difficult—probability theorists use these numbers because it makes some of the mathematics much easier to work with.) You may be more familiar with speaking of probabilities in terms of percentages and/or fractions, but it is easy to translate between these different ways of speaking. For example:

- ▷ $\mathcal{A}$ has a probability of 0.1 = $\mathcal{A}$ has a probability of 1/10 = $\mathcal{A}$ is 10% probable
- ▷ $\mathcal{A}$ has a probability of 0.25 = $\mathcal{A}$ has a probability of 1/4 = $\mathcal{A}$ is 25% probable
- ▷ $\mathcal{A}$ has a probability of 0.5 = $\mathcal{A}$ has a probability of 1/2 = $\mathcal{A}$ is 50% probable
- ▷ $\mathcal{A}$ has a probability of 0.6 = $\mathcal{A}$ has a probability of 3/5 = $\mathcal{A}$ is 60% probable

Now, rather than always saying '$\mathcal{A}$ has a probability of ...', we can instead use a **probability function** to represent probabilities. We will label our probability function using '$\mathcal{P}r(\cdot)$'. A probability function takes a proposition as input (represented by the capital letter that goes inside the brackets), and gives us its probability value as output. So, if the probability of the proposition expressed by 'S' is 0.25, we can simply write $\mathcal{P}r(S) = 0.25$; and if the probability of '$S \vee W$' is 0.5, we write $\mathcal{P}r(S \vee W) = 0.53$.

In what follows, I'll sometimes talk about probabilities as ordinary numbers, sometimes as fractions, and sometimes as percentages. This shouldn't be seen as implying anything significant—they are just different ways of saying the same thing.

Any given proposition can only be assigned a *single* probability value at a time. This is what it means for $\mathcal{P}r(\cdot)$ to be a **function**: for any proposition it takes as input, it spits out one and only one value as its output. Probabilities may change over time—for example, if you get new evidence that the proposition is true—but it doesn't make sense to say that the probability of any proposition $\mathcal{A}$ is, say, both 0.3 and 0.5 *at the same time*.

> A PROBABILITY is a number between 0 and 1 which attaches to whole propositions, which corresponds to how confident you should be that the proposition in question is true. In general, we write '$\mathcal{P}r(\mathcal{A}) = n$' to mean 'The probability of $\mathcal{A}$ is n.'

9.2 The Probability Calculus

In mathematics, an **axiom** is a usually quite simple statement upon which a more complex mathematical structure is defined. In this section we're going to look at the three basic axioms of probability theory. As you'll soon see, they are really very simple indeed—but from such small beginnings, we can do quite a lot!

The original version of the following axioms were described by the Russian mathematician Andrey Kolmogorov, in his 1933 book *Foundations of the Theory of Probability*. Here, we'll work with a lightly simplified version of the axioms that foregoes some of the mathematical details. A summary of the axioms is given in appendix C.

Axiom 1
For any $\mathcal{A}$, $\mathcal{P}r(\mathcal{A}) \geq 0$.

That is, the probability value of any proposition whatsoever can never be less than 0. This should be very intuitive—after all, it doesn't make sense to say that the probability of something being true is *less than* 0%! Axiom 1 is usually called the **non-negativity** axiom, because it ensures that probability values are never negative.

Axiom 2
For any tautology $\mathcal{T}$, $\mathcal{P}r(\mathcal{T}) = 1$.

Recall from §4.2 that a tautology is a proposition which must be true as a matter of logic, regardless of how the world turns out to be. For example, 'If snow is white then snow is white' is a tautology, as is 'Either snow is white or snow isn't white.' We can know that each of these is true without doing any empirical work, purely by our logical reasoning.

Given this, the second probability axiom is very intuitive. It effectively says that tautologies have the highest possible probability value, 1. Axiom 2 is usually called the **normalization** axiom, because (in combination with Axiom 1) it ensures that probabilities are always within the 'normal' range from 0 to 1.

Axiom 3
If $\mathcal{A}$ and $\mathcal{B}$ are jointly inconsistent, then $\mathcal{P}r(\mathcal{A} \vee \mathcal{B}) = \mathcal{P}r(\mathcal{A}) + \mathcal{P}r(\mathcal{B})$

Recall from §4.4 that a pair of propositions are jointly inconsistent just in case it's not possible for both to be true at the same time. So, Axiom 3 says that the probability of a *disjunction* of any two *jointly inconsistent* propositions $\mathcal{A}$ and $\mathcal{B}$ is equal to the sum of the probabilities of $\mathcal{A}$ and $\mathcal{B}$ individually. Axiom 3 is called the **additivity** axiom, because it says that the probabilities of non-overlapping parts of a complex proposition 'add up' to give us the probability of the whole.

And that's it. Those three simple axioms will be the fundamental basis of everything we discuss in this chapter.

The laws of probability

Let's apply our three axioms to an example. Along the way, we'll derive four useful **laws of probability**. These are general truths which can be proven to follow logically as a result of the axioms. A summary of these laws and their proofs are given in appendix C.

Suppose we toss a 4-sided die, with sides labelled '1' through '4'. There are then five salient possibilities to focus on:

D_1 : The die will land on side 1
D_2 : The die will land on side 2
D_3 : The die will land on side 3
D_4 : The die will land on side 4
D_5 : None of the above will happen

D_5 is a 'catch-all'; it captures the possibility that the die doesn't land on any side at all. (Perhaps it is caught, lands on its edge, or maybe even teleported into space by an evil gambling-hating demon.) Effectively, D_5 is the negation of the disjunction of D_1 through to D_4. So, either the die lands on one side or other, or it doesn't land on any side at all.

Exactly one of D_1, D_2, D_3, D_4, D_5 must be true—no more, no less. Each one is incompatible with the truth of each of the others, and at least one of them *must* be true. In this case, we will say that $\{D_1, D_2, D_3, D_4, D_5\}$ is a **partition**.

> A set of propositions is a PARTITION just in case exactly one member of the set must be true.

We can also create more complex propositions out of D_1, D_2, D_3, D_4, D_5. For instance, '$D_1 \vee D_2$' is the proposition 'The die lands on side 1 or on side 2', '$D_1 \wedge D_2$' is the proposition that 'The die lands on side 1 and on side 2', and so on. Whether these propositions are true or false then depends on which of the five atomic propositions turns out to be true.

Axiom 1 tells us that the probability of every one of these propositions must be at least 0.

Axiom 2 tells us that the probability of the tautology, '$(D_1 \vee D_2 \vee D_3 \vee D_4 \vee D_5)$,' is equal to 1. ('$(D_1 \vee D_2 \vee D_3 \vee D_4 \vee D_5)$' is a tautology because exactly one of its disjuncts must be true—and if any one of the disjuncts is true, then the disjunction as a whole is true.) Likewise, every other tautology that we can construct out of our original five propositions also has a probability of 1. For example, $\mathcal{P}r(D_1 \vee \neg D_1) = 1$, and $\mathcal{P}r(D_2 \rightarrow D_2) = 1$, and so on.

Axiom 3 is where things start to get interesting. We can use it to quickly derive the following law of probability:

Law of Probability 1
For any $\mathcal{A}$, $\mathcal{P}r(\mathcal{A}) + \mathcal{P}r(\neg\mathcal{A}) = 1$.
Alternative version: $\mathcal{P}r(\mathcal{A}) = 1 - \mathcal{P}r(\neg\mathcal{A})$.

Now, since D_5 is equivalent to '$\neg(D_1 \vee D_2 \vee D_3 \vee D_4)$,' we can know using Law 1 that $\mathcal{P}r(D_5)$ is equal to 1 minus $\mathcal{P}r(D_1 \vee D_2 \vee D_3 \vee D_4)$. So, let's make two simplifying assumptions. First, we'll assume that the probability of D_5 is zero. That is, $\mathcal{P}r(D_5) = 0$—there are no evil gambling-hating demons, and we are pretending that we are certain of this. From this assumption we know that $\mathcal{P}r(D_1 \vee D_2 \vee D_3 \vee D_4) = 1$. And second, we'll assume that the die is **fair**, in the sense that every side has an equal chance of coming up top.

And now, thanks to Axiom 3, we can now know what the probability of D_1 through to D_4 must be: $n = 1/4$, or 0.25. Here is the reasoning. We know that

D_1 and D_2 are pairwise inconsistent—they cannot both be true at the same time. So, by Axiom 3, we know:

$$\mathcal{P}r(D_1 \vee D_2) = \mathcal{P}r(D_1) + \mathcal{P}r(D_2)$$

Because we are assuming that $\mathcal{P}r(D_1)$ and $\mathcal{P}r(D_2)$ have the same value, we know that $\mathcal{P}r(D_1 \vee D_2) = n + n$, for some probability value n. Furthermore, we know that D_3 is inconsistent with '$D_1 \vee D_2$': if either of D_1 or D_2 are true, then D_3 can't be true; and if D_3 is true, then neither of D_1 nor D_2 can be true. Therefore,

$$\mathcal{P}r(D_1 \vee D_2 \vee D_3) = n + n + \mathcal{P}r(D_3)$$

In other words, $\mathcal{P}r(D_1 \vee D_2 \vee D_3) = n + n + n$. And, likewise, we know that D_4 is inconsistent with $(D_1 \vee D_2 \vee D_3)$, for the same reasons as above. So:

$$\mathcal{P}r(D_1 \vee D_2 \vee D_3 \vee D_4) = n + n + n + \mathcal{P}r(D_4)$$

So all in all, $\mathcal{P}r(D_1 \vee D_2 \vee D_3 \vee D_4) = 4 \times n$. Now, we said earlier that $\mathcal{P}r(D_1 \vee D_2 \vee D_3 \vee D_4) = 1$. So,

$$\begin{aligned}
&4 \times n = 1 \\
&n = 1/4
\end{aligned}$$

Now there's another law of probability that we can learn from this example (the proof of which is implicit in the reasoning we've just gone through):

> **Law of Probability 2**
> If it is not possible for more than one member of a (finite) set n of propositions $\{\mathcal{A}_1, \mathcal{A}_2, \ldots, \mathcal{A}_n\}$ to be true at the same time, then:
>
> $$\mathcal{P}r(\mathcal{A}_1 \vee \mathcal{A}_2 \vee \ldots \vee \mathcal{A}_n) \ = \ \mathcal{P}r(\mathcal{A}_1) + \mathcal{P}r(\mathcal{A}_2) + \ldots + \mathcal{P}r(\mathcal{A}_n)$$
>
> Furthermore, if $\{\mathcal{A}_1, \mathcal{A}_2, \ldots, \mathcal{A}_n\}$ is a partition, then:
>
> $$\mathcal{P}r(\mathcal{A}_1 \vee \mathcal{A}_2 \vee \ldots \vee \mathcal{A}_n) \ = \ 1$$

This is a helpful generalization of Axiom 3. Axiom 3 tells us what to do when we're trying to work out the probability of the disjunction of *two* jointly inconsistent propositions. Law 2, on the other hand, tells us what to do when we're trying to work out the probability of the disjunction of any finite number of jointly inconsistent propositions. (Things are very slightly more complicated if the set of propositions is infinite, but in general the point is the same in the infinite case.)

So, suppose we want to know the probability of the die landing either 1, 2, or 3, but not 4. Law 2 then lets us calculate this easily: it's just the sum of

$\mathcal{P}r(D_1)$, $\mathcal{P}r(D_2)$, and $\mathcal{P}r(D_3)$. In this case, given our assumption that the coin is unbiased, the probability is 0.75, or 75%.

The next law of probability is also very important. Suppose that two propositions, $\mathcal{A}$ and $\mathcal{B}$, are logically equivalent (see §4.3 for a recap on logical equivalence). Then, just as we should expect that $\mathcal{A}$ is true if and only if $\mathcal{B}$ is true, we should also expect that the *probability of* $\mathcal{A}$ is n just in case the *probability of* $\mathcal{B}$ is also n.

Law of Probability 3
If $\mathcal{A}$ and $\mathcal{B}$ are logically equivalent, then $\mathcal{P}r(\mathcal{A}) = \mathcal{P}r(\mathcal{B})$.

For example, the probability of '$D_1 \vee D_2$,' that the die lands on side 1 or side 2, is exactly equal to the probability of '$\neg(\neg D_1 \wedge \neg D_2)$,' which says that the die *doesn't* land on something which is not side 1 nor side 2—i.e., that it doesn't land on either side 3 or 4. In both cases, the probability is 0.5.

The final law of probability that we'll look at in this chapter is also going to be another generalization of Axiom 3. So far, we know how to work out the probabilities of disjunctions of jointly inconsistent propositions. But what if we want to know the probability of the disjunction of a pair of propositions that *can* both be true at the same time? For this, we can apply the following law:

Law of Probability 4
For any two $\mathcal{A}$ and $\mathcal{B}$,

$$\mathcal{P}r(\mathcal{A} \vee \mathcal{B}) = \mathcal{P}r(\mathcal{A}) + \mathcal{P}r(\mathcal{B}) - \mathcal{P}r(\mathcal{A} \wedge \mathcal{B})$$

From an intuitive perspective, Law 4 makes a lot of sense. Suppose we want to know the probability that the die will either land on something other than side 4, or on side 1. That is, we're after the probability of '$\neg D_4 \vee D_1$.' We know the probability that it won't land on side 4, which is 0.75. (i.e., if it doesn't land on side 4, then it will land on side 1, 2, or 3, and each has a 0.25 probability.) And we know the probability that it will land on side 1, which is 0.25.

However, if we simply added $\mathcal{P}r(\neg D_4 \vee D_1)$ and $\mathcal{P}r(D_1)$ together, we'd get 1. In this case, we'd be **double counting** the event that *both* $\mathcal{P}r(\neg D_4 \vee D_1)$ and $\mathcal{P}r(D_1)$ are true. That's why we should subtract away the probability of *both* '$\neg D_4 \vee D_1$' and 'D_1' occurring at the same time.

Now the probability of the conjunction '$(\neg D_4 \vee D_1) \wedge D_1$' is just the probability of 'D_1.' This is because 'D_1' is logically equivalent to '$(\neg D_4 \vee D_1) \wedge D_1$.' (You can check this for yourself using truth tables.) Hence, we can work out the probability of '$(\neg D_4 \vee D_1) \vee D_1$' to be:

$$\mathcal{P}r((\neg D_4 \vee D_1) \vee D_1) \;=\; 0.75 + 0.25 - 0.25 \;=\; 0.75$$

This is the right result: the probability of '$(\neg D_4 \vee D_1) \vee D_1$' is just the probability that the die lands either on side 1, 2, or 3, which is 0.75.

There are many other laws of probability than the four summarised here. Some of these are set as practice problems below.

Practice Exercises

Part A
Suppose you have a standard deck of 52 cards. What is:

1. The probability that the card is hearts?
2. The probability that the card is not a jack?
3. The probability that the card is hearts or clubs?
4. The probability that the card is an ace or a king?
5. The probability that the card is hearts or an ace?
6. The probability that the card is hearts and not an ace?
7. The probability that the card is an ace or a red queen?
8. The probability that the card is either not hearts, or an ace?

$\star$ **Part B**
Suppose that $\mathcal{P}r(S) = 0.25$, $\mathcal{P}r(V) = 0.25$, and $\mathcal{P}r(S \wedge V) = 0.125$.

1. Are S and V consistent with one another?
2. Are S and V logically equivalent to one another?
3. What is the value of $\mathcal{P}r(S \vee V)$?
4. What is the value of $\mathcal{P}r((S \vee V) \wedge S)$?

Part C
Without checking the appendix, see if you can use the three axioms of the probability calculus to derive the four laws of probability outlined in this chapter.

Part D
Explain how one might try to derive the following additional laws of probability using the three axioms of this chapter, which hold for any pair of propositions $\mathcal{A}$ and $\mathcal{B}$.

1. $\mathcal{P}r(\mathcal{A}) \geq \mathcal{P}r(\mathcal{A} \wedge \mathcal{B})$
2. $\mathcal{P}r(\mathcal{A}) \leq \mathcal{P}r(\mathcal{A} \vee \mathcal{B})$
3. $\mathcal{P}r(\mathcal{A}) = \mathcal{P}r(\mathcal{A} \vee \mathcal{B})$ only if $\mathcal{P}r(\mathcal{B}) = \mathcal{P}r(\mathcal{A} \wedge \mathcal{B})$

Chapter 10

[*Req.*] Conditional Probabilities

In Chapter 9, we developed the basic mathematics for dealing with unconditional probability statements. These are statements like 'The probability of that the coin will land tails is 50%'. However, when we're thinking probabilistically, we also have to deal with **conditional probabilities**. This chapter will develop the theory of conditional probability, and highlight some very important characteristics of good probabilistic reasoning.

10.1 Notation

Suppose you come across a coin, and decide to toss it. Knowing nothing much about the coin, you probably expect it to be just like every other coin you've seen. In particular, you figure that there's a roughly 50% chance that when you toss it, the coin will land heads, and a roughly 50% chance that it will land tails. These are unconditional probability values.

On the other hand, what is the probability that a coin will land heads *given that* it is biased 70% towards tails? In this case, the correct answer would be 0.3, or 30%. This value isn't the probability that the coin will land heads, because you may not *know* whether the coin is biased towards tails or not. Instead, it's the probability that you would say the proposition has, *if* you were to suppose that the coin were biased in such a way.

Likewise, we could ask: what is the probability that the coin will land heads, *given that* it is biased 80% towards heads? In this case, the correct answer would be 0.8. We could even ask: what is the probability that the coin will land heads, *given that* it lands heads. Here, the correct answer is obvious: 1. And the probability that it lands heads, *given that* it lands tails, is 0.

Conditional probabilities are probabilities *under a condition*. When we want to say that the probability of a proposition $\mathcal{A}$ under a particular condition $\mathcal{B}$ is n, we write '$\mathcal{P}r(\mathcal{A}|\mathcal{B}) = n$'. You can read this as 'The probability of $\mathcal{A}$, *given that* $\mathcal{B}$ is true, is n.'

By way of example, suppose you have an ordinary, well-shuffled deck of 52 cards. You select a card at random from the deck—what is the probability that it is a heart? The answer is 1/4. That is,

$$\mathcal{P}r(\text{The card is a heart}) = 1/4$$

On the other hand, what is the probability that the card is a heart, *given that* it is a red card? This time, the answer is 1/2. So,

$$\mathcal{P}r(\text{The card is a heart} \mid \text{The card is red}) = 1/2$$

Alternatively, what is the probability that the card is a heart, *given that* you pick up a black card? Obviously, the answer is 0. Hence,

$$\mathcal{P}r(\text{The card is a heart} \mid \text{The card is black}) = 0$$

In these examples, the conditional probability of 'The card is a heart' is different than its unconditional probability, depending on what the relevant condition is, and different conditions led to different conditional probabilities.

10.2 Independence and the Gambler's fallacy

Ok, so what if we asked about the probability that the card is a heart, *given that* the cards in the deck have rounded rather than sharp corners? Presumably, the sharpness of the corners makes no difference to the probability that the card is a heart, so:

$$\mathcal{P}r(\text{The card is a heart} \mid \text{The card's corners are rounded}) = 1/4$$

In the jargon of probability theory, we say in this case that whether the card is a heart is **probabilistically independent** (or just 'independent') of the roundness of the card's corners. If two propositions are not probabilistically independent of one another, then we can say that they are **probabilistically correlated**. Generally:

> A proposition $\mathcal{A}$ is PROBABILISTICALLY INDEPENDENT of another proposition $\mathcal{B}$ if and only if $\mathcal{P}r(\mathcal{A}) = \mathcal{P}r(\mathcal{A}|\mathcal{B})$. Otherwise, they are PROBABILISTICALLY CORRELATED.

Essentially, $\mathcal{A}$ is independent of $\mathcal{B}$ if assuming $\mathcal{B}$ makes no difference to the probability of $\mathcal{A}$. There are plenty of things which are probabilistically independent of whether a particular card is a heart: the number of dust motes on Mars, the day of the week that Christmas will be on this year, the average lifespan of a porcupine, and so on.

Many very important results in the mathematics of probabilities depend on this notion of independence, and it is an extremely useful notion in the sciences which have to deal with statistics. However, people are apt to forget about independence frequently. This forgetfulness is the basis of what's known as the **Gambler's fallacy**.

Suppose you have a fair coin—you know it's fair—and you decide to start tossing it over and over again. The first time it lands, it comes up heads. It had a 50% chance of coming up heads, so nothing strange there. Then it lands heads again, and again, and again. The coin has now landed heads four times in a row. The chances of this happening are fairly low (6.25%). You decide to flip it one more time. What are the chances that it will land heads this time?

You might be tempted by the following thought:

> The chances of getting five heads in a row are very small (3.125%), so—surely—the next toss will probably come up tails. The coin is supposed to be fair, so I should be seeing about as many heads come up as tails. At this point, I'm *due* for a tails to show up.

The problem with this kind of thinking is that it fails to accommodate the fact that each successive coin toss is probabilistically independent of the outcome of any previous toss. That is, the probability that the next coin toss will land heads is independent of whether it landed heads (or tails) on the last toss, or the last four tosses, or the last ten thousand tosses. In this case, if we use 'H' to mean 'The coin will land heads', and 'P' to mean 'The coin landed heads previously', then:

$$\mathcal{P}r(H \mid P) \;=\; \mathcal{P}r(H) \;=\; 1/2$$

This may seem very obvious, but failing to account for independence is very common—and not only amongst gamblers! In general, people often seem to think that coins can remember how they landed on previous occasions, and then try to land differently the next time (all in the interests of trying to preserve the appearance of randomness).

Likewise, at the roulette wheel, people are more likely to vote red if the ball has landed black a few times in a row—as if the ball knew what it was doing, and wanted to keep things fair. This sort of thinking is absurd, of course, but it's a fact about us that we are often tempted to think like this.

10.3 Conditional probability and conjunctions

So far, we've introduced unconditional and conditional probabilities, but we haven't said much about how they're related to one another. Here, we'll introduce one more axiom to the three introduced in Chapter 9, which will allow us to connect the two kinds of probability to one another.

Axiom 4
For any propositions $\mathcal{A}$ and $\mathcal{B}$, $\mathcal{P}r(\mathcal{A} \wedge \mathcal{B}) = \mathcal{P}r(\mathcal{A}|\mathcal{B}) \times \mathcal{P}r(\mathcal{B})$

To return to the cards example, what is the probability that a card drawn at random is both *red* and a *jack*?

Let 'R' stand for 'The card is red', and 'J' stand for 'The card is a jack'. We can then symbolize 'The card is a red jack' as '$R \wedge J$.' Since there are 2 red jacks in a deck of 52, $\mathcal{P}r(R \wedge J)$ is $2/52 = 1/26$. This is exactly what Axiom 4 tells us as well. Making the required substitutions, it says:

$$\mathcal{P}r(R \wedge J) = \mathcal{P}r(R|J) \times \mathcal{P}r(J)$$

Now, the probability that the card is red *given that* it is a jack is $1/2$: in a standard deck, there are 2 red jacks, and 2 black jacks. Of the 4 possible jacks, only 2 are red, so $2/4 = 1/2$. And the probability that the card is a jack is $1/13$. When we multiply $1/2$ and $1/13$ together, we get $1/26$.

With Axiom 4 in hand, we can now develop some new laws of probability. This next one is particularly useful:

Law of Probability 5
If $\mathcal{P}r(\mathcal{B}) > 0$, then $\mathcal{P}r(\mathcal{A}|\mathcal{B}) \; = \; \mathcal{P}r(\mathcal{A} \wedge \mathcal{B}) \; / \; \mathcal{P}r(\mathcal{B})$

In many developments of probability theory, Law 5 is actually taken as a *definition* of conditional probability. (In this case, Axiom 4 is also left out, being derivable from the earlier three axioms plus the definition.)

Note the importance of the condition, $\mathcal{P}r(\mathcal{B}) > 0$. It makes no mathematical sense to divide by zero, so we need to make sure we never do so. Philosophers are undecided on what to say about $\mathcal{P}r(\mathcal{A}|\mathcal{B})$ when $\mathcal{P}r(\mathcal{B}) = 0$, and the question is a matter of much current debate. For our part, we don't have to worry about the case where $\mathcal{P}r(\mathcal{B}) = 0$

Next, recall the definition of probabilistic independence above: $\mathcal{A}$ is probabilistically independent of $\mathcal{B}$ if and only if $\mathcal{P}r(\mathcal{A}) = \mathcal{P}r(\mathcal{A}|\mathcal{B})$. This plus Axiom 4 lets us derive another law of probability:

> **Law of Probability 6**
> If $\mathcal{A}$ and $\mathcal{B}$ are probabilistically independent, then $\mathcal{P}r(\mathcal{A} \wedge \mathcal{B}) = \mathcal{P}r(\mathcal{A}) \times \mathcal{P}r(\mathcal{B})$

We've already seen an instance of this law in action: the probability that the card is red is probabilistically independent of whether it is a jack, so:

$$\mathcal{P}r(R|J) = \mathcal{P}r(R) = 1/2$$

So, by Law 6,

$$\mathcal{P}r(R \wedge J) = \mathcal{P}r(R) \times \mathcal{P}r(J)$$
$$\mathcal{P}r(R \wedge J) = 1/2 \times 1/13$$
$$\mathcal{P}r(R \wedge J) = 1/26$$

This is exactly the right result, and it's easier to use than Axiom 4. But be careful: the probability of a conjunction is only equal to the product of the probabilities of its conjuncts *if* those conjuncts are probabilistically independent of one another. In any other case, you should use Axiom 4 to work out the probability of a conjunction.

10.4 Bayes' theorem and the base rate fallacy

We are now in a position to introduce one of the most important equations in the theory of probability, **Bayes' theorem**. A useful formulation of the theorem is this (a proof is given in appendix C):

Law of Probability 7 (Bayes' Theorem)
For any pair of propositions $\mathcal{A}$ and $\mathcal{B}$, if $\mathcal{P}r(\mathcal{B}) > 0$, then:

$$\mathcal{P}r(\mathcal{A}|\mathcal{B}) = \frac{\mathcal{P}r(\mathcal{A})\times\mathcal{P}r(\mathcal{B}|\mathcal{A})}{\mathcal{P}r(\mathcal{B})}$$

Bayes' theorem tells us that the conditional probability $\mathcal{P}r(\mathcal{A}|\mathcal{B})$ depends on three things (assuming $\mathcal{P}r(\mathcal{B}) > 0$):

- The unconditional probability of $\mathcal{A}$,
- The probability of $\mathcal{B}$ conditional on $\mathcal{A}$, and
- The unconditional probability of $\mathcal{B}$.

Why is this important? Because it turns out that we humans are very bad at reasoning in accordance with Bayes' theorem. Here is an example.

Suppose there is a terrible disease going around, we'll call it disease X. This disease has no particular outward symptoms, but after three days it kills you. You know that about 1 in 1000 people have gotten disease X, but you don't know if you have it. So 1/1000 is the background probability of the disease being present in a randomly selected member of the population. Because the disease shows no symptoms, you have no particular reason to think that you're more likely to have the disease than anyone else.

Luckily, scientists have developed a test for the disease. If you have the disease, then the test *will* tell you that you have it with 100% accuracy. In the jargon, we say that the test has 0 probability of giving a **false negative**—the probability that the test will say you *don't* have the disease when you *do* is 0.

However, in some rare cases—let's say, 5%—the test will say that you have the disease even if you don't. That is, 95% of the time, if the test comes out positive for the disease, then the test subject actually has the disease; but in 5% of cases the test returns a **false positive**.

Being worried about your health, you go in for a test. It comes back positive. What's the probability that you have the disease, given this test result? At this point, most people will say that the probability is very high—many will say that it is 95%. However, the actual probability is closer to 2%.

It's helpful to think about this in terms of proportions. Let's say that a town of 100,000 people all decided to take this test. Since the prevalence of the disease is 1 in 1000, we should expect that the test will show a true positive for 100 people. However, of the remaining 99900 people in the population—all of whom do not have the disease—the test will *also* say that 5% of *them* have the disease as well. So the test will return a false positive for 4995 people on average. So now you've received a positive test result. You're either one of the unlucky 100 who actually have the disease, or one of the 4995 people for whom the test failed to work properly. So it's about 50 times more likely that you don't actually have the disease!

This is the same result we get from applying Bayes' theorem. Let 'D' stand for 'You have the disease', and 'T' stand for 'The test result is positive' Then, the unconditional probability of D (i.e., its probability *before* you take into account the evidence of the test result) is 0.001, or 1/1000. And the probability of T *given that* D is 1. So we can substitute these values into our formula for Bayes' theorem:

$$\mathcal{P}r(D|T) = \frac{\mathcal{P}r(D) \times \mathcal{P}r(T|D)}{\mathcal{P}r(T)}$$
$$\mathcal{P}r(D|T) = \frac{0.001 \times 1}{\mathcal{P}r(T)}$$

The only thing we need now is the unconditional probability of T. This is the background probability that the test will return a positive result regardless of whether the disease is present or not. We can easily work this out. In the population of 100,000, we've said that the test will return on average a true positive for 10 people and a false positive for 4995 people. So, on average it will return a positive result for 5095 people in total, out of 100,000. So:

$$\mathcal{P}r(T) = 5095/100000 = 0.05095$$

We can then work out the value of $\mathcal{P}r(D|T)$:

$$\mathcal{P}r(D|T) = (0.001 \times 1)\ /\ 0.05095$$
$$\mathcal{P}r(D|T) = 0.0196$$

In other words, even given the highly accurate testing procedure, the probability you have the disease given the positive test result is just under 2%.

The base rate fallacy is extremely important, and unfortunately very common. As the example shows, committing the fallacy can lead us to over-estimate the importance of test results. This is obviously problematic in situations where a correct understanding of the test results is crucial. In medical situations it can lead us to over-estimate the likelihood of a disease being present, leading to a misdiagnosis (with potentially life-threatening consequences).

Similarly, in legal situations the base rate fallacy can cause us to wrongly judge the importance of a piece of evidence or testimony. Indeed, sometimes the base rate fallacy is also called the *prosecutor's fallacy*, after hypothetical unscrupulous prosecutors who may try to use it to make their evidence appear stronger than it really is.

For instance, it is illegal to drink and drive, and police use breathalyzers to determine whether drivers are over the legal limit. However, no breathalyzer is perfect, and it's important to take this into account when we're considering the importance of breathalyzer results.

Suppose someone—let's call him 'Fred'—has been pulled over by police one night and given a breathalyzer test. The breathalyzer indicates drunkeness. Let 'D' now mean 'Fred was drink driving', and 'T' mean 'The breathalyzer indicates drunkeness'. Say that the breathalyzer is highly accurate, in that:

$$\mathcal{P}r(T|D) = 0.99$$
$$\mathcal{P}r(T|\neg D) = 0.01$$

That is, the breathalyzer correctly picks up on drunkeness 99% of the time, and only returns a false positive 1% of the time. So, what is the probability that Fred was drink driving, *given that* the breathalyzer indicates that he was?

By now you will know better than to answer '99%,' which would be the common intuitive answer. As it turns out, we haven't yet been given enough information to determine what $\mathcal{P}r(D|T)$ actually is. This is because we don't yet know the base rate. That is, we do not know the unconditional probability that Fred was drink driving, $\mathcal{P}r(D)$. If we knew that value, we could work out the value of $\mathcal{P}r(D|T)$.

Suppose that only 1 in every 100 drivers are drink driving on this particular night. This gives us a plausible value for the base rate $\mathcal{P}r(D)$. For a population of 10,000 drivers, there will be 100 drunks on the road. If every one of them were tested, the breathalyzer test would catch 99 of them. However, the test would also wrongly show that 99 out of the 9,900 *non*-drunk drivers were driving drunk. So, of the 198 people in total for whom the test indicates drunkeness, only 50% of them were actually drink driving.

In terms of Bayes' theorem, the relevant values are as follows:

- ▷ $\mathcal{P}r(D) = 0.01$
- ▷ $\mathcal{P}r(T|D) = 0.99$
- ▷ $\mathcal{P}r(T) = 0.0198$

You can work out $\mathcal{P}r(T)$ as just the total number of people for whom the test indicates drunkeness, divided by the total number of people in the population.

Once we know the values of $\mathcal{P}r(D)$, $\mathcal{P}r(T|D)$, and $\mathcal{P}r(T|\neg D)$, it's easy to work out the value of $\mathcal{P}r(T)$.

Plugging all this in to the equation gets us:

$$\mathcal{P}r(D|T) = \frac{\mathcal{P}r(D)\times\mathcal{P}r(T|D)}{\mathcal{P}r(T)}$$
$$\mathcal{P}r(D|T) = \frac{0.01\times 0.99}{0.0198}$$
$$\mathcal{P}r(D|T) = 0.5$$

The upshot here is: *don't ignore base rates*! Doing so could land you (or someone else) in deep trouble.

Practice Exercises

Part A
You have a standard deck of 52 cards. Of a card drawn at random, what is:

1. The probability that the card is hearts given that it's red?
2. The probability that the card is not a jack, given that it's a face card?
3. The probability that the card is hearts or clubs, given that it's an ace?
4. The probability that the card is an ace or a king, given it's not a 2 or a 3?
5. The probability that the card is a black jack, given it's either red or black?
6. The probability that the card is hearts and not an ace, given that water is wet?

⋆ **Part B**
Which of the above cases involve probabilistic independence?

⋆ **Part C**
Your friend is worried she might have caught the flu. This season, about 1 in every 100 people will catch it. As yet, she hasn't shown any particular symptoms, nor any particular reason to think that she's more likely to have caught it than anyone else. Nevertheless, she has a test which boasts high levels of accuracy. If you have the flu, the test will return a positive result 99% of the time. If you don't have the flu, it will return a negative result 90% of the time.

Letting 'F' stand for 'She has the flu', and 'T' stand for 'The test returns positive', what are the values of the following (rounded to the 4th decimal):

1. $\mathcal{P}r(F)$
2. $\mathcal{P}r(\neg F)$
3. $\mathcal{P}r(T|F)$
4. $\mathcal{P}r(\neg T|\neg F)$
5. $\mathcal{P}r(T|\neg F)$
6. $\mathcal{P}r(T)$
7. $\mathcal{P}r(\neg T)$
8. $\mathcal{P}r(F|T)$
9. $\mathcal{P}r(\neg F|\neg T)$

Part D
Explain how one might try to derive the following additional laws of probability, which hold for any pair of propositions $\mathcal{A}$ and $\mathcal{B}$. You can use any of the axioms and laws of probability already given in this chapter and the last chapter.

1. $\mathcal{P}r(\mathcal{A}|\mathcal{B}) \times \mathcal{P}r(\mathcal{B}) = \mathcal{P}r(\mathcal{B}|\mathcal{A}) \times \mathcal{P}r(\mathcal{A})$
2. $\mathcal{P}r(\mathcal{A}) = (\mathcal{P}r(\mathcal{A}|\mathcal{B}) \times \mathcal{P}r(\mathcal{B})) + (\mathcal{P}r(\mathcal{A}|\neg\mathcal{B}) \times \mathcal{P}r(\neg\mathcal{B}))$

Part IV

Appendix

Appendix A

Solutions to Selected Exercises

Chapter 1 Part A

1. Declarative
2. Declarative
3. Not declarative
4. Declarative
5. Declarative
6. Declarative
7. Not declarative
8. Not declarative
9. Declarative
10. Not declarative

Chapter 1 Part B

1. True
2. True
3. True
4. False
5. True
6. False
7. True
8. False
9. False
10. True

Chapter 2 Part A

1. $\neg M$
2. $G \land \neg C$
3. $M \lor \neg M$
4. $G \lor C$
5. $\neg(G \lor C)$
6. $\neg M \land \neg G$
7. $(G \lor C) \land \neg(G \land C)$

Chapter 2 Part C
Symbolization key:

A: Alice lifted the couch.
B: Bob lifted the couch.
C: Alice and Bob together lifted the couch.

1. $A \land B$
2. $(A \lor B) \land \neg(A \land B)$
3. C
4. $\neg A \land B$
5. $\neg(A \lor B)$

Chapter 3 Part B

1. $E_1 \land E_2$
2. $F_1 \to S_1$
3. $F_1 \lor E_1$
4. $E_2 \land \neg S_2$
5. $\neg E_1 \land \neg E_2$
6. $(E_1 \land E_2) \land \neg(S_1 \lor S_2)$
7. $S_2 \to F_2$
8. $(\neg E_1 \to \neg E_2) \land (E_1 \to E_2)$
9. $S_1 \leftrightarrow \neg S_2$
10. $(E_2 \land F_2) \to S_2$
11. $\neg(E_2 \land F_2)$
12. $(F_1 \land F_2) \leftrightarrow (\neg E_1 \land \neg E_2)$

Chapter 3 Part C

A: Alice is a spy.
B: Bob is a spy.

C: The code has been broken.
G: The German embassy will be in an uproar.

1. $A \wedge B$
2. $(A \vee B) \rightarrow C$
3. $\neg(A \vee B) \rightarrow \neg C$
4. $G \vee C$
5. $(C \vee \neg C) \wedge G$
6. $(A \vee B) \wedge \neg(A \wedge B)$

Chapter 4 Part A

1. tautology
2. contradiction
3. contingent
4. tautology
5. tautology
6. contingent
7. tautology
8. contradiction
9. tautology
10. contradiction
11. tautology
12. contingent
13. contradiction
14. contingent
15. tautology
16. tautology
17. contingent
18. contingent

Chapter 4 Part B 2, 3, 5, 6, 8, and 9 are logically equivalent.

Chapter 4 Part C 1, 3, 6, 7, and 8 are consistent.

Chapter 4 Part D 3, 5, 8, and 10 are valid.

Chapter 4 Part E

1. $\mathcal{A}$ and $\mathcal{B}$ have the same truth value on every line of a complete truth table, so $\mathcal{A} \leftrightarrow \mathcal{B}$ is true on every line. It is a tautology.
2. The proposition is false on some line of a complete truth table. On that line, $\mathcal{A}$ and $\mathcal{B}$ are true and $\mathcal{C}$ is false. So the argument is invalid.

3. Since there is no line of a complete truth table on which all three propositions are true, the conjunction is false on every line. So it is a contradiction. The status of $\mathcal{B} \wedge (\mathcal{A} \wedge \mathcal{C})$ is exactly the same: the order of the conjuncts makes no difference.
4. Since $\mathcal{A}$ is false on every line of a complete truth table, there is no line on which $\mathcal{A}$ and $\mathcal{B}$ are true and $\mathcal{C}$ is false. So the argument is valid.
5. Since $\mathcal{C}$ is true on every line of a complete truth table, there is no line on which $\mathcal{A}$ and $\mathcal{B}$ are true and $\mathcal{C}$ is false. So the argument is valid.
6. Not much. $\mathcal{A} \vee \mathcal{B}$ is a tautology if $\mathcal{A}$ and $\mathcal{B}$ are tautologies; it is a contradiction if they are contradictions; it is contingent if they are contingent.
7. $\mathcal{A}$ and $\mathcal{B}$ have different truth values on at least one line of a complete truth table, and $\mathcal{A} \vee \mathcal{B}$ will be true on that line. On other lines, it might be true or false. So $\mathcal{A} \vee \mathcal{B}$ is either a tautology or it is contingent; it is *not* a contradiction.

Chapter 4 Part F

1. $\neg(\neg A \wedge \neg B)$
2. $\neg(A \wedge \neg B) \wedge \neg(\neg A \wedge B)$
3. $\neg A \vee B$
4. $\neg(\neg A \vee \neg B)$
5. $\neg(\neg A \vee \neg B) \vee \neg(A \vee B)$

Chapter 5 Part A

1	$(W \rightarrow \neg B)$	
2	$(A \wedge W)$	
3	$(B \vee (J \wedge K))$	
4	W	$\wedge$E 2
5	$\neg B$	$\rightarrow$E 1, 4
6	$(J \wedge K)$	$\vee$E 3, 5
7	K	$\wedge$E 6

1	$(L \leftrightarrow \neg O)$		
2	$(L \vee \neg O)$		
3		$\neg L$	
4		$\neg O$	$\vee$E 2, 3
5		L	$\leftrightarrow$E 1, 4
6		$\neg L$	R 3
7	L		$\neg$E 3–6

1	$(Z \to (C \land \neg N))$	
2	$(\neg Z \to (N \land \neg C))$	
3	$\vert\ \neg(N \lor C)$	
4	$\vert\ (\neg N \land \neg C)$	DeM 3
5	$\vert\ \vert\ Z$	
6	$\vert\ \vert\ (C \land \neg N)$	$\to$E 1, 5
7	$\vert\ \vert\ C$	$\land$E 6
8	$\vert\ \vert\ \neg C$	$\land$E 4
9	$\vert\ \neg Z$	$\neg$I 5–8
10	$\vert\ (N \land \neg C)$	$\to$E 2, 9
11	$\vert\ N$	$\land$E 10
12	$\vert\ \neg N$	$\land$E 4
13	$(N \lor C)$	$\neg$E 3–12

Chapter 5 Part B

1.

1	$(K \land L)$	want $(K \leftrightarrow L)$
2	$\vert\ K$	want L
3	$\vert\ L$	$\land$E 1
4	$\vert\ L$	want K
5	$\vert\ K$	$\land$E 1
6	$(K \leftrightarrow L)$	$\leftrightarrow$I 2–3, 4–5

2.

1	$(A \to (B \to C))$	want $((A \land B) \to C)$
2	$\vert\ (A \land B)$	want C
3	$\vert\ A$	$\land$E 2
4	$\vert\ (B \to C)$	$\to$E 1, 3
5	$\vert\ B$	$\land$E 2
6	$\vert\ C$	$\to$E 4, 5
7	$((A \land B) \to C)$	$\to$I 2–6

3.

1	$(P \wedge (Q \vee R))$	
2	$(P \to \neg R)$	want $(Q \vee E)$
3	P	$\wedge$E 1
4	$\neg R$	$\to$E 2, 3
5	$(Q \vee R)$	$\wedge$E 1
6	Q	$\vee$E 5, 4
7	$(Q \vee E)$	$\vee$I 6

4.

1	$((C \wedge D) \vee E)$	want $E \vee D$
2	$\quad \mid \neg E$	want D
3	$\quad \mid (C \wedge D)$	$\vee$E 1, 2
4	$\quad \mid D$	$\wedge$E 3
5	$(\neg E \to D)$	$\to$I 2–4
6	$(E \vee D)$	MC 5

5.

1	$(\neg F \to G)$	
2	$(F \to H)$	want $(G \vee H)$
3	$\quad \mid \neg G$	want H
4	$\quad \mid \neg\neg F$	MT 1, 3
5	$\quad \mid F$	DN 4
6	$\quad \mid H$	$\to$E 2, 5
7	$(\neg G \to H)$	$\to$I 3–6
8	$(G \vee H)$	MC 7

6.

1	$((X \wedge Y) \vee (X \wedge Z))$	
2	$\neg(X \wedge D)$	
3	$(D \vee M)$	want M
4	$\quad \neg X$	for reductio
5	$\quad (\neg X \vee \neg Y)$	$\vee$I 4
6	$\quad \neg(X \wedge Y)$	DeM 5
7	$\quad (X \wedge Z)$	$\vee$E 1, 6
8	$\quad X$	$\wedge$E 7
9	$\quad \neg X$	R 4
10	X	$\neg$E 4–9
11	$\quad \neg M$	for reductio
12	$\quad D$	$\vee$E 3, 11
13	$\quad (X \wedge D)$	$\wedge$I 10, 12
14	$\quad \neg(X \wedge D)$	R 2
15	M	$\neg$E 11–14

Chapter 6 Part A

1. $Za \wedge (Zb \wedge Zc)$
2. $Rb \wedge \neg Ab$
3. $Lcb \rightarrow Mb$
4. $(Ab \wedge Ac) \rightarrow (Lab \wedge Lac)$
5. $\exists x(Rx \wedge Zx)$
6. $\forall x(Ax \rightarrow Rx)$
7. $\exists x(Rx \wedge \neg Ax)$
8. $\exists x(Rx \wedge Lcx)$
9. $\forall x\big((Mx \wedge Zx) \rightarrow Lbx)\big)$

Chapter 6 Part B

1. $\neg\exists x(Tx)$
2. $\forall x(Mx \rightarrow Sx)$
3. $\exists x(\neg Sx)$
4. $\forall x(Sx \rightarrow Mx)$
5. $\neg\exists x(Bxx)$

6. $\neg\exists x(Cx \wedge (\neg Sx \wedge Tx))$

Chapter 7 Part B

1. $\forall x(Cxp \to Dx)$
2. $Cjp \wedge Fj$
3. $\exists x(Cxp \wedge Fx)$
4. $\neg\exists x(Sxj)$
5. $\forall x\big((Cxp \wedge Fx) \to Dx\big)$
6. $\neg\exists x(Cxp \wedge Mx)$
7. $\exists x(Cjx \wedge Sxe \wedge Fj)$
8. $Spe \wedge Mp$
9. $\forall x\big((Sxp \wedge Mx) \to \neg\exists y(Cyx)\big)$
10. $\exists x(Sxj \wedge \exists y(Cyx) \wedge Fj)$
11. $\forall x\big(Dx \to \exists y(Sxy \wedge Fy \wedge Dy)\big)$
12. $\forall x\big((Mx \wedge Dx) \to \exists y(Cxy \wedge Dy)\big)$

Chapter 7 Part D

1. $\forall x(Cx \to Bx)$
2. $\neg\exists x(Wx)$
3. $\exists x\exists y(Cx \wedge Cy \wedge x \neq y)$
4. $\exists x\exists y(Jx \wedge Ox \wedge Jy \wedge Oy \wedge x \neq y)$
5. $\forall x\forall y\forall z\big((Jx \wedge Ox \wedge Jy \wedge Oy \wedge Jz \wedge Oz) \to (x = y \vee x = z \vee y = z)\big)$
6. $\exists x\exists y\big(Jx \wedge Bx \wedge Jy \wedge By \wedge x \neq y \wedge \forall z((Jz \wedge Bz) \to (x = z \vee y = z))\big)$
7. $\exists x_1\exists x_2\exists x_3\exists x_4\big(Dx_1 \wedge Dx_2 \wedge Dx_3 \wedge Dx_4 \wedge x_1 \neq x_2 \wedge x_1 \neq x_3 \wedge x_1 \neq x_4 \wedge x_2 \neq x_3 \wedge x_2 \neq x_4 \wedge x_3 \neq x_4 \wedge \neg\exists y(Dy \wedge y \neq x_1 \wedge y \neq x_2 \wedge y \neq x_3 \wedge y \neq x_4)\big)$
8. $\exists x\big(Dx \wedge Cx \wedge \forall y((Dy \wedge Cy) \to x = y) \wedge Bx\big)$
9. $\forall x\big((Ox \wedge Jx) \to Wx\big) \wedge \exists x\big(Mx \wedge \forall y(My \to x = y) \wedge Wx\big)$
10. $\exists x\big(Dx \wedge Cx \wedge \forall y((Dy \wedge Cy) \to x = y) \wedge Wx\big) \to \exists x\forall y(Wx \leftrightarrow x = y)$
11. Wide scope: $\neg\exists x\big(Mx \wedge \forall y(My \to x = y) \wedge Jx\big)$
    Narrow scope: $\exists x\big(Mx \wedge \forall y(My \to x = y) \wedge \neg Jx\big)$
12. Wide scope: $\neg\exists x\exists z\big(Dx \wedge Cx \wedge Mz \wedge \forall y((Dy \wedge Cy) \to x = y) \wedge \forall y((My \to z = y) \wedge x = z)\big)$
    Narrow scope: $\exists x\exists z\big(Dx \wedge Cx \wedge Mz \wedge \forall y((Dy \wedge Cy) \to x = y) \wedge \forall y((My \to z = y) \wedge x \neq z)\big)$

Chapter 8 Part A

1	$\forall x \exists y(Rxy \vee Ryx)$	
2	$\forall x(\neg Rmx)$	
3	$\exists y(Rmy \vee Rym)$	$\forall$E 1
4	$(Rma \vee Ram)$	
5	$\neg Rma$	$\forall$E 2
6	Ram	$\vee$E 4, 5
7	$\exists x Rxm$	$\exists$I 6
8	$\exists x(Rxm)$	$\exists$E 3, 4–7

1	$(\forall x(\exists y(Lxy) \to \forall z(Lzx)))$	
2	Lab	
3	$(\exists y(Lay) \to \forall z(Lza))$	$\forall$E 1
4	$\exists y(Lay)$	$\exists$I 2
5	$\forall z(Lza)$	$\to$E 3, 4
6	Lca	$\forall$E 5
7	$(\exists y(Lcy) \to \forall z(Lzc))$	$\forall$E 1
8	$\exists y(Lcy)$	$\exists$I 6
9	$\forall z(Lzc)$	$\to$E 7, 8
10	Lcc	$\forall$E 9
11	$\forall x(Lxx)$	$\forall$I 10

1	$\forall x(Jx \to Kx)$	
2	$\exists x \forall y(Lxy)$	
3	$\forall x(Jx)$	
4	$\forall y(Lay)$	
5	Ja	$\forall$E 3
6	$(Ja \to Ka)$	$\forall$E 1
7	Ka	$\to$E 6, 5
8	Laa	$\forall$E 4
9	$(Ka \wedge Laa)$	$\wedge$I 7, 8
10	$\exists x(Kx \wedge Lxx)$	$\exists$I 9
11	$\exists x(Kx \wedge Lxx)$	$\exists$E 2, 4–10

1	$(\neg(\exists x(Mx) \vee \forall x(\neg Mx)))$	
2	$(\neg\exists x(Mx) \wedge \neg\forall x(\neg Mx))$	DeM 1
3	$\neg\exists x(Mx)$	$\wedge$E 2
4	$\forall x(\neg Mx)$	QN 3
5	$\neg\forall x(\neg Mx)$	$\wedge$E 2
6	$(\exists x(Mx) \vee \forall x(\neg Mx))$	$\neg$E 1–5

Chapter 8 Part B

1.

1	$(\neg(\forall x(Fx) \vee \neg\forall x(Fx)))$	for reductio
2	$(\neg\forall x(Fx) \wedge \neg\neg\forall x(Fx))$	DeM 1
3	$\neg\forall x(Fx)$	$\wedge$E 2
4	$\neg\neg\forall x(Fx)$	$\wedge$E 2
5	$(\forall x(Fx) \vee \neg\forall x(Fx))$	$\neg$E 1–5

2.

1	$\forall x(Mx \leftrightarrow Nx)$	
2	$(Ma \wedge \exists x(Rxa))$	want $\exists x(Nx)$
3	$(Ma \leftrightarrow Na)$	$\forall$E 1
4	Ma	$\wedge$E 2
5	Na	$\leftrightarrow$E 3, 4
6	$\exists x(Nx)$	$\exists$I 5

3.

1	$\forall x(\neg Mx \vee Ljx)$	
2	$\forall x(Bx \rightarrow Ljx)$	
3	$\forall x(Mx \vee Bx)$	want $\forall x(Ljx)$
4	$(\neg Ma \vee Lja)$	$\forall$E 1
5	$(Ma \rightarrow Lja)$	MC 4
6	$(Ba \rightarrow Lja)$	$\forall$E 2
7	$(Ma \vee Ba)$	$\forall$E 3
8	Lja	DIL 7, 5, 6
9	$\forall x(Ljx)$	$\forall$I 8

4.

1	$\forall x(Cx \wedge Dt)$	want $(\forall x(Cx) \wedge Dt)$
2	$(Ca \wedge Dt)$	$\forall$E 1
3	Ca	$\wedge$E 2
4	$\forall x(Cx)$	$\forall$I 3
5	Dt	$\wedge$E 2
6	$(\forall x(Cx) \wedge Dt)$	$\wedge$I 4, 5

5.

1	$\exists x(Cx \vee Dt)$	want $(\exists x(Cx) \vee Dt)$
2	$\quad (Ca \vee Dt)$	for $\exists$E
3	$\quad\quad \neg(\exists x(Cx) \vee Dt)$	for reductio
4	$\quad\quad (\neg\exists x(Cx) \wedge \neg Dt)$	DeM 3
5	$\quad\quad \neg Dt$	$\wedge$E 4
6	$\quad\quad Ca$	$\vee$E 2, 5
7	$\quad\quad \exists x(Cx)$	$\exists$I 6
8	$\quad\quad \neg\exists x(Cx)$	$\wedge$E 4
9	$\quad (\exists x(Cx) \vee Dt)$	$\neg$E 3–8
10	$(\exists x(Cx) \vee Dt)$	$\exists$E 1, 2–9

Chapter 8 Part I Regarding the translation of this argument, see p. 97.

1	$\exists x \forall y(\forall z(Lxz \to Lyz) \to Lxy)$	
2	$\quad \forall y(\forall z(Laz \to Lyz) \to Lay)$	
3	$\quad (\forall z(Laz \to Laz) \to Laa)$	$\forall$E 2
4	$\quad\quad \neg\exists x(Lxx)$	for reductio
5	$\quad\quad \forall x(\neg Lxx)$	QN 4
6	$\quad\quad \neg Laa$	$\forall$E 5
7	$\quad\quad \neg\forall z(Laz \to Laz)$	MT 5, 6
8	$\quad\quad\quad Lab$	
9	$\quad\quad\quad Lab$	R 8
10	$\quad\quad (Lab \to Lab)$	$\to$I 8—9
11	$\quad\quad \forall z(Laz \to Laz)$	$\forall$I 10
12	$\quad\quad \neg\forall z(Laz \to Laz)$	R 7
13	$\quad \exists x(Lxx)$	$\neg$E 4—12
14	$\exists x(Lxx)$	$\exists$E 1, 2—13

Chapter 8 Part K 2, 3, and 5 are logically equivalent.

Chapter 8 Part L 2, 4, 5, 7, and 10 are valid. Here is a complete proof of 2:

1	$\exists y \forall x (Rxy)$	want $\forall x \exists y (Rxy)$
2	$\quad \forall x (Rxa)$	
3	$\quad Rba$	$\forall$E 2
4	$\quad \exists y (Rby)$	$\exists$I 3
5	$\quad \forall x \exists y (Rxy)$	$\forall$I 4
6	$\forall x \exists y (Rxy)$	$\exists$E 1, 2–5

Chapter 9 Part B

1. Yes: $\mathcal{P}r(S \wedge V)$ is not 0, which it would be if S and V were inconsistent with one another.
2. No: $\mathcal{P}r(S \wedge V)$ does not equal $\mathcal{P}r(S)$ (or $\mathcal{P}r(V)$), which it would if S and V were equivalent to one another.
3. $\mathcal{P}r(S \vee V) = 0.375$.
4. $\mathcal{P}r((S \vee V) \wedge S) = 0.25$.

Chapter 10 Part B

3, 5 and 6 are all cases of probabilistic independence.

Chapter 10 Part C

1. $\mathcal{P}r(F) = 0.01$
2. $\mathcal{P}r(\neg F) = 0.99$
3. $\mathcal{P}r(T|F) = 0.99$
4. $\mathcal{P}r(\neg T|\neg F) = 0.9$
5. $\mathcal{P}r(T|\neg F) = 0.1$
6. $\mathcal{P}r(T) = 0.1089$
7. $\mathcal{P}r(\neg T) = 0.8911$
8. $\mathcal{P}r(F|T) = 0.0909$
9. $\mathcal{P}r(\neg F|\neg T) = 0.9999$

Quick Reference: Symbolization

Translations

Translating Connectives (chapters 2 and 3)

It is not the case that P	$\neg P$
Either P, or Q	$P \vee Q$
Neither P, nor Q	$\neg(P \vee Q)$ or $\neg P \wedge \neg Q$
Both P, and Q	$P \wedge Q$
If P, then Q	$P \rightarrow Q$
P only if Q	$P \rightarrow Q$
P if and only if Q	$P \leftrightarrow Q$
Unless P, Q	$P \vee Q$
P unless Q	$P \vee Q$

Predicates and Quantifiers (chapter 6)

All Fs are Gs	$\forall x(Fx \rightarrow Gx)$
Some Fs are Gs	$\exists x(Fx \wedge Gx)$
Not all Fs are Gs	$\neg\forall x(Fx \rightarrow Gx)$ or $\exists x(Fx \wedge \neg Gx)$
No Fs are Gs	$\forall x(Fx \rightarrow \neg Gx)$ or $\neg\exists x(Fx \wedge Gx)$

Identity (chapter 7, section 7.4)

a is b	$a = b$
a is not b	$a \neq b$
Only j is G	$\forall x(Gx \leftrightarrow x = j)$
Everything besides j is G	$\forall x(x \neq j \rightarrow Gx)$

Definite Descriptions (chapter 7, section 7.6)

The F is G.	$\exists x(Fx \wedge \forall y(Fy \rightarrow x = y) \wedge Gx)$

'The F is not G' can be translated two ways:

[**Wide**] It is not the case that the F is G	$\neg\exists x(Fx \wedge \forall y(Fy \rightarrow x = y) \wedge Gx)$
[**Narrow**] The F is non-G	$\exists x(Fx \wedge \forall y(Fy \rightarrow x = y) \wedge \neg Gx)$

There are at least ______ Fs (chapter 7, section 7.5)

1: $\exists x(Fx)$
2: $\exists x_1 \exists x_2(Fx_1 \wedge Fx_2 \wedge x_1 \neq x_2)$
3: $\exists x_1 \exists x_2 \exists x_3(Fx_1 \wedge Fx_2 \wedge Fx_3 \wedge x_1 \neq x_2 \wedge x_1 \neq x_3 \wedge x_2 \neq x_3)$
***n*:** $\exists x_1 \cdots \exists x_n(Fx_1 \wedge \cdots \wedge Fx_n \wedge x_1 \neq x_2 \wedge \cdots \wedge x_{n-1} \neq x_n)$

There are at most ______ Fs (chapter 7, section 7.5)

One way to say 'at most n things are F' is to put a negation sign in front of one of the symbolizations above and say $\neg$'at least $n+1$ things are F.' Equivalently:

1: $\forall x_1 \forall x_2\big((Fx_1 \wedge Fx_2) \rightarrow x_1 = x_2\big)$
2: $\forall x_1 \forall x_2 \forall x_3\big((Fx_1 \wedge Fx_2 \wedge Fx_3) \rightarrow (x_1 = x_2 \vee x_1 = x_3 \vee x_2 = x_3)\big)$
3: $\forall x_1 \forall x_2 \forall x_3 \forall x_4\big((Fx_1 \wedge Fx_2 \wedge Fx_3 \wedge Fx_4) \rightarrow (x_1 = x_2 \vee x_1 = x_3 \vee x_1 = x_4 \vee x_2 = x_3 \vee x_2 = x_4 \vee x_3 = x_4)\big)$
***n*:** $\forall x_1 \cdots \forall x_{n+1}\big((Fx_1 \wedge \cdots \wedge Fx_{n+1}) \rightarrow (x_1 = x_2 \vee \cdots \vee x_n = x_{n+1})\big)$

There are exactly ______ Fs (chapter 7, section 7.5)

One way to say 'exactly n things are F' is to conjoin two of the symbolizations above and say 'at least n things are F' $\wedge$ 'at most n things are F.' The following equivalent formulae are shorter:

0: $\forall x(\neg Fx)$
1: $\exists x\big(Fx \wedge \neg\exists y(Fy \wedge x \neq y)\big)$
2: $\exists x_1 \exists x_2\big(Fx_1 \wedge Fx_2 \wedge x_1 \neq x_2 \wedge \neg\exists y\big(Fy \wedge y \neq x_1 \wedge y \neq x_2\big)\big)$
3: $\exists x_1 \exists x_2 \exists x_3\big(Fx_1 \wedge Fx_2 \wedge Fx_3 \wedge x_1 \neq x_2 \wedge x_1 \neq x_3 \wedge x_2 \neq x_3 \wedge \neg\exists y(Fy \wedge y \neq x_1 \wedge y \neq x_2 \wedge y \neq x_3)\big)$
***n*:** $\exists x_1 \cdots \exists x_n\big(Fx_1 \wedge \cdots \wedge Fx_n \wedge x_1 \neq x_2 \wedge \cdots \wedge x_{n-1} \neq x_n \wedge \neg\exists y(Fy \wedge y \neq x_1 \wedge \cdots \wedge y \neq x_n)\big)$

Quick Reference: Probability Theory

Axioms of Probability

Axiom 1
For any $\mathcal{A}$, $\mathcal{P}r(\mathcal{A}) \geq 0$.

Axiom 2
For any tautology $\mathcal{T}$, $\mathcal{P}r(\mathcal{T}) = 1$.

Axiom 3
If $\mathcal{A}$ and $\mathcal{B}$ are jointly inconsistent, then $\mathcal{P}r(\mathcal{A} \vee \mathcal{B}) = \mathcal{P}r(\mathcal{A}) + \mathcal{P}r(\mathcal{B})$

Axiom 4
For any $\mathcal{A}$ and $\mathcal{B}$, $\mathcal{P}r(\mathcal{A} \wedge \mathcal{B}) = \mathcal{P}r(\mathcal{A}|\mathcal{B}) \times \mathcal{P}r(\mathcal{B})$

Laws of Probability

Law of Probability 1
For any $\mathcal{A}$, $\mathcal{P}r(\mathcal{A}) + \mathcal{P}r(\neg\mathcal{A}) \ = \ 1$.
Alternative version: $\mathcal{P}r(\mathcal{A}) \ = \ 1 - \mathcal{P}r(\neg\mathcal{A})$.

▷ *Proof:*

By Axiom 1, every tautology has a probability of 1. Therefore, for any $\mathcal{A}$,

$$\mathcal{P}r(\mathcal{A} \vee \neg\mathcal{A}) \ = \ 1$$

Now, because $\mathcal{A}$ and $\neg\mathcal{A}$ are always inconsistent with one another, Axiom 3 says that:

$$\mathcal{P}r(\mathcal{A} \vee \neg\mathcal{A}) \ = \ \mathcal{P}r(\mathcal{A}) + \mathcal{P}r(\neg\mathcal{A})$$

Putting the two together, we see that:

$$\mathcal{P}r(\mathcal{A}) + \mathcal{P}r(\neg\mathcal{A}) \ = \ 1$$

The alternative version of the law then follows immediately by subtracting $\mathcal{P}r(\neg\mathcal{A})$ from both sides of the last equation.

Law of Probability 2
If *no more than one* member of a (finite) set of n propositions $\{\mathcal{A}_1, \mathcal{A}_2, \ldots, \mathcal{A}_n\}$ can be true, then:

$$\mathcal{P}r(\mathcal{A}_1 \vee \mathcal{A}_2 \vee \ldots \vee \mathcal{A}_n) \;=\; \mathcal{P}r(\mathcal{A}_1) + \mathcal{P}r(\mathcal{A}_2) + \ldots + \mathcal{P}r(\mathcal{A}_n)$$

▷ *Proof:*

The proof of this is effectively given in the main text (§9.2). In general, if every member of $\{\mathcal{A}_1, \mathcal{A}_2, \ldots, \mathcal{A}_n\}$ is inconsistent with every other member, then

$$\mathcal{P}r(\mathcal{A}_1 \vee \mathcal{A}_2 \vee \mathcal{A}_3 \vee \ldots \vee \mathcal{A}_n) \;=\; \mathcal{P}r(\mathcal{A}_1) + \mathcal{P}r(\mathcal{A}_2 \vee \mathcal{A}_3 \vee \ldots \vee \mathcal{A}_n)$$

Continuing, we know:

$$\mathcal{P}r(\mathcal{A}_2 \vee \mathcal{A}_3 \vee \ldots \vee \mathcal{A}_n) \;=\; \mathcal{P}r(\mathcal{A}_2) + \mathcal{P}r(\mathcal{A}_3 \vee \ldots \vee \mathcal{A}_n)$$

From there, it's straightforward to see how to continue. If $\{\mathcal{A}_1, \mathcal{A}_2, \ldots, \mathcal{A}_n\}$ is a partitin, then $(\mathcal{A}_2 \vee \mathcal{A}_3 \vee \ldots \vee \mathcal{A}_n)$ is a tautology, and the second part of Law 2 follows from the first part plus Axiom 2.

Law of Probability 3
If $\mathcal{A}$ and $\mathcal{B}$ are logically equivalent, then $\mathcal{P}r(\mathcal{A}) = \mathcal{P}r(\mathcal{B})$.

▷ *Proof:*

Suppose that $\mathcal{A}$ and $\mathcal{B}$ are logically equivalent. (For example, you might suppose that $\mathcal{A} = D_1$ and $\mathcal{B} = \neg\neg D_1$.) Then, two things are true. First, $(\mathcal{A} \vee \neg\mathcal{B})$ is a tautology. (You can check this using truth tables if you're uncertain.) Therefore, by applying Axiom 2, we know:

$$\mathcal{P}r(\mathcal{A} \vee \neg\mathcal{B}) = 1$$

Second, $\mathcal{A}$ and $\neg\mathcal{B}$ are jointly inconsistent. Therefore, by applying Axiom 3, we know:

$$\mathcal{P}r(\mathcal{A} \vee \neg\mathcal{B}) = \mathcal{P}r(\mathcal{A}) + \mathcal{P}r(\neg\mathcal{B})$$

Putting the two together yields:

$$\mathcal{P}r(\mathcal{A}) + \mathcal{P}r(\neg\mathcal{B}) = 1$$

Now we know from Law 1 that the following things must be true:

$$\mathcal{P}r(\mathcal{A}) + \mathcal{P}r(\neg\mathcal{A}) = 1$$

$$\mathcal{P}r(\mathcal{B}) + \mathcal{P}r(\neg\mathcal{B}) = 1$$

It then follows by simple algebra that $\mathcal{P}r(\neg\mathcal{A}) = \mathcal{P}r(\neg\mathcal{B})$, and $\mathcal{P}r(\mathcal{A}) = \mathcal{P}r(\mathcal{B})$. Thus, Law 3 is proved.

Law of Probability 4
If $\mathcal{A}$ and $\mathcal{B}$ are any two propositions, then:

$$\mathcal{P}r(\mathcal{A} \vee \mathcal{B}) = \mathcal{P}r(\mathcal{A}) + \mathcal{P}r(\mathcal{B}) - \mathcal{P}r(\mathcal{A} \wedge \mathcal{B})$$

▷ *Proof:*

To begin with, it's helpful to note that any sentence of the form $(\mathcal{A} \vee \mathcal{B})$ is logically equivalent to $(\mathcal{A} \wedge \mathcal{B}) \vee (\mathcal{A} \wedge \neg\mathcal{B}) \vee (\neg\mathcal{A} \wedge \mathcal{B})$. That is, $(\mathcal{A} \vee \mathcal{B})$ is true just in case both disjuncts are true, only the first disjunct is true, or only the second disjunct is true. Each of these three possibilities are inconsistent with the two others. Therefore, according to Law 2:

$$\mathcal{P}r(\mathcal{A} \vee \mathcal{B}) = \mathcal{P}r((\mathcal{A} \wedge \mathcal{B}) \vee (\mathcal{A} \wedge \neg\mathcal{B}) \vee (\neg\mathcal{A} \wedge \mathcal{B}))$$

And then according to Law 3:

$$\mathcal{P}r(\mathcal{A} \vee \mathcal{B}) = \mathcal{P}r((\mathcal{A} \wedge \mathcal{B}) \vee (\mathcal{A} \wedge \neg\mathcal{B})) + \mathcal{P}r(\neg\mathcal{A} \wedge \mathcal{B})$$

But now note that $(\mathcal{A} \wedge \mathcal{B}) \vee (\mathcal{A} \wedge \neg\mathcal{B})$ is logically equivalent to just $\mathcal{A}$. Hence, we can replace the $\mathcal{P}r((\mathcal{A} \wedge \mathcal{B}) \vee (\mathcal{A} \wedge \neg\mathcal{B}))$ on the right-hand side of the previous equation with $\mathcal{P}r(\mathcal{A})$:

$$\mathcal{P}r(\mathcal{A} \vee \mathcal{B}) = \mathcal{P}r(\mathcal{A}) + \mathcal{P}r(\neg\mathcal{A} \wedge \mathcal{B})$$

Now we're going to pull a little trick. First of all, we'll add $\mathcal{P}r(\mathcal{A} \wedge \mathcal{B})$ to both sides of this equation:

$$\mathcal{P}r(\mathcal{A} \vee \mathcal{B}) + \mathcal{P}r(\mathcal{A} \wedge \mathcal{B}) = \mathcal{P}r(\mathcal{A}) + \mathcal{P}r(\neg\mathcal{A} \wedge \mathcal{B}) + \mathcal{P}r(\mathcal{A} \wedge \mathcal{B})$$

And, just as before we said that $(\mathcal{A} \wedge \mathcal{B}) \vee (\mathcal{A} \wedge \neg\mathcal{B})$ was logically equivalent to $\mathcal{A}$, so too is $((\mathcal{A} \wedge \mathcal{B}) \vee (\neg\mathcal{A} \wedge \mathcal{B}))$ logically equivalent to $\mathcal{B}$. Hence,

$$\mathcal{P}r(\mathcal{B}) = \mathcal{P}r(\neg\mathcal{A} \wedge \mathcal{B}) + \mathcal{P}r(\mathcal{A} \wedge \mathcal{B})$$

This means we can substitute $\mathcal{P}r(\mathcal{B})$ for $\mathcal{P}r(\neg\mathcal{A} \wedge \mathcal{B}) + \mathcal{P}r(\mathcal{A} \wedge \mathcal{B})$ in the equation above to get:

$$\mathcal{P}r(\mathcal{A} \vee \mathcal{B}) + \mathcal{P}r(\mathcal{A} \wedge \mathcal{B}) = \mathcal{P}r(\mathcal{A}) + \mathcal{P}r(\mathcal{B})$$

From this point, we just subtract $\mathcal{P}r(\mathcal{A} \wedge \mathcal{B})$ from both sides of the equation:

$$\mathcal{P}r(\mathcal{A} \vee \mathcal{B}) = \mathcal{P}r(\mathcal{A}) + \mathcal{P}r(\mathcal{B}) - \mathcal{P}r(\mathcal{A} \wedge \mathcal{B})$$

And so concludes our proof of Law 4.

Law of Probability 5
If $\mathcal{P}r(\mathcal{B}) > 0$, then $\mathcal{P}r(\mathcal{A}|\mathcal{B}) \ = \ \mathcal{P}r(\mathcal{A} \wedge \mathcal{B}) \ / \ \mathcal{P}r(\mathcal{B})$

▷ *Proof:*

Given Axiom 4, the proof is very straightforward. That axiom says:

$$\mathcal{P}r(\mathcal{A} \wedge \mathcal{B}) = \mathcal{P}r(\mathcal{A}|\mathcal{B}) \times \mathcal{P}r(\mathcal{B})$$

Assume $\mathcal{P}r(\mathcal{B}) > 0$; then, we can divide both sides by $\mathcal{P}r(\mathcal{B})$ to get the equation in Law 5.

Law of Probability 6
If $\mathcal{A}$ and $\mathcal{B}$ are probabilistically independent, then $\mathcal{P}r(\mathcal{A} \wedge \mathcal{B}) \ = \ \mathcal{P}r(\mathcal{A}) \times \mathcal{P}r(\mathcal{B})$

▷ *Proof:*

This follows immediately from Axiom 4 and the definition of independence (§10.2). If $\mathcal{A}$ and $\mathcal{B}$ are independent, then $\mathcal{P}r(\mathcal{A}|\mathcal{B}) = \mathcal{P}r(\mathcal{A})$. So we can substitute $\mathcal{P}r(\mathcal{A})$ for $\mathcal{P}r(\mathcal{A}|\mathcal{B})$ in Axiom 4 to get Law 6.

Law of Probability 7
For any pair $\mathcal{A}$ and $\mathcal{B}$, if $\mathcal{P}r(\mathcal{B}) > 0$, then:

$$\mathcal{P}r(\mathcal{A}|\mathcal{B}) = \frac{\mathcal{P}r(\mathcal{A}) \times \mathcal{P}r(\mathcal{B}|\mathcal{A})}{\mathcal{P}r(\mathcal{B})}$$

▷ *Proof:*

We know that $(\mathcal{A} \wedge \mathcal{B})$ and $(\mathcal{B} \wedge \mathcal{A})$ are logically equivalent, so they have the same probability (Law 3). So, by Axiom 4,

$$\mathcal{P}r(\mathcal{A} \wedge \mathcal{B}) = \mathcal{P}r(\mathcal{A}|\mathcal{B}) \times \mathcal{P}r(\mathcal{B})$$

And:

$$\mathcal{P}r(\mathcal{B} \wedge \mathcal{A}) = \mathcal{P}r(\mathcal{B}|\mathcal{A}) \times \mathcal{P}r(\mathcal{A})$$

Putting all this together, we get:

$$\mathcal{P}r(\mathcal{A}|\mathcal{B}) \times \mathcal{P}r(\mathcal{B}) = \mathcal{P}r(\mathcal{B}|\mathcal{A}) \times \mathcal{P}r(\mathcal{A})$$

Now, assuming that $\mathcal{P}r(\mathcal{B}) > 0$, we can divide both sides by $\mathcal{P}r(\mathcal{B})$. Doing so nets us:

$$\mathcal{P}r(\mathcal{A}|\mathcal{B}) = \frac{\mathcal{P}r(\mathcal{A}) \times \mathcal{P}r(\mathcal{B}|\mathcal{A})}{\mathcal{P}r(\mathcal{B})}$$

Quick Reference: Proof Rules

Basic Rules of Proof

Reiteration

$$\begin{array}{r|ll} m & \mathcal{A} & \\ & \mathcal{A} & \text{R } m \end{array}$$

Conjunction Introduction

$$\begin{array}{r|ll} m & \mathcal{A} & \\ n & \mathcal{B} & \\ & (\mathcal{A} \wedge \mathcal{B}) & \wedge\text{I } m, n \end{array}$$

Conjunction Elimination

$$\begin{array}{r|ll} m & (\mathcal{A} \wedge \mathcal{B}) & \\ & \mathcal{A} & \wedge\text{E } m \\ & \mathcal{B} & \wedge\text{E } m \end{array}$$

Disjunction Introduction

$$\begin{array}{r|ll} m & \mathcal{A} & \\ & (\mathcal{A} \vee \mathcal{B}) & \vee\text{I } m \\ & (\mathcal{B} \vee \mathcal{A}) & \vee\text{I } m \end{array}$$

Disjunction Elimination

$$\begin{array}{r|ll} m & (\mathcal{A} \vee \mathcal{B}) & \\ n & \neg\mathcal{B} & \\ & \mathcal{A} & \vee\text{E } m, n \end{array}$$

$$\begin{array}{r|ll} m & (\mathcal{A} \vee \mathcal{B}) & \\ n & \neg\mathcal{A} & \\ & \mathcal{B} & \vee\text{E } m, n \end{array}$$

Conditional Introduction

m	$\quad\vert\ \mathcal{A}$	want $\mathcal{B}$
n	$\quad\vert\ \mathcal{B}$	
$\mathcal{A} \to \mathcal{B})$	(	$\to$I m–n

Conditional Elimination

m	$(\mathcal{A} \to \mathcal{B})$	
n	$\mathcal{A}$	
	$\mathcal{B}$	$\to$E m, n

Biconditional Introduction

m	$\quad\vert\ \mathcal{A}$	want $\mathcal{B}$
n	$\quad\vert\ \mathcal{B}$	
p	$\quad\vert\ \mathcal{B}$	want $\mathcal{A}$
q	$\quad\vert\ \mathcal{A}$	
	$(\mathcal{A} \leftrightarrow \mathcal{B})$	$\leftrightarrow$I m–n, p–q

Biconditional Elimination

m	$(\mathcal{A} \leftrightarrow \mathcal{B})$	
n	$\mathcal{B}$	
	$\mathcal{A}$	$\leftrightarrow$E m, n

m	$(\mathcal{A} \leftrightarrow \mathcal{B})$	
n	$\mathcal{A}$	
	$\mathcal{B}$	$\leftrightarrow$E m, n

Negation Introduction

m	$\quad\vert\ \mathcal{A}$	for reductio
$n-1$	$\quad\vert\ \mathcal{B}$	
n	$\quad\vert\ \neg\mathcal{B}$	
	$\neg\mathcal{A}$	$\neg$I m–n

Negation Elimination

$$
\begin{array}{l|l|l}
m & \quad \underline{\neg\mathcal{A}} & \text{for reductio} \\
n-1 & \quad \mathcal{B} & \\
n & \quad \neg\mathcal{B} & \\
 & \mathcal{A} & \neg\text{E } m\text{–}n
\end{array}
$$

Quantifier Rules

Existential Introduction

$$
\begin{array}{l|ll}
m & \mathcal{A} & \\
 & \exists\mathcal{X}(\mathcal{A}\boxed{\boxed{c^*\Rightarrow\mathcal{X}}}) & \exists\text{I } m
\end{array}
$$

* $\mathcal{X}$ may replace some or all occurrences of c in $\mathcal{A}$.

Existential Elimination

$$
\begin{array}{l|ll}
m & \exists\mathcal{X}(\mathcal{A}) & \\
n & \quad \underline{\mathcal{A}\boxed{\mathcal{X}\Rightarrow c^*}} & \\
p & \quad \mathcal{B} & \\
 & \mathcal{B} & \exists\text{E } m,\ n\text{–}p
\end{array}
$$

* c must not appear outside the subproof.

Universal Introduction

$$
\begin{array}{l|ll}
m & \mathcal{A} & \\
 & \forall\mathcal{X}(\mathcal{A}\boxed{c\Rightarrow\mathcal{X}}) & \forall\text{I } m
\end{array}
$$

c must not occur in any undischarged assumptions.

Universal Elimination

$$
\begin{array}{r|ll}
m & \forall \chi(\mathcal{A}) & \\
 & \mathcal{A}\boxed{\chi \Rightarrow c} & \forall\mathrm{E}\ m
\end{array}
$$

Identity Rules

$$
\begin{array}{r|ll}
 & c = c & =\mathrm{I}
\end{array}
$$

$$
\begin{array}{r|ll}
m & c = d & \\
n & \mathcal{A} & \\
 & \mathcal{A}\boxed{\boxed{c \Rightarrow d}} & =\mathrm{E}\ m,\ n \\
 & \mathcal{A}\boxed{\boxed{d \Rightarrow c}} & =\mathrm{E}\ m,\ n
\end{array}
$$

One constant may replace some or all occurrences of the other.

Derived Rules

Dilemma

$$
\begin{array}{r|ll}
m & (\mathcal{A} \vee \mathcal{B}) & \\
n & (\mathcal{A} \rightarrow \mathcal{C}) & \\
p & (\mathcal{B} \rightarrow \mathcal{C}) & \\
 & \mathcal{C} & \mathrm{DIL}\ m,\ n,\ p
\end{array}
$$

Modus Tollens

$$
\begin{array}{r|ll}
m & (\mathcal{A} \rightarrow \mathcal{B}) & \\
n & \neg\mathcal{B} & \\
 & \neg\mathcal{A} & \mathrm{MT}\ m,\ n
\end{array}
$$

Hypothetical Syllogism

$$\begin{array}{r|ll} m & (\mathcal{A} \to \mathcal{B}) & \\ n & (\mathcal{B} \to \mathcal{C}) & \\ & (\mathcal{A} \to \mathcal{C}) & \text{HS } m, n \end{array}$$

Replacement Rules

Commutivity (Comm)
$(\mathcal{A} \wedge \mathcal{B}) \iff (\mathcal{B} \wedge \mathcal{A})$
$(\mathcal{A} \vee \mathcal{B}) \iff (\mathcal{B} \vee \mathcal{A})$
$(\mathcal{A} \leftrightarrow \mathcal{B}) \iff (\mathcal{B} \leftrightarrow \mathcal{A})$

DeMorgan (DeM)
$\neg(\mathcal{A} \vee \mathcal{B}) \iff (\neg\mathcal{A} \wedge \neg\mathcal{B})$
$\neg(\mathcal{A} \wedge \mathcal{B}) \iff (\neg\mathcal{A} \vee \neg\mathcal{B})$

Double Negation (DN)
$\neg\neg\mathcal{A} \iff \mathcal{A}$

Material Conditional (MC)
$(\mathcal{A} \to \mathcal{B}) \iff (\neg\mathcal{A} \vee \mathcal{B})$
$(\mathcal{A} \vee \mathcal{B}) \iff (\neg\mathcal{A} \to \mathcal{B})$

Biconditional Exchange ($\leftrightarrow$ex)
$((\mathcal{A} \to \mathcal{B}) \wedge (\mathcal{B} \to \mathcal{A})) \iff (\mathcal{A} \leftrightarrow \mathcal{B})$

Quantifier Negation (QN)
$\neg\forall\chi(\mathcal{A}) \iff \exists\chi(\neg\mathcal{A})$
$\neg\exists\chi(\mathcal{A}) \iff \forall\chi(\neg\mathcal{A})$