forall$x$

BRISTOL

With-Answers

For PHIL10032 Logic and Critical Thinking
University of Bristol

In-progress version for 2025–26
Last updated October 10, 2025. Please check for updates.

Current maintainer: Catrin Campbell-Moore

This book is provided on a best-endeavours basis.
Not all questions are available and there will be mistakes.

It is only recommended to use these answers when you have already tried it yourself.
Use the standard version for learning!

Links for the most recent versions of the 2025–26 edition:

- forall$x$: Bristol 2025–26.
- forall$x$: Bristol Accessible 2025–26.
  With formatting designed with students in dyslexia in mind.
  Adopted from Richard Zach (see discussion).
- forall$x$: Bristol With-Answers 2025–26.
  Is provided on a best endeavours basis.
  Has gaps and mistakes.

Links to other editions can be found at

catrincm.github.io/forallx-bris

forall x: Bristol

By **P. D. Magnus**
Tim Button
with additions by
J. Robert Loftis
Robert Trueman
remixed and revised by
Aaron Thomas-Bolduc
Richard Zach
and by
Catrin Campbell-Moore
Johannes Stern

`forall`x: *Bristol* is based on `forall`x: *Calgary* by

Aaron Thomas-Bolduc & Richard Zach
University of Calgary

used under a CC BY-SA 4.0 license, which is based on `forall`x: *Cambridge*, by

Tim Button
University of Cambridge

used under a CC BY-SA 3.0 license, which is based in turn on `forall`x, by

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

used under a CC BY-SA 3.0 license.

The `forall`x: *Bristol* version is by

Catrin Campbell-Moore & Johannes Stern
University of Bristol

Further contributors: Stuart Presnell, Richard Pettigrew.

Preface

The book is divided into nine parts. Part I introduces the topic and notions of logic in an informal way, without introducing a formal language yet. Parts II–IV concern truth-functional languages. In it, sentences are formed from basic sentences using a number of connectives ('or', 'and', 'not', 'if ... then') which just combine sentences into more complicated ones. We discuss logical notions in two ways: semantically, using the method of truth tables (in Part II) and proof-theoretically, using a system of formal derivations (in Part IV). Parts V–VII deal with a more complicated language, that is, the language of first-order logic. It includes, in addition to the connectives of truth-functional logic, also names, predicates, identity, and the so-called quantifiers. These additional elements of the language makes it much more expressive than the truth-functional language, and we'll spend a fair amount of time investigating just how much one can express in it. Again, logical notions for the language of first-order logic are defined semantically, using interpretations, and proof-theoretically, using a more complex version of the formal derivation system introduced in Part IV.

In the appendices you'll find a discussion of alternative notations for the languages we discuss in this text, of alternative derivation systems, and a quick reference listing most of the important rules and definitions. The central terms are listed in a glossary at the very end.

This book is based on a text originally written by P. D. Magnus in the version revised and expanded by Tim Button. It also includes some material (mainly exercises) by J. Robert Loftis. Aaron Thomas-Bolduc and Richard Zach have combined elements of these texts into the present version, changed some of the terminology and examples, rewritten some sections, and added material of their own. Catrin Campbell-Moore and Johannes Stern have then both made substantial

Contents

Appendices 308

PART I

Arguments

CHAPTER 1

Arguments

1.1 Introduction and Overview

In philosophy we think, reason, and argue. We intellectually scrutinize theories and view points. We reflect on arguments in support or in opposition of some philosophical position. In other words arguments and argumentation take centerstage in philosophy. But what is an argument and what are good arguments?

An argument, as we will understand it, is something like this:

- If I acted of my own free will, then I could have acted otherwise.
- I could not have acted otherwise.
- Therefore: I did not act of my own free will.

We here have a series of sentences which may either be true or false. The final sentence, "I did not act of my own free will." expresses the *conclusion* of the argument. The two sentences before that are the *premises* of the argument. In a good argument, the conclusion follows from the premises. If you believe the premises then the argument should lead you to believing the conclusion.

On the face of it, our sample argument appears to be a good argument, but can one develop a general theory of what a good argument is, that is, a theory that abstracts away from particular arguments and equips us with general tools for evaluating arguments? It turns out that one can and this general theory is what we call 'logic'.

In logic we want to abstract away from specific arguments and their particular content, and focus on specific structural features of arguments. To achieve this logic is spelled out in its own particular language like, say, latin. This means that learning logic amounts to learning a new language and it can be helpful to approach the course with this mindset. As we want to make sure that there are no ambiguities or misunderstandings the language of logic happens to be a formal language. Now to learn a language one needs to

VOC: learn its vocabulary (what are the basic expressions of language?);

GRA: learn its grammar (what is a well-formed sentence of the language?);

SEM: learn what the sentences of the new language say or mean;

PFTH: learn how to use the new language and how to reason with it.

This course is about learning logic and thus learning the language of logic. Accordingly, we will work ourselves through the different items of the above list. But in addition we need to think about how this helps us with assessing philosophical arguments and reasoning, which is carried out in English or some other natural language. To this effect we discuss how one can symbolize (or formalize) English arguments in the language of logic, which will allow us to evaluate English arguments in the logical language. Summing up, the structure of the course will be as follows: we first introduce the SYNTAX of the language, which subsumes both VOC and GRA, and then turn to its SEMANTICS (SEM), that is, the meaning of the sentences of the language. The semantics of the language will also allow us to spell out whether an argument is a good argument from the logical perspective. Such an argument is called a *valid* argument. We then explain how to symbolize English sentences in the logical language, which enables us to check whether an English argument is valid or not. Finally, we turn to PfTh and introduce a reasoning system for the logical language, that makes it possible for us to argue or reason within the logical language. In the context of our formal language, this means that we give a proof system, that is, we specify a number of rules that allows us to derive a sentence (the conclusion) from other sentences (the premises) in the same way that in natural language a good argument licenses us to infer the conclusion of the argument from its premises. Importantly, the rules will only allow us to produce valid argument. If we apply the rules correctly our reasoning cannot go wrong (this doesn't meant that our premises were true however).

But we are getting ahead of ourselves and before we introduce the language of logic, we should first better understand what an argument is and which arguments are good ones. To this effect we discuss arguments in natural languages and more specifically arguments in English. As we mentioned at the beginning, understanding and evaluating arguments is a crucial aspect of philosophy. Reflecting and investigating arguments will thus help you to analyse some of the texts you are reading during your philosophy studies.

1.2 Finding the components of an argument

Arguments consist of a list of *premises* along with a *conclusion*. In a good argument, the conclusion will follow from the premises. Our standard way to present them is:

1. Premise 1
2. Premise 2

   . . .

n. Premise *n*

∴ Therefore: Conclusion

For example

1. If I acted of my own free will, then I could have acted otherwise.
2. I could not have acted otherwise.

∴ Therefore: I did not act of my own free will.

The three dots in this final line can be read "therefore". Really then we're duplicating things by also adding the word "Therefore". But we do this to really carefully highlight that this is the conclusion.

Often arguments are presented simply in a paragraph of text, or in a speech or article, and we first have to work out what the premises and conclusions are. Sometimes it's easy, for example:

> If I acted of my own free will, then I could have acted otherwise. But, I could not have acted otherwise. So, I did not act of my own free will.

But often it is a significant piece of work to work out the premises and conclusion of an argument.

Many arguments start with premises, and end with a conclusion, but not all of them. It might start with the conclusion:

> **We should not have a second Brexit referendum**. A second Brexit referendum would erode the very basis of

democracy by suggesting that rule by the majority is an insufficient condition for democratic legitimacy.

Or it might have been presented with the conclusion in the middle:

> Since the first Brexit referendum was made under false pretences, **the voters deserve a further say on any final deal agreed with Brussels**. After all, decisions as big as this need to have the public support, which has to come from a referendum.

Sometimes premises or the conclusion may be clauses in a sentence. A complete argument may even be contained in a single sentence:

> The butler has an alibi; so they cannot have done it.

This argument has one premise, followed immediately by its conclusion.

One particular kind of sentence can be confusing. Consider:

- If the murder weapon was a gun, then Prof. Plum did it.

These conditional, or "if-then", statements might look like it expresses the argument, but in itself it does not. It's just stating a fact, albeit a conditional fact. It might also be used in an argument, even as the conclusion of the argument:

1. If I have free will, then there is some event that I could have caused to go differently.
2. If determinism is true, then there is no event that I could have caused to go differently.

∴ Therefore: If determinism is true, I do not have free will.

As a guideline, there are some words you can look for which are often used to indicate whether something is a premise or conclusion:

Words often used to indicate an argument's conclusion:

so, therefore, hence, thus, accordingly, consequently

Words often used to indicate a premise:

since, because, as, given that, recalling that, after all

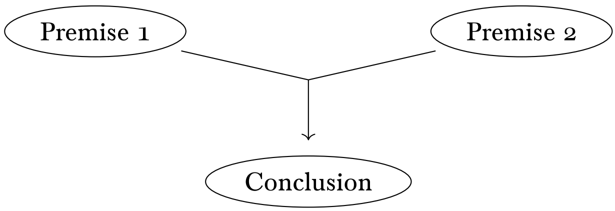
In analysing an argument, there is no substitute for a good nose. Whenever you come across an argument in a piece of philosophy you read, be it lecture notes, primary text, or secondary text, or in a newspaper article or on the internet, practice identifying the premises and conclusion.

Sometimes, though, people aren't giving arguments but are simply presenting facts or stating their opinion. For example, the following do not contain arguments, they're not trying to convince us of anything.

- I don't like cats. I think they're evil.
- Hundreds of vulnerable children as young as 10, who have spent most of their lives in the UK, are having their applications for British citizenship denied for failing to pass the government's controversial 'good character' test.

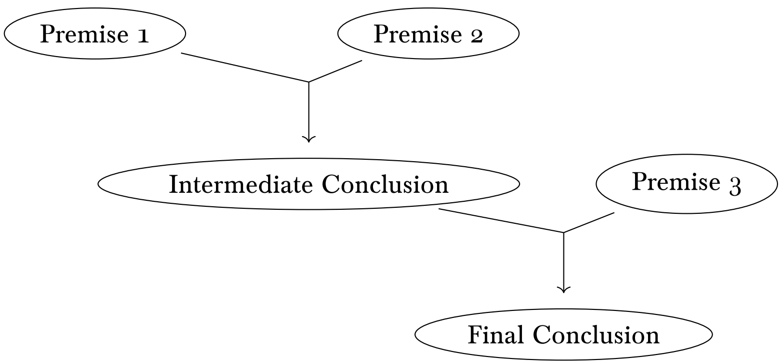
1.3 Intermediate Conclusions

We said an argument is given by a collection of *premises* along with a single *conclusion*. We might represent this as something like:



The premises are working together to lead to the conclusion.

But sometimes in the process of someone making an argument someone will make use of *intermediate conclusions*. Such arguments might have a structure more like:



However, we say that an argument is only something of the first kind. So what do we say about the second kind of thing? We can consider it two ways. We could consider it as an argument from premise 1, 2 and 3 to the conclusion. Or alternatively we can think of it as two arguments of the first kind chained together, one from premise 1 and 2 to the intermediate conclusion, and the second from the intermediate conclusion and premise 3 to the final conclusion.

1.4 Sentences

What kinds of things are the premises and conclusions of arguments? They are sentences which can either be true or false. Such sentences are called DECLARATIVE SENTENCES.

There are many other kinds of sentences, for example:

Questions 'Are you sleepy yet?' is an interrogative sentence. Although you might be sleepy or you might be alert, the question itself is neither true nor false. For this reason, questions will not count as declarative sentences. Suppose you answer the question: 'I am not sleepy.' This is either true or false, and so it is a declarative sentences. Generally, *questions* will not count as declarative sentences, but *answers* will.

'What is this course about?' is not a declarative sentence (in our sense). 'No one knows what this course is about' is a declarative sentence.

Imperatives Commands are often phrased as imperatives like 'Wake up!', 'Sit up straight', and so on. These are imperative sentences. Although it might be good for you to sit up straight or it might not, the command is neither true nor false and it is thus not a declarative sentence. Note, however, that commands are not always phrased as imperatives. 'You will respect my authority' *is* either true or false— either you will or you will not— and so it counts as a declarative sentences.

Exclamations 'Ouch!' is sometimes called an exclamatory sentence, but it is not the sort of thing which is true or false. 'That hurt!', however, is a declarative sentence.

Arguments are formed of *declarative sentences* — those sentences which can be true or false — for example 'spiders have eight legs'.

> An ARGUMENT consists of a collection of declarative sentences of which one is marked as the conclusion of the argument.

The unmarked declarative sentences are of course the premises of the argument. We typically drop the term 'declarative' and simply call them sentences, but bear in mind that it is only these sorts of sentences that are relevant in this textbook.

You should not confuse the idea of a sentence that can be true or false with the difference between fact and opinion. Often, sentences in logic will express things that would count as facts— such as 'spiders have eight legs' or 'Kierkegaard liked almonds.' They can also express things that you might think of as matters of opinion—such as, 'Almonds are tasty.' In other words, a sentence is not disqualified from being part of an argument because we don't know if it is true or false, or because its truth or falsity seems to be a purely subjective matter. All that matters is whether what the sentence expresses it is the sort of thing that could be true or false. If it is, it can play the role of premise or conclusion.

When you are reading a text and putting it in our standard form you should make sure that your premises and conclusions are declarative sentences. You should also make them as clear as possible. Each premise and the conclusion should be able to be read and understood independently. Any context from the original paragraph should be copied over to each of the premises and conclusions. For example:

> Donating to charity no strings attached is the most effective way to do so. So if you are going to donate to charity, you should do it this way.

When presenting this we should fill out "this way" with the relevant way. So I'd write:

1. Donating to charity no strings attached is the most effective way to do so.

∴ Therefore: If you are going to donate to charity, you should do so no strings attached.

Practice exercises

At the end of some chapters, there are exercises that review and explore the material covered in the chapter. There is no substitute for actually working through some problems. This course isn't about memorizing facts but about developing a way of thinking.

So here's the first exercise.

A.

1. Are arguments always presented in our standard form?
2. Do conclusions always come after the premises in arguments in texts?
3. Might premises and conclusions be clauses within sentences?
4. Can questions be premises?

B. Write down the conclusion of each of these arguments:

1. It is sunny. So I should take my sunglasses.
   I should take my sunglasses.
2. It must have been sunny. I did wear my sunglasses, after all.
   It was sunny.
3. No one but you has had their hands in the cookie-jar. And the scene of the crime is littered with cookie-crumbs. You're the culprit!
4. Miss Scarlett and Professor Plum were in the study at the time of the murder. Reverend Green had the candlestick in the ballroom, and we know that there is no blood on his hands. Hence Colonel Mustard did it in the kitchen with the lead-piping. Recall, after all, that the gun had not been fired.
5. Since I do not know that I am not under the spell of a malicious demon, I do not know that this table exists. After all, if I know that this table exists, then I know that I am not under the spell of a malicious demon.
6. Cutting the interest rate will have no effect on the stock market this time round as people have been expecting a rate cut all along. This factor has already been reflected in the market.
7. Virgin would then dominate the rail system. Is that something the government should worry about? Not necessarily. The industry is regulated, and one powerful company might at least offer a more coherent schedule of services than the present arrangement has produced. The reason the industry was broken up into more than 100 companies at privatisation was not operational, but political: the Conservative government thought it would thus be harder to renationalise. *The Economist 16.12.2000; used on critical thinking web*

8. The idea that being vegetarian is better for the environment has, over the last forty years, become a piece of conventional wisdom. But it is simply wrong. A paper from Carnegie Mellon University researchers published this week finds that the diets recommended by the Dietary Guidelines for Americans, which include more fruits and vegetables and less meat, exacts a greater environmental toll than the typical American diet. Shifting to the diets recommended by Dietary Guidelines for American would increase energy use by 38 percent, water use by ten percent and greenhouse gas emissions by six percent, according to the paper.
9. There are no hard numbers, but the evidence from Asia's expatriate community is unequivocal. Three years after its handover from Britain to China, Hong Kong is unlearning English. The city's gweilos (Cantonese for "ghost men") must go to ever greater lengths to catch the oldest taxi driver available to maximize their chances of comprehension. Hotel managers are complaining that they can no longer find enough English- speakers to act as receptionists. Departing tourists, polled at the airport, voice growing frustration at not being understood.
   The Economist 20.1.2001, used in Critical Thinking Web

C. Write each of the following arguments in the standard form.

x. It might surprise you, but denoting to charity no strings attached is the most effective way to do so. So if you are going to donate to charity, you should do it this way.

1. Answer:

   1. Denoting to charity no strings attached is the most effective way to do so.

   ∴ Therefore: If you are going to donate to charity, you should do so no strings attached.

2. It is sunny. So I should take my sunglasses.
   I should take my sunglasses

3. It must have been sunny. I did wear my sunglasses, after all.
   It was sunny

4. No one but you has had their hands in the cookie-jar. And the scene of the crime is littered with cookie-crumbs. You're the culprit!
   You're the culprit

5. Kate didn't write it. If Kate or David wrote it, it will be reliable; and it isn't.

6. Since I do not know that I am not under the spell of a malicious demon, I do not know that this table exists. After all, if I know that this table exists, then I know that I am not under the spell of a malicious demon.

7. Miss Scarlett and Professor Plum were in the study at the time of the murder. And Reverend Green had the candlestick in the ballroom, and we know that there is no blood on his hands. Hence Colonel Mustard did it in the kitchen with the lead-piping. Recall, after all, that the gun had not been fired.
   Colonel Mustard did it in the kitchen with the lead-piping

CHAPTER 2

The scope of logic

2.1 Consequence and validity

In §1, we talked about arguments, i.e., a collection of sentences (the premises), followed by a single sentence (the conclusion). We said that some words, such as "therefore," indicate which sentence is supposed to be the conclusion. "Therefore," of course, suggests that there is a connection between the premises and the conclusion, namely that the conclusion *follows from*, or *is a consequence of*, the premises.

This notion of consequence is one of the primary things logic is concerned with and we will ultimately define a precise notion of consequence for our formal language. One might even say that logic is the science of what follows from what. Logic develops a general account of and general tools that tell us when a sentence follows from some other sentences.

What about the following argument:

1. The butler or the gardener did it.
2. The butler did not do it.

∴ Therefore: The gardener did it.

We don't have any context for what the sentences in this argument refer to. Perhaps you suspect that "did it" here means "was the perpetrator" of some unspecified crime. You might imagine that the argument occurs in a mystery novel or TV show, perhaps spoken by a detective working through the evidence. But even without having any of this information, you probably agree that the argument is a good one in the sense that whatever the premises refer to, if they are both true, the conclusion is guaranteed to be true as well. If the first premise is true, i.e., it's true that "the butler did it or the gardener did it," then at least one of them "did it," whatever that means. And if the second premise is true, then the butler did not "do it." That leaves only one option: "the gardener did it" must be true. Here, the conclusion follows from the premises. We call arguments that have this property VALID.

By way of contrast, consider the following argument:

1. If the driver did it, the maid didn't do it.
2. The maid didn't do it.

∴ Therefore: The driver did it.

We still have no idea what is being talked about here. But, again, you probably agree that this argument is different from the previous one in an important respect. If the premises are true, it is not guaranteed that the conclusion is also true. The premises of this argument do not rule out, by themselves, that someone other than the maid or the driver "did it." In this second argument, the conclusion does not follow from the premises. If, like in this argument, the conclusion does not follow from the premises, we say it is INVALID.

We said the first argument was valid because if the premises will be true, we are guaranteed that the conclusion is true. In fact in this argument the premises guarantee the truth independently of whether we are talking about butlers, crocodiles, murderers or cake thieves:

1. The crocodile or the kangaroo stole the cake.
2. The crocodile did not steal the cake.

∴ Therefore: The kangaroo stole the cake.

It is irrelevant for the validity of the argument what the premises and the conclusion are about. The argument is valid on all ways of interpreting the premises as long as we understand the sentential connectives 'or' and 'not'. To put it in less abstract terms an argument is valid if and only if no COUNTEREXAMPLE to the argument can be produced.

> An argument is VALID if and only if there is no interpretation such that all the premises are true and the conclusion false. Otherwise the argument is INVALID.

We said that the following argument was invalid.

1. If the driver did it, the maid didn't do it.
2. The maid didn't do it.

∴ Therefore: The driver did it.

Can we find a counterexample to the argument? The answer is yes. For example, if we understand 'did it' as 'mowed the lawn', then the premises of the argument are both true, but the conclusion is false, as it was the gardener who mowed the lawn (at least on one interpretation/scenario).

Earlier we introduced the idea that the conclusion of an argument is meant to follow from the premises; that it is a consequence of the premises. This motivates the following definition:

> A sentence Y is a LOGICAL CONSEQUENCE of sentences $X_1, \ldots, X_n$ if and only if there is no interpretation such that $X_1, \ldots, X_n$ are all true and Y is not true. (We then also say that Y LOGICALLY FOLLOWS FROM $X_1, \ldots, X_n$.)

Another way of saying that Y is a logical consequence of sentences $X_1, \ldots, X_n$ is to say that the argument with premises $X_1, \ldots, X_n$ and conclusion Y is valid.

Valid arguments are arguments for which there is no interpretation such that all premises are true but the conclusion is not. It is irrelevant whether such an interpretation is "reasonable": no matter how unreasonable an interpretation is that can be used to give a counterexample to an argument, if there exists such an interpretation the argument will not be valid.

Is there a straightforward way of telling whether an argument is logically valid? Is there some feature that sets apart all valid arguments from other (possibly convincing) arguments? We have already seen that validity should not depend on the content of the premises and conclusion. Rather it should only depend on their (logical) form. For instance, consider the valid argument

1. Either Priya is an ophthalmologist or a dentist.
2. Priya isn't a dentist.

∴ Therefore: Priya is an ophthalmologist.

We can describe the "form" of this argument as the following pattern:

1. Either a is an F or a G.
2. a isn't an F.

∴ Therefore: a is a G.

Here, a, F, and G are placeholders for appropriate expressions that, when substituted for a, F, and G, turn the pattern into an argument consisting of sentences (at a first approximation this is also one way of understanding the "interpretation" talk). For instance,

1. Either Mei is a mathematician or a botanist.
2. Mei isn't a botanist.

∴ Therefore: Mei is a mathematician.

is an argument of the same form and it is also valid. However, the following argument is not of the same form:

1. Either Priya is an ophthalmologist or a dentist.
2. Priya isn't a dentist.

∴ Therefore: Priya is an eye doctor.

we would have to replace F by different expressions (once by "ophthalmologist" and once by "eye doctor") to obtain it from the pattern. This argument is not valid. To see that the conclusion follows from the premises we need the additional information that an ophthalmologist is indeed an eye doctor, that is, we need information that "ophthalmologist" and "eye doctor" mean the same thing.

To see more clearly that the latter argument cannot be deemed valid solely on the basis of its logical form let's consider *its* form:

1. Either a is an F or a G.
2. a isn't an F.

∴ Therefore: a is a H.

In this pattern we can replace F by "ophthalmologist" and H by "eye doctor" to obtain the original argument. But here is another argument of the same form which can be obtained by replacing F by "is a mathematician", G by "is a botanist", and. H by "is an acrobat":

1. Either Mei is a mathematician or a botanist.
2. Mei isn't a botanist.

∴ Therefore: Mei is an acrobat.

This argument is clearly not valid. The conclusion does not follow from the premises of the argument.

In valid arguments the conclusion follows from the premises of the argument solely in virtue of its logical form, that is, the logical structure of the premises and the conclusion. This feature is an aspect of the so-called FORMALITY of logic. Much of the present logic course will be devoted to studying and determining valid argument forms and structures, and to make precise the idea of *interpretation* we used in discussing the validity of arguments.

2.2 Sound arguments

Arguments in our sense, as conclusions which (supposedly) follow from premises, are of course used all the time in everyday reasoning, but also philosophical and scientific discourse. When they are, arguments are given to support or even prove their conclusions. Now, if an argument is valid, it will support its conclusion, but *only if* its premises are all true: validity rules out that the premises are true and the conclusion false. It does not, by itself, rule out that the conclusion is false, as the premises can be false. An argument can be valid, but have false premises. In short, a valid argument may have a conclusion that is not true!

Consider this example:

1. Oranges are either fruit or musical instruments.
2. Oranges are not fruit.

∴ Therefore: Oranges are musical instruments.

The conclusion of this argument is ridiculous. Nevertheless, it logically follows from the premises due to the logical form of the argument. For what the argument is concerned, it is not relevant whether the oranges are musical instruments (of course, they are not!). What is relevant is that if according to a (weird) interpretation Oranges are either fruit or musical instruments, but not fruit, then oranges are musical instruments according to that interpretation: *If* both premises are true, *then* the conclusion just has to be true independently of the content of the premises and the conclusion. The argument is valid.

Conversely, having true premises and a true conclusion does not guarantee that the argument is valid. Consider this example:

1. London is in England.

2. Beijing is in China.
∴ Therefore: Paris is in France.

The premises and conclusion of this argument are all true, but the argument is invalid. The logical form of premises and conclusion do not guarantee that the conclusion is true whenever the premises are true: on an interpretation on which 'France' is interpreted to mean Great Britain, the conclusion is not true, even though both of the premises would remain true. So the argument is invalid.

The important thing to remember is that validity is not about the truth or falsity of the sentences in the argument. It is about whether the conclusion follows from the premises of the argument in virtue of their logical form; about whether the conclusion is true whenever the premises are true, that is, whether for all interpretations on which all premises are true, the conclusion is true likewise. Nothing about the way things are—whether something is true or false—can by itself determine if an argument is valid. It is often said that logic doesn't care about feelings. Actually, it doesn't care about facts, either.

When we use an argument to prove that its conclusion *is true*, then, we need two things. First, we need the argument to be valid, i.e., we need the conclusion to logically follow from the premises. But we also need the premises to be true. We will say that an argument is SOUND if and only if it is both valid and all of its premises are true.

The flip side of this is that when you want to rebut an argument, you have two options: you can show that (one or more of) the premises are not true, or you can show that the argument is not valid. Logic, however, will only help you with the latter!

2.3 Missing premises

Most arguments we make and evaluate in everyday reasoning are not valid simpliciter. We are often interested in whether the conclusion follows from the premises given certain implicit or explicit background assumptions which the interlocutor has failed to explicitly mention. If the missing background assumption is explicitly added as a premise, the argument may turn out to be valid after all. For example, we already discussed that the argument

1. Either Priya is an ophthalmologist or a dentist.
2. Priya isn't a dentist.
∴ Therefore: Priya is an eye doctor.

is strictly speaking not a valid argument. Still it seems to be a good argument in the sense that the truth of the premises seems to guarantee the truth of the conclusion. One explanation for why we think it is a good argument, is that it can easily be turned into a valid argument by adding the (true) premise

1. If Priya is an ophthalmologist, then Priya is an eye doctor.

Arguably this premise is one we all implicitly assume, which explains why an interlocutor might not feel the need of mentioning it.

Sometimes it is not obvious to tell what kind of implicit underlying assumption are assumed in the formulation of an argument. If someone you disagree with makes an invalid argument it's often more useful (and more charitable) to consider whether there are missing premises rather than to simply dismiss the argument. Perhaps the author or interlocutor was assuming that an additional premise was so obvious that it didn't need to be stated.

For example an author might make the following argument:

1. I could not have acted otherwise.

∴ Therefore: I did not act of my own free will.

This argument is invalid. But, it can be made valid by addition of the premise:

1. If I could not have acted otherwise, I did not act of my own free will.

But be careful when you're filling in 'missing' premises. The aim is to help improve the argument, to make it more convincing, so you can assess it fairly. Only add extra premises that seem reasonable, or that you think the original author would agree with. There's no point in adding absurd or unreasonable premises, or premises that the author wouldn't endorse. Then you just create a *strawman* argument – a caricature of the original argument.

> "Just how charitable are you supposed to be when criticizing the views of an opponent? If there are obvious contradictions in the opponent's case, then of course you should point them out, forcefully. If there are somewhat hidden contradictions, you should carefully expose them to view—and then dump on them. But the search for hidden

> contradictions often crosses the line into nitpicking, sea-lawyering, and—as we have seen—outright parody. The thrill of the chase and the conviction that your opponent has to be harboring a confusion somewhere encourages uncharitable interpretation, which gives you an easy target to attack. But such easy targets are typically irrelevant to the real issues at stake and simply waste everybody's time and patience, even if they give amusement to your supporters."
> *Daniel C. Dennett (2013). "Intuition Pumps And Other Tools for Thinking".*

Dennett formulates the following four rules (named after Anatol Rapoport) for "how to compose a successful critical commentary":

1. You should attempt to re-express your target's position so clearly, vividly, and fairly that your target says, "Thanks, I wish I'd thought of putting it that way."

2. You should list any points of agreement (especially if they are not matters of general or widespread agreement).

3. You should mention anything you have learned from your target.

4. Only then are you permitted to say so much as a word of rebuttal or criticism

2.4 Ampliative Arguments

There is further reason why many arguments of everyday reasoning are not strictly speaking valid: not all arguments of everyday reasoning are so-called DEDUCTIVE arguments. In deductive arguments the truth of the premises is supposed to guarantee the truth of the conclusion. Not all good arguments are deductive and sometimes there are no plausible missing premises you could add to someone's argument to make it valid.

However, this doesn't necessarily mean that the interlocutor was wrong or mistaken. Deductively valid arguments with plausible premises are good arguments, but they aren't the only good arguments there are. This is just as well, since many arguments we give in our everyday lives are not deductively valid, even after filling in plausible missing premises. Here's an example:

1. Every winter we have so far observed, it has rained in Bristol.

∴ Therefore: It will rain in Bristol this coming winter.

This argument generalises from observations about several cases to a conclusion about all cases. Such arguments are called INDUCTIVE arguments. The argument could be made stronger by adding additional premises before drawing the conclusion: In January 2001, it rained in London; In January 2002.... But, however many premises of this form we add, the argument will remain invalid. Even if it has rained in London in every January thus far, it remains possible that London will stay dry next January. The point of all this is that inductive arguments—even good inductive arguments—are not (deductively) valid. They are not watertight. The premises might make the conclusion very likely, but they don't absolutely guarantee its truth. Unlikely though it might be, it is possible for their conclusion to be false, even when all of their premises are true.

Inductive arguments of the sort just given belong to a species of argument called AMPLIATIVE ARGUMENTS. This means that the conclusion goes beyond what you find in the premises. That is, the premises don't guarantee, or entail, the conclusion. They do, however, provide some support for it. These arguments are deductively invalid. They may be good and useful, however it is important to know the difference.

In this book, we will set aside the question of what makes for a good ampliative argument and focus instead on sorting the deductively valid arguments from the deductively invalid ones. But we pause here to mention some further forms of ampliative argument.

Inductive arguments, like the one we saw above, allow one to infer from a series of observed cases to a generalization that covers them: from all observed Fs have been Gs, we infer all Fs are Gs. We use these all the time. Every time I've drunk water from my tap, it's quenched my thirst; therefore, every time I ever drink water from my tap, it will quench my thirst. Every time I've stroked my neighbour's cat, it hasn't bitten me; therefore, every time I ever stroke my neighbour's cat, it won't bite me. And it's a form of arguments much beloved by scientists. Every time we've measured the acceleration of a body falling, it's matched Newton's theory, therefore, all bodies are governed by Newton's theory. The premises of these argument seem to make their conclusions likely without guaranteeing them. The areas of philosophy called inductive logic or confirmation theory try to make precise what that means and why it's true. And of course inductive arguments can go wrong. Before I visited Australia, every swan I'd seen was white, and so I concluded that all swans were white; but when I visited Australia, I realised my conclusion was wrong, because some swans there are black.

A closely related, but different form of argument, is STATISTICAL. Here, we start with an observation about the proportion of Fs that are Gs in a sample that we've observed, and we infer that the same proportion of Fs are Gs in general. So, for instance, if I poll 1,000 people in Scotland eligible to vote in a second independence referendum, and 600 say that they'll vote yes, I might infer that 60% of all eligible voters will vote yes. Or if I test 1,000,000 people in England for an active infection, and 20,000 test positive, I might infer that 2% of the whole population has an active infection. How good these argument are depends on a number of things, and these are studied by statisticians. For instance, suppose you picked the 1,000 Scottish voters entirely at random from an anonymised version of the electoral register. But suppose that, when you deanonymised, you learned that, by chance, all of the people you'd picked were over 65, or they all lived on the Isle of Skye. Then you might worry that your sample, though random, was unrepresentative of the population as a whole. This question is a genuine concern for randomised controlled trials in medicine.

ABDUCTIVE ARGUMENTS provide an inference from a phenomenon you've observed to the *best explanation* of that phenomenon: from E, and the best explanation of E is H, you might conclude H. Again, this is extremely widespread. A classic sort of example would be the inferences that detectives draw during their investigations. They look at the evidence and the possible explanations of it, and they tend to conclude in favour of the best one. And similarly for doctors looking at a patient's suite of symptoms and trying to discover what ails them. Another important example comes from science. Here is Charles Darwin explaining what convinces him of his theory of natural selection:

> "It can hardly be supposed that a false theory would explain, in so satisfactory a manner as does the theory of natural selection, the several large classes of facts above specified. It has recently been objected that this is an unsafe method of arguing; but it is a method used in judging of the common events of life, and has often been used by the greatest natural philosophers."
> (Charles Darwin, On the origin of species by means of natural selection (6th ed.). London: John Murray)

2.5 Beyond Validity

As we mentioned, many arguments we make and evaluate in everyday reasoning are not strictly speaking valid. As discussed sometimes important implicit premises are not made explicit. However, the fact that implicit premises are not made explicit point to a more general phenomenon, namely, that in everyday reasoning we take certain conceptual or meaning relations for granted. Going back to the argument involving Priya, we found the the conclusion "Priya is an eye doctor" not to be a logical consequence of the premises despite the fact that intuitively one might be inclined to say the the conclusion follows from the premises of the argument. While the argument is not strictly speaking valid, the conclusion follows from the premises of the argument once we acknowledge that 'ophthalmologist' is just a fancy word for an eye doctor. More generally, there is no interpretation that respects all conceptual relations between expressions of the language on which all premises are true but the conclusion false, that is, there is no counterexample to the argument involving Priya which acknowledges that 'ophthalmologist' and 'eye doctor' mean the same thing.

Arguments for which there is no interpretation that respects all conceptual/meaning connections between the various words of our language are sometimes called CONCEPTUALLY VALID and sometimes you'll find the term 'validity' used in this sense in the literature . For example, the arguments

1. Priya is an ophthalmologist.

∴ Therefore: Priya is an eye doctor.

1. Jonas is a bachelor.

∴ Therefore: Jonas is an unmarried man.

are both conceptually valid but not (logically) valid according to our definition of validity. All (logically) valid arguments are also conceptually valid, but not the other way around.

Perhaps in everyday reasoning we are often judging arguments whether they are conceptually valid as opposed to valid simpliciter. However, while the cases of conceptual validity discussed have been reasonably clear, it is sometimes not that easy to make precise and agree on the exact underlying conceptual relations. Consequently, while, as we shall see, it is relatively straightforward to decide whether an argument is (logically) valid, this becomes much more tricky turning to conceptual validity. For this reason it is preferable to focus on (logical) validity and focus on what additional premises are needed to turn an intuitively convincing argument into a valid argument. In a second step one can then ask whether the additional premises are conceptual truths. If the answer is yes, then we can deem the argument conceptually valid despite being invalid in the strict sense of validity.

Practice exercises

A.

1. What kind of things are valid or invalid?
2. When is an argument said to be valid?
3. When is an argument said to be sound?

B. Are the following valid? If it is invalid, describe a counterexample.

x. 1. Every good zoo has a giraffe.
   2. It is a zoo.
   ∴ Therefore: It has a giraffe.

1. Invalid. It is a zoo, but not a good one. (And has a giraffe.)

2. 1. Everyone in group 1 handed in their homework.
   2. Jenny is in group 1.
   ∴ Therefore: Jenny handed in her homework.

   valid

   1. If she won the lottery then she is rich.
   2. She is rich.
   ∴ Therefore: She won the lottery.

   Invalid, there are many ways to get rich...

3. 1. Most people are scared of spiders.
   ∴ Therefore: Oscar is scared of spiders.

   Invalid

4. 1. She is a donkey.
   ∴ Therefore: She does not talk.

C. Which of the following arguments is valid? Which is invalid?

1. Socrates is a man.
2. All men are carrots.

∴ Socrates is a carrot. Valid

1. Abe Lincoln was either born in Illinois or he was president.
2. Abe Lincoln was not president.

∴ Abe Lincoln was born in Illinois. Valid

1. If I pull the trigger, Abe Lincoln will die.
2. I do not pull the trigger.

∴ Abe Lincoln will not die. Invalid
Abe Lincoln might die for some other reason: someone else might pull the trigger; he might die of old age.

1. Abe Lincoln was either from France or from Luxembourg.
2. Abe Lincoln was not from Luxembourg.

∴ Abe Lincoln was from France. Valid

1. If the world were to end today, then I would not need to get up tomorrow morning.
2. I will need to get up tomorrow morning.

∴ The world will not end today. Valid

1. Joe is now 19 years old.
2. Joe is now 87 years old.

∴ Bob is now 20 years old. Not valid (although arguably conceptually valid. To make it valid one would need to add the premise that one cannot be both 19 years old and 87 years old at the same time.

D. Could there be:

1. A valid argument that has one false premise and one true premise? Yes.
   Example: the first argument, above.
2. A valid argument that has only false premises? Yes.
   Example: Socrates is a frog, all frogs are excellent pianists, therefore Socrates is an excellent pianist.
3. A valid argument with only false premises and a false conclusion?
   Yes.
   The same example will suffice.

4. An invalid argument that can be made valid by the addition of a new premise? Yes.
   Plenty of examples, but let me offer a more general observation. We can *always* make an invalid argument valid, by adding a contradiction into the premises. For an argument is valid if and only if it is impossible for all the premises to be true and the conclusion false. If the premises are contradictory, then it is impossible for them all to be true (and the conclusion false).
5. A valid argument that can be made invalid by the addition of a new premise? No.
   An argument is valid if and only if it is impossible for all the premises to be true and the conclusion false. Adding another premise will only make it harder for the premises all to be true together.

In each case: if so, give an example; if not, explain why not.

PART II

Truth-functional logic

CHAPTER 3

A Prolegomenon to TFL

In this part the lecture notes we start developing our theory of validity and logical consequence, that is, we start introducing the logical language. In §2 we already introduced the idea that an argument is valid if and only if the truth of the premises guarantees the truth of the conclusion in virtue of their logical form. It is this idea that suggest to spell out an account of validity in a formal language, that is, to conceive of the logical language as a formal language. This will enable us to single out arguments that are valid in virtue of their form and eventually make sense of the notion of an interpretation we used in our definition of validity in §2. We can then give a rigorous formal definition of validity of arguments in the formal language we shall devise. This language will be the language of Truth-functional logic (TFL).

Before we introduce the language of TFL, let us take a look at why a formal language may be helpful for capturing validity of arguments, i.e., the validity of arguments in virtue of their form.

Consider this argument:

1. It is raining outside.
2. If it is raining outside, then Jenny is miserable.

∴ Therefore: Jenny is miserable.

and another argument:

1. Jenny is an anarcho-syndicalist.
2. If Jenny is an anarcho-syndicalist, then Dipan is an avid reader of Tolstoy.

∴ Therefore: Dipan is an avid reader of Tolstoy.

Both arguments are valid, and there is a straightforward sense in which we can say that they share a common structure. We might express the structure thus:

1. A
2. If A, then B

∴ Therefore: B

This looks like an excellent argument *structure*. Indeed, surely any argument with this *structure* will be valid.

What about:

1. Jenny is miserable.
2. If it is raining outside, then Jenny is miserable.

∴ Therefore: It is raining outside.

The form of this argument is:

1. B
2. If A then B

∴ Therefore: A

Arguments of this form are generally invalid.

Be careful, though, not every argument of this form is sure to be invalid. It's possible to have an argument of this form that's valid – see if you can work out how! But most arguments of this form are invalid.

There a lot more valid argument forms. For example the argument form

1. A or B
2. not-A

∴ Therefore: B

as well as the form

1. not-(A and B)
2. A

∴ Therefore: not-B

lead to valid arguments independently of what expressions we substitute for 'A' and 'B': we can understand (interpret) 'A' and 'B' in whatever way we want, as long as we take them to be place holder for sentences the resulting argument will be valid. These examples illustrate the important idea that the validity of the arguments just considered has nothing to do with the meanings of English expressions like 'Jenny is miserable', 'Dipan is an avid reader of Tolstoy', or any other sentence. If it has to do with meanings at all, it is with the meanings of conjunction-words like 'and', 'or', 'not,' and 'if..., then...'. The language of truth-functional logic is built to single out characteristic feature of these conjunction words and this will enable us to fruitfully study the idea of validity of an argument in virtue of its form.

When one introduces a language there are (at least) two task: the first is to specify the vocabulary of the language and equip the language with a grammar, that is, one has to specify how well-formed sentences of the language look like. This aspect of the language is called its SYNTAX. The second task is to specify the SEMANTICS of the language. The semantics specifies how we are to understand the expressions of the language, what the sentences of the language mean etc. Part II develops both the syntax and the semantics of the language of TFL. Once this has been established we can consider how TFL may be useful for thinking about arguments in English. This will lead to the idea of symbolizing arguments in TFL and will be picked up in Part 25.

CHAPTER 4

Syntax of TFL

4.1 Atomic sentences

We started isolating the form of an argument by replacing *subsentences* of sentences with individual letters. Thus in the first example of this section, 'it is raining outside' is a subsentence of 'If it is raining outside, then Jenny is miserable', and we replaced this subsentence with 'A'.

Our artificial language, TFL, pursues this idea absolutely ruthlessly. We start with some *atomic sentences*. These will be the basic building blocks out of which more complex sentences are built. We will use uppercase Roman letters for atomic sentences of TFL (except for X, Y, and Z which we reserve for metavariables). There are only twenty-three letters A–W, but there is no limit to the number of atomic sentences that we might want to consider. By adding subscripts to letters, we obtain new atomic sentences. So, here are some different atomic sentences of TFL:

$$A, B, P, P_1, P_2, A_{234}$$

You can think of atomic sentences as representing certain English sentences but for now this is simply a heuristic (in Part ?? we shall take atomic sentences to *symbolize* certain English sentence). For example, you can think of A as representing the English sentence 'It is raining outside', and the atomic sentence of TFL, C, as representing the English sentence 'Jenny is miserable'.

However, if you think of the letter P as representing a particular English sentence it is important to understand that whatever structure the English sentence has, atomic sentence P will not reflect this structure. From the point of view of TFL, an atomic sentence is just a letter. It can be used to build more complex sentences, but it cannot be taken apart.

4.2 Connectives

In the previous section, we introduced the atomic sentences of TFL. In TFL we have counterparts to the conjunction-words that play an important role for spelling out arguments in English, that is, in TFL we have expression that play a similar role to the role expressions like 'and', 'or' and 'not' play in English. These are the *connectives*—they can be used to form new sentences out of old ones. In TFL, we will make use of logical connectives to build complex sentences from atomic components. There are four logical connectives in TFL. This table summarises them, and they are explained throughout this section.

symbol	**what it is called**	**rough meaning**
$\neg$	negation	'It is not the case that...'
$\wedge$	conjunction	'... and ...',
$\vee$	disjunction	'... or ...'
$\rightarrow$	conditional	'If ... then ...'

If we were to substitute declarative sentences for '...' in the right hand column of the table above, we obtain new English sentences. The language of truth-functional logic works in the same way: the connectives '$\neg$','$\wedge$','$\vee$' and '$\rightarrow$' combine with sentences as introduced in §4.1 to form new sentences. For example, the atomic sentences 'A' and 'C' combine with '$\wedge$' to form the sentence '$A \wedge C$'. If think of 'A' and 'C' as representing the English sentences

A: It is raining outside
C: Jenny is miserable

'$A \wedge C$' can be read as:

- It is raining *and* Jenny is miserable.

Similarly, on this understanding we get the following readings:

- ▹ $\neg A$: It is not the case that it is raining outside. (Alternatively, it is not raining outside).
- ▹ $A \vee C$: It is raining outside *or* Jenny is miserable.
- ▹ $A \rightarrow C$: *If* it is raining outside, *then* Jenny is miserable.

We shall go back to studying the connection between the connectives of truth-functional logic and various conjunction-words of English in §6, when we discuss the precise meaning of the connectives in truth-functional logic. For now we focus on completing the description of the formal language of truth-functional logic.

To conclude our discussion of the connectives we focused on the application of the connectives to atomic sentences, but connectives can be applied to all sorts of sentences not only atomic sentences. For example, as '$\neg A$' is a TFL sentence we can use '$\wedge$' to conjoin it the sentence 'C' to form the new sentence '$(\neg A \wedge C)$'. Connectives can be applied to all TFL sentences, not only atomic sentences. It is time to say precisely what TFL sentences are.

4.3 Sentences

We have introduced the basic building blocks of truth-functional logic, the atomic sentences, and the connectives, which allow us to conjoin different sentences to form new sentences. In terms of the VOCABULARY of a (written) language we have introduced all the important parts save the punctuation marks. In the language of truth-functional logic we use brackets for this purpose.

What is still missing is to equip the language with a GRAMMAR. The purpose of the grammar is to distinguish wellformed sentences from nonsense, but also to avoid ambiguity.

We wish to have rules that guarantee that

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

and

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

are wellformed sentences of the language of truth-functional logic, while

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

is not. It is merely a sequence of symbols of TFL — an *expression* but not a *sentence*. It is gibberish. In this respect it is like the following sequence of English words:

- Green quickly the without ran ideas.
- Not; John is happy and, not or Sue is tall only if

These are nonsense: a string of meaningful words, but not a grammatical English sentence.

The second purpose is to avoid ambiguity. In English we use commas to distinguish between two sentences

1. John's tired, and Sue's tall or Rob's short.
2. John's tired and Sue's tall, or Rob's short.

and without a comma it would be unclear which of the two sentences we intend to convey. In TFL this job of punctuation marks is assumed by brackets, that is, we distinguish between:

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

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

You can think of the former TFL-sentence as representing the English sentence 1, whereas the latter as representing the English sentence 2.

You might know this use of brackets from mathematics:

3. $9 + 3 \times 4$

can either be read as:

4. $9 + (3 \times 4) \qquad (= 9 + 12 = 21)$
5. $(9 + 3) \times 4 \qquad (= 12 \times 4 = 48)$

Importantly, the language of TFL is designed to exclude any form of ambiguity. For example, $A \wedge B \vee C$ will not be a sentence of TFL, as it would require disambiguation. Rather the syntactic/grammatical rules of the language of TFL will be such that only expressions that have a *unique* reading can be sentences of TFL. To make this precise we now provide a formal definition of what it is to be a sentence in TFL.

1. Every atomic sentence is a sentence.
2. If X is a sentence, then $\neg X$ is a sentence.
3. If X and Y are sentences, then $(X \wedge Y)$ is a sentence.
4. If X and Y are sentences, then $(X \vee Y)$ is a sentence.
5. If X and Y are sentences, then $(X \rightarrow Y)$ is a sentence.
6. Nothing else is a sentence.

The definition specifies rules according to which sentences of the language can be formed. To understand this definition let us pick it apart and consider the rules individually.

1. Tells us that atomic sentences as discussed in §4.1 are sentence of TFL.

Recall that any uppercase Roman letters A–W, or with subscripts, e.g., $A_1, B_3, A_{100}, J_{375}$, are atomic sentences of TFL. Notice X, Y, and Z are not atomic sentences. They are so-called METAVARIABLES and used as place holders for sentences of TFL (see more on metavairables in §5).

Our second rule says:

2. If X is a sentence of TFL, then so is $\neg X$.

By rule 1, we know that A is a sentence. Rule 2 then allows us to conclude that $\neg A$ is also a sentence. We could then apply it again and conclude that $\neg\neg A$ is also a sentence. More generally, if, by whatever rule we have constructed a sentence X, Rule 2 tells us that $\neg X$ will also be a sentence of TFL.

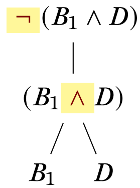
FORMATION TREES help us keep track of this process. For the case of $\neg\neg A$ this would be:

$$\begin{array}{c} \neg\neg A \\ | \\ \neg A \\ | \\ A \end{array}$$

Our third rule says:

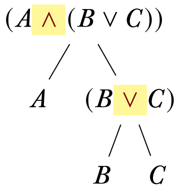
3. If X and Y are sentences, then so is $(X \wedge Y)$.

By rule 1, B_1 and D are both sentences. So rule 3 allows us to conclude that $(B_1 \wedge D)$ is a sentence. We might then apply rule 2 to conclude that $\neg(B_1 \wedge D)$ is also a sentence.

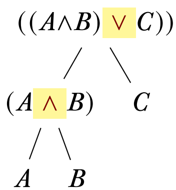


The rules 4 and 5 then tell us how the $\vee$- and $\rightarrow$-connective respectively can be used to produce new sentences of TFL. Rule 6, in contrast, tells us that sentences of TFL must be formed using the rules 1-5: if an expression cannot be obtained by consecutively applying rules 1-5, then the expression is not a sentence of TFL. Again formation trees are helpful to understand this: rule 7 tells us that all nodes of the formation tree must be sentences of TFL.

For example, consider $(A \wedge (B \vee C))$ we can check this is a sentence by drawing the following formation tree:

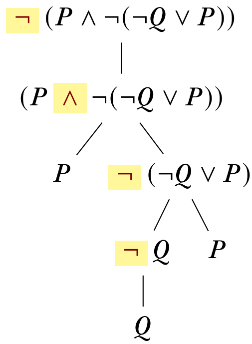


Each of the steps here tracks one of the rules of what it is to be a sentence. So we can conclude that this is a sentence of TFL. This also helps us see how to read it. It has a different formation tree from $((A \wedge B) \vee C)$:



The different formations will be important when we describe truth-tables for these sentences (§6). $((A \wedge B) \vee C)$ and $(A \wedge (B \vee C))$ will differ in when they are true.

One more example: consider $\neg(P \wedge \neg(\neg Q \vee P))$ we can check this is a sentence by drawing the following formation tree:



each of the steps here tracks one of the rules of what it is to be a sentence. So we can conclude that this is a sentence of TFL. The sentences further up the tree are formed by one of the formation rules from the sentences further down the tree.

When drawing these trees we have highlighted a particular connective on each of our nodes. We call that connective the MAIN CONNECTIVE of the sentence.

> The MAIN CONNECTIVE of sentence is the last connective that was introduced in the construction of the sentence.

In the case of $((\neg E \vee F) \to \neg\neg G)$, the main connective is $\to$. Here we can see that the whole sentence can be described in the form $(X \to Y)$ with both X and Y being complete sentences (put $X = (\neg E \vee F)$ and $Y = \neg\neg G$). That's enough to see that $\to$ is the main connective. In the case of $\neg\neg\neg D$, the main connective is the very first $\neg$ sign. This is because we can see the sentence as having the form $\neg X$ with X being the complete sentence $\neg\neg D$. In the case of $(P \wedge \neg(\neg Q \vee R))$, the main connective is $\wedge$: it's an $(X \wedge Y)$ with X as P and Y as $\neg(\neg Q \vee R)$.

Inductive Definition

The definition of a TFL-sentence is a so-called *inductive* definition. Inductive 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 an inductive definition is, we can give an inductive 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 inductive definition of sentences of TFL. Just as the inductive definition allows complex sentences to be built up from simpler parts, the definition allows us to decompose sentences into their simpler parts. Once we get down to atomic sentences, then we know we are ok.

4.4 Bracketing conventions

Strictly speaking, $A \wedge B$ is not a sentence of TFL. When we introduce a connective $\wedge, \vee$ or $\rightarrow$, strictly speaking, we must include brackets. Only $(A \wedge B)$ is strictly speaking a sentence of TFL. The reason for this rule is that we might use $(A \wedge B)$ as a subsentence in a more complicated sentence. For example, we might want to negate $(A \wedge B)$, obtaining $\neg(A \wedge B)$. If we just had $A \wedge B$ without the brackets and put a negation in front of it, we would have $\neg A \wedge B$. It is most natural to read this as meaning the same thing as $(\neg A \wedge B)$, but this may be very different from $\neg(A \wedge B)$.

When working with TFL, however, it will make our lives easier if we are sometimes a little less than strict. So, here are two convenient conventions.

1. We can remove *outermost* brackets of a sentence. Thus we allow ourselves to write $A \wedge B$ instead of the sentence $(A \wedge B)$. However, we must remember to put the brackets back in, when we want to embed the sentence into a more complicated sentence!

2. It can be a bit painful to stare at long sentences with many nested pairs of brackets. To make things a bit easier on the eyes, we will allow ourselves to use square brackets, '[' and ']', instead of rounded ones. So there is no logical difference between $(P \vee Q)$ and $[P \vee Q]$, for example.

Combining these two conventions, we can rewrite the unwieldy sentence

$$(((H \to I) \vee (I \to H)) \wedge (J \vee K))$$

rather more clearly as follows:

$$\big[(H \to I) \vee (I \to H)\big] \wedge (J \vee K)$$

The scope of each connective is now much easier to pick out.

Practice exercises

A. Add brackets to form a sentence of TFL, strictly speaking (i.e., no dropped brackets), which satisfies the given condition.

$$\neg \quad P \quad \vee \quad Q \quad \to \neg \quad P_2 \quad \wedge \quad R$$

1. Where the '$\to$' is the main connective
2. Where the '$\vee$' is the main connective.
3. Where the left '$\neg$' is the main connective.
4. Can it be done so that the right '$\neg$' is the main connective?

B. For each of the following: (a) Is it a sentence of TFL, strictly speaking? (b) Is it a sentence of TFL, allowing for our relaxed bracketing conventions? (c) If the answer to (b) is yes, write down the formation tree of each sentence and determin the main connective at each node (if there is one). Is there a main connective for every node of the formation tree of a sentence.

1. (A) — (a) no (b) no
2. $J_{374} \vee \neg J_{374}$ — (a) no (b) yes
3. $\neg\neg\neg\neg F$ — (a) yes (b) yes
4. $\neg \wedge S$ — (a) no (b) no
5. $(G \wedge \neg G)$ — (a) yes (b) yes
6. $(A \to (A \wedge \neg F)) \vee (D \to E)$ — (a) no (b) yes
7. $[(Z \to S) \to W] \wedge [J \vee X]$ — (a) no (b) yes
8. $(F \to \neg D \to J) \vee (C \wedge D)$ — (a) no (b) no

C.

1. Does $\neg A \vee B$ have the form $X \vee Y$?
2. Does $\neg(A \vee B)$ have the form $X \vee Y$?
3. Does $(A \vee B) \vee C$ have the form $X \vee Y$?

D.

1. Can there be a sentence of TFL of the form $(X \wedge Y)$ that includes three different connectives (the connectives are $\neg, \wedge, \vee, \rightarrow$)? If so, give an example, if not explain why not.
   yes, for example $(\neg A \wedge (B \vee C))$

2. Are there any sentences of TFL that contain no atomic sentences? Explain your answer.
   No. Atomic sentences contain atomic sentences (trivially). And every more complicated sentence is built up out of less complicated sentences, that were in turn built out of less complicated sentences, ..., that were ultimately built out of atomic sentences.

3. Are there sentences of TFL that don't have a main connective? Explain your answer.Yes, atomic sentences don't have a main connective.

CHAPTER 5

Use and mention

We have talked a lot *about* sentences. So we should pause to explain an important, and very general, point.

5.1 Quotation conventions

Consider these two sentences:

- Justin Trudeau is the Prime Minister.
- The expression 'Justin Trudeau' is composed of two uppercase letters and eleven lowercase letters

When we want to talk about the Prime Minister, we *use* his name. When we want to talk about the Prime Minister's name, we *mention* that name, which we do by putting it in quotation marks.

There is a general point here. When we want to talk about things in the world, we just *use* words. When we want to talk about words, we typically have to *mention* those words. We need to indicate that we are mentioning them, rather than using them. To do this, some convention is needed. We can put them in quotation marks, or display them centrally in the page (say). So this sentence:

- 'Justin Trudeau' is the Prime Minister.

says that some *expression* is the Prime Minister. That's false. The *man* is the Prime Minister; his *name* isn't. Conversely, this sentence:

▹ Justin Trudeau is composed of two uppercase letters and eleven lowercase letters.

also says something false: Justin Trudeau is a man, made of flesh rather than letters. One final example:

▹ " 'Justin Trudeau' " is the name of 'Justin Trudeau'.

On the left-hand-side, here, we have the name of a name. On the right hand side, we have a name. Perhaps this kind of sentence only occurs in logic textbooks, but it is true nonetheless.

Those are just general rules for quotation, and you should observe them carefully in all your work! To be clear, the quotation-marks here do not indicate reported speech. They indicate that you are moving from talking about an object, to talking about a name of that object.

5.2 Object language and metalanguage

These general quotation conventions are very important for us. After all, we are describing a formal language here, TFL, and so we must often *mention* expressions from TFL.

When we talk about a language, the language that we are talking about is called the OBJECT LANGUAGE. The language that we use to talk about the object language is called the METALANGUAGE.

For the most part, the object language in this chapter has been the formal language that we have been developing: TFL. The metalanguage is English. Not conversational English exactly, but English supplemented with some additional vocabulary to help us get along.

Now, we have used uppercase letters as sentence letters of TFL:

$$A, B, C, Z, A_1, B_4, A_{25}, J_{375}, \ldots$$

These are sentences of the object language (TFL). They are not sentences of English. So we must not say, for example:

▹ D is a sentence letter of TFL.

Obviously, we are trying to come out with an English sentence that says something about the object language (TFL), but 'D' is a sentence of TFL, and not part of English. So the preceding is gibberish, just like:

▹ Schnee ist weiß is a German sentence.

What we surely meant to say, in this case, is:

▹ ‘Schnee ist weiß’ is a German sentence.

Equally, what we meant to say above is just:

▹ ‘D’ is a sentence letter of TFL.

The general point is that, whenever we want to talk in English about some specific expression of TFL, we need to indicate that we are *mentioning* the expression, rather than *using* it. We can either deploy quotation marks, or we can adopt some similar convention, such as placing it centrally in the page.

5.3 Metavariables

However, we do not just want to talk about *specific* expressions of TFL. We also want to be able to talk about *any arbitrary* sentence of TFL. Indeed, we had to do this in §4.3, when we presented the recursive definition of a sentence of TFL. We used uppercase script letters to do this, namely:

$$\mathscr{X}, \mathscr{Y}, \mathscr{Z}, \mathscr{X}_1, \mathscr{Y}_1, \mathscr{Z}_1 \ldots$$

These symbols do not belong to TFL. Rather, they are part of our (augmented) metalanguage that we use to talk about *any* expression of TFL. To explain why we need them, recall the second clause of the recursive definition of a sentence of TFL:

2. If $\mathscr{X}$ is a sentence, then $\neg\mathscr{X}$ is a sentence.

This talks about *arbitrary* sentences. If we had instead offered:

2′. If ‘A’ is a sentence, then ‘$\neg A$’ is a sentence.

this would not have allowed us to determine whether ‘$\neg B$’ is a sentence. To emphasize:

> ‘$\mathscr{X}$’ is a symbol (called a METAVARIABLE) in augmented English, which we use to talk about expressions of TFL. ‘A’ is a particular sentence letter of TFL.

But this last example raises a further complication, concerning quotation conventions. We did not include any quotation marks in the second clause of our inductive definition. Should we have done so?

The problem is that the expression on the right-hand-side of this rule, i.e., '$\neg X$', is not a sentence of English, since it contains '$\neg$'. So we might try to write:

2''. If X is a sentence, then '$\neg X$' is a sentence.

But this is no good: '$\neg X$' is not a TFL sentence, since 'X' is a symbol of (augmented) English rather than a symbol of TFL.

What we really want to say is something like this:

2'''. If X is a sentence, then the result of placing the symbol '$\neg$' in front of the sentence X is a sentence.

This is impeccable, but rather long-winded. But we can avoid long-windedness by creating our own conventions. We can perfectly well stipulate that an expression like '$\neg X$' should simply be read *directly* in terms of rules for concatenation. So, *officially*, the metalanguage expression '$\neg X$' simply abbreviates:

> the result of placing the symbol '$\neg$' in front of the sentence X

and similarly, for expressions like '$(X \wedge Y)$', '$(X \vee Y)$', etc.

5.4 Quotation conventions for arguments

One of our main purposes for using TFL is to study arguments, and that will be our concern in §6. In English, the premises of an argument are often expressed by individual sentences, and the conclusion by a further sentence. Since we can symbolize English sentences, we can symbolize English arguments using TFL.

Or rather, we can use TFL to symbolize each of the *sentences* used in an English argument. However, TFL itself has no way to flag some of them as the *premises* and another as the *conclusion* of an argument. (Contrast this with natural English, which uses words like 'so', 'therefore', etc., to mark that a sentence is the *conclusion* of an argument.)

So, we need another bit of notation. Suppose we want to symbolize the premises of an argument with $X_1, \ldots, X_n$ and the conclusion with Z. Then we will write:

$$X_1, \ldots, X_n \therefore Z$$

The role of the symbol '∴' is simply to indicate which sentences are the premises and which is the conclusion.

Strictly, the symbol '∴' will not be a part of the object language, but of the *metalanguage*. As such, one might think that we would need to put quote-marks around the TFL-sentences which flank it. That is a sensible thought, but adding these quote-marks would make things harder to read. Moreover—and as above—recall that *we* are stipulating some new conventions. So, we can simply stipulate that these quote-marks are unnecessary. That is, we can simply write:

$$A, A \to B \therefore B$$

without any quotation marks, to indicate an argument whose premises are (symbolized by) 'A' and '$A \to B$' and whose conclusion is (symbolized by) 'B'.

Practice exercises

A. Add quotation marks to the following sentences where necessary:

1. '$\neg$', '$\wedge$', '$\vee$' and '$\to$' are the basic logical connectives of the language of TFL.
2. In a well-formed sentence for every opening bracket '(' there must be a closing bracket ')'.
3. The sentence 'snow is white' is true if and only if snow is white.
4. Kevin is a lecturer, whereas 'Kevin' names Kevin and the latter is an expression of the language (and not a lecturer).

B. Given our convention on metavariables what do the following expressions convey:

1. $X \wedge Y$
   the result of taking the sentence Y followed by '$\wedge$' followed by Z
2. $Y \to Z$
   the result of taking the sentence Y followed by '$\to$' followed by Z
3. $\neg(Y \vee Z)$
   the result of placing '$\neg$(' before the sentence Y, then adding '$\vee$', then Z and finally ')'.

C. Consider the sentence

▹ Sharky is so-called because of his teeth.

Is ‘Sharky’ used or mentioned in the above sentence? Discuss.

CHAPTER 6

Truth rules for the connectives

We have completed introducing the syntax of the language of TFL. It is now time to turn to the semantics of TFL. The idea underlying the semantics is to specify conditions under which sentences of TFL are true. In §4 we introduced the idea that an argument is valid if there is no interpretation on which all premises of the argument are true but the conclusion false. Accordingly, if we wish to make a start on making this idea more precise, we need to say when a sentence of TFL is true according to an interpretation and when it is false. The target of this chapter is to give precise rules to this effect. The important feature of truth functional logic is that the truth value of a complex sentence, such as '$A \vee (B \wedge C)$' is fully determined by the truths of is component parts, that is 'A', 'B' and 'C'. If we're told whether 'A', 'B' and C are true or false, then we will be able to say whether '$A \vee (B \wedge C)$' is true or false.

To spell out this idea, we need to describe how the truth values of different sentences (e.g., 'A', 'B', and 'C') are to be combined to obtain the truth value of a sentence that has been obtained via the formation rules 2-5 (e.g., '$A \vee (B \wedge C)$'). To do this we work through each of our connectives describing the rules governing it.

6.1 Negation

The '$\neg$'-connective is called negation. When we introduced the '$\neg$'-connective we said it should roughly be understood as 'it is not the case' or, perhaps, simply 'not'. Let's make that official:

> If a sentence can be paraphrased as 'it is not the case that …' it can be symbolised as $\neg X$.

What does that mean for the truth rules? Consider:

1. Bristol is not in France.
2. Bristol is not in England.

'Bristol is in France' is false, so 'Bristol is not in France' is true. 'Bristol is in England' is true, so 'Bristol is not in England' is false.

In general, to determine whether a sentence of the form $\neg X$ is true we check whether X is false. If the answer is yes, then $\neg X$ is indeed true. If to the contrary X is true, then $\neg X$ is false:

- ▹ If X is true, then $\neg X$ is false.
- ▹ If X is false, then $\neg X$ is true.

We record this in shorthand:

If X is:		then $\neg X$ is:
T	$\rightsquigarrow$	F
F	$\rightsquigarrow$	T

We have abbreviated 'True' with 'T' and 'False' with 'F'. (But just to be clear, the two truth values are True and False; the truth values are not *letters*!)

6.2 Conjunction

The $\wedge$-connective, called conjunction, is meant to be understood of the English word 'and':

> If a sentence can be paraphrased as '…and…' it can be symbolised as $(X \wedge Y)$.

What is the appropriate truth rule for conjunction? Consider:

3. She can speak German and she can speak French.

If she can speak German and she can speak French, then this is true, but otherwise it is false.

More generally, the rule governing $\wedge$ is:

- ▹ If X and Y are both true, then $X \wedge Y$ is true.
- ▹ Otherwise, $X \wedge Y$ is false.

Which we summarise

If X is:	and Y is:		then $(X \wedge Y)$ is:
T	T	$\rightsquigarrow$	T
T	F	$\rightsquigarrow$	F
F	T	$\rightsquigarrow$	F
F	F	$\rightsquigarrow$	F

Note that conjunction is *symmetrical*. The truth value for $(X \wedge Y)$ is always the same as the truth value for $(Y \wedge X)$.

6.3 Disjunction

The '$\vee$'-connective is called disjunction and is meant to be understood in terms of the English 'or'.

> If a sentence can be paraphrased as '. . . or . . . '
> it can be symbolised as $(X \vee Y)$.

Whereas the truth rules for the negation and conjunction were relatively straightforward, the rule for disjunction is a bit more subtle. Consider:

4. She can speak German or she can speak French.

If she cannot speak either German or French, then this is false. If she can speak German but not French, then it is true, and if she can speak French but not German it is also true. We have the general rules:

- ▹ If X and Y are both false, then $(X \vee Y)$ is false.
- ▹ If X is true and Y is false, then $(X \vee Y)$ is true.
- ▹ If X is false and Y is true, then $(X \vee Y)$ is true.

But what if she can speak both? Is it true or false? It seems that in English there are two kinds of disjunctions: an INCLUSIVE and an EXCLUSIVE one. For the inclusive *or*, we might whisper a "or both" after it; whereas for the exclusive *or*, we'd want to whisper a "but not both":

5. She can speak German or she can speak French (or both).
6. She can speak German or she can speak French (but not both).

In logic there can be no ambiguity. We choose that ∨ stands for the *inclusive or*. That is, we give the final rule:

▹ If X and Y are both true, then $(X \vee Y)$ is true.

Once we have completed our presentation of the semantics for. TFL we will see that more complex sentence: $((X \vee Y) \wedge \neg(X \wedge Y))$ has the truth conditions of the *exclusive or*, that is it is true if X is true or Y is true but not both X and Y are true.

To summarise the rules for ∨:

If X is:	and Y is:		then $(X \vee Y)$ is:
T	T	$\rightsquigarrow$	T
T	F	$\rightsquigarrow$	T
F	T	$\rightsquigarrow$	T
F	F	$\rightsquigarrow$	F

Like conjunction, disjunction is symmetrical.

6.4 Conditional

The '→'-connective is called the conditional-connective and it is meant to be related to our understanding of *if…, then…* sentences. Unlike conjunction and disjunction, the conditional is asymmetric, it matters which order the sentences come in, and $(X \to Y)$ will be different to $(Y \to X)$. It is thus valuable to give a name to indicate which part goes where:

> In $X \to Y$, X is called the ANTECEDENT, and Y is called the CONSEQUENT.

Recall our guidance on the rough meaning of →:

> If a sentence can be paraphrased as
> 'If ..., then ...' it can be symbolised as $(X \rightarrow Y)$.

What are the truth rules for the conditional? Consider the sentence:

7. If she is drinking a beer, then she is over eighteen.

What are the circumstances under which this conditional is false? Here is what we'll say:

- ▹ If she's drinking beer and is under age, then it is false.
- ▹ It is true in all other circumstances.

To understand the rationale for this, let us think about when a bartender would get into trouble (clearly the conditional should be true for everyone drinking beer in a bar). They will get into trouble in case there a woman in drinking beer and it turns out that the woman is under age. They will not get into trouble if the woman is over 18 years old, i.e., if the consequent of the conditional is true. They will also not get in trouble if it turns out that the woman is not drinking beer (but orange juice), i.e., if the antecedent of the conditional is false. In that case it is irrelevant whether she is 18 or not. The barkeeper doesn't have to check her age. The conditional is true for trivial reasons.

We are led to the following rules:

If X is:	and Y is:		then $(X \rightarrow Y)$ is:
T	T	$\rightsquigarrow$	T
T	F	$\rightsquigarrow$	F
F	T	$\rightsquigarrow$	T
F	F	$\rightsquigarrow$	T

In this case, it's very important to remember which way around it goes. The TF-line is different to the FT-line.

This is why the terms 'antecedent' and 'consequent' are so useful. In $X \rightarrow Y$, X is called the ANTECEDENT, and Y the CONSEQUENT. Using the antecedent/consequent terminology we can summarize the truth rule as follows: If the antecedent is true and the consequent false, then the conditional sentence is false, otherwise it is true.

The TFL connective $\rightarrow$ is *stipulated* to be governed by these rules. This sometimes marked by calling it the MATERIAL CONDITIONAL. But the truth rules of the $\rightarrow$-connective only tell us part of the story of *if…, then…* sentences in English, as the truth rules do not seem to work well for all these sentences. For example, truth rules for material conditional do not seem to work well with our understanding of the sentence

8. If Kangaroos had no tails, they would topple over.

For our purposes $\rightarrow$-connective will be understood in terms of truth rules given in this section.

6.5 On Truth-functional connectives

Let's introduce an important idea.

> A connective is TRUTH-FUNCTIONAL iff the truth value of a sentence with that connective as its main connective is uniquely determined by the truth value(s) of the constituent sentence(s).

Every connective in TFL is truth-functional. We were able to give rules to determine what the truth value of a sentence $\neg X$ is depending only on the truth value of X. The truth value of X uniquely determines the truth value of $\neg X$. The same was true for all the other connectives of TFL ($\wedge, \vee, \rightarrow$), as is evidenced by the truth rules and truth tables we gave for these connectives. This is what gives TFL its name: it is *truth-functional logic.*

In plenty of languages, e.g. English, there are connectives that are not truth-functional. We here describe just two (Exercise: find further non-truth-functional connectives of English):

Necessarily

In English, for example, we can form a new sentence from any simpler sentence by prefixing it with 'Necessarily, …' or 'It is necessary that…'. The truth value of this new sentence is not fixed solely by the truth value of the original sentence. For consider two true sentences:

9. $2 + 2 = 4$
10. Shostakovich wrote fifteen string quartets

Whereas it is necessary that $2 + 2 = 4$, it is not *necessary* tthat Shostakovich wrote fifteen string quartets. If Shostakovich had died earlier, he would have failed to finish Quartet no. 15; if he had lived longer, he might have written a few more. So ‘It is necessary that…’ is a connective of English, but it is not *truth-functional.*

Subjunctive conditionals

We said that $\rightarrow$ was pretty bad at capturing some uses of ‘if…, then…’ in Englihs. In particular, it is bad at capturing use of *subjunctive conditionals* of English. The problem is that a subjunctive conditional is not truth functional. Consider the two sentences:

11. If Mitt Romney had won the 2012 election, then he would have been the 45th President of the USA.
12. If Mitt Romney had won the 2012 election, then he would have turned into a helium-filled balloon and floated away into the night sky.

Sentence 11 is true; sentence 12 is false, but both have false antecedents and false consequents. So the truth value of the whole sentence is not uniquely determined by the truth value of the constituent sentences.

However, the material conditional, which is what we specify $\rightarrow$ to be is the best that can be done at symbolising subjunctive conditionals of English in TFL. TFL just doesn't have the required resources as the subjunctive conditional is not truth-functional.

Practice exercises

A.

1. Suppose X is true, what can you say about $X \vee Y$?
2. Suppose X is false, what can you say about $X \rightarrow Y$?

B.

1. Is ‘I know that’ truth functional?

   No.

   ‘I know that I have hands’

   is true (true $\rightsquigarrow$ true). But there are some sentences that are true but I don't know them. E.g., ‘I know that France won the FIFA world cup’

(well, I now know it but I didn’t 20 minutes ago.)

CHAPTER 7

Truth on a Valuation

In the previous section we have learned how truth values of constituent sentences determine the truth value of complex sentences. This means that if we are presented with the truth value of the relevant atomic sentences that appear in a complex sentence, we can determine whether that sentence is true or false.

But how do we determine whether a given atomic sentence is true? This is where the notion of an *interpretation* comes into the picture. An interpretation will stipulate (assign) truth values of particular atomic sentences. Let 'B' stand for the English sentence 'Ben is happy'. Then there is one interpretation according to which this is true, that is, B will be assigned the value "True" on this interpretation. There is another interpretation according to which Ben is not happy, that is, B is assigned the value "False" on this interpretation. In TFL an interpretation of the atomic sentences is called a VALUATION:

> A VALUATION is any assignment of truth values (true or false) to the atomic sentences of TFL.

For example a valuation v might assign:

	A	H	I	B	C_{35}	...	
v:	F	T	F	F	F	T	...

Sentences of TFL are true or false *relative to a valuation.* The valuation determines the truth values of the atomic sentences, and complex sentences are determined using the truth rules described in the previous section.

Let v be a valuation. Then

1. An atomic sentence X is true relative to v, if and only if v assigns the value T to X.

2. For a sentence $\neg X$, $(X \wedge Y)$, $(X \vee Y)$, $(X \to Y)$, whether it is true or false relative to v can be determined by applying the appropriate truth rule:

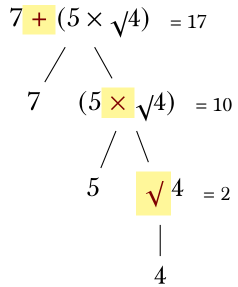
If X is:		then $\neg X$ is:
T	$\rightsquigarrow$	F
F	$\rightsquigarrow$	T

If X is:	and Y is:		then $(X \wedge Y)$ is:
T	T	$\rightsquigarrow$	T
T	F	$\rightsquigarrow$	F
F	T	$\rightsquigarrow$	F
F	F	$\rightsquigarrow$	F

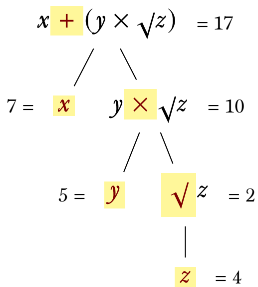
If X is:	and Y is:		then $(X \vee Y)$ is:
T	T	$\rightsquigarrow$	T
T	F	$\rightsquigarrow$	T
F	T	$\rightsquigarrow$	T
F	F	$\rightsquigarrow$	F

If X is:	and Y is:		then $(X \to Y)$ is:
T	T	$\rightsquigarrow$	T
T	F	$\rightsquigarrow$	F
F	T	$\rightsquigarrow$	T
F	F	$\rightsquigarrow$	T

Consider '$(\neg I \wedge H) \to H$'. We will determine whether it is true or false on the valuation given. The idea is that we draw the formation tree and then work our way up the formation tree to, in stages, determine the value of our complex sentence. This is similar to as one might do in mathematics to determine the value of a complex equation in stages starting from the bottom of the tree:

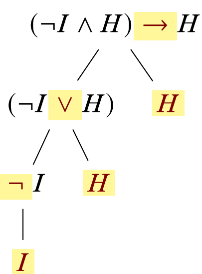


Or, it might be more similar to this maths question: Setting $x = 7$, $y = 5$ and $z = 4$, determine the value of $x + (y \times \sqrt{z})$.



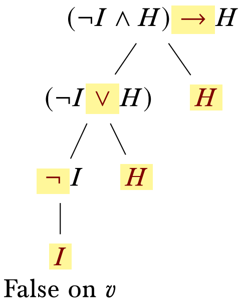
Similarly for TFL, we use the valuation to set the values for our atomic sentences and then determine the values of our complex sentences by applying our truth rules.

Let us see an example. Consider '$(\neg I \wedge H) \rightarrow H$'. The formation tree will help us know what to do (see §4.3):

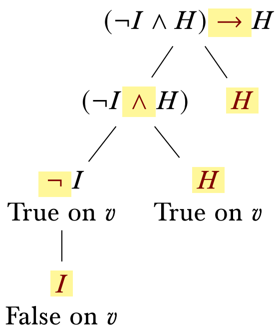


The lowest point on the tree is I. For the truth of this, we just consult the valuation. On this valuation:

	A	H	I	B	C_{35}	$\dots$	
v:	F	T	F	F	F	T	$\dots$



The truth rule for $\neg$ tells us how the truth value of $\neg I$ depends on the truth of I. Since I is false on valuation v, by consulting the truth rule for $\neg$ we can se that $\neg I$ is true on the valuation v:



Then the rule for $\wedge$ tells us how the truth value of $\neg I \wedge H$ depends on the truths of $\neg I$ and H: as $\neg I$ is true (which we determined) and H is true (consult the valuation), then $\neg I \wedge H$ is also true on the valuation v.

$(\neg I \wedge H) \to H$

$(\neg I \wedge H)$ True on v — H True on v

$\neg I$ True on v — H True on v

I False on v

Finally, the rule for $\to$ tells us how the truth value of $(\neg I \wedge H) \to H$ depends on those of $(\neg I \wedge H)$ and H.

$(\neg I \wedge H) \to H$
True on v

$(\neg I \wedge H)$ True on v — H True on v

$\neg I$ True on v — H True on v

I False on v

To keep track of the answer, we will present it on the truth table.

			(	$\neg$	I	$\wedge$	H	)	$\to$	H
			(	$\neg$	I	$\wedge$	H	)		H
				$\neg$	I		H			
Valuation	I	H			I					
v :	F	T		T	F	T	T		T	T

We do this because when we consider multiple valuations, we'll need to track where their answers are. For example:

$$\neg(C \vee (D \to (A \vee \neg D)))$$
$$(C \vee (D \to (A \vee \neg D)))$$
$$C \qquad (D \to (A \vee \neg D))$$
$$D \qquad A \vee \neg D$$
$$A \qquad \neg D$$
$$D$$

Valuation	A	C	D	$\neg$	C	$\vee$	D	$\to$	A	$\vee$	$\neg$	D
v_1 :	T	F	F	F	F	T	F	T	T	T	T	F
v_2 :	F	T	T	F	T	T	T	F	F	F	F	T

Practice exercises

A. Determine the truth values for these sentences on this valuation:

	A	B	C	D	E	H	I	J	C_{35}
v:	F	F	F	T	F	T	F	T	F

1. $\neg A \vee (B \to \neg(C \vee A))$
2. $(A \wedge H) \vee (I \wedge \neg B)$
3. $H \to (A \vee C)$
4. $\neg(D \wedge E)$
5. $(B \to J) \wedge (J \to H)$
6. $(C \vee C_{35}) \to H$
7. $\neg\neg A \vee \neg I$
8. $(A \to B) \to ((H \wedge D) \vee \neg E)$
9. $((J \wedge \neg C) \to D) \wedge \neg(J \wedge E)$
10. $(A \vee B) \wedge (\neg A \vee \neg B)$
11. $(H \to D) \vee (I \to A)$

CHAPTER 8

Truth Tables

8.1 Validity and truth tables, the idea

In §2 we said that an argument was valid, if and only if there is no interpretation such that all premises are true but the conclusion false. At that point the definition was suggestive, but we lacked a clear understanding of 'interpretation' and when a sentence is true relative to an interpretation. But for the language of TFL we can now turn the informal definition in a precise and rigorous definition as, for TFL, interpretations are given by valuations, assignments of truth values to the atomic sentences.

Recall that an argument consisted of a number of premises together with a conclusion. A TFL-argument then can be written as $X_1, \ldots, X_n \therefore Y$ where $X_1, \ldots, X_n$ and Y are sentences of TFL, and $X_1, \ldots, X_n$ are the premises and Y the conclusion of the argument.

> A TFL-argument is valid if and only if there is no valuation on which all the premises are true but the conclusion is false.
> That is, '$X_1, \ldots, X_n \therefore Y$' is VALID if and only if there is no valuation v such that X_1 is true relative to v and X_2 is true relative to v, and so on, up to X_n being true relative to v, but Y is false relative to v.

So to determine whether a TFL-argument is valid or not, we should show whether or not there is a valuation on which all premises are true but the conclusion false.

We have discussed how to determine the truth of a sentence relative to a single valuation. To determine whether an argument is valid or not we need to consider all valuations and ask whether there is at least one valuation where the premises are all true but the conclusion is false.

There are many different valuations. But to determine validity, we only need to consider the valuations' assignment to the atomic sentences that occur in the premises and the conclusion of the argument. Consider the argument:

$$\neg C, D \to C \therefore \neg D$$

To determine its validity, we only care about the truth values of the three sentence $\neg C$, $D \to C$ and $\neg D$. So all that matters for our purposes is the valuation's assignment of truth values to the atomic sentences C and D.

There will be four possible such valuations.

Valuation	C	D
v_1 :	T	T
v_2 :	T	F
v_3 :	F	T
v_4 :	F	F

We need to determine the truth values of the three sentences, $\neg C$, $D \to C$ and $\neg D$ on each valuation to determine whether the argument is valid.

This will result in:

Valuation	C	D	$\neg D$	$C \to D$	$\neg C$
v_1 :	T	T	F	T	F
v_2 :	T	F	T	F	F
v_3 :	F	T	F	T	T
v_4 :	F	F	T	T	T

v_4 is the only valuation on which both the premises ('$\neg D$' and '$C \to D$') are true; and on this valuation, the conclusion ('$\neg C$') is also true. There is thus no valuation where all the premises are true and the conclusion is false. So the argument is valid.

To allow you to determine this we will show how to construct so-called (complete) truth tables.

8.2 Determining all the valuations

Suppose we are interested in determining the validity of an argument. We need to consider all the valuations for the atomic sentences that appear in the argument.

If we are only interested in the atomic sentences A there are two ways of assigning truth values to the atomic sentence A:

A
T
F

That is, there are two valuations for the atomic sentence A.

If we are interested in two atomic sentences, say A and B, we need to consider all possible assignments of truth values to A and to B.

Whatever the valuation assigns to A, it can assign either true or false to B. So we "split" each valuation we have so far into two versions, one which assigns B as true and another assigning B as false.

A	B
T	T
T	F
F	T
F	F

If we are also interested in a further atomic sentence, say C, we further split each of our valuations into two, one which also assigns C as true, and one which assigns C as false.

A	B	C
T	T	T
T	T	F
T	F	T
T	F	F
F	T	T
F	T	F
F	F	T
F	F	F

What if we add a further atomic sentence, say P_{235}? We want to consider the valuations for atomic sentences A, B, C, P_{235}. Again we need to split each of our valuations for A, B, C into one which assigns P_{235} as true and another which assigns it as false.

A	B	C	P_{235}
T	T	T	T
T	T	T	F
T	T	F	T
T	T	F	F
T	F	T	T
T	F	T	F
T	F	F	T
T	F	F	F
F	T	T	T
F	T	T	F
F	T	F	T
F	T	F	F
F	F	T	T
F	F	T	F
F	F	F	T
F	F	F	F

The list of all possible valuations gets very big very fast.

There are 2^n ways to assign truth values to n atomic sentences. Using truth tables to determine validity thus becomes quite difficult and it is one reason we consider an alternative method: that of proofs, which we will see in IV.

Here's one strategy that some people find useful for listing all the valuations.

- For n atomic sentences, there are 2^n valuations.
- Draw a table with the n atomic sentences in the header row and count 2^n rows (make a mark).
- We'll now fill out the 'T' and 'F''s in these n-columns and 2^n rows:
  - Start with the last atomic sentence, the right-most column and *alternate* between 'T' and 'F' until you've filed out all 2^n rows.
  - Next consider the penultimate atomic sentence, the column to the left of the one just completed, write *two* 'T's, write two 'F's, and repeat.
  - For the third-last atomic sentence letter, write *four* 'T's followed by four 'F's. This would yields an eight line truth table.
  - For the fourth-last atomic sentence, you have *eight* 'T's followed by eight 'F's.
  - And so on until you get to the first atomic sentence.

8.3 Complete truth tables for a sentence

A COMPLETE TRUTH TABLE for a sentence has a line for each valuation for the atomic sentences appearing in the sentence, and displays the truth value of the sentence on each of the valuations.

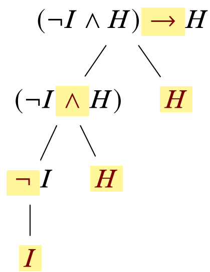
Consider the sentence '$(\neg I \wedge H) \rightarrow H$'. We will give a *truth table* which lists all the valuations and says whether this sentence is true or false on each of them.

The valuations assign either the value "True" or "False" to each atomic sentence. In this case we have two atomic sentences, I and H, so we have four (2^2) valuations ($v_1, \ldots, v_4$) each of which is a line in the truth table:

Valuation	I	H	$(\neg I \wedge H) \to H$
v_1	T	T	
v_2	T	F	
v_3	F	T	
v_4	F	F	

Our job is to fill out the truth values of '$(\neg I \wedge H) \to H$' on each of the four valuations.

We have already shown how to do this on a particular valuation: one first writes down the formation tree of the sentence. (see §4.3):



The idea is that we work ourselves from LEAVES of the tree (the atomic sentence) to the ROOT of the formation tree (the sentence that has been constructed). The truth rule for $\neg$ tells us how the truth value of $\neg I$ depends on the truth of I. Then the rule for $\wedge$ tells us how the truth value of $\neg I \wedge H$ depends on the truths of $\neg I$ and H; and finally, the rule for $\to$ tells us how the truth value of $(\neg I \wedge H) \to H$ depends on those of $\neg I \wedge H$ and H.

We will record the truth value of each subsentence underneath its main connective. So we will somewhat recreate the formation tree in the header of the truth table.

Valuation	H	I	$(\neg I \wedge H) \to H$
			$(\neg I \vee H) \quad H$
			$\neg I \quad H$
			I
v_1 :	T	T	
v_2 :	T	F	
v_3 :	F	T	
v_4 :	F	F	

We will work our way up the formation tree, as displayed in the header of the table. In the bottom line we just have 'I', whose truth value is just specified by the valuation. So we can fill that out

			(	$\neg$	I	$\wedge$	H	)	$\to$	H	
Valuation	H	I	(	$\neg$	I	$\vee$	H	)		H	
				$\neg$	I		H				
					I						★
v_1 :	T	T			T						
v_2 :	T	F			F						
v_3 :	F	T			T						
v_4 :	F	F			F						
					★						

The next line is to include the truth value for '$\neg I$' and the truth value for 'H'. The truth value for 'H' is just specified by the valuation so that can be copied across.

			(	$\neg$	I	$\wedge$	H	)	$\to$	H	
Valuation	H	I	(	$\neg$	I	$\vee$	H	)		H	
				$\neg$	I		H				★
					I						
v_1 :	T	T			T		T				
v_2 :	T	F			F		T				
v_3 :	F	T			T		F				
v_4 :	F	F			F		F				
							★				

The truth value for '$\neg I$' is determined using the truth rule for negation:

If X is		then $\neg X$ is
T	$\rightsquigarrow$	F
F	$\rightsquigarrow$	T

So we take the truth value of 'I' as recorded underneath that 'I', and apply this rule to determine the truth value for '$\neg I$'. We record the answer underneath the main connective of '$\neg I$', the '$\neg$'. This results in:

			(	$\neg$	I	$\wedge$	H	)	$\rightarrow$	H	
Valuation	H	I	(	$\neg$	I	$\vee$	H	)		H	
				$\neg$	I		H				$\star$
					I						
v_1 :	T	T		F	T		T				
v_2 :	T	F		T	F		T				
v_3 :	F	T		F	T		F				
v_4 :	F	F		T	F		F				
				$\star$							

The next line we need to add the truth value for 'H' again, which is specified by the valuation, and the truth value for '$\neg I \wedge H$'. For '$\neg I \wedge H$' we will use the truth rule for $\wedge$:

If X is	and Y is		then $X \wedge Y$ is
T	T	$\rightsquigarrow$	T
T	F	$\rightsquigarrow$	F
F	T	$\rightsquigarrow$	F
F	F	$\rightsquigarrow$	F

So we take the truth value for '$\neg I$', as recorded underneath '$\neg$', and the truth value for 'H', and apply this truth rule, writing the answer under its main connective, '$\wedge$'. So we can fill out:

			(	$\neg$	I	$\wedge$	H	)	$\rightarrow$	H	
Valuation	H	I	(	$\neg$	I	$\wedge$	H	)		H	$\star$
				$\neg$	I		H				
					I						
v_1 :	T	T		F	T	F	T			T	
v_2 :	T	F		T	F	T	T			T	
v_3 :	F	T		F	T	F	F			F	
v_4 :	F	F		T	F	F	F			F	
						$\star$					

Our final stage is to determine the truth value for '$(\neg I \wedge H) \rightarrow H$'

So we take our truth value we have determined for '$(\neg I \wedge H)$', which is recorded underneath the '$\wedge$', and the truth value for 'H', and apply the truth rule for '$\rightarrow$'

If X is	and Y is		then $X \to Y$ is
T	T	$\rightsquigarrow$	T
T	F	$\rightsquigarrow$	F
F	T	$\rightsquigarrow$	T
F	F	$\rightsquigarrow$	T

recording the answer underneath '$\to$'. As this is the final stage, we have determined the truth value for the target sentence, '$(\neg I \wedge H) \to H$', so we will boldface this answer to indicate that it is our final answer, the truth value for '$(\neg I \wedge H) \to H$'.

			(	$\neg$	I	$\wedge$	H	)	$\to$	H	$\star$
Valuation	H	I	(	$\neg$	I	$\wedge$	H	)		H	
				$\neg$	I		H				
					I						
v_1 :	T	T		F	T	F	T		T	T	
v_2 :	T	F		T	F	T	T		T	T	
v_3 :	F	T		F	T	F	F		T	F	
v_4 :	F	F		T	F	F	F		T	F	
									$\star$		

Here's another example.

Valuation	A	C	D	$\neg$	$(C$	$\vee$	$(D$	$\rightarrow$	$(A$	$\vee$	$\neg$	$D)))$
					$(C$	$\vee$	$(D$	$\rightarrow$	$(A$	$\vee$	$\neg$	$D))$
					C		$(D$	$\rightarrow$	$(A$	$\vee$	$\neg$	$D)$
							D		A	$\vee$	$\neg$	D
									A		$\neg$	D
												D
v_1 :	T	T	T	F	T	T	T	T	T	T	F	T
v_2 :	T	T	F	F	T	T	F	T	T	T	T	F
v_3 :	T	F	T	F	F	T	T	T	T	T	F	T
v_4 :	T	F	F	F	F	T	F	T	T	T	T	F
v_5 :	F	T	T	F	T	T	T	F	F	F	F	T
v_6 :	F	T	F	F	T	T	F	T	F	F	T	F
v_7 :	F	F	T	T	F	F	T	F	F	F	F	T
v_8 :	F	F	F	F	F	T	F	T	F	F	T	F

How to do truth tables

With this example in mind let us try to give a general instruction for constructing a truth table for a sentence X.

1. Write down the formation tree of X. Set up the header of the truth table with lines matching the formation tree.

2. Consider any atomic sentences which appear in X. If you are doing a truth table for multiple sentences, you'll need to also consider those appearing in the other sentences. Write all these atomic sentences in the header row, at the beginning of the table.

3. We now need to determine all the valuations, the different ways of assigning truth values to all these atomic sentences. If there are n atomic sentences, there will be 2^n such valuations. Write them as the rows of the truth table.

4. Work up through the formation tree, which has now been written as lines in the header of the truth table to determine the truth value of the consituant sentences of X. Record its truth value underneath the main connective of that sentence.

   - If it is an atomic sentence, just copy the truth values from the valuation.
   - If it is a complex sentence, apply the relevant truth rule to its consituatnt parts, remembering that their truth values are written underneath their main connectives.

5. Highlight or boldface the column under the main connective of the overall sentence, which provides our final answer.

If you follow these outlines, you should be able to construct truth tables for arbitrary TFL sentences. Of course, the more complicated the sentences are, and the more atomic sentence letters they contain the longer and tedious the truth table—but the more important it becomes to painstakingly stick to the guidelines we have given.

You can provide your answer without explicitly providing the header row matching the formation tree, unless explicitly requested.

Valuation	A	C	D	$\neg\ (C \vee (D \rightarrow (A \vee \neg D)))$
v_1 :	T	T	T	F TT T T T T F T
v_2 :	T	T	F	F TT F T T T T F
v_3 :	T	F	T	F FT T T T T F T
v_4 :	T	F	F	F FT F T T T T F
v_5 :	F	T	T	F TT T F F F F T
v_6 :	F	T	F	F TT F T F F T F
v_7 :	F	F	T	T FF T F F F F T
v_8 :	F	F	F	F FT F T F F T F

There are tricks that mean you can miss out some of the gaps by using strategies such as: knowing that X is false is already enough to see that $X \wedge Y$ is false, so we don't need to continue to work out the truth value of Y. These "shortcuts" are discussed in §12, but we won't read it in this course.

Practice exercises

A. Complete truth tables for each of the following:

1. $A \rightarrow A$

A	A	$\rightarrow$	A
T	T	T	T
F	F	T	F

2. $C \rightarrow \neg C$

C	C	$\rightarrow$	$\neg$	C
T	T	F	F	T
F	F	T	T	F

3. $(A \rightarrow B) \rightarrow (\neg A \vee B)$
4. $(A \rightarrow B) \vee (B \rightarrow A)$

A	B	$(A$	$\rightarrow$	$B)$	$\vee$	$(B$	$\rightarrow$	$A)$
T	T	T	T	T	T	T	T	T
T	F	T	F	F	T	F	T	T
F	T	F	T	T	T	T	F	F
F	F	F	T	F	T	F	T	F

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

A	B	$(A$	$\wedge$	$B)$	$\rightarrow$	$(B$	$\vee$	$A)$
T	T	T	T	T	T	T	T	T
T	F	T	F	F	T	F	T	T
F	T	F	F	T	T	T	T	F
F	F	F	F	F	T	F	F	F

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

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

7. $(\neg A \wedge \neg B) \rightarrow \neg(A \vee B)$
8. $\big[(A \wedge B) \wedge \neg(A \wedge B)\big] \wedge C$

A	B	C	$\big[(A$	$\wedge$	$B)$	$\wedge$	$\neg$	$(A$	$\wedge$	$B)\big]$	$\wedge$	C
T	T	T	T	T	T	F	F	T	T	T	F	T
T	T	F	T	T	T	F	F	T	T	T	F	F
T	F	T	T	F	F	F	T	T	F	F	F	T
T	F	F	T	F	F	F	T	T	F	F	F	F
F	T	T	F	F	T	F	T	F	F	T	F	T
F	T	F	F	F	T	F	T	F	F	T	F	F
F	F	T	F	F	F	F	T	F	F	F	F	T
F	F	F	F	F	F	F	T	F	F	F	F	F

9. $[(A \wedge B) \wedge C] \rightarrow B$

A	B	C	$[(A$	$\wedge$	$B)$	$\wedge$	$C]$	$\rightarrow$	B
T	T	T	T	T	T	T	T	T	T
T	T	F	T	T	T	F	F	T	T
T	F	T	T	F	F	F	T	T	F
T	F	F	T	F	F	F	F	T	F
F	T	T	F	F	T	F	T	T	T
F	T	F	F	F	T	F	F	T	T
F	F	T	F	F	F	F	T	T	F
F	F	F	F	F	F	F	F	T	F

10. $\neg\big[(C \vee A) \vee B\big]$

A	B	C	$\neg$	$[(C$	$\vee$	$A)$	$\vee$	$B]$
T	T	T	F	T	T	T	T	T
T	T	F	F	F	T	T	T	T
T	F	T	F	T	T	T	T	F
T	F	F	F	F	T	T	T	F
F	T	T	F	T	T	F	T	T
F	T	F	F	F	F	F	T	T
F	F	T	F	T	T	F	T	F
F	F	F	T	F	F	F	F	F

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

(A	$\rightarrow$	B)	$\boxed{\rightarrow}$	(¬	B	$\rightarrow$	¬	A)
T	T	T	T	F	T	T	F	T
T	F	F	T	T	F	F	F	T
F	T	T	T	F	T	T	T	F
F	T	F	T	T	F	T	T	F

12. $[C \rightarrow (D \vee E)] \wedge \neg C$
13. $\neg(G \wedge (B \wedge H)) \rightarrow (G \vee (B \vee H))$

B. Write complete truth tables for the following sentences and mark the column that represents the possible truth values for the whole sentence.

1. $(D \wedge \neg D) \rightarrow G$

(D	$\wedge$	$\neg$	D)	$\boxed{\rightarrow}$	G
T	F	F	T	T	T
T	F	F	T	T	F
F	F	T	F	T	T
F	F	T	F	T	F

2. $(\neg P \vee \neg M) \rightarrow M$

3. $\neg\neg(\neg A \wedge \neg B)$

$\boxed{\neg}$	$\neg$	($\neg$	A	$\wedge$	$\neg$	B)
F	T	F	T	F	F	T
F	T	F	T	F	T	F
F	T	T	F	F	F	T
T	F	T	F	T	T	F

4. $[(D \wedge R) \rightarrow I] \rightarrow \neg(D \vee R)$

[(D	$\wedge$	R)	$\rightarrow$	I]	$\boxed{\rightarrow}$	$\neg$	(D	$\vee$	R)
T	T	T	T	T	F	F	T	T	T
T	T	T	F	F	T	F	T	T	T
T	F	F	T	T	F	F	T	T	F
T	F	F	T	F	F	F	T	T	F
F	F	T	T	T	F	F	F	T	T
F	F	T	T	F	F	F	F	T	T
F	F	F	T	T	T	T	F	F	F
F	F	F	T	F	T	T	F	F	F

C. Can you think of sentences with the following truth table:

1.

A	B	?
T	T	T
T	F	T
F	T	T
F	F	T

$A \vee \neg A$

2\.

A	B	?
T	T	F
T	F	F
F	T	T
F	F	F

$\neg A \wedge B$

3\.

A	B	?
T	T	T
T	F	F
F	T	T
F	F	T

$\neg(A \wedge \neg B)$;
or a more systematic answer by simply writing out the different options when the sentence is true:

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

D. Suppose X is TFL sentence containing two atomic sentences. Then there are in fact sixteen different possible columns for X.

1. Can you explain why?
   The first line might have T or F, the second line might have T or F, etc. There are 4 valuations, so there are $2 \times 2 \times 2 \times 2 = 2^4 = 16$ different ways of putting Ts and Fs to these 4 lines. I.e. 16 different columns.

2. Can you show for each of these combinations that there is a sentence of TFL with that column describing its truth.

3. Can you show there's always a formula just using $\neg$ and $\wedge$ with that column describing its truth.

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

CHAPTER 9

Validity and the method of truth tables

9.1 Validity

A TFL-argument $X_1, \ldots, X_n \therefore Y$ is VALID if and only if there is no valuation v such that all the premises, $X_1, \ldots, X_n$ are true relative to v but Y is false relative to v.

This can be determined by the method of truth tables.

Consider the argument:

$$\neg L \to (J \lor L), \neg J \therefore (\neg L \to J)$$

We need to check whether there is any valuation which makes both $\neg L \to (J \lor L)$ and $\neg J$ true whilst making $(\neg L \to J)$ false.

We so far have only done complete truth tables for single sentences. Now we need to do a complete truth table which simultaneously works for the three sentences, $\neg L \to (J \lor L)$, $\neg J$ and $(\neg L \to J)$.

The atomic sentences that appear in the premises and conclusion are J and L. So we will need to fill out a truth table

Valuation	J	L	$\neg L \to (J \vee L)$	$\neg J$	$(\neg L \to J)$
v_1 :	T	T			
v_2 :	T	F			
v_3 :	F	T			
v_4 :	F	F			

We will work up through the formation tree of each of our three sentences, starting with the leaves, the atomic sentences, for which we consult the valuation, and applying the truth rules to determine the truth values of higher up sentences on that valuation.

The result will be:

Valuation	J	L	$(\neg$	L	$\to$	$(J$	$\vee$	$L))$	$\neg$	J	$(\neg$	L	$\to$	$J)$
v_1 :	T	T	F	T	T	T	T	T	F	T	F	T	T	T
v_2 :	T	F	T	F	T	T	T	F	F	T	T	F	T	T
v_3 :	F	T	F	T	T	F	T	T	T	F	F	T	T	F
v_4 :	F	F	T	F	F	F	F	F	T	F	T	F	F	F

The only valuation on which both '$\neg L \to (J \vee L)$' and '$\neg J$' are true is v_3, and on v_3, '$\neg L \to J$' is true. So there is no valuation where the premises are true and the conclusion is false. So the argument is valid.

> To check validity with truth tables, draw a complete truth table for all the sentences simultaneously and look for a row like:
>
> | Valuation | Premise 1 | Premise 2 | $\cdots$ | Conclusion | |
> |---|---|---|---|---|---|
> | v_1 | | | | | |
> | ★ v_2 | T | T | T | F | ⇜ counterexample |
> | v_3 | | | | | |
> | v_4 | | | | | |
>
> **Invalid**
>
> If there is no such row, then it is *valid*.

(In fact, for full marks to it is invalid, one only needs to give such a counterexample valuation, not complete truth tables.)

We will also say:

Y is a LOGICAL CONSEQUENCE of $X_1, \ldots, X_n$ if and only if $X_1, \ldots, X_n \therefore Y$ is valid

One needs to check every valuation to see if there is one which is a counterexample valuation.

Consider all assignments of truth values to all the atomic sentences that appear in the premises or the conclusion.

Example

$D \to V, \neg D \;\therefore\; \neg V$

			Premise 1			Premise 2		Conclusion	
Valuation	D	V	D	$\to$	V	$\neg$	D	$\neg$	V
v_1 :	T	T	T	T	T	F	T	F	T
v_2 :	T	F	T	F	F	F	T	T	F
★ v_3 :	F	T	F	**T**	T	**T**	F	**F**	T
v_4 :	F	F	F	T	F	T	F	T	F

On v_3, the premises are true and the conclusion is false. It is invalid.

Example

Compare: $D \to V, D \;\therefore\; V$

			Premise 1			Premise 2	Conclusion
Valuation	D	V	D	$\to$	V	D	V
v_1 :	T	T	T	T	T	T	T
v_2 :	T	F	T	F	F	T	F
v_3 :	F	T	F	T	T	F	T
v_4 :	F	F	F	T	F	F	F

v_1 is the only valuation where premises are both true, and on it, conclusion is also true. So is valid.

Example

$P \to (C \lor D), \neg C \;\therefore P \to D$

				Premise 1					Premise 2		Conclusion		
Valuation	P	C	D	P	$\to$	$(C$	$\lor$	$D)$	$\neg$	C	$(P$	$\to$	$D)$
v_1 :	T	T	T	T	T	T	T	T	F	T	T	T	T
v_2 :	T	T	F	T	T	T	T	F	F	T	T	F	F
v_3 :	T	F	T	T	T	F	T	T	T	F	T	T	T
v_4 :	T	F	F	T	F	F	F	F	T	F	T	F	F
v_5 :	F	T	T	F	T	T	T	T	F	T	F	T	T
v_6 :	F	T	F	F	T	T	T	F	F	T	F	T	F
v_7 :	F	F	T	F	T	F	T	T	T	F	F	T	T
v_8 :	F	F	F	F	T	F	F	F	T	F	F	T	F

This is valid.

Example

$A \lor B, \neg(B \land \neg C) \therefore \neg C \lor A$

				Premise 1			Premise 2					Conclusion			
Val	A	B	C	A	$\lor$	B	$\neg$	$(B$	$\land$	$\neg$	$C)$	$(\neg$	C	$\lor$	$A)$
v_1 :	T	T	T	T	T	T	T	T	F	F	T	F	T	T	T
v_2 :	T	T	F	T	T	T	F	T	T	T	F	T	F	T	T
v_3 :	T	F	T	T	T	F	T	F	F	F	T	F	T	T	T
v_4 :	T	F	F	T	T	F	T	F	F	T	F	T	F	T	T
★ v_5 :	F	T	T	F	T	T	T	T	F	F	T	F	T	F	F
v_6 :	F	T	F	F	T	T	F	T	T	T	F	T	F	T	F
v_7 :	F	F	T	F	F	F	T	F	F	F	T	F	T	F	F
v_8 :	F	F	F	F	F	F	T	F	F	T	F	T	F	T	F

On v_5, premises are both true but conclusion is false. The argument is invalid.

Inconsistent premises

1. Jenny is happy.
2. Jenny is not happy.

$\therefore$ Therefore: The moon is made of cheese.

$$H, \neg H \therefore C$$

Valuation	H	C	Premise 1 H	Premise 2 $\neg H$	Conclusion C
v_1 :	T	T	T	F	T
v_2 :	T	F	T	F	F
v_3 :	F	T	F	T	T
v_4 :	F	F	F	T	F

Valid. There is no counterexample valuation!

We can make a more general observation: any argument with premises that are *inconsistent* is valid. We will now carefully introduce this notion of consistency and other logical notions.

CHAPTER 10

Other logical notions

10.1 Consistency

We have just seen an example where the premises are inconsistent: Here's the careful definition of this:

> A collection of TFL-sentences $X_1, \ldots, X_n$ (with possibly $n = 1$) is CONSISTENT iff there is a valuation relative to which $X_1, \ldots, X_n$ are true. Otherwise the collection is INCONSISTENT.

> To check if a collection is CONSISTENT with truth tables, look for a row like:

Valns	X_1	X_2	$\cdots$	X_n
v_1				
★ v_2	T	T	T	T
v_3				
v_4				

consistent

> If there is no such row then it is inconsistent.

'H' and '$\neg H$' are inconsistent.

Inconsistent collections have truth tables more like:

Valns	X_1	X_2	$\cdots$	X_n
v_1	F			
v_2		F		
v_3				F
v_4			F	

inconsistent

10.2 Tautologies and contradictions

A sentence X of TFL is a LOGICAL TRUTH or TAUTOLOGY if and only if it is true on every valuation.

"Tautology" is a name which is specific to TFL. The notion of being a "logical truth" is broader and can also apply in the case of First Order Logic.

We can determine whether a sentence is a tautology just by using truth tables. A complete truth table had a line for each valuation, so we just need to check if it is true on every line of the truth table.

Our main example of a tautology is $A \vee \neg A$. Whatever the weather is like you know it's either raining or not. Similarly here, whether or not A is true, we already know $A \vee \neg A$. We can check this with a truth table.

A	$A \vee \neg A$
T	T **T** F T
F	F **T** T F

Since there is a T under the main connective on every line of the truth table, $A \vee \neg A$ is true, whatever the valuation is, i.e. whether A is true or false.

X is a CONTRADICTION iff it is false on every valuation.

We can determine whether a sentence is a contradiction just by using truth tables. If the sentence is false on every line of a complete truth table, then it is false on every valuation, so it is a contradiction.

Our core example of a contradiction is $A \wedge \neg A$. Whether A is true or false, $A \wedge \neg A$ is false. This can again be checked using truth tables, observing that there is an F under the main connective in each line of the truth table.

A	$A \wedge \neg A$
T	T **F** F T
F	F **F** T F

> X is CONTINGENT iff it is neither a tautology nor a contradiction.

> To check these notions on a truth table, see if the value recoreded underneath the main connective is always T, always F or is mixed:
>
> | Valns | X |
> |---|---|
> | v_1 | T |
> | v_2 | T |
> | $\vdots$ | $\vdots$ |
> | v_n | T |
>
> **logical truth / tautology**
>
> | Valns | X |
> |---|---|
> | v_1 | F |
> | v_2 | F |
> | $\vdots$ | $\vdots$ |
> | v_n | F |
>
> **contradiction**
>
> | Valns | X |
> |---|---|
> | v_1 | |
> | v_2 | T |
> | $\vdots$ | |
> | v_n | F |
>
> **contingent**

10.3 Logical equivalence

Here is another useful notion:

> X and Y are LOGICALLY EQUIVALENT iff, on each valuation they are either both true or both false.

Again, it is easy to test for logical equivalence using truth tables. Consider the sentences '$\neg(P \vee Q)$' and '$\neg P \wedge \neg Q$'. Are they logically equivalent? To find out, we construct a truth table.

P	Q	$\neg\ (P \vee Q)$	$\neg P\ \wedge\ \neg Q$
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 logical operators; negation for the first sentence, conjunction for the second. On the first three rows, both are false. On the final row, both are true. Since they match on every row, the two sentences are logically equivalent.

One can show that '$(A \wedge B) \wedge C$' and '$A \wedge (B \wedge C)$' are logically equivalent.

> To check for logical equivalence with truth tables, look for a row where one of them is true and the other false, for example:
>
> | Valns | X | Y |
> |---|---|---|
> | v_1 | | |
> | ★ v_2 | T | F |
> | v_3 | | |
> | v_4 | | |
>
> **Not logically equivalent**

Practice exercises

A. Revisit your answers to A. Determine which sentences were tautologies, which were contradictions, and which were neither tautologies nor contradictions.

B. Use truth tables to determine whether these sentences are jointly consistent, or jointly inconsistent:

1. $A \to A$, $\neg A \to \neg A$, $A \wedge A$, $A \vee A$
2. $A \vee B$, $A \to C$, $B \to C$
3. $B \wedge (C \vee A)$, $A \to B$, $\neg(B \vee C)$
4. $A \leftrightarrow (B \vee C)$, $C \to \neg A$, $A \to \neg B$

C. Determine whether each sentence is a tautology, a contradiction, or a contingent sentence, using a complete truth table.

1. $\neg B \wedge B$
2. $\neg D \vee D$
3. $(A \wedge B) \vee (B \wedge A)$
4. $\neg[A \to (B \to A)]$
5. $A \leftrightarrow [A \to (B \wedge \neg B)]$
6. $[(A \wedge B) \leftrightarrow B] \to (A \to B)$

D. Determine whether each the following sentences are logically equivalent using complete truth tables. If the two sentences really are logically equivalent, write "equivalent." Otherwise write, "Not equivalent."

1. A and $\neg A$
2. $A \wedge \neg A$ and $\neg B \leftrightarrow B$
3. $[(A \vee B) \vee C]$ and $[A \vee (B \vee C)]$
4. $A \vee (B \wedge C)$ and $(A \vee B) \wedge (A \vee C)$
5. $[A \wedge (A \vee B)] \rightarrow B$ and $A \rightarrow B$

E. Determine whether each the following sentences are logically equivalent using complete truth tables. If the two sentences really are equivalent, write "equivalent." Otherwise write, "not equivalent."

1. $A \rightarrow A$ and $A \leftrightarrow A$
2. $\neg(A \rightarrow B)$ and $\neg A \rightarrow \neg B$
3. $A \vee B$ and $\neg A \rightarrow B$
4. $(A \rightarrow B) \rightarrow C$ and $A \rightarrow (B \rightarrow C)$
5. $A \leftrightarrow (B \leftrightarrow C)$ and $A \wedge (B \wedge C)$

F. Determine whether each collection of sentences is jointly consistent or jointly inconsistent using a complete truth table.

1. $A \wedge \neg B$, $\neg(A \rightarrow B)$, $B \rightarrow A$
2. $A \vee B$, $A \rightarrow \neg A$, $B \rightarrow \neg B$
3. $\neg(\neg A \vee B)$, $A \rightarrow \neg C$, $A \rightarrow (B \rightarrow C)$
4. $A \rightarrow B$, $A \wedge \neg B$
5. $A \rightarrow (B \rightarrow C)$, $(A \rightarrow B) \rightarrow C$, $A \rightarrow C$

G. Determine whether each collection of sentences is jointly consistent or jointly inconsistent, using a complete truth table.

1. $\neg B$, $A \rightarrow B$, A
2. $\neg(A \vee B)$, $A \leftrightarrow B$, $B \rightarrow A$
3. $A \vee B$, $\neg B$, $\neg B \rightarrow \neg A$
4. $A \leftrightarrow B$, $\neg B \vee \neg A$, $A \rightarrow B$
5. $(A \vee B) \vee C$, $\neg A \vee \neg B$, $\neg C \vee \neg B$

H. Answer each of the questions below and justify your answer.

1. Suppose X is true and Y is true
   Is $X \vee Y$ true.
   true

2. Suppose X is false
Is $X \vee Y$ true?
false. If Y is also false, then $X \vee Y$ is false.
3. Suppose X is false
Is $X \rightarrow Y$ true.
true
4. Suppose that X, Y and Z are jointly inconsistent.
What can you say about $X \wedge Y \wedge Y$?
Since the sentences are jointly inconsistent, there is no valuation on which they are all true. So their conjunction is false on every valuation. It is a contradiction
5. Suppose X is a tautology and $X \therefore Y$ is a valid argument.
What can you say about Y?
6. Suppose that $(X \wedge Y) \rightarrow Z$ is not a tautology.
What can you say about whether $X, Y \therefore Z$ is valid?
Since the sentence $(X \wedge Y) \rightarrow Z$ is not a tautology, there is some line on which it is false. Since it is a conditional, on that line, X and Y are true and Z is false. So the argument is invalid.
7. Suppose that X is a contradiction.
What can you say about whether $X, Y \vDash Z$?
Since X is false on every line of a complete truth table, there is no line on which X and Y are true and Z is false. So the entailment holds.
8. Suppose that X and Y are logically equivalent.
What can you say about $(X \vee Y)$?
Not much. Since X and Y are true on exactly the same lines of the truth table, their disjunction is true on exactly the same lines. So, their disjunction is logically equivalent to them.
9. Suppose that X and Y are *not* logically equivalent.
What can you say about $X \vee Y$?
X and Y have different truth values on at least one line of a complete truth table, and $(X \vee Y)$ will be true on that line. On other lines, it might be true or false. So $X \vee Y$ is either a tautology or it is contingent; it is *not* a contradiction.

I. Are the following statements true? Why?

- if Y is a logical consequence of X, then $X \rightarrow Y$ is a logical truth.Yes

- if $X \rightarrow Y$ is a logical truth, then Y is a logical consequence of X.Yes

- ▹ if $X \to Y \wedge \neg Y$ is a logical truth, then X is a logical contradiction. Yes
- ▹ if $X \to Y \wedge \neg Y$ is true relative to a valuation, then X is a logical contradiction. No
- ▹ if $X \vee \neg X \to Y$ is a logical truth, then Y is a tautology. Yes
- ▹ if $X \vee \neg X \to Y$ is true relative to a valuation, then Y is a tautology. No

J. Consider the following principle:

- Suppose X and Y are logically equivalent. Suppose an argument contains X (either as a premise, or as the conclusion). The validity of the argument would be unaffected, if we replaced X with Y.

Is this principle correct? Explain your answer.

CHAPTER 11

Truth table shortcuts

Not taught in the course. You may use these techniques if you wish.

Sometimes shortcuts can be taken when doing truth tables. In this section, we want to give you some permissible shortcuts to help you along the way. These will never be required but they can speed things up.

11.1 Working through truth tables

You will quickly find that you do not need to copy the truth value of each atomic sentence, but can simply refer back to them. So you can speed things up by writing:

P	Q	$(P \vee Q)$	$\leftrightarrow$	$\neg P$
T	T	T	F	F
T	F	T	F	F
F	T	T	T	T
F	F	F	F	T

You also know for sure that a disjunction is true whenever one of the disjuncts is true. So if you find a true disjunct, there is no need to work out the truth values of the other disjuncts. Thus you might offer:

P	Q	$(\neg P$	$\vee$	$\neg Q)$	$\vee$	$\neg P$
T	T	F	F	F	F	F
T	F	F	T	T	T	F
F	T				T	T
F	F				T	T

Equally, you know for sure that a conjunction is false whenever one of the conjuncts is false. So if you find a false conjunct, there is no need to work out the truth value of the other conjunct. Thus you might offer:

P	Q	$\neg$	$(P$	$\wedge$	$\neg Q)$	$\wedge$	$\neg P$
T	T					F	F
T	F					F	F
F	T	T		F		T	T
F	F	T		F		T	T

A similar short cut is available for conditionals. You immediately know that a conditional is true if either its consequent is true, or its antecedent is false. Thus you might present:

P	Q	$((P$	$\rightarrow$	$Q)$	$\rightarrow$	$P)$	$\rightarrow$	P
T	T						T	
T	F						T	
F	T		T		F		T	
F	F		T		F		T	

So '$((P \rightarrow Q) \rightarrow P) \rightarrow P$' is a tautology. In fact, it is an instance of *Peirce's Law*, named after Charles Sanders Peirce.

11.2 Testing for validity and logical consequence

When we use truth tables to test for validity or logical consequence, we are checking for *bad* lines: lines where the premises are all true and the conclusion is false. Note:

- Any line where the conclusion is true is not a bad line.
- Any line where some premise is false is not a bad line.

Since *all* we are doing is looking for bad lines, we should bear this in mind. So: if we find a line where the conclusion is true, we do not need to evaluate anything else on that line: that line definitely isn't bad.

Likewise, if we find a line where some premise is false, we do not need to evaluate anything else on that line.

With this in mind, consider how we might test the following for validity:

$$\neg L \to (J \vee L), \neg L \therefore J$$

The *first* thing we should do is evaluate the conclusion. If we find that the conclusion is *true* on some line, then that is not a bad line. So we can simply ignore the rest of the line. So at our first stage, we are left with something like:

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

where the blanks indicate that we are not going to bother doing any more investigation (since the line is not bad) and the question-marks indicate that we need to keep investigating.

The easiest premise to evaluate is the second, so we next do that:

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

Note that we no longer need to consider the third line on the table: it will not be a bad line, because (at least) one of premises is false on that line. Finally, we complete the truth table:

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

The truth table has no bad lines, so the argument is valid. (Any valuation on which all the premises are true is a valuation on which the conclusion is true.)

It might be worth illustrating the tactic again. Let us check whether the following argument is valid

$$A \vee B, \neg(A \wedge C), \neg(B \wedge \neg D) \therefore (\neg C \vee D)$$

At the first stage, we determine the truth value of the conclusion. Since this is a disjunction, it is true whenever either disjunct is true, so we can speed things along a bit. We can then ignore every line apart from the few lines where the conclusion is false.

A	B	C	D	$A \vee B$	$\neg(A \wedge C)$	$\neg(B \wedge \neg D)$	$(\neg C$	$\vee$	$D)$
T	T	T	T					T	
T	T	T	F	?	?	?	F	F	
T	T	F	T					T	
T	T	F	F				T	T	
T	F	T	T					T	
T	F	T	F	?	?	?	F	F	
T	F	F	T					T	
T	F	F	F				T	T	
F	T	T	T					T	
F	T	T	F	?	?	?	F	F	
F	T	F	T					T	
F	T	F	F				T	T	
F	F	T	T					T	
F	F	T	F	?	?	?	F	F	
F	F	F	T					T	
F	F	F	F				T	T	

We must now evaluate the premises. We use shortcuts where we can:

A	B	C	D	$A \lor B$	$\lnot$	$(A \land C)$	$\lnot$	$(B \land \lnot D)$	$(\lnot C$	$\lor\ D)$
T	T	T	T							T
T	T	T	F	T	F	T			F	F
T	T	F	T							T
T	T	F	F						T	T
T	F	T	T							T
T	F	T	F	T	F	T			F	F
T	F	F	T							T
T	F	F	F						T	T
F	T	T	T							T
F	T	T	F	T	T	F	F	T T	F	F
F	T	F	T							T
F	T	F	F						T	T
F	F	T	T							T
F	F	T	F	F					F	F
F	F	F	T							T
F	F	F	F						T	T

If we had used no shortcuts, we would have had to write 256 'T's or 'F's on this table. Using shortcuts, we only had to write 37. We have saved ourselves a *lot* of work.

We have been discussing shortcuts in testing for logically validity, but exactly the same shortcuts can be used in testing for logical consequence. By employing a similar notion of *bad* lines, you can save yourself a huge amount of work.

Practice exercises

A. Using shortcuts, determine whether each sentence is a tautology, a contradiction, or neither.

1. $\lnot B \land B$ — Contradiction

B	$\lnot B$	$\land\ B$
T	F	F
F		F

2. $\lnot D \lor D$ — Tautology

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

3. $(A \wedge B) \vee (B \wedge A)$ Neither

A	B	$(A \wedge B)$	$\vee$	$(B \wedge A)$
T	T	T	T	
T	F	F	F	F
F	T	F	F	F
F	F	F	F	F

4. $\neg[A \rightarrow (B \rightarrow A)]$ Contradiction

A	B	$\neg$	$[A$	$\rightarrow$	$(B$	$\rightarrow$	$A)]$
T	T	F		T		T	
T	F	F		T		T	
F	T	F		T			
F	F	F		T			

5. $A \leftrightarrow [A \rightarrow (B \wedge \neg B)]$ Contradiction

A	B	A	$\leftrightarrow$	$[A$	$\rightarrow$	$(B$	$\wedge$	$\neg$	$B)]$
T	T		F		F		F	F	
T	F		F		F		F		
F	T		F		T				
F	F		F		T				

6. $\neg(A \wedge B) \leftrightarrow A$ Neither

A	B	$\neg$	$(A$	$\wedge$	$B)$	$\leftrightarrow$	A
T	T	F		T		F	
T	F	T		F		T	
F	T	T		F		F	
F	F	T		F		F	

7. $A \rightarrow (B \vee C)$ Neither

A	B	C	$A \to$	$(B \vee C)$
T	T	T	T	T
T	T	F	T	T
T	F	T	T	T
T	F	F	F	F
F	T	T	T	
F	T	F	T	
F	F	T	T	
F	F	F	T	

8. $(A \wedge \neg A) \to (B \vee C)$ Tautology

A	B	C	$(A \wedge$	$\neg A)$	$\to (B \vee C)$
T	T	T	F	F	T
T	T	F	F	F	T
T	F	T	F	F	T
T	F	F	F	F	T
F	T	T	F		T
F	T	F	F		T
F	F	T	F		T
F	F	F	F		T

9. $(B \wedge D) \leftrightarrow [A \leftrightarrow (A \vee C)]$ Neither

A	B	C	D	$(B \wedge D)$	$\leftrightarrow$	$[A \leftrightarrow$	$(A \vee C)]$
T	T	T	T	T	T	T	T
T	T	T	F	F	F	T	T
T	T	F	T	T	T	T	T
T	T	F	F	F	F	T	T
T	F	T	T	F	F	T	T
T	F	T	F	F	F	T	T
T	F	F	T	F	F	T	T
T	F	F	F	F	F	T	T
F	T	T	T	T	F	F	T
F	T	T	F	F	T	F	T
F	T	F	T	T	T	T	F
F	T	F	F	F	F	T	F
F	F	T	T	F	T	F	T
F	F	T	F	F	T	F	T
F	F	F	T	F	F	T	F
F	F	F	F	F	F	T	F

CHAPTER 12

Partial truth tables

Not taught in course. This knowledge will not be needed for the exam. (Although you can use these techniques in the exam if you wish.)

Sometimes, we do not need to know what happens on every line of a truth table. Sometimes, just a line or two will do.

Tautology. In order to show that a sentence is a tautology, we need to show that it is true on every valuation. That is to say, we need to know that it comes out true on every line of the truth table. So we need a complete truth table.

To show that a sentence is *not* a tautology, however, we only need one line: a line on which the sentence is false. Therefore, in order to show that some sentence is not a tautology, it is enough to provide a single valuation—a single line of the truth table—which makes the sentence false.

Suppose that we want to show that the sentence '$(U \wedge T) \rightarrow (S \wedge W)$' is *not* a tautology. We set up a PARTIAL TRUTH TABLE:

S	T	U	W	$(U \wedge T) \rightarrow (S \wedge W)$
				F

We have only left space for one line, rather than 16, since we are only looking for one line on which the sentence is false. For just that reason, we have filled in 'F' for the entire sentence.

The main logical operator of the sentence is a conditional. In order for the conditional to be false, the antecedent must be true and the consequent must be false. So we fill these in on the table:

S	T	U	W	$(U$	$\wedge$	$T)$	$\rightarrow$	$(S$	$\wedge$	$W)$
					T		F		F	

In order for the '$(U \wedge T)$' to be true, both 'U' and 'T' must be true.

S	T	U	W	$(U$	$\wedge$	$T)$	$\rightarrow$	$(S$	$\wedge$	$W)$
	T	T		T	T	T	F		F	

Now we just need to make '$(S \wedge W)$' false. To do this, we need to make at least one of 'S' and 'W' false. We can make both 'S' and 'W' false if we want. All that matters is that the whole sentence turns out false on this line. Making an arbitrary decision, we finish the table in this way:

S	T	U	W	$(U$	$\wedge$	$T)$	$\rightarrow$	$(S$	$\wedge$	$W)$
F	T	T	F	T	T	T	F	F	F	F

We now have a partial truth table, which shows that '$(U \wedge T) \rightarrow (S \wedge W)$' is not a tautology. Put otherwise, we have shown that there is a valuation which makes '$(U \wedge T) \rightarrow (S \wedge W)$' false, namely, the valuation which makes 'S' false, 'T' true, 'U' true and 'W' false.

Contradiction. Showing that something is a contradiction requires a complete truth table: we need to show that there is no valuation which makes the sentence true; that is, we need to show that the sentence is false on every line of the truth table.

However, to show that something is *not* a contradiction, all we need to do is find a valuation which makes the sentence true, and a single line of a truth table will suffice. We can illustrate this with the same example.

S	T	U	W	$(U$	$\wedge$	$T)$	$\rightarrow$	$(S$	$\wedge$	$W)$
							T			

To make the sentence true, it will suffice to ensure that the antecedent is false. Since the antecedent is a conjunction, we can just make one of them false. For no particular reason, we choose to make 'U' false; and then we can assign whatever truth value we like to the other atomic sentences.

S	T	U	W	$(U$	$\wedge$	$T)$	$\rightarrow$	$(S$	$\wedge$	$W)$
F	T	F	F	F	F	T	T	F	F	F

Truth functional equivalence. To show that two sentences are logically equivalent, we must show that the sentences have the same truth value on every valuation. So this requires a complete truth table.

To show that two sentences are *not* logically equivalent, we only need to show that there is a valuation on which they have different truth values. So this requires only a one-line partial truth table: make the table so that one sentence is true and the other false.

Consistency. To show that some sentences are jointly consistent, we must show that there is a valuation which makes all of the sentences true,so this requires only a partial truth table with a single line.

To show that some sentences are jointly inconsistent, we must show that there is no valuation which makes all of the sentence true. So this requires a complete truth table: You must show that on every row of the table at least one of the sentences is false.

Validity. To show that an argument is valid, we must show that there is no valuation which makes all of the premises true and the conclusion false. So this requires a complete truth table. (Likewise for logical consequence.)

To show that argument is *invalid*, we must show that there is a valuation which makes all of the premises true and the conclusion false. So this requires only a one-line partial truth table on which all of the premises are true and the conclusion is false. (Likewise for a failure of logical consequence.)

This table summarises what is required:

	Yes	**No**
tautology?	complete truth table	one-line partial truth table
contradiction?	complete truth table	one-line partial truth table
equivalent?	complete truth table	one-line partial truth table
consistent?	one-line partial truth table	complete truth table
valid?	complete truth table	one-line partial truth table
logical consequence?	complete truth table	one-line partial truth table

PART III

Symbolizations in TFL

CHAPTER 13

First steps towards Symbolization

In Part II we have introduced syntax and semantics of the language of Truth Functional Logic. Now we wish to put the formal apparatus to use. We wish to show how we can symbolize arguments of English (or some other language) and then test them for validity.

Recall the arguments we discussed at the beginning of Part II, e.g., the argument

1. It is raining outside.
2. If it is raining outside, then Jenny is miserable.

$\therefore$ Therefore: Jenny is miserable.

We determined that the argument was of the form:

1. A
2. If A then B

$\therefore$ Therefore: B

We might symbolize the argument in TFL as follows:

$$A, A \rightarrow B \therefore B.$$

And, on the face of it, this seems to be a pretty good symbolization. We us the sentence letters 'A' and 'B' to stand for the declarative sentences 'it is raining outside' and 'Jenny is miserable' respectively, and the 'if…, then …' in terms of the conditional connective. Recall that we introduced the conditional with the 'if…, then…' reading in mind and that we motivated the truth table for the conditional using this understanding. Of course, now that we have a symbolization of the argument it is straightforward to test whether the argument is valid or not: we simply build a truth table.

Now, let's look at the argument

1. It is raining outside.
2. It is not raining outside or Jenny is miserable.

∴ Therefore: Jenny is miserable.

The form of the argument seems to be the following:

1. A
2. not A or B

∴ Therefore: B

If we use the sentence letters 'A' and 'B' to stand for same declarative sentences as in the previous argument, the argument can be aptly symbolized in TFL by:

$$A, \neg A \vee B \therefore B.$$

Recall that we introduced the negation connective in terms of 'not' or 'it is not the case', and understood the disjunction connective in terms of the English conjunction word 'or'. Given these assumptions the symbolization of the English argument turns out to be a valid TFL argument.

Our informal procedure for symbolizing English arguments can be summarized as follows: identify the form of the argument; identify the English (sub)sentence appearing in the argument with sentence letters, that is, atomic sentences of TFL and symbolize the conjunction words by the matching connective. But what precisely is the matching connective? The answer to this question might have been straightforward in the arguments above, but that's not always the case. It is worth tackling the idea of symbolization more systematically.

CHAPTER 14

Symbolizing Arguments

14.1 Atomic sentences

The starting point of symbolizing arguments in TFL was to identify (sub)sentences that were the *building blocks* of the argument and to symbolize them in TFL using certain sentence letters, e.g. in the previous chapter we used:

A: It's raining outside.

B: Jenny is miserable.

Such an assignment of declarative sentences to atomic sentences is called a SYMBOLIZATION KEY. Specifying a symbolization key is the first step in formalizing an argument in TFL. In doing this, we are not fixing this symbolization *once and for all.* We are just saying that, for the time being, we will think of the atomic sentence of TFL, 'A', as symbolizing the English sentence 'It is raining outside', and the atomic sentence of TFL, 'B', as symbolizing the English sentence 'Jenny is miserable'. Later, when we are dealing with different sentences or different arguments, we can provide a new symbolization key; as it might be:

A: Jenny is an anarcho-syndicalist
B: Dipan is an avid reader of Tolstoy

It is important to understand that whatever structure an English sentence might have is lost when it is symbolized by an atomic sentence of TFL. Recall that from the point of view of TFL, an atomic sentence is just a letter. It can be used to build more complex sentences, but it cannot be taken apart.

Once we have fixed a symbolization key, the next step is to symbolize the complex sentences of the argument, that is, the sentences that conjoin different sentences into a new more complex sentence via conjunction words. This requires a closer look at the relation between specific conjunction words and the connectives of TFL.

14.2 Symbolizing Complex Sentences

We introduced our connectives giving their rough English meaning. This gives us some tools to know how to symbolise English sentences. As a general instruction:

- ▹ If a sentence can be paraphrased as 'it is not the case that . . . '
  it can be symbolised as $\neg X$.
- ▹ If a sentence can be paraphrased as '. . . and . . . '
  it can be symbolised as $X \wedge Y$.
- ▹ If a sentence can be paraphrased as '. . . or . . . '
  it can be symbolised as $X \vee Y$.
- ▹ If a sentence can be paraphrased as 'if . . . , then . . . '
  it can be symbolised as $X \rightarrow Y$.

Usually, this process needs to be applied a number of times. Let's go through an example to see how to apply the strategy. Consider:

1. You won't get both soup and salad.

Let's start with the symbolization procedure.

You won't get both soup and salad.

It's not the case that you will get both soup and salad

$\neg \ldots$

|

You will get both soup and salad

Our next stage is to aim to symbolise 'you will get both soup and salad'

You won't get both soup and salad.
It's not the case that you will get both soup and salad
$\neg \ldots$

You will get both soup and salad
You will get soup and you will get salad
$\ldots \wedge \ldots$

You will get soup　　You will get salad

Both 'You will get soup' and 'You will get salad' cannot be paraphrased in any of our standard forms so we have to use an atomic TFL-sentence. As we haven't already introduced any atomic TFL-sentences, we need to introduce new ones. Let's say

S_1: You will get soup
S_2: You will get salad

So we can replace these leaves of the symbolisation tree with the atomic sentences, and then work the way back up to determine the symbolisation of the original sentence.

You won't get both soup and salad.
It's not the case that you will get both soup and salad
$\neg(S_1 \wedge S_2)$

You will get both soup and salad
You will get soup and you will get salad
$S_1 \wedge S_2$

You will get soup　　You will get salad
S_1　　S_2

More generally, here's our strategy:

1. Consider whether the sentence can be paraphrased as sentences in one of our standard forms.

paraphrased as	**symbolised as**
'It is not the case that …'	$\neg X$
'…and…'	$(X \wedge Y)$
'…or…'	$(X \wedge Y)$
'If $\dots_X$ then $\dots_Y$'	$(X \rightarrow Y)$

   If the answer is no, go to Step 3. If the answer is yes, go to Step 2.

2. Symbolize the sentence in accordance with the symbolization guideline with the component sentences in the place of the metavariables. For each of the component sentences, repeat the procedure, that is, go back to Step 1.

3. Check the symbolization key to see if we already have already chosen an atomic TFL-sentence to symbolise that English sentence.

   ▹ If so, replace the English sentence by that atomic TFL-sentence.

   ▹ Otherwise, choose a new atomic TFL-sentence to symbolize the sentence, extend the symbolization key accordingly and replace the English sentence by the atomic TFL sentence.

This process may seem very tedious and, of course, very often we can determine a correct symbolization of an English sentence without sticking painstakingly to this step by step process. However, by sticking to the process we make sure not to move too quickly. So if you are unsure of how to symbolize a sentence, the process provides a safety net. Yet before starting with the symbolization process one should determine the precise sentential structure of the sentence under consideration. Consider a slight variant of Sentence 1:

▹ It is not the case that you will get soup and you will get salad.

In this form the sentence displays a potential ambiguity between:

- ▹ It is not the case that: you will get soup and you will get salad.
- ▹ It is not the case that you will get soup, and you will get salad.

As we have seen the first reading gets symbolized as $\neg(S_1 \wedge S_2)$, while the second reading should be symbolized as $(\neg S_1 \wedge S_2)$ (Exercise: check that this is the correct symbolization). These two symbolizations are of course very different and, thus, at the beginning of the symbolization process it is important to determine the precise structure of the sentence one wishes to symbolize. We discuss this issue some more in Chapter 16.

14.3 Symbolizing Arguments

We have learned how to symbolize English sentences in TFL. But how does one symbolize arguments in TFL? The answer should not be to surprising: one simply needs to symbolize the premises and conclusion in TFL with the caveat that one must use one and the same symbolization key in this symbolization process. Let's look at a simple example (notice that the argument may not be sound):

1. Rishi Sunak will make Great Britain great again or he will loose the election.
2. Rishi Sunak will not make Great Britain great again.

∴ Therefore: Rishi Sunak will loose the election.

with the following symbolization key:

R: Rishi Sunak will make Great Britain great again.
L: Rishi Sunak will loose the election.

Using our symbolization procedure we then obtain the following TFL-argument

$$R \vee L, \neg R \therefore L$$

which can be checked for logical validity (Check!).

CHAPTER 15

Conjunction-words and TFL-Connectives

Consider these sentences:

1. Although Alice will come to the party, Beth will not.
2. Alice will come to the party, but Beth will not.
3. Either Alice or Beth will come to the party *(and not both)*.
4. Unless Alice will come to the party, Beth will not.
5. Neither Alice nor Beth will come to the party.
6. Alice will come to the party only if Beth will too.
7. Alice will come to the party if Beth will too.
8. Alice will come to the party if and only if Beth will.

In this chapter we will discuss how they should be symbolised in TFL. The feature of TFL is that it can capture any *truth functional* connective (recall 6.5), possibly by a more complex sentence.

We will also discuss in more detail when it is appropriate to paraphrase in the cannonical ways and when it is not.

15.1 Negation

Let's start by recalling the guideline we proposed in Section 6.

> If a sentence can be paraphrased as 'it is not the case that ...' it can be symbolised as $\neg X$.

How are we to understand this guideline? Consider the following sentences:

1. The information is retrievable.
2. The information is not be retrievable.
3. The information is irretrievable.
4. The information is not irretrievable.

Let us use the following representation key:

R: The information is retrievable.

Sentence 1 can now be symbolized by R. Moving on to sentence 3: saying that the information is irretrievable means that it is not the case that the information is retrievable. So even though sentence 3 does not contain the word 'not', we will symbolize it as '$\neg R$'.

Sentence 4 can be paraphrased as 'It is not the case that the information is irretrievable.' Which can again be paraphrased as 'It is not the case that it is not the case that the information is retrievable'. So we might symbolize this English sentence with the TFL sentence '$\neg\neg R$'. In other words, in English we can also express negation using prefixes such as '*ir*' or '*un*'. There are still further ways of expressing negation, for example, sometimes this may be done using the prefix '*dis*' as in honest/dishonest.

But some care is needed when handling negations. For example, one might think that sentences 5 and 6 negate each other and that if we symbolize sentence 5 with G, then we should symbolize sentence 6 as $\neg G$.

5. Stealing is good.
6. Stealing is bad.

There are two reasons why that would not be a good symbolization. For one, while we may all agree that stealing is not a good thing to do, we may think that it is perhaps not exactly bad, if someone who is on the verge of starvation steels some food. In other words, there may be things that are neither good nor bad. However, in TFL if 'G' is false, then '$\neg G$' is true and we would not have the option so say that something, e.g. stealing, is neither good nor bad. If we symbolize sentence 6 using a new sentence letter, say 'B', then we can allow for that option. For another, there is no syntactic marker such as a prefix, a word like 'not', etc. that suggests that the sentences 5 and 6 are negations of each other. There is nothing in the structure of the two sentences that suggests that one negates the other. But recall that one principal goal for symbolization is to check whether a given argument is valid and consider the argument:

1. Steeling is good or donating for charity is good.
2. Steeling is bad.

∴ Therefore: Donating for charity is good.

While this is a good argument and perhaps a conceptually valid argument it is not valid in the strict sense because it not valid in virtue of its form: it requires understanding 'bad' as 'not good'. So, if we were to symbolize 6 as '$\neg G$ we would stipulate some logical structure that the English sentence does not display and would wrongly deem the argument to be valid.

> When symbolizing sentences do only introduce logical structure that is displayed by the English sentence

If you wish to make the information that in the specific circumstances 'bad' means that something is not good and vice versa available in TFL, you can add this as an additional premises when symbolizing the relevant argument in TFL. That, is if we think that 6 is the negation of 5 we should add the following premise (using truth tables check that the premise is true, if and only if, '$\neg G$' is true whenever 'B' is true):

$$(\neg G \rightarrow B) \wedge (B \rightarrow \neg G).$$

If this premise is added to the above argument, then the argument is valid and rightly so!

Sometimes even in symbolizing sentences that very much look like they are negations of each other one needs to be careful. Consider:

7. Jane is happy.
8. Jane is unhappy.

To some it may feel like Jane can be neither happy nor unhappy, and this may even be assumed in some argument. Jane is without emotions and in a state of blank indifference: she does not feel happy nor does she feel unhappy. If we were to symbolize 7 by 'H' and then 8 by '$\neg H$', then we would rule out the possibility of Jane being neither happy nor unhappy. If we try to remain faithful to the idea that Jane can be neither happy nor unhappy then we need to symbolize 8 using a new atomic sentence of TFL.

15.2 Conjunction

The symbolization guideline for conjunction we introduced Section 6 was:.

> If a sentence can be paraphrased as '... and ...'
> it can be symbolised as $(X \wedge Y)$.

Let's look at some more examples:

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

Let's fix the following symbolization key:

A: Adam is athletic.
B: Barbara is athletic.
C: Barbara ie energetic.

Sentences 9-11 are obviously conjunctions. Notice that we make no attempt to symbolize the word 'also' in sentence 9 and 'both' in sentence 11. Words like 'both' and 'also' function to draw our attention to the fact that two things are being conjoined. Maybe they affect the emphasis of a sentence, but we will not (and cannot) symbolize such things in TFL. With this caveat, given the above symbolization key, we can symbolize the Sentence 9 by the TFL-sentence '$A \wedge B$'.

Sentence 10 says two things (about Barbara). In English, it is permissible to refer to Barbara only once. It *might* be tempting to think that we need to symbolize sentence 10 with something along the lines of 'B and energetic'. This would be a mistake. Once we symbolize part of a sentence as B, any further structure is lost, as B is an atomic sentence of TFL. Conversely, 'energetic' is not an English sentence at all. What we are aiming for is something like 'B and Barbara is energetic'. Given our symbolization key we should thus symbolize 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, even though in English we use the word 'athletic' only once. The sentence can be paraphrased as 'Barbara is athletic, and Adam is athletic'. We can symbolize this in TFL as $(B \wedge A)$, using the same symbolization key that we have been using.

Are Sentences 12 and 13 conjuctions? The word 'although' sets up a contrast between the first part of the sentence and the second part. Nevertheless, the sentence tells us 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 sentence 12 as, 'Barbara is energetic, *and* Barbara is not athletic'. The second conjunct contains a negation, so we paraphrase further: 'Barbara is energetic *and it is not the case that* Barbara is athletic'. Now we can symbolize this with the TFL sentence $(C \wedge \neg B)$. Note that we have lost all sorts of nuance in this symbolization. There is a distinct difference in tone between sentence 12 and 'Both Barbara is energetic and it is not the case that Barbara is athletic'. TFL does not (and cannot) preserve these nuances.

Sentence 13 raises similar issues. There is a contrastive structure, but this is not something that TFL can deal with. So we can paraphrase the sentence as 'Adam is athletic, *and* Barbara is more athletic than Adam'. (Notice that we once again replace the pronoun 'him' with 'Adam'.) How should we deal with the second conjunct? We already have the sentence letter A, which is being used to symbolize 'Adam is athletic', and the sentence B which is being used to symbolize 'Barbara is athletic'; but neither of these concerns their relative athleticity. So, to symbolize the entire sentence, we need a new sentence letter. Let the TFL sentence R symbolize the English sentence 'Barbara is more athletic than Adam'. Now we can symbolize sentence 13 by $(A \wedge R)$.

We can add these to our toolbox for symbolisation:

If a sentence can be paraphrased as

- ▹ '... and ...'
- ▹ '... but ...'
- ▹ 'Both... and ...'
- ▹ 'Although ..., ...'

it can be symbolised as $X \wedge Y$.

15.3 Disjunction

In Section 6 proposed that:

If a sentence can be paraphrased as '... or ...'
it can be symbolised as $(X \vee Y)$.

Let's consider some examples again

14. Fatima will play videogames, or she will watch movies.
15. Fatima or Omar will play videogames.

with the symbolization key

F: Fatima will play videogames.
O: Omar will play videogames.
M: Fatima will watch movies.

In Chapter 6 we already point out that here are two different reading available for the English word 'or': an *inclusive* one and an *exclusive* one. On the inclusive reading of 'or' we take the Sentence 15 to be true, if either Fatima or Omar play videogames, or both play videogames. In contrast, on the exclusive reading we take 15 to be true if either Fatima or Omar play videogames, but false if both play video games. It is the inclusive reading of 'or' that we symbolize using the TFL-connective $\vee$.

For Sentence 15 the inclusive reading is available and using the above symbolization key it can be symbolized as $F \vee O$. For Sentence 14 it seems that the salient understanding of the sentence points to the exclusive reading: presumably Fatima cannot play videogames and watch movies at the same time. In fact the reason why we tend to understand Sentence 14 in that way is, arguably, precisely because we think that Fatima cannot do two things at once, and not because of our understanding of 'or'. As a consequence when, in the context of an argument, we symbolize the Sentence 14 it is still preferable to symbolize it using the inclusive 'or' as '$F \vee M$' and to add the additional premise that she does not do both, that is '$\neg(F \wedge M)$'.

In contrast consider the sentence:

16. Either Fatima will play videogames, or she will watch movies.
17. Either Fatima will play videogames or Omar will play videogames.

It seems that '*either... or...*' in contrast to simple disjunction with '*or*', favors an exclusive reading. So unless, there is strong evidence to suspect that sentences 16 and 17 are to be understood in an inclusive way, they should be symbolized using the exclusive 'or'.

How do we symbolize the exclusive 'or' in TFL? Understood exclusively Sentence 16 can be paraphrased as follows:

▹ Fatima will play videogames or Omar will play videogames, but not both of them will play videogames.

Using our symbolization key we can symbolize the first part of sentence, that is, the part up to the comma as '$F \vee O$'. Since by our previous discussion we know that 'but' should be symbolized by a conjunction it remains to see how to symbolize 'not both of them will play videogames'. Going back to our discussion of negation and conjunction we see that the latter sentence is appropriately symbolized by the sentence $\neg(F \wedge O)$. Putting everything together Sentence 17 should be symbolized by the sentence

$$(F \vee O) \wedge \neg(F \wedge O).$$

> If a sentence can be paraphrased in English as '*either ... or ...*', it can be symbolised as $((X \vee Y) \wedge \neg(X \wedge Y))$.

Finally notice that in English '*neither... nor...*' is used to negate a disjunction, that is, using our symbolization key the sentences

▹ Neither Fatima nor Omar will play videogames.

▹ Neither Fatima will play videogames nor Omar will play videogames.

▹ Neither Fatima will play videogames nor will Omar.

should all be symbolized by the TFL sentence '$\neg(F \vee O)$'.

> If a sentence can be paraphrased in English as '*neither ... nor ...*', it can be symbolised as $\neg(X \vee Y)$.

15.4 Conditional

Let us again recall the symbolization guideline that we introduced in Section 6:

> If a sentence can be paraphrased as '*If ..., then ...*' it can be symbolised as $(X \rightarrow Y)$.

Now consider the sentences

18. If the number can be divided by 4, then it is an even number.
19. If the number can be divided by 4, it is an even number.
20. it is an even number, if the number can be divided by 4.

and use the following symbolization key:

P: The number can be divided by 4.
F: The number is an even number.

Sentence 19 and 20 are just a rephrasing of 18. So we will again symbolise them as $(P \rightarrow F)$.

Now consider

21. The number is an even number only if the number can be divided by 4.

21 is also a conditional. But, intuitively, while 18–19 are true 20 is false, as, e.g., 6 is an even number but cannot be divided by 4. This suggest that 20 needs to be symbolized differently than 18–19. Indeed, 'only if' inverses the order of the conditional: whereas 18–19 are symbolized as $(P \rightarrow F)$, 20 needs to be symbolized as $(F \rightarrow P)$. Reflecting on the truth table of the conditional, $F \rightarrow P$ says, assuming our symbolization key, that if F is true, that if it is true that the number is even, then P must be true for the conditional to be true, i.e., the number

must be divisible by 4. That's of course not always the case, so the truth conditions for 21 fits with our intuitive judgement.

In fact, sentence 21 can be paraphrased as 'If Jean is in France, then Jean is in Paris'. So we can symbolize it by $(F \rightarrow P)$: the other way around to 18.

> If a sentence can be paraphrased as
>
> - 'If $\ldots_X$, then $\ldots_Y$'
> - 'If $\ldots_X$, $\ldots_Y$'
> - '$\ldots_Y$ if $\ldots_X$'
> - '$\ldots_X$ only if $\ldots_Y$'
>
> it can be symbolized as $(X \rightarrow Y)$ such that X assumes the position of $\ldots_X$ and Y that of $\ldots_Y$.

Following this guidance, we symbolise all the following as '$A \rightarrow B$':

22. If Alice is on shift then Beth is on shift.
23. Beth is on shift if Alice is.
24. Alice is on shift only if Beth is.

At this point, a word of warning about the connective '$\rightarrow$' seems required: while the connectives like '$\wedge$' and '$\vee$' arguable closely track our understanding of '*and*' and '*or*' in natural language, the situation is slightly more complicated with respect to '$\rightarrow$' and '*if*…, *then*…' as we discussed in §6.5.

15.5 Biconditional

In this textbook we have frequently used the English expression '*if and only if*', as in: an atomic sentence X is true relative to a valuation v *if and only if* v assigns the value T to X. '*if and only if*' seems to be another conjunction word of English, that is, a way of composing two sentences into a new sentence. Is there a way to symbolize sentences of the form '… *if and only if*…' in TFL?

Consider the sentences:

25. Laika is a dog only if she is a mammal

26. Laika is a dog if she is a mammal
27. Laika is a dog if and only if she is a mammal

We will use the following symbolization key:

D: Laika is a dog
M: Laika is a mammal

Sentence 25, for reasons discussed above, can be symbolized by '$D \rightarrow M$'. in contrast Sentence 26 can be paraphrased as, 'If Laika is a mammal then Laika is a dog'. So it can be symbolized by '$M \rightarrow D$'.

Sentence 27 says something stronger than both 25 and 26. It can be paraphrased as 'Laika is a dog if Laika is a mammal, and Laika is a dog only if Laika is a mammal'. This is just the conjunction of sentences 25 and 26. So we can symbolize it as '$((D \rightarrow M) \wedge (M \rightarrow D))$'. We call this a BICONDITIONAL, because it entails the conditional in both directions. This leads to the following symbolization guideline:

> If a sentence can be paraphrased as '$\ldots_x$ if and only if $\ldots_y$' it can be symbolised as $((X \rightarrow Y) \wedge (Y \rightarrow X))$.

The expression 'if and only if' occurs a lot especially in philosophy, mathematics, and logic. For brevity, we can abbreviate it with the snappier word 'iff'. We will follow this practice. So 'if' with only *one* 'f' is the English conditional. But 'iff' with *two* 'f's is the English biconditional. Because the biconditional occurs so often, we will sometimes abbreviate the lengthy '$(X \rightarrow Y) \wedge (Y \rightarrow X)$' and write $X \leftrightarrow Y$ instead. However, officially the symbol '$\leftrightarrow$' is not a symbol of the language of TFL. It is merely used as a convenient way to state a biconditional and it is good to keep that in mind.

> $(X \leftrightarrow Y)$ is an abbreviation for $((X \rightarrow Y) \wedge (Y \rightarrow X))$.

A word of caution. Ordinary speakers of English often use 'if ..., then...' when they really mean to use something more like '...if and only if ...'. Perhaps your parents told you, when you were a child: 'if you don't eat your greens, you won't get any dessert'. Suppose you ate your greens, but that your parents refused to give you any dessert, on the grounds that they were only committed to the *conditional* (roughly 'if you get dessert, then you will have eaten your greens'), rather than the biconditional (roughly, 'you get dessert iff you eat your greens'). Well, a tantrum would rightly ensue. So, be aware of this when interpreting people; but in your own writing, make sure you use the biconditional iff you mean to.

15.6 Unless

A difficult case is when we use the conjunction word 'unless':

28. Unless you wear a jacket, you will catch a cold.
29. You will catch a cold unless you wear a jacket.

These two sentences are equivalent. They are also equivalent to the following:

30. If you do not wear a jacket, then you will catch a cold.
31. If you do not catch a cold, then you wore a jacket.
32. Either you will wear a jacket or you will catch a cold.

And we know how to symbolise these sentences. We will use the symbolization key:

J: You will wear a jacket.
D: You will catch a cold.

and can then give the symbolizations '$\neg J \rightarrow D$', '$\neg D \rightarrow J$' and '$J \vee D$'.

All three are correct symbolizations. Indeed, in you may wish to check that all three symbolizations are equivalent in TFL.

> If a sentence can be paraphrased as '*Unless* ..., ...',
> then it can be symbolized as $(X \vee Y)$.

Again, though, there is a little complication. 'Unless' can be symbolized as a conditional; but as we said above, people often use the conditional (on its own) when they mean to use the biconditional. Equally, 'unless' can be symbolized as a disjunction; but there are two kinds of disjunction (exclusive and inclusive). So it will not surprise you to discover that ordinary speakers of English often use 'unless' to mean something more like the biconditional, or like exclusive disjunction. Suppose someone says: 'I will go running unless it rains'. They probably mean something like 'I will go running iff it does not rain' (i.e. the biconditional), or 'either I will go running or it will rain, but not both' (i.e. exclusive disjunction). Again: be aware of this when interpreting what other people have said, but be precise in your writing.

15.7 More on Connectives in English

We have discussed several conjunction-words and sentence constructions that can be aptly symbolized in TFL.

In fact, TFL has the resources to express any *truth-functional* connective of English, possibly by a more complex sentence, as we did for, for example, exclusive 'or' or 'if and only if'.

However, there are of course many more conjunction-words in English and some of them can be adequately symbolized in TFL. However, there are also many conjunction words that cannot be adequately symbolized in TFL. We shall discuss some examples in Chapter 6.5.

Practice exercises

A. Using the symbolization key given, symbolize each English sentence in TFL.

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.
   $\neg M$
2. Those creatures are men in suits, or they are not.
   $(M \vee \neg M)$
3. Those creatures are either gorillas or chimpanzees.
   $(G \vee C)$
4. Those creatures are neither gorillas nor chimpanzees.
   $\neg(C \vee G)$
5. If those creatures are chimpanzees, then they are neither gorillas nor men in suits.
   $(C \rightarrow \neg(G \vee M))$
6. Unless those creatures are men in suits, they are either chimpanzees or they are gorillas.
   $(M \vee (C \vee G))$

B. Using the symbolization key given, symbolize each English sentence in TFL.

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.
   $(A \vee E)$
2. If Mister Ace was murdered, then the cook did it.
   $(A \to C)$
3. If Mister Edge was murdered, then the cook did not do it.
   $(E \to \neg C)$
4. Either the butler did it, or the Duchess is lying.
   $(B \vee D)$
5. The cook did it only if the Duchess is lying.
   $(C \to D)$
6. If the murder weapon was a frying pan, then the culprit must have been the cook.
   $(F \to C)$
7. If the murder weapon was not a frying pan, then the culprit was either the cook or the butler.
   $(\neg F \to (C \vee B))$
8. Mister Ace was murdered if and only if Mister Edge was not murdered.
   $(A \leftrightarrow \neg E)$
9. The Duchess is lying, unless it was Mister Edge who was murdered.
   $(D \vee E)$
10. If Mister Ace was murdered, he was done in with a frying pan.
    $(A \to F)$
11. Since the cook did it, the butler did not.
    $(C \wedge \neg B)$
12. Of course the Duchess is lying!
    D

C. Using the symbolization key given, symbolize each English sentence in TFL.

E_1: Ava is an electrician.
E_2: Harrison is an electrician.
F_1: Ava is a firefighter.
F_2: Harrison is a firefighter.
S_1: Ava is satisfied with her career.

S_2: Harrison is satisfied with his career.

1. Ava and Harrison are both electricians.
   $(E_1 \wedge E_2)$
2. If Ava is a firefighter, then she is satisfied with her career.
   $(F_1 \rightarrow S_1)$
3. Ava is a firefighter, unless she is an electrician.
   $(F_1 \vee E_1)$
4. Harrison is an unsatisfied electrician.
   $(E_2 \wedge \neg S_2)$
5. Neither Ava nor Harrison is an electrician.
   $\neg(E_1 \vee E_2)$
6. Both Ava and Harrison are electricians, but neither of them find it satisfying.
   $((E_1 \wedge E_2) \wedge \neg(S_1 \vee S_2))$
7. Harrison is satisfied only if he is a firefighter.
   $(S_2 \rightarrow F_2)$
8. If Ava is not an electrician, then neither is Harrison, but if she is, then he is too.
   $((\neg E_1 \rightarrow \neg E_2) \wedge (E_1 \rightarrow E_2))$
9. Ava is satisfied with her career if and only if Harrison is not satisfied with his.
   $(S_1 \leftrightarrow \neg S_2)$
10. If Harrison is both an electrician and a firefighter, then he must be satisfied with his work.
    $((E_2 \wedge F_2) \rightarrow S_2)$
11. It cannot be that Harrison is both an electrician and a firefighter.

    $\neg(E_2 \wedge F_2)$
12. Harrison and Ava are both firefighters if and only if neither of them is an electrician.
    $((F_2 \wedge F_1) \leftrightarrow \neg(E_2 \vee E_1))$

D. Give a symbolization key and symbolize the following English sentences in TFL.

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. Alice and Bob are both spies.
   $(A \land B)$
2. If either Alice or Bob is a spy, then the code has been broken.
   $((A \lor B) \to C)$
3. If neither Alice nor Bob is a spy, then the code remains unbroken.

   $(\neg(A \lor B) \to \neg C)$
4. The German embassy will be in an uproar, unless someone has broken the code.
   $(G \lor C)$
5. Either the code has been broken or it has not, but the German embassy will be in an uproar regardless.
   $((C \lor \neg C) \land G)$
6. Either Alice or Bob is a spy, but not both.
   $((A \lor B) \land \neg(A \land B))$

E. Give a symbolization key and symbolize the following English sentences in TFL.

F: There is food to be found in the pridelands.
R: Rafiki will talk about squashed bananas.
A: Simba is alive.
K: Scar will remain as king.

1. If there is food to be found in the pridelands, then Rafiki will talk about squashed bananas.
   $(F \to R)$
2. Rafiki will talk about squashed bananas unless Simba is alive.
   $(R \lor A)$
3. Rafiki will either talk about squashed bananas or he won't, but there is food to be found in the pridelands regardless.
   $((R \lor \neg R) \land F)$
4. Scar will remain as king if and only if there is food to be found in the pridelands.
   $(K \leftrightarrow F)$
5. If Simba is alive, then Scar will not remain as king.
   $(A \to \neg K)$

F. For each argument, write a symbolization key and symbolize the argument in TFL. Check whether these symbolizations are valid arguments. If not, give a valuation that shows that the argument is invalid. If the argument is invalid, are premises and conclusion jointly consistent?

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.

   P: Dorothy plays the Piano in the morning.
   C: Roger wakes up cranky.
   D: Dorothy is distracted.

   $(P \to C)$, $(P \vee D)$, $(\neg C \to D)$
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.

   T_1: It rains on Tuesday
   T_2: It snows on Tuesday
   S: Neville is sad on Tuesday
   C: Neville is cold on Tuesday

   $(T_1 \vee T_2)$, $(T_1 \to S)$, $(T_2 \to C)$, $(S \vee C)$
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.

   Z: Zoog remembered to do his chores
   C: Things are clean
   N: Things are neat

   $(Z \to (C \wedge \neg N))$, $(\neg Z \to (N \wedge \neg C))$, $((N \vee C) \wedge \neg(N \wedge C))$.

G. For each argument, write a symbolization key and translate the argument as well as possible into TFL. The part of the passage in italics is there to provide context for the argument, and doesn't need to be symbolized. Check for validity. Do these arguments use English connectives that cannot be symbolized appropriately in TFL (cf. Chapter 6.5)

1. It is going to rain soon. I know because my leg is hurting, and my leg hurts if it's going to rain.
2. *Spider-man tries to figure out the bad guy's plan.* If Doctor Octopus gets the uranium, he will blackmail the city. I am certain of this because if Doctor Octopus gets the uranium, he can make a dirty

bomb, and if he can make a dirty bomb, he will blackmail the city.

3. *A westerner tries to predict the policies of the Chinese government.* If the Chinese government cannot solve the water shortages in Beijing, they will have to move the capital. They don't want to move the capital. Therefore they must solve the water shortage. But the only way to solve the water shortage is to divert almost all the water from the Yangzi river northward. Therefore the Chinese government will go with the project to divert water from the south to the north.

CHAPTER 16

Ambiguity

In English, sentences can be AMBIGUOUS, i.e., they can have more than one meaning. There are many sources of ambiguity. One is *lexical ambiguity:* a sentence can contain words which have more than one meaning. For instance, 'bank' can mean the bank of a river, or a financial institution. So I might say that 'I went to the bank' when I took a stroll along the river, or when I went to deposit a check. Depending on the situation, a different meaning of 'bank' is intended, and so the sentence, when uttered in these different contexts, expresses different meanings.

A different kind of ambiguity is *structural ambiguity.* This arises when a sentence can be interpreted in different ways, and depending on the interpretation, a different meaning is selected. A famous example due to Noam Chomsky is the following:

1. Flying planes can be dangerous.

There is one reading in which 'flying' is used as an adjective which modifies 'planes'. In this sense, what's claimed to be dangerous are airplanes which are in the process of flying. In another reading, 'flying' is a gerund: what's claimed to be dangerous is the act of flying a plane. In the first case, you might use the sentence to warn someone who's about to launch a hot air balloon. In the second case, you might use it to counsel someone against becoming a pilot.

When the sentence is uttered, usually only one meaning is intended. Which of the possible meanings an utterance of a sentence intends is determined by context, or sometimes by how it is uttered (which parts of the sentence are stressed, for instance). Often one interpretation is much more likely to be intended, and in that case it will even be difficult to “see” the unintended reading. This is often the reason why a joke works, as in this example from Groucho Marx:

1. One morning I shot an elephant in my pajamas.
2. How he got in my pajamas, I don’t know.

Ambiguity is related to, but not the same as, vagueness. An adjective, for instance ‘rich’ or ‘tall,’ is VAGUE when it is not always possible to determine if it applies or not. For instance, a person who’s 6 ft 4 in (1.9 m) tall is pretty clearly tall, but a building that size is tiny. Here, context has a role to play in determining what the clear cases and clear non-cases are (‘tall for a person,’ ‘tall for a basketball player,’ ‘tall for a building’). Even when the context is clear, however, there will still be cases that fall in a middle range.

In TFL, we generally aim to avoid ambiguity. We will try to give our symbolization keys in such a way that they do not use ambiguous words or disambiguate them if a word has different meanings. So, e.g., your symbolization key will need two different sentence letters for ‘Rebecca went to the (money) bank’ and ‘Rebecca went to the (river) bank.’ Vagueness is harder to avoid. Since we have stipulated that every case (and later, every valuation) must make every basic sentence (or sentence letter) either true or false and nothing in between, we cannot accommodate borderline cases in TFL.

It is an important feature of sentences of TFL that they *cannot* be structurally ambiguous. Every sentence of TFL can be read in one, and only one, way. This feature of TFL is also a strength. If an English sentence is ambiguous, TFL can help us make clear what the different meanings are. Although we are pretty good at dealing with ambiguity in everyday conversation, avoiding it can sometimes be terribly important. Logic can then be usefully applied: it helps philosopher express their thoughts clearly, mathematicians to state their theorems rigorously, and software engineers to specify loop conditions, database queries, or verification criteria unambiguously.

Stating things without ambiguity is of crucial importance in the law as well. Here, ambiguity can, without exaggeration, be a matter of life and death. Here is a famous example of where a death sentence hinged on the interpretation of an ambiguity in the law. Roger Casement (1864–1916) was a British diplomat who was famous in his time for publicizing human-rights violations in the Congo and Peru (for which he was knighted in 1911). He was also an Irish nationalist. In 1914–16, Casement secretly travelled to Germany, with which Britain was at war at the time, and tried to recruit Irish prisoners of war to fight against Britain and for Irish independence. Upon his return to Ireland, he was captured by the British and tried for high treason.

The law under which Casement was tried is the *Treason Act of 1351*. That act specifies what counts as treason, and so the prosecution had to establish at trial that Casement's actions met the criteria set forth in the Treason Act. The relevant passage stipulated that someone is guilty of treason

> if a man is adherent to the King's enemies in his realm, giving to them aid and comfort in the realm, or elsewhere.

Casement's defense hinged on the last comma in this sentence, which is not present in the original French text of the law from 1351. It was not under dispute that Casement had been 'adherent to the King's enemies', but the question was whether being adherent to the King's enemies constituted treason only when it was done in the realm, or also when it was done abroad. The defense argued that the law was ambiguous. The claimed ambiguity hinged on whether 'or elsewhere' attaches only to 'giving aid and comfort to the King's enemies' (the natural reading without the comma), or to both 'being adherent to the King's enemies' and 'giving aid and comfort to the King's enemies' (the natural reading with the comma). Although the former interpretation might seem far fetched, the argument in its favor was actually not unpersuasive. Nevertheless, the court decided that the passage should be read with the comma, so Casement's antics in Germany were treasonous, and he was sentenced to death. Casement himself wrote that he was 'hanged by a comma'.

We can use TFL to symbolize both readings of the passage, and thus to provide a disambiguiation. First, we need a symbolization key:

> A: Casement was adherent to the King's enemies in the realm.
> G: Casement gave aid and comfort to the King's enemies in the realm.
> B: Casement was adherent to the King's enemies abroad.
> H: Casement gave aid and comfort to the King's enemies abroad.

The interpretation according to which Casement's behavior was not treasonous is this:

1. $A \vee (G \vee H)$

The interpretation which got him executed, on the other hand, can be symbolized by:

1. $(A \vee B) \vee (G \vee H)$

Remember that in the case we're dealing with Casement, was adherent to the King's enemies abroad (B is true), but not in the realm, and he did not give the King's enemies aid or comfort in or outside the realm (A, G, and H are false).

One common source of structural ambiguity in English arises from its lack of parentheses. For instance, if I say 'I like movies that are not long and boring', you will most likely think that what I dislike are movies that are long and boring. A less likely, but possible, interpretation is that I like movies that are both (a) not long and (b) boring. The first reading is more likely because who likes boring movies? But what about 'I like dishes that are not sweet and flavorful'? Here, the more likely interpretation is that I like savory, flavorful dishes. (Of course, I could have said that better, e.g., 'I like dishes that are not sweet, yet flavorful'.) Similar ambiguities result from the interaction of 'and' with 'or'. For instance, suppose I ask you to send me a picture of a small and dangerous or stealthy animal. Would a leopard count? It's stealthy, but not small. So it depends whether I'm looking for small animals that are dangerous or stealthy (leopard doesn't count), or whether I'm after either a small, dangerous animal or a stealthy animal (of any size).

These kinds of ambiguities are called *scope ambiguities*, since they depend on whether or not a connective is in the scope of another. For instance, the sentence, '*Avengers: Endgame* is not long and boring' is ambiguous between:

1. *Avengers: Endgame* is not: both long and boring.
2. *Avengers: Endgame* is both: not long and boring.

Sentence 2 is certainly false, since *Avengers: Endgame* is over three hours long. Whether you think 1 is true depends on if you think it is boring or not. We can use the symbolization key:

B: *Avengers: Endgame* is boring.
L: *Avengers: Endgame* is long.

Sentence 1 can now be symbolized as '$\neg(L \wedge B)$', whereas sentence 2 would be '$\neg L \wedge B$'. In the first case, the '$\wedge$' is in the scope of '$\neg$', in the second case '$\neg$' is in the scope of '$\wedge$'.

The sentence 'Tai Lung is small and dangerous or stealthy' is ambiguous between:

3. Tai Lung is either both small and dangerous or stealthy.
4. Tai Lung is both small and either dangerous or stealthy.

We can use the following symbolization key:

D: Tai Lung is dangerous.
S: Tai Lung is small.
T: Tai Lung is stealthy.

The symbolization of sentence 3 is '$(S \land D) \lor T$' and that of sentence 4 is '$S \land (D \lor T)$'. In the first, $\land$is in the scope of $\lor$, and in the second $\lor$is in the scope of $\land$.

Practice exercises

A. The following sentences are ambiguous. Give symbolization keys for each and symbolize the different readings.

1. Haskell is a birder and enjoys watching cranes.
2. The zoo has lions or tigers and bears.
3. The flower is not red or fragrant.

16.1 Limits of Symbolization in TFL

All of the connectives of TFL are truth-functional, but more than that: they really do nothing *but* map us between truth values. When we symbolize a sentence or an argument in TFL, we ignore everything *besides* the contribution that the truth values of a component might make to the truth value of the whole. There are subtleties to our ordinary claims that far outstrip their mere truth values. Sarcasm; poetry; snide implicature; emphasis; these are important parts of everyday discourse, but none of this is retained in TFL. As remarked in §15, TFL cannot capture the subtle differences between the following English sentences:

1. Dana is a logician and Dana is a nice person
2. Although Dana is a logician, Dana is a nice person
3. Dana is a logician despite being a nice person
4. Dana is a nice person, but also a logician
5. Dana's being a logician notwithstanding, he is a nice person

All of the above sentences will be symbolized with the same TFL sentence, perhaps '$L \wedge N$'. This does not mean that there are no subtle differences between those sentences. It just means that from the perspective of TFL, they should be symbolized by the same TFL-sentence and that TFL is ignorant of these subtle differences.

PART IV

Natural deduction for TFL

CHAPTER 17

The very idea of natural deduction

Way back in §2, we said that an argument is valid iff it is impossible to make all of the premises true and the conclusion false. In the case of TFL, this led us to develop truth tables. Each line of a complete truth table corresponds to a valuation. So, when faced with a TFL argument, we have a very direct way of assessing whether it is possible to make all of the premises true and the conclusion false: just thrash through the truth table. However, using truth tables means that we are not reasoning or arguing within the language of TFL. Rather, we reasoning *about* TFL-arguments rather than with them.

When you actually use arguments, for example in philosophical essays, you typically instead will break down an argument into smaller compelling steps, which your interlocutor or reader is forced to accept. It is good that at the end we may check, using truth tables, that your argument was indeed valid, but that's not how reasoning and argumentation works. The goal of this section is to provide a "reasoning system" for TFL. This reasoning system should produce valid arguments only and, ideally, also enable us to produce, that is, *prove* all valid arguments of TFL.

The reasoning systems for TFL is called NATURAL DEDUCTION. The idea of natural deduction is to indeed to break down an argument into smaller steps that are obviously valid and piecing these together, and thereby arriving at the intended conclusion. Suppose you're trying to convince someone that

$$A \to (B \wedge C), A \therefore B$$

is valid, and they don't see it. You can help them by breaking it up into two steps: first see that $B \wedge C$ follows and then note that B follows from $B \wedge C$.

In more detail: Grant me that $A \to (B \wedge C)$ and A are true. Then what else do we know to be true? Here's something: $B \wedge C$. Why? Because in general Modus Ponens is an excellent, compelling argument pattern: any argument of the form $X,\ X \to Y \ \therefore\ Y$ is valid: there are no valuations where X and $X \to Y$ are true but Y is false. So if our premises $A \to (B \wedge C)$ and A are true, then $B \wedge C$ must also be true: that's just Modus Ponens. So now from our supposition of $A \to (B \wedge C)$ and A, we now also know $B \wedge C$. What else follows from these three statements? Here's something: B. Why? Well, $B \wedge C$ is true; so certainly B must be true. So we know that from $A \to (B \wedge C)$ and A we can conclude B by walking someone through these two steps. This will be enough to show that $A \to (B \wedge C)$, $A \therefore B$ is valid.

To keep track of what assumptions have been made and steps of the argument we will give precise forms that this argument should be written:

1	$A \to (B \wedge C)$	
2	A	
3	$B \wedge C$	From, 1, 2
4	B	From, 3

Our premises are written above the horizontal line. They have to be granted without justification. Then each new line follows from the previous lines. The vertical line is there to highlight that everything coming below is within the context of the premises that have been assumed; that we are looking for consequences of the premises.

We can also use this presentation to be clear about arguments that we make in English:

1	If Alice came to the party, then Beth and Cath came	
2	Alice came to the party	
3	Beth and Cath came to the party	From, 1, 2
4	Beth came to the party	From, 3

You might think of it as a bag you're collecting things to be accepted in. You have to grant the premises, they go in the bag for free, then we give certain rules that allow us to add additional statements which must be true so long as the other things already in the bag are true.

Suppose I provide you with the following argument:

1	$P \to (\neg Q \to \neg R)$	
2	$P \to \neg Q$	
3	$P \wedge S$	
4	$\neg R$	From, 1, 2, 3

I concluded line 3 as a logical consequence of lines 1, 2 and 3. It does follow, i.e, the argument is valid; but this is not very helpful to someone who doesn't yet see that it's valid.

Instead, we will be describing various rules which we propose that have to be accepted as valid reasoning steps, and all more complicated steps should be broken up into simpler ones. So we should break this argument up into the steps:

1	$P \to (\neg Q \to \neg R)$	
2	$P \to \neg Q$	
3	$P \wedge S$	
4	P	From, 3
5	$\neg Q \to R$	From, 1, 4
6	$\neg Q$	From, 2, 4
7	R	From, 5, 6

The other thing we should do is to give a name for the steps that we use. Here we've just said 'From 1,4', but someone might ask: *how* does it follow from lines 1 and 4. We will give names for the various simple steps of reasoning we use and say: “well, it follows from lines 1 and 4 by the rule *conditional elimination*, which we abbreviate by“$\rightarrow$E”.”

1	$P \rightarrow (\neg Q \rightarrow \neg R)$	
2	$P \rightarrow \neg Q$	
3	$P \wedge S$	
4	P	$\wedge$E, 3
5	$\neg Q \rightarrow R$	$\rightarrow$E, 1, 4
6	$\neg Q$	$\rightarrow$E, 2, 4
7	R	$\rightarrow$E, 5, 6

We will provide various rules, and describe why they are acceptable. We should then break any other valid arguments should be broken up into these steps of reasoning.

17.1 More reasons for natural deduction

Using truth tables to show validity does not necessarily give us much *insight*. Consider two arguments in TFL:

$$P \vee Q, \neg P \therefore Q$$
$$P \rightarrow Q, P \therefore Q$$

Clearly, these are valid arguments. You can confirm that they are valid by constructing four-line truth tables, but we might say that they make use of different *forms* of reasoning. It might be nice to keep track of these different forms of inference.

One aim of a *natural deduction system* is to show that particular arguments are valid, in a way that allows us to understand the reasoning that the arguments might involve.

This is a very different way of thinking about arguments.

With truth tables, we directly consider different ways to make sentences true or false. With natural deduction systems, we manipulate sentences in accordance with rules that we have set down as good rules. The latter promises to give us a better insight—or at least, a different insight—into how arguments work.

The move to natural deduction might be motivated by more than the search for insight. It might also be motivated by *necessity*. Once our arguments involve 5 atomic sentences, a truth table test for validity will require 32 lines of truth table. That's quite a lot to check. But sometimes we might want to check such arguments.

1. Alice, or Betty, or Carys, or Dan, or Ella stole the teacher's pen.
2. It wasn't Alice.
3. It wasn't Betty,
4. It wasn't Carys,
5. It wasn't Dan

∴ It was Ella.

$$A \vee (B \vee (C \vee (D \vee E))), \neg A, \neg B, \neg C, \neg D \therefore E$$

And that will increase exponentially as more atomic sentences get added. Once an argument involves 20 atomic sentences,

1. Alice, or Betty, or Carys, ..., or Uli or Volker stole the teacher's pen.
2. It wasn't Alice.
3. It wasn't Betty,
4. $\vdots$
5. It wasn't Uli,

∴ Therefore: It was Volker.

This argument is also valid—as you might be able to tell—but to test it requires a truth table with $2^{20} = 1048576$ lines. In principle, we can set a machine to grind through truth tables and report back when it is finished. In practice, complicated arguments in TFL can become *intractable* if we use truth tables.

When we get to first-order logic (FOL) (beginning in chapter 25) the problem gets dramatically worse. There is nothing like the truth table test for FOL. To assess whether or not an argument is valid, we have to reason about *all* interpretations, but, as we will see, there are infinitely many possible interpretations. We cannot even in principle set a machine to grind through infinitely many possible interpretations and report back when it is finished: it will *never* finish. We either need to come up with some more efficient way of reasoning about all interpretations, or we need to look for something different. We will be looking for something different; and we will develop natural deduction.[1]

The modern development of natural deduction dates from simultaneous and unrelated papers by Gerhard Gentzen and Stanisław Jaśkowski (both in 1934). However, the natural deduction system that we will consider is based largely around work by Frederic Fitch (first published in 1952).

Natural deduction selects a few basic rules of inference and natural forms of reasoning and encodes these into a proof system. We will now see natural deduction for TFL. This system will form the basis also for natural deduction for FOL, which will also add rules for the quantifiers.

[1]There are, in fact, systems that codify ways to reason about all possible interpretations which can be used for FOL in a similar way to the way we use truth tables for TFL. They were developed in the 1950s by Evert Beth and Jaakko Hintikka, but we will not follow this path.

CHAPTER 18

The First Basic Rules for TFL: the basic rules without subproofs

We will now describe the various rules one can use. All other valid arguments should be broken up into steps using these rules. We will give a particular list of rules, other systems will choose other particular rules.

The rules we give will often be attached to particular connectives. This will help guide finding proofs.

The full list of the rules can be found in Appendix B.

18.1 Reiteration

The very first rule is so breathtakingly obvious that it is surprising we bother with it at all.

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 example:

$$
\begin{array}{r|ll}
4 & A \wedge B & \\
\vdots & \vdots & \\
10 & A \wedge B & \text{R, } 4
\end{array}
$$

This indicates that we have written '$A \wedge B$' on line 4. Now, at some later line—line 10, for example—we have decided that we want to repeat this. So we write it down again. We also add a citation which justifies what we have written. In this case, we write 'A', to indicate that we are using the reiteration rule, and we write '4', to indicate that we have applied it to line 4.

Here is a general expression of the rule:

$$
\begin{array}{r|ll}
 & \vdots & \\
m & X & \\
 & \vdots & \\
 & X & \text{R, } m
\end{array}
$$

The point is that, if any sentence X occurs on some line, then we can repeat X on later lines. Each line of our proof must be justified by some rule, and here we have 'R m'. This means: Reiteration, applied to line m.

Two things need emphasising. First 'X' is not a sentence of TFL. Rather, it a symbol in the metalanguage, which we use when we want to talk about any sentence of TFL (see §5). Second, and similarly, 'm' is not a numeral that will appear on a proof. Rather, it is a symbol in the metalanguage, which we use when we want to talk about any line number of a proof. In an actual proof, the lines are numbered '1', '2', '3', and so forth. But when we define the rule, we use variables to underscore the point that the rule may be applied at any point.

Why might this be useful? For example, we can now show $A \therefore A$ is valid using the proof:

1	A	
2	A	R, 2

The rule really becomes useful, though, once we are dealing with subproofs, which we will see in the next chapter.

18.2 Modus Ponens

Consider the following argument:

> If Jane is smart then she is fast. Jane is smart. ∴ Jane is fast.

This argument is certainly valid. In fact any argument of the form

$$X \to Y, X \;\therefore Y$$

is valid. We introduce a rule of natural deduction that encodes this idea. This is called *Modus Ponens.*

We introduce a rule of Natural Deduction which allows us to make this reasoning step. We will call it the "Conditional Elimination" rule ($\to$E). This choice of name is because we start with something including the connective $\to$ and we derive something without the connective, that is we have *eliminated* the $\to$ connective. For each connective we will have introduction and elimination rules, however we will wait until the next chapter to see Conditional Introduction.

In a simple use of this rule, we might just use it to derive from the premises $S \to F$ and S the conclusion F:

1	$S \to F$	
2	S	
3	F	$\to$E, 1, 2

This would then be a natural deduction proof that $S \to F, S \therefore F$ is valid.

Each line, except for the premsies which are taken as assumptions, has to be labelled with the rule it used. So here, we write "$\to$E 1,2" to say that we obtained line 3 by use of this rule $\rightarrow$ Elimination applied to lines 1 and 2.

We can also apply the rule when our $X \rightarrow Y$ and X are not themselves premises but have themselves been derived in the course of the proof.

1	Premise 1	
2	Premise 2	
	$\vdots$	
8	$S \rightarrow F$	some rule
	$\vdots$	
15	S	another rule
	$\vdots$	
23	F	$\rightarrow$E, 8, 15

It also can be that they appear in a different order, or that one appears in the premises, for example:

1	S	
	$\vdots$	
8	$S \rightarrow F$	
	$\vdots$	
23	F	$\rightarrow$E, 8, 1

We write our general rule as:

$$
\begin{array}{r|ll}
 & \vdots & \\
m & X \to Y & \\
 & \vdots & \\
n & X & \\
 & \vdots & \\
 & Y & \to\text{E}, m, n
\end{array}
$$

We can apply it to any X and Y. For example,

$$
\begin{array}{r|ll}
1 & (A \vee B) \to \neg F & \\
2 & (A \vee B) & \\
\hline
3 & \neg F & \to\text{E}, 1, 2
\end{array}
$$

In this, X is $(A \vee B)$, and Y is $\neg F$.

We would typically now move to introducing Conditional Introduction. However, we will first do all the other rules of the system, because Conditional Introduction involves additional complexity.

18.3 Conjunction Introduction

Suppose we want to show that Alice and Beth both came to the party. One obvious way to do this would be as follows: first we show that Alice came to the party; then Beth came to the party; then we put these two demonstrations together, to obtain the conjunction.

Our natural deduction system will capture this thought straightforwardly. In the example given, we might adopt the following symbolization key:

A: Alice came to the party
B: Beth came to the party

Perhaps we are working through a proof, and we have obtained 'A' on line 8 and 'B' on line 15. Then on any subsequent line we can obtain '$A \wedge B$'. For example our proof might contain the following lines:

$$
\begin{array}{r|ll}
8 & A & \\
15 & B & \\
23 & A \wedge B & \wedge\text{I}, 8, 15
\end{array}
$$

Note that every line of our proof must either be an assumption, or must be justified by some rule. We cite '$\wedge$I 8, 15' here to indicate that the line is obtained by the rule of conjunction introduction ($\wedge$I) applied to lines 8 and 15. More generally, here is our conjunction introduction rule:

$$
\begin{array}{r|ll}
 & \vdots & \\
m & X & \\
 & \vdots & \\
n & Y & \\
 & \vdots & \\
 & X \wedge Y & \wedge\text{I}, m, n
\end{array}
$$

Two things need emphasising.

First 'X' and 'Y' are metavariables. They are not particular sentences of TFL but are there to play the role of any particular sentence (see §5).

Similarly, 'm' is not a numeral that will appear on a proof. Rather, it is a symbol in the metalanguage, which we use when we want to talk about any line number of a proof. In an actual proof, the lines are numbered '1', '2', '3', and so forth. But when we define the rule, we use variables to underscore the point that the rule may be applied at any point.

To be clear, the statement of the rule is *schematic*. It is not itself a proof. 'X' and 'Y' are not sentences of TFL. Rather, they are symbols in the metalanguage, which we use when we want to talk about any sentence of TFL (see §5). Similarly, 'm' and 'n' are not a numerals that will appear on any actual proof. Rather, they are symbols in the metalanguage, which we use when we want to talk about any line number of any proof. In an actual proof, the lines are numbered '1', '2', '3', and so forth, but when we define the rule, we use variables to emphasize that the rule may be applied at any point. The rule requires only that we have both conjuncts available to us somewhere in the proof. They can be separated from one another, and they can appear in any order.

The rule is called 'conjunction *introduction*' because it introduces the symbol '$\wedge$' into our proof where it may have been absent.

18.4 Conjunction Elimination

Correspondingly, we have a rule that *eliminates* that symbol. Suppose you have shown that Alice and Beth both came to the party. You are entitled to conclude that Alice came to the party. Equally, you are entitled to conclude that Beth came to the party. Putting this together, we obtain our conjunction elimination rule(s):

	$\vdots$	
m	$X \wedge Y$	
	$\vdots$	
	X	$\wedge$E, m

and equally:

	$\vdots$	
m	$X \wedge Y$	
	$\vdots$	
	Y	$\wedge$E, m

The point is simply that, when you have a conjunction on some line of a proof, you can obtain either of the conjuncts by $\wedge$E. (One point, might be worth emphasising: you can only apply this rule when conjunction is the main logical operator. So you cannot infer 'D' just from '$C \vee (D \wedge E)$'!)

Even with just these two rules, we can start to see some of the power of our formal proof system. Consider:

1. $[(A \vee B) \rightarrow (C \vee D)] \wedge \neg(E \vee F)$

∴ Therefore: $\neg(E \vee F) \wedge [(A \vee B) \rightarrow (C \vee D)]$

The main logical operator in both the premise and conclusion of this argument is '$\wedge$'. In order to provide a proof, we begin by writing down the premise, which is our assumption. We draw a line below this: everything after this line must follow from our assumptions by (repeated applications of) our rules of inference. So the beginning of the proof looks like this:

1	$[(A \vee B) \rightarrow (C \vee D)] \wedge \neg(E \vee F)$	

From the premise, we can get each of the conjuncts by $\wedge$E. The proof now looks like this:

1	$[(A \vee B) \rightarrow (C \vee D)] \wedge \neg(E \vee F)$	
2	$[(A \vee B) \rightarrow (C \vee D)]$	$\wedge$E, 1
3	$\neg(E \vee F)$	$\wedge$E, 1

So by applying the $\wedge$I rule to lines 3 and 2 (in that order), we arrive at the desired conclusion. The finished proof looks like this:

1	$[(A \vee B) \rightarrow (C \vee D)] \wedge \neg(E \vee F)$	
2	$[(A \vee B) \rightarrow (C \vee D)]$	$\wedge$E, 1
3	$\neg(E \vee F)$	$\wedge$E, 1
4	$\neg(E \vee F) \wedge [(A \vee B) \rightarrow (C \vee D)]$	$\wedge$I, 3, 2

This is a very simple proof, but it shows how we can chain rules of proof together into longer proofs. In passing, note that investigating this argument with a truth table would have required a staggering 256 lines; our formal proof required only four lines.

It is worth giving another example. Way back in §??, we noted that this argument is valid:

$$A \wedge (B \wedge C) \therefore (A \wedge B) \wedge C$$

To provide a proof corresponding with this argument, we start by writing:

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

From the premise, we can get each of the conjuncts by applying $\wedge$E twice. We can then apply $\wedge$E twice more, so our proof looks like:

1	$A \wedge (B \wedge C)$	
2	A	$\wedge$E, 1
3	$B \wedge C$	$\wedge$E, 1
4	B	$\wedge$E, 3
5	C	$\wedge$E, 3

But now we can merrily reintroduce conjunctions in the order we wanted them, so that our final proof is:

1	$A \wedge (B \wedge C)$	
2	A	$\wedge$E, 1
3	$B \wedge C$	$\wedge$E, 1
4	B	$\wedge$E, 3
5	C	$\wedge$E, 3
6	$A \wedge B$	$\wedge$I, 2, 4
7	$(A \wedge B) \wedge C$	$\wedge$I, 6, 5

Recall that our official definition of sentences in TFL only allowed conjunctions with two conjuncts. The proof just given suggests that we could drop inner brackets in all of our proofs. However, this is not standard, and we will not do this. Instead, we will maintain our more austere bracketing conventions. (Though we will still allow ourselves to drop outermost brackets, for legibility.)

Let me offer one final illustration. When using the $\wedge$I rule, there is no requirement that it is applied to two different sentences. So we can formally prove 'A' from 'A' as follows:

1	A	
2	$A \wedge A$	$\wedge$I, 1, 1
3	A	$\wedge$E, 2

Simple, but effective. In fact this shows that we didn't need to have the rule of Reiteration as we could always argue by $\wedge$I then $\wedge$E. But for ease we will allow Reiteration as a basic rule so you don't have to argue this way.

18.5 Disjunction Introduction

Suppose Alice came to the party. Then Alice or Beth came to the party. After all, to say that Alice or Beth came to the party is to say something weaker than to say that Alice came to the party.

Let me emphasize this point. Suppose Alice came to the party. It follows that Alice came to the party *or* I am the Queen of England. Equally, it follows that Alice or the Queen came to the party. Equally, it follows that Alice came to the party *or* that God is dead. Many of these are strange inferences to draw, but there is nothing *logically* wrong with them (even if they maybe violate all sorts of implicit conversational norms).

Armed with all this, we present the disjunction introduction rule(s):

$$
\begin{array}{r|ll}
 & \vdots & \\
m & X & \\
 & \vdots & \\
 & X \vee Y & \vee\text{I}, m
\end{array}
$$

and

$$
\begin{array}{r|ll}
 & \vdots & \\
m & X & \\
 & \vdots & \\
 & Y \vee X & \vee\text{I}, m
\end{array}
$$

Notice that Y can be *any* sentence whatsoever, so the following is a perfectly acceptable proof:

1	M		
2	$M \lor ([(A \leftrightarrow B) \to (C \land D)] \leftrightarrow [E \land F])$	$\lor$I, 1	

Using a truth table to show this would have taken 128 lines.

18.6 Law of Excluded Middle

We will actually add another rule for how to introduce a disjunction. There are special kinds of disjunctions that don't need further justification: sentences of the form $X \lor \neg X$. In §10.2 we saw that $A \lor \neg A$ is a tautology: it is true on all valuations. More generally, any sentence of the form $X \lor \neg X$ is a tautology. The rule *Law of Excluded Middle* encodes this fact: it simply says that you are always allowed to write $X \lor \neg X$:

$\vdots$	
$X \lor \neg X$	LEM

As always, X can be whatever you want, e.g. $(A \land (B \to C))$. Then the rule tells us, e.g. that you can write $(A \land (B \to C)) \lor \neg(A \land (B \to C))$ on any line of the proof.

The law of Excluded Middle is often used in combination with Disjunction Elimination. However, we will not yet introduce Disjunction Elimination. It will be introduced in §19.2. That's when we'll see the full power of the LEM rule.

18.7 Strategies

We have a few more rules to introduce, but at this point we pause to mention a strategies to come up with proofs yourself:

Work backwards from what you want. The ultimate goal is to obtain 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 get to this new goal.

For example: If your conclusion is a conjunction $X \wedge Y$, plan to use the $\wedge$I rule. This requires finding both X and Y.

Work forwards from what you have. When you are starting a proof, look at the premises; later, look at the sentences that you have obtained so far. Think about the elimination rules for the main operators of these sentences. These will often 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.

Sometimes, though, this won't yet be possible. For that, we need to finish off our list of the basic rules of natural deduction.

Practice exercises

A. The following proofs are missing their citations (rule and line numbers). Add them, to turn them into *bona fide* proofs. Additionally, write down the argument that corresponds to each proof.

1	$P \wedge S$	
2	$S \rightarrow R$	
3	S	$\wedge$E, 1
4	R	$\rightarrow$E, 2, 3
5	$A \vee E$	$\vee$I, 4

B. Give a proof corresponding to each of the following arguments:

1. $A \wedge (B \vee C)$ $\therefore$ $(B \vee C) \wedge A$

1	$A \wedge (B \vee C)$	
2	A	$\wedge$E, 1
3	$B \vee C$	$\wedge$E, 1
4	$(B \vee C) \wedge A$	$\wedge$I, 3, 2

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

1	$(A \vee \neg A) \to B$	
2	$A \vee \neg A$	LEM
3	B	$\to$E, 1, 2

CHAPTER 19

More Basic Rules for TFL: the basic rules with subproofs

To introduce a conditional we need to introduce another kind of thing: a sub-proof.

19.1 Conditional Introduction

The following argument is valid:

> Alice came to the party. Therefore if Beth came to the party, then both Alice and Beth came.

If someone doubted that this was valid, we might try to convince them otherwise by explaining ourselves as follows:

> Assume that Alice came to the party. Now, *additionally* assume that Beth came to the party. Then by conjunction introduction—which we just discussed—both Alice and Beth came. Of course, that's conditional on the assumption that Beth came to the party. But this just means that, if Beth came to the party, then both Alice and Beth came.

We might write this in a form that is closer to our natural deduction format:

1	Alice came to the party	
2	Beth came to the party	
3	Both Alice and Beth came to the party	$\wedge$I, 1, 2
4	Thus, *if* Beth came to the party, both Alice and Beth came.	$\rightarrow$I, 2–3

The natural deduction format of lines and indentations is there to replace the words and context like "suppose that". The "suppose that" used on line 2 is represented in the formal system by the indentation and additional line. And like the premises, this line does not need to be justified, it is taken as an assumption, and a line underneath it is drawn. What comes underneath this, on line 3, is still indented, and is within the context of the supposition that Beth came to the party. But once we move to line 4, the additional assumption is no longer in place. It has been *discharged.*

Now let's present this again a little more formally: We started with one premise, 'Alice came to the party', thus:

1	A

The next thing we did is to make an *additional* assumption ('Beth came to the party'), for the sake of argument. To indicate that we are no longer dealing *merely* with our original assumption ('A'), but with some additional assumption, we continue our proof as follows:

1	A
2	B

Note that we are *not* claiming, on line 2, to have proved 'B' from line 1, so we do not need to write in any justification for the additional assumption on line 2. We do, however, need to mark that it is an additional assumption. We do this by drawing a line under it (to indicate that it is an assumption) and by indenting it with a further vertical line (to indicate that it is additional).

With this extra assumption in place, we are in a position to use $\wedge$I. So we can continue our proof:

$$
\begin{array}{r|l l}
1 & A & \\
\cline{2-2}
2 & \quad \begin{array}{|l} B \\ \hline A \wedge B \end{array} & \\
3 & & \wedge\text{I}, 1, 2
\end{array}
$$

So we have now shown that, on the additional assumption, 'B', we can obtain '$A \wedge B$'. We can therefore conclude that, if 'B' obtains, then so does '$A \wedge B$'. Or, to put it more briefly, we can conclude '$B \rightarrow (A \wedge B)$':

$$
\begin{array}{r|l l}
1 & A & \\
\cline{2-2}
2 & \quad \begin{array}{|l} B \\ \hline A \wedge B \end{array} & \\
3 & & \wedge\text{I}, 1, 2 \\
4 & B \rightarrow (A \wedge B) & \rightarrow\text{I}, 2\text{–}3
\end{array}
$$

Observe that we have dropped back to using one vertical line. We have *discharged* the additional assumption, 'B', since the conditional itself follows just from our original assumption, 'A'.

The general pattern at work here is the following. We first make an additional assumption, X; and from that additional assumption, we prove Y. In that case, we know the following: If X, then Y. This is wrapped up in the rule for conditional introduction:

$$
\begin{array}{r|l l}
m & \quad \begin{array}{|l} X \\ \hline \vdots \\ Y \end{array} & \\
n & & \\
 & X \rightarrow Y & \rightarrow\text{I}, m\text{–}n
\end{array}
$$

There can be as many or as few lines as you like between lines m and n.

It will help to offer a second illustration of →I in action. Suppose we want to consider the following:

$$P \to Q, Q \to R \therefore P \to R$$

We start by listing *both* of our premises. Then, since we want to arrive at a conditional (namely, '$P \to R$'), we additionally assume the antecedent to that conditional. Thus our main proof starts:

$$\begin{array}{r|l}
1 & P \to Q \\
2 & \underline{Q \to R} \\
3 & \quad \begin{array}{|l} \underline{P} \end{array}
\end{array}$$

Note that we have made 'P' available, by treating it as an additional assumption, but now, we can use →E on the first premise. This will yield 'Q'. We can then use →E on the second premise. So, by assuming 'P' we were able to prove 'R', so we apply the →I rule—discharging 'P'—and finish the proof. Putting all this together, we have:

$$\begin{array}{r|lll}
1 & P \to Q & & \\
2 & \underline{Q \to R} & & \\
3 & \quad \begin{array}{|l} \underline{P} \end{array} & & \\
4 & \quad \begin{array}{|l} Q \end{array} & & \to\text{E, 1, 3} \\
5 & \quad \begin{array}{|l} R \end{array} & & \to\text{E, 2, 4} \\
6 & P \to R & & \to\text{I, 3–5}
\end{array}$$

The subproof also doesn't need to start immediately. For example:

1	$(P \to Q) \land O$	
2	$Q \to R$	
3	$P \to Q$	$\land$E, 1
4	$\quad P$	
5	$\quad Q$	$\to$E, 3, 4
6	$\quad R$	$\to$E, 2, 5
7	$P \to R$	$\to$I, 4–6

19.2 Disjunction Elimination

The disjunction elimination rule also makes use of subproofs.

Suppose that Alice came to the party or Beth came to the party. What can you conclude? Not that Alice came to the party; it might be that Beth came to the party instead. Equally, not that Beth came to the party; for it might be that only Alice came. Disjunctions, just by themselves, are hard to work with.

But suppose that we could somehow show both of the following: first, that Alice coming to the party entails that it was fun: second, that Beth coming to the party entails that it was fun. Then if we know that Alice or Beth came to the party, then we know that either way, it was fun. This insight can be expressed in the following rule, which is our disjunction elimination ($\lor$E) rule:

$$
\begin{array}{r|l}
m & X \lor Y \\
i & \quad \begin{array}{|l} X \\ \hline \vdots \\ Z \end{array} \\
j & \\
k & \quad \begin{array}{|l} Y \\ \hline \vdots \\ Z \end{array} \\
l & \\
 & Z \qquad\qquad\qquad \lor\text{E}, m, i\text{–}j, k\text{–}l
\end{array}
$$

This is obviously a bit clunkier to write down than our previous rules, but the point is fairly simple. Suppose we have some disjunction, $X \lor Y$. Suppose we have two subproofs, showing us that Z follows from the assumption that X, and that Z follows from the assumption that Y. Then we can infer Z itself. As usual, there can be as many lines as you like between i and j, and as many lines as you like between k and l. Moreover, the subproofs and the disjunction can come in any order, and do not have to be adjacent.

Some examples might help illustrate this. Consider this argument:

$$(P \land Q) \lor (P \land R) \therefore P$$

A proof corresponding to this argument is:

1	$(P \land Q) \lor (P \land R)$	
2	$\quad$ $P \land Q$	
3	$\quad$ P	$\land$E, 2
4	$\quad$ $P \land R$	
5	$\quad$ P	$\land$E, 4
6	P	$\lor$E, 1, 2–3, 4–5

Consider the following brain teaser:[1]

> Three people are standing in a row looking at each other.
>
> 
>
> Alice Bob Charlie
>
> Alice is happy. Charlie is not happy. Is there someone who is happy who is looking at someone who is not happy?

...Think about it!
...Answer: Yes. Our Disjunction Elimination rule along with the Law of Excluded Middle allow us to show this. We can demonstrate this in the following argument, which we present in a pseudo-formal style.[2]

1	Bob is either happy or he's not happy	
2		Suppose Bob is happy
3		Then happy Bob is looking at not-happy Charlie
4		So someone who is happy is looking at someone who is not
5		Suppose Bob is not happy
6		Then happy Alice is looking at not-happy Bob
7		So someone who is happy is looking at someone who is not
8	Therefore, someone who is happy is looking at someone who is not.	

Are you convinced now that someone who is happy is looking at someone who is not happy? If not, find a friend and work through it together. Sometimes it can really help to try walking through the argument together.

Coming up with this sort of argument does just take that moment of inspiration to see how this argument will go (that's why it's a brain

[1]Originally by Hector Levesque.

[2]Though, actually, this is most naturally formulated as a validity claim of First Order Logic. We'll walk through the formal proof as formulated in First Order Logic in §36.2.

teaser). This is often the case with arguments that involve the law of excluded middle. We pick it out of nowhere and have to use our inspiration to see how it might be useful. But hopefully, with more examples you'll become familiar with cases where it might be of use. A strategy that might help is: it's a backup option if everything else fails. If it doesn't look like there's any elimination rules to use on your premises or any introduction rules that can get you to your conclusion, then perhaps LEM is the way forwards.

One more example:

$$P \therefore (P \wedge D) \vee (P \wedge \neg D)$$

Here is a proof corresponding with the argument:

1	P	
2	$D \vee \neg D$	LEM
3	$\quad D$	
4	$\quad P \wedge D$	$\wedge$I, 1, 3
5	$\quad (P \wedge D) \vee (P \wedge \neg D)$	$\vee$I, 4
6	$\quad \neg D$	
7	$\quad P \wedge \neg D$	$\wedge$I, 1, 6
8	$\quad (P \wedge D) \vee (P \wedge \neg D)$	$\vee$I, 7
9	$(P \wedge D) \vee (P \wedge \neg D)$	$\vee$E, 2, 3–5, 6–8

19.3 Negation

Negation Introduction

If assuming something leads you to a contradiction, then the assumption must be wrong. This thought motivates the following rule:

m	$\quad \vert\ X$	
	$\quad \vert\ \vdots$	
k	$\quad \vert\ Y \wedge \neg Y$	
	$\neg X$	$\neg$I, m–k

It does not matter whether in the subproof you first derive Y and then derive $\neg Y$ or whether you first derive $\neg Y$ and the derive Y. It is only important that in the subproof you derive both Y and $\neg Y$.

To see this in practice, and interacting with negation, consider this proof:

1	D	
2	$\quad \vert\ \neg D$	
3	$\quad \vert\ D \wedge \neg D$	$\wedge$I, 1, 2
4	$\neg\neg D$	$\neg$I, 2–3

Negation Elimination

The negation elimination rule is quite similar. We motivate the rule by the following thought: If assuming that something is false leads you to a contradiction, then that assumption must be wrong — and so that 'something' must in fact be true. In other words, if we assume $\neg X$ and derive a contradiction, then we can conclude X: we eliminate the negation.

m	$\quad \vert\ \neg X$	
	$\quad \vert\ \vdots$	
k	$\quad \vert\ Y \wedge \neg Y$	
	X	$\neg$E, m–k

Formally, the rule is very similar to ¬I, but they are different.[3] They switch X and $\neg X$. ¬I introduced a negation, but ¬E eliminates a negation.

Sometimes this is called the rule of *indirect proof*, *reductio ad absurdum* or *proof by contradiction*. It allows us to prove something by showing that its negation leads to contradiction. This technique is very common in mathematics.[4]

To see this rule in action, here is an example proof using it:

1	$\neg\neg D$	
2	$\quad\vert\ \neg D$	
3	$\quad\vert\ \neg D \wedge \neg\neg D$	∧I, 1, 2
4	D	¬E, 2–3

Here X is D and Y is $\neg D$.

This rule also allows us to prove that anything follows from a contradiction. For example:

1	A	
2	$\neg A$	
3	$\quad\vert\ \neg B$	
4	$\quad\vert\ A \wedge \neg A$	∧I, 1, 2
5	B	¬E, 3–4

19.4 Additional assumptions and subproofs

The rules we have just seen involve the idea of making additional assumptions. These need to be handled with some care.

[3] There are logicians who accept ¬I but reject ¬E. They are called "intuitionists." Intuitionists reject the law of excluded middle. They don't buy our basic assumption that every sentence has one of two truth values, true or false. For intuitionists, X and $\neg\neg X$ are not equivalent.

[4] For example it is used to show there are infinitely many prime numbers. We assume, for contradiction, that there are not infinitely many primes, so there must be a biggest prime number. We can then use this assumption to construct a further prime number which is larger than the biggest prime number. This is a contradiction. We can thus conclude that there are infinitely many primes.

Consider this proof:

$$
\begin{array}{r|l l}
1 & A & \\
\hline
2 & \quad \begin{array}{|l} B \\ \hline B \end{array} & \\
3 & & \text{R, 2} \\
4 & B \rightarrow B & \rightarrow\text{I, 2–3}
\end{array}
$$

This is perfectly in keeping with the rules we have laid down already, and it should not seem particularly strange. Since '$B \rightarrow B$' is a tautology, no particular premises should be required to prove it.

But suppose we now tried to continue the proof as follows:

$$
\begin{array}{r|l l}
1 & A & \\
\hline
2 & \quad \begin{array}{|l} B \\ \hline B \end{array} & \\
3 & & \text{R, 2} \\
4 & B \rightarrow B & \rightarrow\text{I, 2–3} \\
5 & B & \text{naughty attempt to invoke } \rightarrow\text{E, 4, 3}
\end{array}
$$

If we were allowed to do this, it would be a disaster. It would allow us to prove any atomic sentence letter from any other atomic sentence letter. However, if you tell me that Anne is fast (symbolized by 'A'), we shouldn't be able to conclude that Queen Boudica stood twenty-feet tall (symbolized by 'B')! We must be prohibited from doing this, but how are we to implement the prohibition?

We can describe the process of making an additional assumption as one of performing a *subproof*: a subsidiary 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 the assumption upon which the subproof will be based. A subproof can be thought of as essentially posing this question: *what could we show, if we also make this additional assumption?*

When we are working within the subproof, we can refer to the additional assumption that we made in introducing the subproof, and to anything that we obtained from our original assumptions. (After all, those original assumptions are still in effect.) At some point though, we will want to stop working with the additional assumption: we will want to return from the subproof to the main proof. To indicate that we have returned to the main proof, the vertical line for the subproof comes to an end. At this point, we say that the subproof is closed:

> A subproof is CLOSED when the vertical line for the subproof comes to an end. At that point we say the assumption has been DISCHARGED

We typically do this when we use one of our rules that involve subproofs, such as $\rightarrow$I. We introduced the assumption X to allow us to conclude Y; and this reasoning allows us to close the subproof and conclude $X \rightarrow Y$, which no longer relies on the assumption X. Having closed a subproof, we have set aside the additional assumption, so it will be illegitimate to draw upon anything that depends upon that additional assumption. Thus we stipulate:

> Any rule whose citation requires mentioning individual lines can mention any earlier lines, *except* for those lines which occur within a closed subproof.
> Put another way: you cannot refer back to anything that was obtained using discharged assumptions

This stipulation rules out the disastrous attempted proof above. The rule of $\rightarrow$E requires that we cite two individual lines from earlier in the proof. In the purported proof, above, one of these lines (namely, line 4) occurs within a subproof that has (by line 6) been closed. This is illegitimate.

Subproofs, then, allow us to think about what we could show, if we made additional assumptions. The point to take away from this is not surprising—in the course of a proof, we have to keep very careful track of what assumptions we are making, at any given moment. Our proof system does this very graphically. (Indeed, that's precisely why we have chosen to use *this* proof system.)

Once we have started thinking about what we can show by making additional assumptions, nothing stops us from posing the question of what we could show if we were to make *even more* assumptions. This might motivate us to introduce a subproof within a subproof. Here is an example which only uses the rules of proof that we have considered so far:

1	A	
2	$\quad B$	
3	$\quad\quad C$	
4	$\quad\quad A \wedge B$	$\wedge$I, 1, 2
5	$\quad C \rightarrow (A \wedge B)$	$\rightarrow$I, 3–4
6	$B \rightarrow (C \rightarrow (A \wedge B))$	$\rightarrow$I, 2–5

Notice that the citation on line 4 refers back to the initial assumption (on line 1) and an assumption of a subproof (on line 2). This is perfectly in order, since neither assumption has been discharged at the time (i.e. by line 4).

Again, though, we need to keep careful track of what we are assuming at any given moment. Suppose we tried to continue the proof as follows:

1	A	
2	$\quad B$	
3	$\quad\quad C$	
4	$\quad\quad A \wedge B$	$\wedge$I, 1, 2
5	$\quad C \rightarrow (A \wedge B)$	$\rightarrow$I, 3–4
6	$B \rightarrow (C \rightarrow (A \wedge B))$	$\rightarrow$I, 2–5
7	$C \rightarrow (A \wedge B)$	naughty attempt to invoke $\rightarrow$I, 3–4

This would be awful. If we tell you that Anne is smart, you should not be able to infer that, if Cath is smart (symbolized by 'C') then *both* Anne is smart and Queen Boudica stood 20-feet tall! But this is just what such a proof would suggest, if it were permissible.

The essential problem is that the subproof that began with the assumption 'C' depended crucially on the fact that we had assumed 'B' on line 2. By line 6, we have *discharged* the assumption 'B': we have stopped asking ourselves what we could show, if we also assumed 'B'. So it is simply cheating, to try to help ourselves (on line 7) to the subproof that began with the assumption 'C'. Thus we stipulate, much as before:

Any rule whose citation requires mentioning an entire subproof can mention any earlier subproof, *except* for those subproofs which occur within some *other* closed subproof.

The attempted disastrous proof violates this stipulation. The subproof of lines 3–4 occurs within a subproof that ends on line 5. So it cannot be invoked in line 7.

It is always permissible to open a subproof with any assumption. However, 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 obtain a conditional by →I, for instance, you must assume the antecedent of the conditional in a subproof.

Equally, it is always permissible to close a subproof and discharge its assumptions. However, it will not be helpful to do so until you have reached something useful.

19.5 Proof Strategies

These are all of the basic rules for the proof system for TFL. For ease of reference, they're listed again in appendix B.

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 obtain 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 get to this new goal.

For example: If your conclusion is a conditional $X \rightarrow Y$, plan to use the →I rule. This requires starting a subproof in which you assume X. The subproof ought to end with Y. So, what can you do to get Y?

Work forwards from what you have. When you are starting a proof, look at the premises; later, look at the sentences that you have obtained so far. Think about the elimination rules for the main operators of these sentences. These will often 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.

Try proceeding indirectly. If you cannot find a way to show X directly, try starting by assuming $\neg X$. If a contradiction follows, then you will be able to obtain X by $\neg$E.

Law of Excluded Middle. If you're hitting a blank, try seeing if there's some instance of LEM that might help you. These arguments do just need some inspiration to see the instance of LEM that'll be helpful.

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

If the argument is actually valid (which is defined using truth-tables) there will be a proof of it somehow…

Practice exercises

A. The following three proofs are missing their citations (rule and line numbers). Add them, to turn them into *bona fide* proofs. Additionally, write down the argument that corresponds to each proof.

1	$P \wedge S$	
2	$\underline{S \to R}$	
3	S	$\wedge$E, 1
4	R	$\to$E, 2, 3
5	$R \vee E$	$\vee$I, 4

1	$\underline{A \to D}$	
2	$\quad \vert\ \underline{A \wedge B}$	
3	$\quad \vert\ A$	$\wedge$E, 2
4	$\quad \vert\ D$	$\to$E, 1, 3
5	$\quad \vert\ D \vee E$	$\vee$I, 4
6	$(A \wedge B) \to (D \vee E)$	$\to$I, 2–5

CHAPTER 20

Proofs and Validity

The system of rules we have set up is not just a game. It helps us understand the validity of arguments.

An argument $X_1, X_2, \ldots, X_n \therefore Y$ may have a proof in the system of natural deduction. Such a proof may look something like:

$$
\begin{array}{r|l}
1 & X_1 \\
2 & X_2 \\
n & X_n \\
\hline
 & \vdots \\
 & Y
\end{array}
$$

That is, it will start with the premises as assumptions, and proceed following the rules we have given and finishing with the conclusion. It might also have subproofs along the way, something like:

$$
\begin{array}{r|l}
1 & X_1 \\
2 & X_2 \\
n & \underline{X_n} \\
 & \vdots \\
m & \quad \begin{array}{|l} \underline{\text{Some additional assumption}} \\ \vdots \\ \begin{array}{|l} \underline{\text{Maybe some more assumptions}} \\ \vdots \end{array} \\ \vdots \end{array} \\
 & \vdots \\
 & Y
\end{array}
$$

(The second m labels the line "Maybe some more assumptions".)

But any subproofs need to have been closed by the time we get to Y. So, for example if we gave

$$
\begin{array}{r|lll}
1 & \underline{A} & & \\
2 & & \begin{array}{|l} \underline{B} \end{array} & \\
3 & & \begin{array}{|l} B \end{array} & \text{R, 2}
\end{array}
$$

this does not count as a proof corresponding to the argument $A \therefore B$. But

$$
\begin{array}{r|lll}
1 & \underline{A} & & \\
2 & & \begin{array}{|l} \underline{B} \end{array} & \\
3 & & \begin{array}{|l} B \end{array} & \text{R, 2} \\
4 & B \rightarrow B & & \rightarrow\text{I, 2–3}
\end{array}
$$

counts as a proof corresponding to $A \therefore B \rightarrow B$ as the subproof has been closed on line 4.

So we have now said when we have a proof in our natural deduction system that corresponds to a particular argument. If we can find a proof then we know that the argument is valid.

> If there is a proof in natural deduction corresponding to the argument $X_1 \ldots X_n \therefore Y$, then this argument $X_1 \ldots X_n \therefore Y$ is valid.

This property of our proof system is called SOUNDNESS. It holds because we only chose rules that matched valid reasoning steps. Recall that $A \to B, B \therefore A$ is invalid. Had we added a rule such as

	$\vdots$	
n	$X \to Y$	
	$\vdots$	
m	Y	
	$\vdots$	
	X	Do not do this. Does not follow from, n, m

we would then have been able to construct a proof corresponding to the invalid argument $A \to B, B \therefore A$. We do not have such a rule in our system. All the rules we gave in our system will result in proofs of valid arguments.

We can actually strengthen the link between proofs corresponding to arguments and those argument's validity:

> If an argument is valid, then there is a proof of it in natural deduction.

This property of our proof system is called COMPLETENESS.

So for every valid argument there will be some proof. This doesn't mean it is always easy to come up with such a proof, but there will be one. Persist!

Practice exercises

A. Show that each of the following arguments are valid using natural deduction:

1. $J \to \neg J \therefore \neg J$

1	$J \to \neg J$	
2	$\quad J$	
3	$\quad \neg J$	$\to$E, 1, 2
4	$\quad J \wedge \neg J$	$\wedge$I, 2, 3
5	$\neg J$	$\neg$I, 2–4

2. $Q \to (Q \wedge \neg Q) \therefore \neg Q$

1	$Q \to (Q \wedge \neg Q)$	
2	$\quad Q$	
3	$\quad Q \wedge \neg Q$	$\to$E, 1, 2
4	$\quad \neg Q$	$\wedge$E, 3
5	$\quad Q \wedge \neg Q$	$\wedge$I, 2, 4
6	$\neg Q$	$\neg$I, 2–5

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

1	$A \to (B \to C)$	
2	$\quad A \wedge B$	
3	$\quad A$	$\wedge$E, 2
4	$\quad B$	$\wedge$E, 2
5	$\quad B \to C$	$\to$E, 1, 3
6	$\quad C$	$\to$E, 5, 4
7	$(A \wedge B) \to C$	$\to$I, 2–6

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

1	$(C \wedge D) \vee E$	
2	$\quad C \wedge D$	
3	$\quad D$	$\wedge$E, 2
4	$\quad E \vee D$	$\vee$I, 3
5	$\quad E$	
6	$\quad E \vee D$	$\vee$I, 5
7	$E \vee D$	$\vee$E, 1, 2–4, 5–6

5. $\neg F \to G, F \to H \therefore G \vee H$

1	$\neg F \to G$	
2	$F \to H$	
3	$F \vee \neg F$	LEM
4	$\quad F$	
5	$\quad H$	$\to$E, 2, 4
6	$\quad G \vee H$	$\vee$I, 5
7	$\quad \neg F$	
8	$\quad G$	$\to$E, 1, 7
9	$\quad G \vee H$	$\vee$I, 8
10	$G \vee H$	$\vee$E, 3, 4–6, 7–9

6. $(Z \wedge K) \vee (K \wedge M), K \to D \therefore D$

1	$(Z \wedge K) \vee (K \wedge M)$	
2	$K \to D$	
3	$\quad Z \wedge K$	
4	$\quad K$	$\wedge$E, 3
5	$\quad D$	$\to$E, 2, 4
6	$\quad K \wedge M$	
7	$\quad K$	$\wedge$E, 6
8	$\quad D$	$\to$E, 2, 7
9	D	$\vee$E, 1, 3–5, 6–8

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

1	$P \wedge (Q \vee R)$	
2	$P \to \neg R$	
3	P	$\wedge$E, 1
4	$Q \vee R$	$\wedge$E, 1
5	$\quad Q$	
6	$\quad Q \vee E$	$\vee$I, 5
7	$\quad R$	
8	$\quad \neg R$	$\to$E, 2, 3
9	$\quad\quad \neg(Q \vee E)$	
10	$\quad\quad R \wedge \neg R$	$\wedge$I, 7, 8
11	$\quad Q \vee E$	$\neg$E, 9–10
12	$Q \vee E$	$\vee$E, 4, 5–6, 7–11

8. $\neg(P \to Q) \therefore \neg Q$

1	$\mid \neg(P \to Q)$ $\underline{}$	
2	$\mid \mid Q$ $\underline{}$	
3	$\mid \mid \mid P$ $\underline{}$	
4	$\mid \mid \mid Q$	R, 2
5	$\mid \mid P \to Q$	$\to$I, 3–4
6	$\mid \mid (P \to Q) \wedge \neg(P \to Q)$	$\wedge$I, 5, 1
7	$\mid \neg Q$	$\neg$I, 2–6

9. $\neg(P \to Q) \therefore P$

1	$\mid \neg(P \to Q)$ $\underline{}$	
2	$\mid \mid \neg P$ $\underline{}$	
3	$\mid \mid \mid P$ $\underline{}$	
4	$\mid \mid \mid \mid \neg Q$ $\underline{}$	
5	$\mid \mid \mid \mid P \wedge \neg P$	$\wedge$I, 3, 2
6	$\mid \mid \mid Q$	$\neg$E, 4–5
7	$\mid \mid P \to Q$	$\to$I, 3–6
8	$\mid \mid (P \to Q) \wedge \neg(P \to Q)$	$\wedge$I, 7, 1
9	$\mid P$	$\neg$E, 2–8

CHAPTER 21

Derived rules for TFL

In §18, we introduced the basic rules of our proof system for TFL. In this section, we will add some additional rules to our system. These will make our system much easier to work with. (However, in §22 we will see that they are not strictly speaking *necessary*.)

21.1 Disjunctive syllogism

Here is a very natural argument form.

> Elizabeth is in Massachusetts or in DC. She is not in DC.
> So, she is in Massachusetts.

This inference pattern is called *disjunctive syllogism*. We add it to our proof system as follows:

$$
\begin{array}{r|ll}
m & X \vee Y & \\
n & \neg X & \\
 & Y & \text{DS}, m, n
\end{array}
$$

and

m	$X \vee Y$	
n	$\neg Y$	
	X	DS, m, n

As usual, the disjunction and the negation of one disjunct may occur in either order and need not be adjacent.

21.2 Modus tollens

Another useful pattern of inference is embodied in the following argument:

> If Mitt has won the election, then he is in the White House. He is not in the White House. So he has not won the election.

This inference pattern is called *modus tollens*. The corresponding rule is:

m	$X \rightarrow Y$	
n	$\neg Y$	
	$\neg X$	MT, m, n

As usual, the premises may occur in either order.

21.3 Double-negation

A sentence $\neg\neg X$ is always logically equivalent to X. We can add rules to our system that encode this idea: allowing us to immediately eliminate or introduce double negations:

m	$\neg\neg X$	
	X	DNE, m

m	X	
	$\neg\neg X$	DNI, m

That said, you should be aware that in ordinary language we can sometimes speak in a way that is similar to, but not quite, a double negation. Consider: 'Jane is not *un*happy'. Arguably, one cannot infer 'Jane is happy'. Perhaps the speaker is using this unusual indirect phrasing to draw attention to the possible difference between 'unhappy' and 'not unhappy'. Perhaps what they mean to suggest is that 'Jane is in a state of profound indifference'. Here, then, 'Jane is unhappy' should not be thought of as equivalent to 'It is not the case that Jane is happy', and it should not be symbolised as $\neg H$ but should rather be a separate atomic sentence. So 'Jane is not unhappy' is not then seen as a double negation.

21.4 Explosion

From a contradiction anything follows. This is called the rule of explosion. It is also sometimes called *ex falso quod libet*.

m	X	
n	$\neg X$	
	Y	Explosion, m, n

21.5 De Morgan Rules

Our final additional rules are called De Morgan's Laws. (These are named after Augustus De Morgan.) The shape of the rules should be familiar from truth tables.

The first De Morgan rule is:

m	$\neg(X \wedge Y)$	
	$\neg X \vee \neg Y$	DeM, m

The second De Morgan is the reverse of the first:

m	$\neg X \vee \neg Y$	
	$\neg(X \wedge Y)$	DeM, m

The third De Morgan rule is the *dual* of the first:

m	$\neg(X \vee Y)$	
	$\neg X \wedge \neg Y$	DeM, m

And the fourth is the reverse of the third:

m	$\neg X \wedge \neg Y$	
	$\neg(X \vee Y)$	DeM, m

There are many more rules one could add to the system as derived rules. But these are all the ones we'll introduce.

Practice exercises

A. The following proofs are missing their citations (rule and line numbers). Add them wherever they are required:

1	$W \to \neg B$	
2	$A \wedge W$	
3	$B \vee (J \wedge K)$	
4	W	
5	$\neg B$	
6	$J \wedge K$	
7	K	

1	$\neg O \to L$	
2	$L \vee \neg O$	
3		$\neg L$
4		$\neg O$
5		L
6	$\neg\neg L$	
7	L	

1	$Z \to (C \wedge \neg N)$	
2	$\neg Z \to (N \wedge \neg C)$	
3	$Z \vee \neg Z$	
4		Z
5		$C \wedge \neg N$
6		C
7		$N \vee C$
8	$Z \to (N \vee C)$	
9		$\neg Z$
10		$N \wedge \neg C$
11		N
12		$N \vee C$
13	$\neg Z \to (N \vee C)$	
14	$N \vee C$	

B. Give a proof for each of these arguments:

1. $E \vee F$, $F \vee G$, $\neg F \therefore E \wedge G$
2. $M \vee (N \to M) \therefore \neg M \to \neg N$
3. $(M \vee N) \wedge (O \vee P)$, $N \to P$, $\neg P \therefore M \wedge O$
4. $A \wedge B) \vee (A \wedge C)$, $\neg(A \wedge D)$, $D \vee M \therefore M$

CHAPTER 22

Derived rules

In this section, we will see why we introduced the rules of our proof system in two separate batches. In particular, we want to show that the additional rules of §21 are not strictly speaking necessary, but can be derived from the basic rules of §18.

22.1 Derivation of Disjunctive syllogism

Suppose that you are in a proof, and you have something of this form:

$$\begin{array}{r|l} m & X \vee Y \\ n & \neg X \end{array}$$

You now want, on line k, to prove Y. You can do this with the rule of DS, introduced in §21, but equally well, you can do this with the *basic* rules of §18:

m	$X \vee Y$	
n	$\neg X$	
k	$\quad X$	
$k+1$	$\quad\quad \neg Y$	
$k+2$	$\quad\quad X$	R, k
$k+3$	$\quad\quad \neg X$	R, n
$k+4$	$\quad Y$	PbC, $k+1$–$k+3$
$k+5$	$\quad Y$	
$k+6$	$\quad Y$	R, $k+5$
$k+7$	Y	$\vee$E, m, k–$k+4$, $k+5$–$k+6$

To be clear: this is not a proof. Rather, it is a proof *scheme*. (This is why we use letters like m and k to label the lines of the proof rather than numbers.) Whatever sentences of TFL we plugged in for 'X' or 'Y', and whatever lines we were working on, we could produce a bona fide proof. So you can think of this as a recipe for producing proofs.

Indeed, it is a recipe which shows us that, anything we can prove using the rule DS, we can prove (with a few more lines) using just the other rules of §18.

22.2 Derivation of Modus Tollens

Suppose in the course of you proof you already have $X \rightarrow Y$, say on line m, and $\neg Y$ on line n. At some later line, k, you want to get $\neg X$. You can do this with the rule of Modus Tollens (MT), introduced in §21. But you could also do this with the *basic* rules of §18:

m	$X \to Y$	
n	$\neg Y$	
k	$\quad X$	
$k+1$	$\quad Y$	$\to$E, m, k
$k+2$	$\quad Y \wedge \neg Y$	$\wedge$I, $k+1$, n
$k+3$	$\neg X$	$\neg$I, k–$k+2$

Again, the rule of MT can be derived from the *basic* rules of §18.

22.3 Derivation of Double-negation rules

Consider the following deduction schema:

m	X	
j	$\quad \neg X$	
$j+1$	$\quad X \wedge \neg X$	$\wedge$I, m, j
$j+2$	$\neg\neg X$	$\neg$I, j–$j+1$

and

m	$\neg\neg X$	
j	$\quad \neg X$	
$j+1$	$\quad \neg X \wedge \neg\neg X$	$\wedge$I, m, j
$j+2$	X	$\neg$E, j–$j+1$

So again, we can derive the double negations rules from the *basic* rules of §18.

22.4 Derivation of Explosion

Here is a demonstration of how we could derive explosion using the basic rules:

m	X	
n	$\neg X$	
$n+1$	$\quad \vert\ \neg Y$	
$n+2$	$\quad \vert\ X \wedge \neg X$	$\wedge$I, m, n
	Y	$\neg$E, $n+1$–$n+2$

22.5 Derivation of De Morgan rules

Here is a demonstration of how we could derive the first De Morgan rule:

m	$\neg(X \wedge Y)$	
j	$X \vee \neg X$	LEM
k	$\quad \vert\ X$	
$k+1$	$\quad \vert\ \quad \vert\ Y$	
$k+2$	$\quad \vert\ \quad \vert\ X \wedge Y$	$\wedge$I, k, $k+1$
$k+3$	$\quad \vert\ \quad \vert\ (X \wedge Y) \wedge \neg(X \wedge Y)$	$\wedge$I, $k+2$, m
$k+4$	$\quad \vert\ \neg Y$	$\neg$I, $k+1$–$k+3$
$k+5$	$\quad \vert\ \neg X \vee \neg Y$	$\vee$I, $k+4$
$k+6$	$\quad \vert\ \neg X$	
$k+7$	$\quad \vert\ \neg X \vee \neg Y$	$\vee$I, $k+6$
$k+8$	$\neg X \vee \neg Y$	$\vee$E, j, k–$k+5$, $k+6$–$k+7$

Here is a demonstration of how we could derive the second De Morgan rule:

m	$\neg X \vee \neg Y$	
$m+1$	$\quad \neg X$	
k	$\quad\quad X \wedge Y$	
$k+1$	$\quad\quad X$	$\wedge$E, k
$k+2$	$\quad\quad \neg X$	R, $m+1$
$k+3$	$\quad \neg(X \wedge Y)$	$\neg$I, k–$k+2$
$k+4$	$\quad \neg Y$	
l	$\quad\quad X \wedge Y$	
$l+1$	$\quad\quad Y$	$\wedge$E, k
$l+2$	$\quad\quad Y \wedge \neg Y$	$\wedge$I, $l+1$, $k+4$
$l+3$	$\quad \neg(X \wedge Y)$	$\neg$I, l–$l+2$
$l+4$	$\neg(X \wedge Y)$	$\vee$E, m, $m+1$–$k+3$, $k+4$–$l+3$

Similar demonstrations can be offered explaining how we could derive the third and fourth De Morgan rules. These are left as exercises.

Practice exercises

A. Provide proof schemes that justify the addition of the third and fourth De Morgan rules as derived rules.

B. The proofs you offered in response to the practice exercises of §§21–24 used derived rules. Replace the use of derived rules, in such proofs, with only basic rules. You will find some 'repetition' in the resulting proofs; in such cases, offer a streamlined proof using only basic rules. (This will give you a sense, both of the power of derived rules, and of how all the rules interact.)

CHAPTER 23

Soundness and completeness

(Non-examinable)A very important result:

> A TFL argument is valid if and only if it can be given a proof in this natural deduction system.

In this chapter, we explain a bit more about how such an argument would go. Soundness is proved in more detail in §??.

Logical Consequence

For this chapter we will make use of a further symbols:

We use the symbol $\vDash$ as shorthand for logical consequence, that is, instead of saying that the TFL-sentence Y. is a logical consequence of the TFL sentences $X_1, X_2, \ldots$ and X_n, we will abbreviate this by:

$$X_1, X_2, \ldots, X_n \vDash Y$$

The symbol '$\vDash$' is known as *the double-turnstile*, since it looks like a turnstile with two horizontal beams.

Let us be clear be clear. '$\vDash$' is not a symbol of TFL. Rather, it is a symbol of our metalanguage, augmented English (recall the difference between object language and metalanguage from §5). So the metalanguage sentence:

- $P, P \rightarrow Q \vDash Q$

is just an abbreviation of the sentence:

- The TFL sentences 'Q' is a logical consequence of 'P' and '$P \rightarrow Q$'.

Note that there is no limit on the number of TFL sentences that can be mentioned before the symbol '$\vDash$'. Indeed, we can even consider the limiting case where there is no sentence (this should ring familiar from §8.1):

$$\vDash Y$$

Proof

The following expression:

$$X_1, X_2, \ldots, X_n \vdash Y$$

means that there is some proof which starts with assumptions among $X_1, X_2, \ldots, X_n$ and ends with Y (and contains no undischarged assumptions other than those we started with). Derivatively, we will write:

$$\vdash X$$

to mean that there is a proof of X with no assumptions.

The symbol '$\vdash$' is called the *single turnstile.* We want to emphasize that this is not the double turnstile symbol ('$\vDash$') that we introduced for 'logical consequence'. The single turnstile, '$\vdash$', concerns the existence of proofs; the double turnstile, '$\vDash$', concerns the existence of valuations (or interpretations, when used for FOL). *They are very different notions.*

23.1 Their equivalence

However, it turns out that they are equivalent. That is:

$$X_1, X_2, \ldots, X_n \vdash Y$$
if and only if
$$X_1, X_2, \ldots, X_n \vDash Y$$

A full proof here goes well beyond the scope of this book. However, we can sketch what it would be like.

Soundness

This argument from $\vdash$ to $\vDash$ is the problem of SOUNDNESS. A proof system is SOUND if there are no derivations of arguments that can be shown invalid by truth tables. Demonstrating that the proof system is sound would require showing that *any* possible proof is the proof of a valid argument. It would not be enough simply to succeed when trying to prove many valid arguments and to fail when trying to prove invalid ones.

The proof that we will sketch depends on the fact that we initially defined a sentence of TFL using a inductive definition (see p. 34). We could have also used inductive definitions to define a proper proof in TFL and a proper truth table. (Although we didn't.) If we had these definitions, we could then use a *inductive proof* to show the soundness of TFL. A inductive proof works the same way as a inductive definition. With the inductive definition, we identified a group of base elements that were stipulated to be examples of the thing we were trying to define. In the case of a TFL sentence, the base class was the set of sentence letters A, B, C, ... We just announced that these were sentences. The second step of a inductive definition is to say that anything that is built up from your base class using certain rules also counts as an example of the thing you are defining. In the case of a definition of a sentence, the rules corresponded to the five sentential connectives (see p. 34). Once you have established a inductive definition, you can use that definition to show that all the members of the class you have defined have a certain property. You simply prove that the property is true of the members of the base class, and then you prove that the rules for extending the base class don't change the property. This is what it means to give a inductive proof.

Even though we don't have a inductive definition of a proof in TFL, we can sketch how a inductive proof of the soundness of TFL would go. Imagine a base class of one-line proofs, one for each of our basic rules of inference. The members of this class would look like this $X, Y \vdash X \wedge Y$; $X \wedge Y \vdash X$; $X \vee Y, \neg X \vdash Y$ ... etc. Since some rules have a couple different forms, we would have to have add some members to this base class, for instance $X \wedge Y \vdash Y$ Notice that these are all statements in the metalanguage. The proof that TFL is sound is not a part of TFL, because TFL does not have the power to talk about itself.

You can use truth tables to prove to yourself that each of these one-line proofs in this base class is valid$_\vDash$. For instance the proof $X, Y \vdash X \wedge Y$ corresponds to a truth table that shows $X, Y \vDash X \wedge Y$ This establishes the first part of our inductive proof.

The next step is to show that adding lines to any proof will never change a valid$_\vDash$ proof into an invalid$_\vDash$ one. We would need to do this for each of our basic rules of inference. So, for instance, for $\wedge$I we need to show that for any proof $X_1, \ldots, X_n \vdash Y$ adding a line where we use $\wedge$I to infer $Z \wedge V$, where $Z \wedge V$ can be legitimately inferred from $X_1, \ldots, X_n$, Y, would not change a valid proof into an invalid proof. But wait, if we can legitimately derive $Z \wedge V$ from these premises, then Z and V must be already available in the proof. They are either already among $X_1, \ldots, X_n$, B, or can be legitimately derived from them. As such, any truth table line in which the premises are true must be a truth table line in which Z and V are true. According to the characteristic truth table for $\wedge$, this means that $Z \wedge V$ is also true on that line. Therefore, $Z \wedge V$ validly follows from the premises. This means that using the $\wedge$E rule to extend a valid proof produces another valid proof.

In order to show that the proof system is sound, we would need to show this for the other inference rules. Since the derived rules are consequences of the basic rules, it would suffice to provide similar arguments for the other basic rules. This tedious exercise falls beyond the scope of this book.

So we have shown that $X \vdash Y$ implies $X \vDash Y$. What about the other direction, that is why think that *every* argument that can be shown valid using truth tables can also be proven using a derivation.

Completeness

This is the problem of completeness. A proof system has the property of COMPLETENESS if and only if there is a derivation of every semantically valid argument. Proving that a system is complete is generally harder than proving that it is sound. Proving that a system is sound amounts to showing that all of the rules of your proof system work the way they are supposed to. Showing that a system is complete means showing that you have included *all* the rules you need, that you haven't left any out. Showing this is beyond the scope of this book. The important point is that, happily, the proof system for TFL is both sound and complete. This is not the case for all proof systems or all formal languages. Because it is true of TFL, we can choose to give proofs or give truth tables—whichever is easier for the task at hand.

23.2 Other Semantic and Proof Theoretic Notions

Now we know that the proof theoretic and truth-table methods are equivalent, we can use them interchangable depending on which is more useful.

We can also use proof theoretic methods for determining other logical notions, such as being consistent. We summarise how one would define them in 23.1

In fact, we can give general guidelines about when it's best to give proofs and when it is best to give truth tables. We do this in 23.2:

Practice exercises

A. Use either a derivation or a truth table for each of the following.

1. Show that $A \rightarrow [((B \wedge C) \vee D) \rightarrow A]$ is a tautology.
2. Show that $A \rightarrow (A \rightarrow B)$ is not a tautology
3. Show that the sentence $A \rightarrow \neg A$ is not a contradiction.
4. Show that the sentence $A \leftrightarrow \neg A$ is a contradiction.
5. Show that the sentence $\neg(W \rightarrow (J \vee J))$ is contingent
6. Show that the sentence $\neg(X \vee (Y \vee Z)) \vee (X \vee (Y \vee Z))$ is not contingent
7. Show that the sentence $B \rightarrow \neg S$ is equivalent to the sentence $\neg\neg B \rightarrow \neg S$
8. Show that the sentence $\neg(X \vee O)$ is not equivalent to the sentence $X \wedge O$
9. Show that the sentences $\neg(A \vee B)$, C, $C \rightarrow A$ are jointly inconsistent.
10. Show that the sentences $\neg(A \vee B)$, $\neg B$, $B \rightarrow A$ are jointly consistent
11. Show that $\neg(A \vee (B \vee C)) \therefore \neg C$ is valid.
12. Show that $\neg(A \wedge (B \vee C)) \therefore \neg C$ is invalid.

B. Use either a derivation or a truth table for each of the following.

Concept	**Truth table (semantic) definition**	**Proof-theoretic (syntactic) definition**
Tautology	A sentence whose truth table only has Ts under the main connective	A sentence that can be derived without any premises.
Contradiction	A sentence whose truth table only has Fs under the main connective	A sentence whose negation can be derived without any premises
Contingent sentence	A sentence whose truth table contains both Ts and Fs under the main connective	A sentence that is not a theorem or contradiction
Equivalent sentences	The columns under the main connectives are identical.	The sentences can be derived from each other
Inconsistent sentences	Sentences which do not have a single line in their truth table where they are all true.	Sentences from which one can derive a contradiction
Consistent sentences	Sentences which have at least one line in their truth table where they are all true.	Sentences which are not inconsistent
Valid argument	An argument whose truth table has no lines where there are all Ts under main connectives for the premises and an F under the main connective for the conclusion.	An argument where one can derive the conclusion from the premises

Table 23.1: Two ways to define logical concepts.

Logical property	To prove it present	To prove it absent
Being a tautology	Derive the sentence	Find the false line in the truth table for the sentence
Being a contradiction	Derive the negation of the sentence	Find the true line in the truth table for the sentence
Contingency	Find a false line and a true line in the truth table for the sentence	Prove the sentence or its negation
Equivalence	Derive each sentence from the other	Find a line in the truth tables for the sentence where they have different values
Consistency	Find a line in truth table for the sentence where they all are true	Derive a contradiction from the sentences
Validity	Derive the conclusion from the premises	Find no line in the truth table where the premises are true and the conclusion false.

Table 23.2: When to provide a truth table and when to provide a proof.

1. Show that $A \rightarrow (B \rightarrow A)$ is a tautology
2. Show that $\neg(((N \leftrightarrow Q) \lor Q) \lor N)$ is not a tautology
3. Show that $Z \lor (\neg Z \leftrightarrow Z)$ is contingent
4. show that $(L \leftrightarrow ((N \rightarrow N) \rightarrow L)) \lor H$ is not contingent
5. Show that $(A \leftrightarrow A) \land (B \land \neg B)$ is a contradiction
6. Show that $(B \leftrightarrow (C \lor B))$ is not a contradiction.
7. Show that $((\neg X \leftrightarrow X) \lor X)$ is equivalent to X
8. Show that $F \land (K \land R)$ is not equivalent to $(F \leftrightarrow (K \leftrightarrow R))$

9. Show that the sentences $\neg(W \rightarrow W)$, $(W \leftrightarrow W) \wedge W$, $E \vee (W \rightarrow \neg(E \wedge W))$ are inconsistent.
10. Show that the sentences $\neg R \vee C$, $(C \wedge R) \rightarrow \neg R$, $(\neg(R \vee R) \rightarrow R)$ are consistent.
11. Show that $\neg\neg(C \leftrightarrow \neg C), ((G \vee C) \vee G) \therefore ((G \rightarrow C) \wedge G)$ is valid.
12. Show that $\neg\neg L, (C \rightarrow \neg L) \rightarrow C) \therefore \neg C$ is invalid.

CHAPTER 24

Proof-theoretic concepts

Armed with our '$\vdash$' symbol, we can introduce some new terminology.

> X is a THEOREM iff $\vdash X$

To illustrate this, suppose we want to prove that '$\neg(A \wedge \neg A)$' is a theorem. So we must start our proof without *any* assumptions. However, since we want to prove a sentence whose main logical operator is a negation, we will want to immediately begin a subproof, with the additional assumption '$A \wedge \neg A$', and show that this leads to contradiction. All told, then, the proof looks like this:

1	$\quad \mid\; \underline{A \wedge \neg A}$	
2	$\quad \mid\; A$	$\wedge$E, 1
3	$\quad \mid\; \neg A$	$\wedge$E, 1
4	$\neg(A \wedge \neg A)$	$\neg$I, 1–3

We have therefore proved '$\neg(A \wedge \neg A)$' on no (undischarged) assumptions. This particular theorem is an instance of what is sometimes called *the Law of Non-Contradiction.*

To show that something is a theorem, you just have to find a suitable proof. It is typically much harder to show that something is *not* a theorem. To do 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. Perhaps you just didn't try hard enough.

Here is another new bit of terminology:

> Two sentences X and Y are PROVABLY EQUIVALENT iff each can be proved from the other; i.e., both $X \vdash Y$ and $Y \vdash X$.

As in the case of showing that a sentence is a theorem, 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 is just as hard as showing that a sentence is not a theorem.

Here is a third, related, bit of terminology:

> The sentences $X_1, X_2, \ldots, X_n$ are PROVABLY INCONSISTENT iff a contradiction can be proved from them, i.e. $X_1, X_2, \ldots, X_n \vdash Y \wedge \neg Y$ for some Y. If they are not INCONSISTENT, we call them PROVABLY CONSISTENT.

It is easy to show that some sentences are provably inconsistent: you just need to prove a contradiction from assuming all the sentences. Showing that some sentences are not provably inconsistent is much harder. It would require more than just providing a proof or two; it would require showing that no proof of a certain kind is *possible*.

This table summarises whether one or two proofs suffice, or whether we must reason about all possible proofs.

	Yes	**No**
theorem?	one proof	all possible proofs
inconsistent?	one proof	all possible proofs
equivalent?	two proofs	all possible proofs
consistent?	all possible proofs	one proof

Practice exercises

A. Show that each of the following sentences is a theorem:

1. $O \to O$
2. $N \lor \neg N$
3. $J \leftrightarrow [J \lor (L \land \neg L)]$
4. $((A \to B) \to A) \to A$

B. Provide proofs to show each of the following:

1. $C \to (E \land G), \neg C \to G \vdash G$
2. $M \land (\neg N \to \neg M) \vdash (N \land M) \lor \neg M$
3. $(Z \land K) \leftrightarrow (Y \land M), D \land (D \to M) \vdash Y \to Z$
4. $(W \lor X) \lor (Y \lor Z), X \to Y, \neg Z \vdash W \lor Y$

C. Show that each of the following pairs of sentences are provably equivalent:

1. $R \leftrightarrow E$, $E \leftrightarrow R$
2. G, $\neg\neg\neg\neg G$
3. $T \to S$, $\neg S \to \neg T$
4. $U \to I$, $\neg(U \land \neg I)$
5. $\neg(C \to D), C \land \neg D$
6. $\neg G \leftrightarrow H$, $\neg(G \leftrightarrow H)$

D. If you know that $X \vdash Y$, what can you say about $(X \land Z) \vdash Y$? What about $(X \lor Z) \vdash Y$? Explain your answers.

E. In this chapter, we claimed that it is just as hard to show that two sentences are not provably equivalent, as it is to show that a sentence is not a theorem. Why did we claim this? (*Hint*: think of a sentence that would be a theorem iff X and Y were provably equivalent.)

PART V

First-order logic

CHAPTER 25

Building blocks of FOL

We have been studying arguments, and in particular their validity. In Part we gave a strategy for checking the validity of an argument by using TFL. That was:

1. Find the structure of the argument.
   Identify the premises and conclusion.

2. Symbolise the argument in TFL.

3. Check if the TFL argument is valid.

    - Using truth tables to look for a valuation providing a counter example. If there is no such valuation, then it is valid.
    - Or, use natural deduction to show that it is valid.

However, this allows you to conclude that the original English language argument is valid provided, if its best TFL-symbolization is. But what if the best TFL symbolisation is invalid? Consider the following arguments:

1. 1. Alice is a logician.
    2. All logicians wear funny hats.
    ∴ Therefore: Alice wears a funny hat.

2. 1. Everyone who loves Manchester United hates Manchester City.
   2. Manchester City is not hated by everyone.
   ∴ Therefore: there is at least one person who doesn't love Manchester United.

We can symbolise these in TFL (follow the strategy as in 103). Since we cannot paraphrase any of these sentences with 'and', 'if', 'or' or 'not', we simply have to use atomic sentences. We thus offer the symbolisation:

$$L, A \therefore H$$

with the symbolisation

L: Alice is a logician.
A: All logicians wear funny hats.
H: Alice wears a funny hat.

And for the second argument we would symbolise this as:

$$P, \neg Q \;\; \therefore R$$

using

P: Everyone who loves Manchester United hates Manchester City.
Q: Manchester City is hated by everyone.
R: There is at least one person who doesn't love Manchester United.

Both of these TFL arguments are invalid. But the original English arguments seem valid. Indeed, they seem valid independently of their subject matter, that is, independently of whether they are talking about Alice, logicians, funny hats or loving Manchester United. They seem to be valid in virtue of the argument form. And it is not that we have made a mistake while symbolizing the argument. The problem lies with TFL itself. The expressive power of TFL is not rich enough to explain why these English arguments are valid. TFL can recognise arguments that are valid because of their truth-functional structure, but these arguments are valid in virtue of something else. In particular, their validity seems to hinge on our understanding of 'all', 'everyone', and 'there is'.

We will introduce a new logical language that will allow us to capture the validity of these arguments. We will call this language *first-order logic*, or *FOL*. The details of FOL will be explained throughout this chapter.

25.1 Names and Predicates

Consider

> Alice is a logician.

In TFL we used an atomic sentence to represent this. In FOL we will break it into two components: a name and a predicate.

$$\overbrace{\text{Alice}}^{\text{Name}}\ \overbrace{\text{is a logician.}}^{\text{Predicate}}$$

A name picks out an individual. The name 'Alice' is picking out some particular person, Alice.

A predicate expresses a property, in this case the property of being a logician. The predicate is:

> _____ is a logician

In First Order Logic, FOL, we can symbolise these different components. We will use lower-case letters like $a, b, c \ldots$ for names (except x, y, z which are used for variables as we will later see), and upper case letters like $A, B, C, \ldots$ for predicates (except X, Y, Z, which are used for metavariables). We can also add numbered subscripts if needed, for example using d_{27} as a name, or H_{386} as a predicate.

Like in TFL, when symbolising we have to give a symbolisation key to specify how to interpret the predicates and names. In this case, we might give:

> a: Alice
> Lx: $\underline{\quad\quad}_x$ is a logician

and we can then symbolise 'Alice is a logician' as

$$La.$$

(We will say more about the "x" subscript later.)

Note that in FOL the name follows the predicate: we have to write it as La. The property of being a logician applies to Alice.

As in TFL our choice of which letter to use for our name or predicate doesn't matter. It would be equally good to give

> a: Alice
> Px: $\underline{\quad\quad}_x$ is a logician

And then symbolise ‘Alice is a logician’ as

$$Pa.$$

Let’s see some other example sentences which have this same form. Each of these sentences could similarly be symbolised as Pa, though the symbolisation key would have to change in each of these instances.

1. Rocky is strong
2. Joe Biden is a Democrat
3. Michael Palin is a member of Monty Python

In each of these cases the relevant symbolisation key would then be:

1
a: Rocky
Px: $\_\_\_\_\__x$ is strong

2
a: Joe Biden
Px: $\_\_\_\_\__x$ is a Democrat

3
a: Michael Palin
Px: $\_\_\_\_\__x$ is a member of Monty Python

Names don’t have to name people, for example we can also symbolise

4. The Tower of London is in England.

as Pa using the symbolisation key:

a: The Tower of London
Px: $\_\_\_\_\__x$ is in England

What is important, though, is that what we are symbolising as a name in FOL refers to a *specific* person, place, or thing.

Consider

5. Buses are red.

You might think that this has the same form and symbolise it as La with the symbolisation key:

a: Buses
Lx: $\_\_\_\_\__x$ is red

But this would be wrong. Do not do this. The reason is that ‘Buses’ does not refer to a specific thing, it refers to a great many objects.

25.2 Names, predicates and connectives

In FOL we will also make use of all of the tools from TFL. We can symbolise

6. Joe is happy and Katie is sad

Joe is happy and Katie is sad

Joe is happy **and** Katie is sad

$Hj \wedge Sk$

Joe is happy — Hj

Katie is sad — Sk

as

$$Hj \wedge Sk$$

with the symbolisation key:

Hx: $\underline{\qquad}_x$ is happy
Sx: $\underline{\qquad}_x$ is sad
j: Joe
k: Katie

To symbolise

7. Joe and Katie are happy

we observe that it can be naturally paraphrased as 'Joe is happy and Katie is happy' and thus symbolised as

$$Hj \wedge Hk$$

To symbolise

8. If Joe is happy, then Katie is too

we observe that it can be naturally paraphrased as 'If Joe is happy then Katie is happy' and thus symbolised as

$$Hj \rightarrow Hk$$

We can also symbolise more complex sentences, for example:

9. If Joe is not happy then Katie or Billy is sad.

If Joe is not happy then Katie or Billy is sad
If Joe is not happy then Katie or Billy is sad.
$\neg Hj \to (Sk \vee Sb)$

Joe is not happy
It is not the case that Joe is happy
$\neg Hj$

Joe is happy
Hj

Katie or Billy is sad
Katie is sad or Billy is sad
$Sk \vee Sb$

Katie is sad
Sk

Billy is sad
Sb

One final example. To symbolise:

10. Herbie is a red car

we might simply offer

$$Ah$$

using

Ax: _____$_x$ is a red car
h: Herbie

But it is more informative to observe that we can naturally paraphrase it as 'Herbie is red and Herbie is a car' so symbolise it as

$$Rh \wedge Ch$$

using

Rx: _____$_x$ is red
Cx: _____$_x$ is a car
h: Herbie

Since this latter symbolisation extracts more of the information from the original sentence, it is generally going to be better.

25.3 Many-placed predicates

All of the predicates that we have considered so far concern properties that objects might have. Those predicates have one gap in them, and to make a sentence, we simply need to slot in one term. They are ONE-PLACE predicates.

However, other predicates concern the *relation* between two things. Here are some examples of relational predicates in English:

_____ loves _____
_____ is to the left of _____
_____ is in debt to _____

These are TWO-PLACE predicates. They need to be filled in with two terms in order to make a sentence. They express a relationship between two objects.

Now there is a little foible with the above. We have used the same symbol, '_____', to indicate a gap formed by deleting a term from a sentence. However (as Frege emphasized), these are *different* gaps. To obtain a sentence, we can fill them in with the same term, but we can equally fill them in with different terms, and in various different orders. The following are all perfectly good sentences, and they all mean very different things:

Karl loves Karl
Karl loves Imre
Imre loves Karl
Imre loves Imre

The point is that we need to keep track of the gaps in predicates, so that we can keep track of how we are filling them in.

To keep track of the gaps, we will label them. The labelling conventions we will adopt are best explained by example. Suppose we want to symbolize the following sentences:

11. Karl loves Imre.
12. Imre loves himself.
13. Karl loves Imre, but not vice versa.
14. Karl is loved by Imre.

We will start with the following symbolisation key:

domain: people

i: Imre
k: Karl
Lxy: $\_\_\_\_\__x$ loves $\_\_\_\_\__y$

Sentence 11 will now be symbolized by Lki.

Sentence 12 can be paraphrased as 'Imre loves Imre'. It can now be symbolized by Lii.

Sentence 13 is a conjunction. We might paraphrase it as 'Karl loves Imre, and Imre does not love Karl'. It can now be symbolized by $Lki \land \neg Lik$.

Sentence 14 might be paraphrased by 'Imre loves Karl'. It can then be symbolized by Lik. Of course, this slurs over the difference in tone between the active and passive voice; such nuances are lost in FOL.

This last example, though, highlights something important. Suppose we add to our symbolization key the following:

Mxy: $\_\_\_\_\__y$ loves $\_\_\_\_\__x$

Here, we have used the same English word ('loves') as we used in our symbolization key for Lxy. However, we have swapped the order of the *gaps* around (just look closely at those little subscripts!) So Mki and Lik now *both* symbolize 'Imre loves Karl'. Mik and Lki now *both* symbolize 'Karl loves Imre'. Since love can be unrequited, these are very different claims.

The moral is simple. When we are dealing with predicates with more than one place, we need to pay careful attention to the order of the places.

Predicates can have more than two places.

For example, consider

15. David bought the necklace for Victoria.

We symbolise this as

$$Bdna$$

using the symbolisation key:

d: David
n: the necklace
a: Victoria
$Rxyz$: $\_\_\_\_\__x$ bought $\_\_\_\_\__y$ for $\_\_\_\_\__z$

There is no limit to the number of places that a predicate may have.

16. The daughter of Gregor and Hilary is a friend of the first daughter of Bill and Michelle.

We symbolise this as

$$Rabcd$$

using:

a: Gregor
b: Hilary
c: Bill
d: Michelle
$Rx_1x_2x_3x_4$: The daughter of $\_\_\_\_\__{x_1}$ and $\_\_\_\_\__{x_2}$ is a friend of the first daughter of $\_\_\_\_\__{x_3}$ and $\_\_\_\_\__{x_4}$.

25.4 Universal Quantifier

Consider

17. Everyone wears a funny hat

This doesn't say of any specific individual that they wear a funny hat, but it says everyone does so. To express this, we introduce the $\forall$ symbol. This is called the *universal quantifier.*

$\forall$ should be read as "everything" or "everyone" (although it is sometimes more convenient to merely read "every" or "all") . If we say that everyone wears a funny what do we want to say of *all* people? We want to say that they wear a funny hat. In this sentence we used the "they". This doesn't refer to any particular person, Harry or Katie, instead it can refer to anyone. That is, we are using it as an INDIVIDUAL VARIABLE. We might then paraphrase "Everyone wears a funny hat" more explicitly as:

For everyone, x: x wears a funny hat.

Here we have made explicit the variable as x. In FOL we can also use y, z, or also, for example, x_{32} as variables. Quantifiers always have to be followed immediately by a variable.

If we wanted to symbolise "Alice wears a funny hat" we would use Fa. To symbolise "Everyone wears a funny hat", we paraphrase it as "For everyone : x wears a funny hat." and then symbolise it as $\forall xFx$.

Whatever we wanted to say of an individual we can now say of everyone using this quantifier. Consider

18. Everyone is happy and wears a funny hat

We can break this up:

$$\begin{array}{c}
\text{Everyone is happy and wears a funny hat} \\
\boxed{\text{For everyone, } x\text{:}}\ x \text{ is happy and } x \text{ wears a funny hat} \\
\boxed{\forall x}\,(Hx \wedge Fx) \\
| \\
x \text{ is happy } \boxed{\text{and}}\ x \text{ wears a funny hat} \\
Hx \boxed{\wedge} Fx \\
\diagup \qquad \diagdown \\
\begin{array}{cc} x \text{ is happy} & x \text{ wears a funny hat} \\ Hx & Fx \end{array}
\end{array}$$

So we can symbolise it as

$$\forall x(Hx \wedge Fx)$$

We have here been using $\forall x$ to be read out-loud as "everyone". But when we usually say "everyone" ("everything") do we always mean to talk about every person (every thing)? No! Suppose I say

Everyone has done the problem sheet.

then I do not mean to say that every person in the world has done the problem sheet. Arguably, I rather want to say that every student of the logic course has done the problem sheet. This suggests that how a quantifier is to be understood depends on the DOMAIN. The domain is the collection of things that we are talking about. Ultimately, using the notion of a domain the quantifier $\forall x$ should be understood as "for all objects in the domain, x". If the domain also contains dogs, or landmarks, then it also says something about those dogs, or landmarks. We say that the quantifiers *range over* the objects in the domain.

If I give

Ex: $\_\_\_\__x$ is energetic
domain: dogs

Then $\forall xEx$ says that all dogs are energetic. More generally, in this case we read $\forall x$ as "For every dog, x: …".

If I have a domain consisting of landmarks, then $\forall x$ is read as "For every landmark, x: …".

It should be immediate why it is important to highlight the domain when symbolizing arguments in FOL. The choice of the domain highlights the implicit assumptions we make when considering the sentence in the given situation. Depending on these implicit assumptions, i.e., the choice of the domain, a sentence like 'Everyone has done the problem sheet.' may come out true or false.

The domain can be chosen however you like. However, in FOL domains have to contain at least one object.

25.5 Existential Quantifier

The Universal Quantifier, $\forall$, allows us to capture English notions like "everything". The final component of FOL is the Existential Quantifier, $\exists$. This allows us to capture "someone" or"something".

To symbolise

19. Someone is angry.

with the symbolisation key

domain: people
Ax: _____$_x$ is angry

We can paraphrase it as

There is someone, x, such that: x is angry.

Someone is angry.
There is someone, x, such that: x is angry.
$\exists x\, Ax$
x is angry
Ax

To symbolise

20. There is a logician who wears glasses

There is a logician who wears glasses

There is someone, x: x is a logician who wears glasses

$\exists x\,(Lx \wedge Gx)$

x is a logician who wears glasses

x is a logician and x wears glasses

$Lx \wedge Gx$

x is a logician — x wears glasses

Lx — Gx

giving our symbolisation key:

domain: people

Lx: _____$_x$ is a logician

Gx: _____$_x$ wears glasses

To symbolise

21. There is a Polish woman who won the Nobel Prize

There is a Polish woman who won the Nobel Prize

There is someone, x: x is a Polish woman who won the Nobel Prize

$\exists x\,((Px \wedge Wx) \wedge Nx)$

x is a Polish woman who won the Nobel Prize

x is a Polish woman and x won the Nobel Prize

$(Px \wedge Wx) \wedge Nx$

x is a Polish woman

x is Polish and x is a woman — x won the Nobel Prize

$Px \wedge Wx$ — Nx

x is Polish — x is a woman

Px — Wx

So we symbolise it as:

$$\exists x((Px \wedge Wx) \wedge Nx)$$

giving our symbolisation key:

domain: people
Px: $\_\_\_\_\_\__x$ is polish
Wx: $\_\_\_\_\_\__x$ is a woman
Nx: $\_\_\_\_\_\__x$ won the Nobel Prize

As for the universal quantifier, how to read "$\exists x$" depends on the domain. We might talk not about people but about dogs. If our domain is dogs, then we understand $\exists x$ as "There is a dog, x, such that:". However, a more neutral way of understanding the existential quantifier, which anticipates the semantics for FOL is as " there exists an object of the domain $_x$", which also makes it obvious why $\exists$ is called the existential quantifier.

CHAPTER 26

Sentences of FOL

We have now informally introduced the basic building blocks of sentences of FOL. We will now carefully introduce what it is to be a sentence of FOL.

26.1 Vocabulary of FOL

We'll start by summarising, a bit more formally, the vocabulary of FOL:

Predicates $A, B, C, \ldots, W$, with subscripts, as needed: $A_1, Z_2, A_{25}, J_{375}, \ldots$.[1]

Names $a, b, c, \ldots, s, t$, or with subscripts, as needed $a_1, b_{224}, h_7, m_{32}, \ldots$

Variables x, y, z, or with subscripts, as needed $x_1, y_1, z_1, x_2, \ldots$. u, v, w may also be used.

---

[1]Each predicate will have a number of places associated with it. We should thus really introduce:

Zero-Place Predicates = Atomic sentences of TFL $A^0, B^0, \ldots, Z^0$, with subscripts, as needed: $A^0_1, Z^0_2, A^0_{25}, J^0_{375}, \ldots$

One-Place Predicates $A^1, B^1, \ldots, Z^1$, with subscripts, as needed: $A^1_1, Z^1_2, A^1_{25}, J^1_{375}, \ldots$

Two-Place Predicates $A^2, B^2, \ldots, Z^2$, with subscripts, as needed: $A^2_1, Z^2_2, A^2_{25}, J^2_{375}, \ldots$

Three-Place Predicates $A^3, B^3, \ldots, Z^3$, with subscripts, as needed: $A^3_1, Z^3_2, A^3_{25}, J^3_{375}, \ldots$

etc. We drop the superscripts for ease.

Connectives $\neg, \wedge, \vee, \rightarrow$

Brackets (,)

Quantifiers $\forall, \exists$

26.2 Formulas

In §4.3, we went straight from the presentation of the vocabulary of TFL to the definition of a sentence of TFL. In FOL, we will have to go via an intermediary stage, that is, we first need to introduce FORMULAS of FOL. The intuitive idea is that a formula is any sentence, or anything which can be turned into a sentence by adding quantifiers out front. But this will take some unpacking.

As we did for TFL, we will present a inductive definition of a formula of FOL. The starting point of this is the notion of an *atomic formula*. In TFL we started our definition with the notion of an atomic sentence, which were just given to us in our vocabulary. In FOL, the starting point of our definition is the notion of an atomic formula. Atomic formulas will be given by the following definition:

> If P is an n-place predicate and $t_1, \ldots t_n$ are either variables or names, then $Pt_1 \ldots t_n$ is an ATOMIC FORMULA.[2]

For example, if D is a one-place predicate (we might have introduced it to symbolise '$\_\_\_\__x$ is a dog'), and L is a two-place predicate (we might have introduced it to symbolise '$\_\_\_\__x$ loves $\_\_\_\__y$'), then the following are atomic formulas:

$$Db,\ Dx,\ Dy,\ Lki,\ Lkx,\ Lyz.$$

Formulas are constructed by starting with these and using either our TFL connectives or our quantifiers.

We now give the inductive definition of what it is to be a formula of FOL.

[2]In FOL variables and names are both called (singular) terms. Notice that in the above definition P and t are metavariables ranging over names of predicates and names of terms respectively.

1. If P is an n-place predicate and $t_1, \ldots t_n$ are either variables or names, then $Pt_1 \ldots t_n$ is a formula.
   These are called ATOMIC FORMULAS.

2. If X is a formula, then $\neg X$ is a formula.

3. If X and Y are formulas, then

   a) $(X \wedge Y)$ is a formula,
   b) $(X \vee Y)$ is a formula,
   c) $(X \rightarrow Y)$ is a formula, and

4. If X is a formula, v is a variable , then

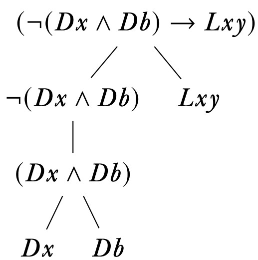
   a) $\forall v X$ is a formula
   b) $\exists v X$ is a formula.

5. Nothing else is a formula.

As for TFL, we start out with some formulas, such as Dx or Db, and we can construct more complicated formulas with our connectives, e.g.

$$(Dx \wedge Db),$$
$$\neg(Dx \wedge Db)$$
$$(\neg(Dx \wedge Db) \rightarrow Lxy)$$

And we can display their construction using our formation trees, as in 4.3.



This is exactly as in the case of TFL, the only difference is that the "leaves" of the tree have more structure to them: they're predicates applied to names or variables rather than simply the single atomic sentences that we had in TFL.

The new clauses here are in 4. This lets us put $\forall x$ in front of a formula, e.g. Bx to construct a formula $\forall x Bx$. We can also add quantifiers when the formula was already more complicated, e.g., we can construct a formula

$$\forall x(\neg(Dx \land Db) \to Lxy).$$

We could also have added an existential quantifier, $\exists$ to construct

$$\exists x(\neg(Dx \land Db) \to Lxy).$$

We can also do it with other variables, e.g.

$$\forall y(\neg(Dx \land Db) \to Lxy).$$

We can then add further quantifiers to *these* new formula, to construct, e.g.

$$\exists y \forall x(\neg(Dx \land Db) \to Lxy).$$

We can again display the structure and construction of the sentence perspicuously by presenting a formation tree:

$$\begin{array}{c}
\exists y\, \forall x(\neg(Dx \land Db) \to Lxy) \\
| \\
\forall x\, (\neg(Dx \land Db) \to Lxy) \\
| \\
(\neg(Dx \land Db) \to Lxy) \\
\diagup \qquad \diagdown \\
\neg\, (Dx \land Db) \qquad Lxy \\
| \\
(Dx \land Db) \\
\diagup \quad \diagdown \\
Dx \qquad Db
\end{array}$$

One more example:

$$\boxed{\exists z}\, \forall y(Ryz \land \exists x Qx)$$

$$\boxed{\forall y}\, (Ryz \land \exists x Qx)$$

$$(Ryz \boxed{\land} \exists x Qx)$$

$$Ryz \qquad \boxed{\exists x}\, Qx$$

$$Qx$$

Moving up the formation tree is following one of the rules of the recursive specification of what it is to be a sentence.

The notions of scope and main logical operators that were given in 4.3 equally applies to FOL but now the main logical operator might be a quantifier. These were:

> The **MAIN LOGICAL OPERATOR** in a sentence is the operator that was introduced last when that sentence was constructed using the recursion rules.
>
> The **SCOPE** of a logical operator in a sentence is the formula for which that operator is the main logical operator.

We can graphically illustrate scopes as follows:

$$\overbrace{\exists z \underbrace{\forall y(\overbrace{\exists x Qx}^{\text{scope of } \exists x} \land Ryz)}_{\text{scope of } \forall y}}^{\text{scope of } \exists z}$$

26.3 Sentences

Recall that we are largely concerned in logic with assertoric sentences: sentences that can be either true or false. Many formulas are not sentences. Consider the following symbolization key:

domain: people

Lxy: _____$_x$ loves _____$_y$
b: Boris

Consider the atomic formula Lzz. Can it be true or false? You might think that it will be true just in case the person named by z loves themself, in the same way that Lbb is true just in case Boris (the person named by b) loves himself. *However, z is a variable, and does not name anyone or any thing.*

Of course, if we put an existential quantifier out front, obtaining $\exists z Lzz$, then this would be true iff someone loves herself. Equally, if we wrote $\forall z Lzz$, this would be true iff everyone loves themself. The point is that we need a quantifier to tell us how to deal with a variable.

Let's make this idea precise.

> A BOUND VARIABLE is an occurrence of the variable v that is within the scope of either $\forall v$ or $\exists v$.
>
> A FREE VARIABLE is any variable that is not bound.

For example, consider the formula

$$\forall x(Ex \lor Dy) \to \exists z(Ex \to Lzx)$$

The scope of the universal quantifier $\forall x$ is $\forall x(Ex \lor Dy)$, so the first x is bound by the universal quantifier. However, the second and third occurrence of x are free. Equally, the y is free. The scope of the existential quantifier $\exists z$ is $(Ex \to Lzx)$, so z is bound.

Finally we can say what a sentence of FOL is

> A SENTENCE of FOL is any formula of FOL that contains no free variables.

26.4 Bracketing conventions

We will adopt the same notational conventions governing brackets that we did for TFL (see §4.3): we may omit the outermost brackets of a formula.

Practice exercises

A. Identify which variables are bound and which are free.

1. $\exists x Lxy \land \forall y Lyx$
2. $\forall x Ax \land Bx$
3. $\forall x(Ax \land Bx) \land \forall y(Cx \land Dy)$
4. $\forall x \exists y[Rxy \rightarrow (Jz \land Kx)] \lor Ryx$
5. $\forall x_1(Mx_2 \rightarrow Lx_2x_1) \land \exists x_2 Lx_3x_2$

CHAPTER 27

FOL-Symbolisations

We have already seem the key idea of FOL symbolisation in Chapter 25 applied to sentences with limited logical complexity. Before moving to symbolise more complex sentences, we explicitly summarise our strategy for symbolising complex sentences. This extends the strategy that we used for TFL in §14.2:

1. See if the sentence can be paraphrased in English in one of the standard forms.
    - ▹ If not, it's an atomic formula: identify the predicate and the variables or names.
2. Use the symbolisation trick for that form.
3. Repeat the procedure with the components. Etc.

Our key forms are:

English paraphrase	Symbolisation
Everything (in the domain), x, is such that:	$\forall x \ldots$
Something (in the domain), x, is such that:	$\exists x \ldots$
It is not the case that X	$\neg X$
X and Y	$(X \wedge Y)$
X or Y	$(X \vee Y)$
If X, then Y	$(X \rightarrow Y)$

Also remember that there were various further tricks from II, such as 'X only if Y' as $(X \rightarrow Y)$ and 'Unless X, Y' as $(X \vee Y)$. These still apply in the FOL setting. We will also see some more such tricks later.

27.1 Clarification on Domains

In FOL, the domain must always include at least one thing. Moreover, in English we can infer 'something is angry' from 'Gregor is angry'. In FOL, then, we will want to be able to infer $\exists xAx$ from Ag. So we will insist that each name must pick out exactly one thing in the domain.

> A domain must have *at least* one member. A name must pick out *exactly* one member of the domain, but a member of the domain may be picked out by one name, many names, or none at all.

Non-referring terms (Further philosophical interest)

In FOL, each name must pick out exactly one member of the domain. A name cannot refer to more than one thing—it is a *singular* term. Each name must still pick out *something*. This is connected to a classic philosophical problem: the so-called problem of non-referring terms.

Medieval philosophers typically used sentences about the *chimera* to exemplify this problem. Chimera is a mythological creature; it does not really exist. Consider these two sentences:

1. Chimera is angry.
2. Chimera is not angry.

It is tempting just to define a name to mean 'chimera.' The symbolization key would look like this:

domain: creatures on Earth
Ax: $\_\_\_\_\_\__x$ is angry.
c: chimera

We could then symbolize sentence 1 as Ac and sentence 2 as $\neg Ac$.

Problems will arise when we ask whether these sentences are true or false.

One option is to say that sentence 1 is not true, because there is no chimera. If sentence 1 is false because it talks about a non-existent thing, then sentence 2 is false for the same reason. Yet this would mean that Ac and $\neg Ac$ would both be false. Given the truth conditions for negation, this cannot be the case, and contradict the Law of Excluded Middle.

Since we cannot say that they are both false, what should we do? Another option is to say that sentence 1 is *meaningless* because it talks about a non-existent thing. So Ac would be a meaningful expression in FOL for some interpretations but not for others. Yet this would make our formal language hostage to particular interpretations. Since we are interested in logical form, we want to consider the logical force of a sentence like Ac apart from any particular interpretation. If Ac were sometimes meaningful and sometimes meaningless, we could not do that.

This is the *problem of non-referring terms* which is important problem in philosophical logic and the philosophy of language. For our purpose the important point is to appreciate that each name of FOL *must* refer to something in the domain (although the domain can contain any things we like). If we want to symbolize arguments about mythological creatures, then we must define a domain that includes them. This option is important if we want to consider the logic of stories. We can symbolize a sentence like 'Sherlock Holmes lived at 221B Baker Street' by including fictional characters like Sherlock Holmes in our domain.

27.2 Symbolisation with Many-Placed Predicates

To symbolise

3. Everyone loves Alice.

We want to paraphrase it in one of our standard forms, which we do as:

$$
\begin{array}{c}
\text{Everyone loves Alice} \\
\boxed{\text{For everyone, } x\text{:}}\ x \text{ loves Alice} \\
\boxed{\forall x}\ Lxa \\
| \\
x \text{ loves Alice} \\
Lxa
\end{array}
$$

So we give the symbolisation

$$\forall x Lxa$$

with the symbolisation key:

domain: people
Lxy: $\_\_\_\_\__x$ loves $\_\_\_\_\__y$
a: Alice

If we instead want to symbolise

4. Alice loves everyone.

We paraphrase this as:

$$
\begin{array}{c}
\text{Alice loves everyone} \\
\boxed{\text{For everyone, } x\text{:}}\ \text{Alice loves } x \\
\boxed{\forall x}\ Lax \\
| \\
\text{Alice loves } x \\
Lax
\end{array}
$$

So we give the symbolisation

$$\forall x Lax$$

To symbolise

5. Someone loves themselves.

Someone loves themselves.
For someone, x: x loves themselves
$\exists x\, Lxx$

x loves themselves
x loves x
Lxx

If we want to symbolise

6. Some dog likes playing with Finley.

we should paraphrase this as:

Some dog likes playing with Finley
There is some thing x: x is a dog who likes playing with Finley
$\exists x\,(Dx \land Pxf)$

x is a dog who likes playing with Finley
x is a dog and x likes playing with Finley
$Dx \land Pxf$

x is a dog — x likes playing with Finley
Dx — Pxf

27.3 Quantifiers inside a sentence

All the sentences we've considered so far have the quantifiers at the beginning of the sentence. But we can also use truth functional connectives to combine sentences of FOL.

7. Finley is not quiet, but some dog is.

We work as follows:

Finley is not quiet, but some dog is
Finley is not quiet and some dog is quiet
$(\neg Qf \wedge \exists x(Dx \wedge Qx))$

Finley is not quiet
It is not the case that Finley is quiet
$\neg Qf$

Finley is quiet
Qf

Some dog is quiet
There is some x: x is a dog and x is quiet
$\exists x (Dx \wedge Qx)$

x is a dog and x is quiet
$Dx \wedge Qx$

x is a dog
Dx

x is quiet
Qx

So we symbolise this sentence as

$$(\neg Qf \wedge \exists x(Dx \wedge Qx))$$

giving the symbolisation key

domain: animals
Dx: $\_\_\_\_\_\__x$ is a dog
Qx: $\_\_\_\_\_\__x$ is quiet
f: Finley

Note, that as per 27.1, this symbolisation is only legitimate assuming that Finley names an animal. Names have to name members of the domain.

Consider also

8. Some dog is quiet but Finley is not

Some dog is quiet but Finley is not

Some dog is quiet and Finley is not quiet

$(\exists x(Dx \wedge Qx) \wedge \neg Qf)$

Some dog is quiet

There is some x: x is a dog and x is quiet

$\exists x(Dx \wedge Qx)$

x is a dog and x is quiet

$Dx \wedge Qx$

x is a dog

Dx

x is quiet

Qx

Finley is not quiet

It is not the case that Finley is quiet

$\neg Qf$

Finley is quiet

Qf

Consider:

9. Not every dog is quiet

We work as follows:

Not every dog is quiet

It is not the case that every dog is quiet

$\neg \forall x(Dx \rightarrow Qx)$

Every dog is quiet

For all x: if x is a dog then x is quiet

$\forall x(Dx \rightarrow Qx)$

If x is a dog then x is quiet

$Dx \rightarrow Qx$

x is a dog

Dx

x is quiet

Qx

So we symbolise this sentence as

$$\neg \forall x(Dx \rightarrow Qx)$$

CHAPTER 28

Common Quantifier Phrases and Domains

28.1 Common quantifier phrases

Consider these sentences:

1. Every coin in my pocket is a quarter.
2. Some coin on the table is a dime.
3. Not all the coins on the table are dimes.
4. None of the coins in my pocket are dimes.

In providing a symbolization key, we need to specify a domain. Since we are talking about coins in my pocket and on the table, the domain must at least contain all of those coins. Since we are not talking about anything besides coins, we let the domain be all coins. Since we are not talking about any specific coins, we do not need to deal with any names. So here is our key:

domain: all coins

Px: $\underline{\qquad}_x$ is in my pocket
Tx: $\underline{\qquad}_x$ is on the table
Qx: $\underline{\qquad}_x$ is a quarter
Dx: $\underline{\qquad}_x$ is a dime

Sentence 1 is most naturally symbolized using a universal quantifier. The universal quantifier says something about everything in the domain, not just about the coins in my pocket. Sentence 1 can be paraphrased as 'for any coin, *if* that coin is in my pocket *then* it is a quarter'. So we can symbolize it as $\forall x(Px \to Qx)$.

Since sentence 1 is about coins that are both in my pocket *and* that are quarters, it might be tempting to symbolize it using a conjunction. However, the sentence $\forall x(Px \land Qx)$ would symbolize the sentence 'every coin is both a quarter and in my pocket'. This obviously means something very different than sentence 1. And so we see:

> If a sentence can be paraphrased in English as
>
> 'every F is G',
> 'all Fs are Gs', or
> 'any F is a G',
>
> it can be symbolised as
>
> $$\forall x(Fx \to Gx).$$

Sentence 2 is most naturally symbolized using an existential quantifier. It can be paraphrased as 'there is some coin which is both on the table and which is a dime'. So we can symbolize it as $\exists x(Tx \land Dx)$.

Notice that we needed to use a conditional with the universal quantifier, but we used a conjunction with the existential quantifier. Suppose we had instead written $\exists x(Tx \to Dx)$. That would mean that there is some object in the domain of which $(Tx \to Dx)$ is true. Recall that, in TFL, $X \to Y$ is logically equivalent (in TFL) to $\neg X \lor Y$. This equivalence will also hold in FOL. So $\exists x(Tx \to Dx)$ is true if there is some object in the domain, such that $(\neg Tx \lor Dx)$ is true of that object. That is, $\exists x(Tx \to Dx)$ 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 lots of other places. So it is *very easy* for $\exists x(Tx \to Dx)$ to be 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 tends to say something very weak indeed. As a general rule of thumb, do not put conditionals in the scope of existential quantifiers unless you are sure that you need one.

> If a sentence can be paraphrased in English as
>
> ‘some F is G’,
> ‘there is some F that is G’,
> ‘An F is G’, or
> ‘there is at least one F that is a G’
>
> it can be symbolised as
>
> $$\exists x(Fx \wedge Gx).$$

Sentence 3 can be paraphrased as, ‘It is not the case that every coin on the table is a dime’. So we can symbolize it by $\neg\forall x(Tx \rightarrow Dx)$. You might look at sentence 3 and paraphrase it instead as, ‘Some coin on the table is not a dime’. You would then symbolize it by $\exists x(Tx \wedge \neg Dx)$. Although it is probably not immediately obvious yet, these two sentences are logically equivalent. (This is due to the logical equivalence between $\neg\forall x X$ and $\exists x\neg X$, mentioned in §25, along with the equivalence between $\neg(X \rightarrow Y)$ and $X \wedge \neg Y$.)

> If a sentence can be paraphrased in English as
>
> ‘not all Fs are Gs’,
>
> it can be symbolised as
>
> $\neg\forall x(Fx \rightarrow Gx)$, or
> $\exists x(Fx \wedge \neg Gx)$.

Sentence 4 can be paraphrased as, ‘It is not the case that there is some dime in my pocket’. This can be symbolized by $\neg\exists x(Px \wedge Dx)$. It might also be paraphrased as, ‘Everything in my pocket is a non-dime’, and then could be symbolized by $\forall x(Px \rightarrow \neg Dx)$. Again the two symbolizations are logically equivalent; both are correct symbolizations of sentence 4.

If a sentence can be paraphrased in English as

'no Fs are Gs',

it can be symbolised as

$$\neg\exists x(Fx \wedge Gx), \text{ or}$$
$$\forall x(Fx \rightarrow \neg Gx).$$

Finally, consider 'only', as in:

5. Only dimes are on the table.

How should we symbolize this? A good strategy is to consider when the sentence would be false. If we are saying that only dimes are on the table, we are excluding all the cases where something on the table is a non-dime. So we can symbolize the sentence the same way we would symbolize 'No non-dimes are on the table.' Remembering the lesson we just learned, and symbolizing 'x is a non-dime' as '$\neg Dx$', the possible symbolizations are: '$\neg\exists x(Tx \wedge \neg Dx)$', or alternatively: '$\forall x(Tx \rightarrow \neg\neg Dx)$'. Since double negations cancel out, the second is just as good as '$\forall x(Tx \rightarrow Dx)$'. In other words, 'Only dimes are on the table' and 'Everything on the table is a dime' are symbolized the same way.

If a sentence can be paraphrased in English as

'only Fs are Gs',

it can be symbolised as

$$\neg\exists x(Gx \wedge \neg Fx), \text{ or}$$
$$\forall x(Gx \rightarrow Fx)$$

28.2 Empty predicates

In §25, we emphasized that a name must pick out exactly one object in the domain. However, a predicate need not apply to anything in the domain. A predicate that applies to nothing in the domain is called an EMPTY PREDICATE. This is worth exploring.

Suppose we want to symbolize these two sentences:

6. Every monkey knows sign language
7. Some monkey knows sign language

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

domain: animals
Mx: _____$_x$ is a monkey.
Sx: _____$_x$ knows sign language.

Sentence 6 can now be symbolized by $\forall x(Mx \rightarrow Sx)$. Sentence 7 can be symbolized as $\exists x(Mx \wedge Sx)$.

It is tempting to say that sentence 6 *entails* sentence 7. That is, we might think that it is impossible for it to be the case that every monkey knows sign language, without its also being the case that some monkey knows sign language, but this would be a mistake. 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 these sentences would be true or false *if there were no monkeys*. If there were no monkeys at all (in the domain), then $\forall x(Mx \rightarrow Sx)$ would be *vacuously* true: take any monkey you like—it knows sign language! But if there were no monkeys at all (in the domain), then $\exists x(Mx \wedge Sx)$ would be false.

Another example will help to bring this home. Suppose we extend the above symbolization key, by adding:

Rx: _____$_x$ is a refrigerator

Now consider the sentence $\forall x(Rx \rightarrow Mx)$. This symbolizes 'every refrigerator is a monkey'. This sentence is true, given our symbolization key, which is counterintuitive, since we (presumably) do not want to say that there are a whole bunch of refrigerator monkeys. It is important to remember, though, that $\forall x(Rx \rightarrow Mx)$ is true iff any member of the domain that is a refrigerator is a monkey. Since the domain is *animals*, there are no refrigerators in the domain. Again, then, the sentence is *vacuously* true.

If you were actually dealing with the sentence 'All refrigerators are monkeys', then you would most likely want to include kitchen appliances in the domain. Then the predicate R would not be empty and the sentence $\forall x(Rx \rightarrow Mx)$ would be false.

When F is an empty predicate, a sentence $\forall x(Fx \to \ldots)$ will be vacuously true.

28.3 Picking a domain

The appropriate symbolization of an English language sentence in FOL will depend on the symbolization key. Choosing a key can be difficult. Suppose we want to symbolize the English sentence:

8. Every rose has a thorn.

We might offer this symbolization key:

Rx: _____$_x$ is a rose
Tx: _____$_x$ has a thorn

According to our symbolization guideline sentence 8 should then be symbolized as $\forall x(Rx \to Tx)$. However, it remains to specify a domain. Without specific contextual information we must assume that the sentence makes a claim about all roses, that is, we should assume that the domain contains all roses. Also, arguable by saying that all roses have thorns we contrast roses to other tings in the domain: for if every object in the domain were a rose, it would suffice to say 'Everything has a thorne'. So in the case of sentence 8 taking the domain to consist of flowers would be a good, plausible choice and picking the domain to consist of roses only would not be satisfactory.

More generally, in absence of specific contextual information about the domain, it is preferable to pick a domain that includes many things. that is, things besides roses: rhododendrons; rats; rifles; whatevers., and we will certainly need to include a more expansive domain if we simultaneously want to symbolize sentences like:

9. Every cowboy sings a sad, sad song.

Our domain must now include both all the roses (so that we can symbolize sentence 8) and all the cowboys (so that we can symbolize sentence 9). So we might offer the following symbolization key:

domain: people and plants
Cx: _____$_x$ is a cowboy
Sx: _____$_x$ sings a sad, sad song
Rx: _____$_x$ is a rose

Tx: _____$_x$ has a thorn

Now we will have to symbolize sentence 8 with $\forall x(Rx \rightarrow Tx)$, since $\forall x Tx$ would symbolize the sentence 'every person or plant has a thorn'. Similarly, we will have to symbolize sentence 9 with $\forall x(Cx \rightarrow Sx)$.

In general, the universal quantifier can be used to symbolize the English expression 'everyone' if the domain only contains people. If there are people and other things in the domain, then 'everyone' must be treated as 'every person'.

28.4 Ambiguous predicates

Suppose we just want to symbolize this sentence:

10. Adina is a skilled surgeon.

Let the domain be people, let Kx mean 'x is a skilled surgeon', and let a mean Adina. Sentence 10 is simply Ka.

Suppose instead that we want to symbolize this argument:

> The hospital will only hire a skilled surgeon. All surgeons are greedy. Billy is a surgeon, but is not skilled. Therefore, Billy is greedy, but the hospital will not hire him.

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

domain: people
Gx: _____$_x$ is greedy.
Hx: The hospital will hire _____$_x$.
Rx: _____$_x$ is a surgeon.
Kx: _____$_x$ is skilled.
b: Billy

Now the argument can be symbolized in this way:

1. $\forall x\big[\neg(Rx \wedge Kx) \rightarrow \neg Hx\big]$
2. $\forall x(Rx \rightarrow Gx)$
3. $Rb \wedge \neg Kb$

$\therefore$ Therefore: $Gb \wedge \neg Hb$

Next suppose that we want to symbolize this argument:

Carol is a skilled surgeon and a tennis player. 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 Tx mean 'x is a tennis player') and a name (let c mean Carol). Then the argument becomes:

1. $(Rc \land Kc) \land Tc$

$\therefore$ Therefore: $Tc \land Kc$

This symbolization is a disaster! It takes what in English is a terrible argument and symbolizes it as a valid argument in FOL. The problem is that there is a difference between being *skilled as a surgeon* and *skilled as a tennis player*. Symbolizing this argument correctly requires two separate predicates, one for each type of skill. If we let K_1x mean 'x is skilled as a surgeon' and K_2x mean 'x is skilled as a tennis player,' then we can symbolize the argument in this way:

1. $(Rc \land K_1c) \land Tc$

$\therefore$ Therefore: $Tc \land K_2c$

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

The moral of these examples is that you need to be careful of symbolizing predicates 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 'x is a skilled surgeon', rather than two predicates 'x is skilled' and 'x is a surgeon'? Sometimes. As sentence 10 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 symbolizing an argument that is just about dogs, it is fine to define a predicate that means 'x is big.' If the domain includes dogs and mice, however, it is probably best to make the predicate mean 'x is big for a dog.'

Practice exercises

A. Here are the syllogistic figures identified by Aristotle and his successors, along with their medieval names:

1. **Barbara**. All G are F. All H are G. So: All H are F
   $\forall x(Gx \rightarrow Fx), \forall x(Hx \rightarrow Gx) \therefore \forall x(Hx \rightarrow Fx)$
2. **Celarent**. No G are F. All H are G. So: No H are F
   $\forall x(Gx \rightarrow \neg Fx), \forall x(Hx \rightarrow Gx) \therefore \forall x(Hx \rightarrow \neg Fx)$
3. **Ferio**. No G are F. Some H is G. So: Some H is not F
   $\forall x(Gx \rightarrow \neg Fx), \exists x(Hx \wedge Gx) \therefore \exists x(Hx \wedge \neg Fx)$
4. **Darii**. All G are H. Some H is G. So: Some H is F.
   $\forall x(Gx \rightarrow Fx), \exists x(Hx \wedge Gx) \therefore \exists x(Hx \wedge Fx)$
5. **Camestres**. All F are G. No H are G. So: No H are F.
   $\forall x(Fx \rightarrow Gx), \forall x(Hx \rightarrow \neg Gx) \therefore \forall x(Hx \rightarrow \neg Fx)$
6. **Cesare**. No F are G. All H are G. So: No H are F.
   $\forall x(Fx \rightarrow \neg Gx), \forall x(Hx \rightarrow Gx) \therefore \forall x(Hx \rightarrow \neg Fx)$
7. **Baroko**. All F are G. Some H is not G. So: Some H is not F.
   $\forall x(Fx \rightarrow Gx), \exists x(Hx \wedge \neg Gx) \therefore \exists x(Hx \wedge \neg Fx)$
8. **Festino**. No F are G. Some H are G. So: Some H is not F.
   $\forall x(Fx \rightarrow \neg Gx), \exists x(Hx \wedge Gx) \therefore \exists x(Hx \wedge \neg Fx)$
9. **Datisi**. All G are F. Some G is H. So: Some H is F.
   $\forall x(Gx \rightarrow Fx), \exists x(Gx \wedge Hx) \therefore \exists x(Hx \wedge Fx)$
10. **Disamis**. Some G is F. All G are H. So: Some H is F.
    $\exists x(Gx \wedge Fx), \forall x(Gx \rightarrow Hx) \therefore \exists x(Hx \wedge Fx)$
11. **Ferison**. No G are F. Some G is H. So: Some H is not F.
    $\forall x(Gx \rightarrow \neg Fx), \exists x(Gx \wedge Hx) \therefore \exists x(Hx \wedge \neg Fx)$
12. **Bokardo**. Some G is not F. All G are H. So: Some H is not F.
    $\exists x(Gx \wedge \neg Fx), \forall x(Gx \rightarrow Hx) \therefore \exists x(Hx \wedge \neg Fx)$
13. **Camenes**. All F are G. No G are H So: No H is F.
    $\forall x(Fx \rightarrow Gx), \forall x(Gx \rightarrow \neg Hx) \therefore \forall x(Hx \rightarrow \neg Fx)$
14. **Dimaris**. Some F is G. All G are H. So: Some H is F.
    $\exists x(Fx \wedge Gx), \forall x(Gx \rightarrow Hx) \therefore \exists x(Hx \wedge Fx)$
15. **Fresison**. No F are G. Some G is H. So: Some H is not F.
    $\forall x(Fx \rightarrow \neg Gx), \exists x(Gx \wedge Hx) \therefore \exists(Hx \wedge \neg Fx)$

Symbolize each argument in FOL.

B. Using the following symbolization key:

domain: people
Kx: $\underline{\quad\quad}_x$ knows the combination to the safe
Sx: $\underline{\quad\quad}_x$ is a spy
Vx: $\underline{\quad\quad}_x$ is a vegetarian
h: Hofthor
i: Ingmar

symbolize the following sentences in FOL:

1. Neither Hofthor nor Ingmar is a vegetarian.
   $\neg Vh \wedge \neg Vi$
2. No spy knows the combination to the safe.
   $\forall x(Sx \rightarrow \neg Kx)$
3. No one knows the combination to the safe unless Ingmar does.
   $\forall x \neg Kx \vee Ki$
4. Hofthor is a spy, but no vegetarian is a spy.
   $Sh \wedge \forall x(Vx \rightarrow \neg Sx)$

C. Using this symbolization key:

domain: all animals
Ax: $\underline{\quad\quad}_x$ is an alligator.
Mx: $\underline{\quad\quad}_x$ is a monkey.
Rx: $\underline{\quad\quad}_x$ is a reptile.
Zx: $\underline{\quad\quad}_x$ lives at the zoo.
a: Amos
b: Bouncer
c: Cleo

symbolize each of the following sentences in FOL:

1. Amos, Bouncer, and Cleo all live at the zoo.
   $Za \wedge Zb \wedge Zc$
2. Bouncer is a reptile, but not an alligator.
   $Rb \wedge \neg Ab$
3. Some reptile lives at the zoo.
   $\exists x(Rx \wedge Zx)$
4. Every alligator is a reptile.
   $\forall x(Ax \rightarrow Rx)$
5. Any animal that lives at the zoo is either a monkey or an alligator.
   $\forall x(Zx \rightarrow (Mx \vee Ax))$
6. There are reptiles which are not alligators.
   $\exists x(Rx \wedge \neg Ax)$
7. If any animal is an reptile, then Amos is.
   $\exists x\, Rx \rightarrow Ra$
8. If any animal is an alligator, then it is a reptile.
   $\forall x(Ax \rightarrow Rx)$

D. For each argument, write a symbolization key and symbolize the argument in FOL.

1. Willard is a logician. All logicians wear funny hats. So Willard wears a funny hat

   domain: people
   Lx: $\_\_\_\_\__x$ is a logician
   Hx: $\_\_\_\_\__x$ wears a funny hat
   i: Willard

   $Li, \forall x(Lx \rightarrow Hx) \therefore Hi$

2. Nothing on my desk escapes my attention. There is a computer on my desk. As such, there is a computer that does not escape my attention.

   domain: physical things
   Dx: $\_\_\_\_\__x$ is on my desk
   Ex: $\_\_\_\_\__x$ escapes my attention
   Cx: $\_\_\_\_\__x$ is a computer

   $\forall x(Dx \rightarrow \neg Ex), \exists x(Dx \wedge Cx) \therefore \exists x(Cx \wedge \neg Ex)$

3. All my dreams are black and white. Old TV shows are in black and white. Therefore, some of my dreams are old TV shows.

   domain: episodes (psychological and televised)
   Dx: $\_\_\_\_\__x$ is one of my dreams
   Bx: $\_\_\_\_\__x$ is in black and white
   Ox: $\_\_\_\_\__x$ is an old TV show

   $\forall x(Dx \rightarrow Bx), \forall x(Ox \rightarrow Bx) \therefore \exists x(Dx \wedge Ox)$.
   Comment: generic statements are tricky to deal with. Does the second sentence mean that *all* old TV shows are in black and white; or that most of them are; or that most of the things which are in black and white are old TV shows? I have gone with the former, but it is not clear that FOL deals with these well.

4. Neither Holmes nor Watson has been to Australia. A person could see a kangaroo only if they had been to Australia or to a zoo. Although Watson has not seen a kangaroo, Holmes has. Therefore, Holmes has been to a zoo.

   domain: people
   Ax: $\_\_\_\_\__x$ has been to Australia
   Kx: $\_\_\_\_\__x$ has seen a kangaroo
   Zx: $\_\_\_\_\__x$ has been to a zoo
   h: Holmes
   a: Watson

   $\neg Ah \wedge \neg Aa, \forall x(Kx \rightarrow (Ax \vee Zx)), \neg Ka \wedge Kh \therefore Zh$

5. No one expects the Spanish Inquisition. No one knows the troubles I've seen. Therefore, anyone who expects the Spanish Inquisition knows the troubles I've seen.

   domain: people
   Sx: $\underline{\quad\quad}_x$ expects the Spanish Inquisition
   Tx: $\underline{\quad\quad}_x$ knows the troubles I've seen
   h: Holmes
   a: Watson

   $\forall x \neg Sx, \forall x \neg Tx \therefore \forall x(Sx \rightarrow Tx)$

6. All babies are illogical. Nobody who is illogical can manage a crocodile. Berthold is a baby. Therefore, Berthold is unable to manage a crocodile.

   domain: people
   Bx: $\underline{\quad\quad}_x$ is a baby
   Ix: $\underline{\quad\quad}_x$ is illogical
   Cx: $\underline{\quad\quad}_x$ can manage a crocodile
   b: Berthold

   $\forall x(Bx \rightarrow Ix), \forall x(Ix \rightarrow \neg Cx), Bb \therefore \neg Cb$

CHAPTER 29

Symbolisation: Multiple quantifiers

We see more power of FOL when quantifiers start stacking on top of one another.

Consider

1. Someone loves everyone.

Before considering how to symbolise that, we start of by symbolising the related sentence:

2. John loves everyone.

This can be symbolised as $\forall x Ljx$, using the symbolisation key:

domain: all people
j: John
Lxy: _____$_x$ loves _____$_y$

This gives us an insight into how to symbolise 1: it's like 2 except it might not be John who loves everyone, 1 just says that there is *someone*, y, such that y loves everyone. We will thus symbolise it $\exists y \forall x Lyx$.

In the earlier examples we always used x as our variables; but we here had to use y because x is already taken. We don't want to say

$\exists x \forall x Lxx$ as it's not clear how one should read this sentence as we need to identify which variables come with which quantifiers.

29.1 The order of quantifiers

Consider the sentence 'everyone loves someone'. This is potentially ambiguous. It might mean either of the following:

3. For every person x, there is some person that x loves
4. There is some particular person whom every person loves

Sentence 3 can be symbolized by $\forall x \exists y Lxy$, and would be true of a love-triangle. For example, suppose that our domain of discourse is restricted to Imre, Juan and Karl. Suppose also that Karl loves Imre but not Juan, that Imre loves Juan but not Karl, and that Juan loves Karl but not Imre. Then sentence 3 is true.

Sentence 4 is symbolized by $\exists y \forall x Lxy$. Sentence 4 is *not* true in the situation just described. Again, suppose that our domain of discourse is restricted to Imre, Juan and Karl. This requires that all of Juan, Imre and Karl converge on (at least) one object of love.

The point of the example is to illustrate that the order of the quantifiers matters a great deal. Indeed, to switch them around is called a *quantifier shift fallacy*. Here is an example, which comes up in various forms throughout the philosophical literature:

1.	For every person, there is some truth they cannot know.	$(\forall\exists)$
∴	Therefore: There is some truth that no person can know.	$(\exists\forall)$

This argument form is obviously invalid. It's just as bad as:[1]

1.	Every dog has its day.	$(\forall\exists)$
∴	Therefore: There is a day for all the dogs.	$(\exists\forall)$

The moral is: take great care with the order of quantification.

29.2 Stepping-stones to symbolization

Once we have the possibility of multiple quantifiers, representation in FOL can quickly start to become a bit tricky. When you are trying to symbolize a complex sentence, we recommend laying down several

[1] Thanks to Rob Trueman for the example.

stepping stones. As usual, this idea is best illustrated by example. Consider this symbolisation key:

domain: people and dogs
Dx: $\_\_\_\_\__x$ is a dog
Fxy: $\_\_\_\_\__x$ is a friend of $\_\_\_\_\__y$
Oxy: $\_\_\_\_\__x$ owns $\_\_\_\_\__y$
g: Geraldo

Now let's try to symbolize these sentences:

5. Geraldo is a dog owner.
6. Someone is a dog owner.
7. All of Geraldo's friends are dog owners.
8. Every dog owner is a friend of a dog owner.
9. Every dog owner's friend owns a dog of a friend.

Sentence 5 can be paraphrased as, 'There is a dog that Geraldo owns'. This can be symbolized by $\exists x(Dx \wedge Ogx)$.

Sentence 6 can be paraphrased as, 'There is some y such that y is a dog owner'. Dealing with part of this, we might write $\exists y(y$ is a dog owner$)$. Now the fragment we have left as 'y is a dog owner' is much like sentence 5, except that it is not specifically about Geraldo. So we can symbolize sentence 6 by:

$$\exists y \exists x(Dx \wedge Oyx)$$

We should pause to clarify something here. In working out how to symbolize the last sentence, we wrote down $\exists y(y$ is a dog owner$)$. To be very clear: this is *neither* an FOL sentence *nor* an English sentence: it uses bits of FOL ($\exists$, y) and bits of English ('dog owner'). It is really is *just a stepping-stone* on the way to symbolizing the entire English sentence with a FOL sentence. You should regard it as a bit of rough-working-out, on a par with the doodles that you might absent-mindedly draw in the margin of this book, whilst you are concentrating fiercely on some problem.

Sentence 7 can be paraphrased as, 'Everyone who is a friend of Geraldo is a dog owner'. Using our stepping-stone tactic, we might write

$$\forall x\big[Fxg \rightarrow x \text{ is a dog owner}\big]$$

Now the fragment that we have left to deal with, 'x is a dog owner', is structurally just like sentence 5. However, it would be a mistake for us simply to write

$$\forall x\big[Fxg \rightarrow \exists x(Dx \wedge Oxx)\big]$$

for we would here have a *clash of variables.* While the scope of the universal quantifier, $\forall x$, is the entire conditional, the variable x does not occur free in the consequent of the conditional : it is bound by the existential quantifier instead. This means that

$$\forall x\big[Fxg \rightarrow \exists x(Dx \wedge Oxx)\big]$$

does not symbolize 'Every friend of Gerald is a dog owner' but rather a rather something like:

▹ If everyone is a friend of Gerald, then there is a dog that owns itself.

That's clearly not what we wanted!

To make sure that we get the intended outcome we need to see to it that the universal quantifier binds the owner position in the consequent of the conditional. To continue our symbolization, then, we must choose some different variable for our existential quantifier. What we want is something like:

$$\forall x\big[Fxg \rightarrow \exists z(Dz \wedge Oxz)\big]$$

This adequately symbolizes sentence 7, as the variable x now occurs free in the consequent of the conditional and can be bound by the universal quatifier, which in turn guarantees that dog-owner must be friends of Gerald.

Sentence 8 can be paraphrased as 'For any x that is a dog owner, there is a dog owner who x is a friend of'. Using our stepping-stone tactic, this becomes

$$\forall x\big[x \text{ is a dog owner} \rightarrow \exists y(y \text{ is a dog owner} \wedge Fxy)\big]$$

Completing the symbolization, we end up with

$$\forall x\big[\exists z(Dz \wedge Oxz) \rightarrow \exists y\big(\exists z(Dz \wedge Oyz) \wedge Fxy\big)\big]$$

Note that we have used the same letter, z, in both the antecedent and the consequent of the conditional, but that these are governed by two different quantifiers. This is ok: there is no clash here, because it is clear which quantifier that variable falls under. We might graphically represent the scope of the quantifiers thus:

$$\overbrace{\forall x\big[\overbrace{\exists z(Dz \land Oxz)}^{\text{scope of 1st `}\exists z\text{'}} \to \overbrace{\exists y(\overbrace{\exists z(Dz \land Oyz)}^{\text{scope of 2nd `}\exists z\text{'}} \land Fxy)}^{\text{scope of `}\exists y\text{'}}\big]}^{\text{scope of `}\forall x\text{'}}$$

This shows that no variable is being forced to serve two masters simultaneously.

Sentence 9 is the trickiest yet. First we paraphrase it as 'For any x that is a friend of a dog owner, x owns a dog which is also owned by a friend of x'. Using our stepping-stone tactic, this becomes:

$$\forall x\big[x \text{ is a friend of a dog owner} \to \\ x \text{ owns a dog which is owned by a friend of } x\big]$$

Breaking this down a bit more:

$$\forall x\big[\exists y(Fxy \land y \text{ is a dog owner}) \to \\ \exists y(Dy \land Oxy \land y \text{ is owned by a friend of } x)\big]$$

And a bit more:

$$\forall x\big[\exists y(Fxy \land \exists z(Dz \land Oyz)) \to \exists y(Dy \land Oxy \land \exists z(Fzx \land Ozy))\big]$$

And we are done!

29.3 Suppressed quantifiers

Logic can often help to get clear on the meanings of English claims, especially where the quantifiers are left implicit or their order is ambiguous or unclear. The clarity of expression and thinking afforded by FOL can give you a significant advantage in argument, as can be seen in the following takedown by British political philosopher Mary Astell (1666–1731) of her contemporary, the theologian William Nicholls. In Discourse IV: The Duty of Wives to their Husbands of his *The Duty of Inferiors towards their Superiors, in Five Practical Discourses* (London 1701), Nicholls argued that women are naturally inferior to men. In the preface to the 3rd edition of her treatise *Some Reflections upon Marriage, Occasion'd by the Duke and Duchess of Mazarine's Case; which is also considered,* Astell responded as follows:

> 'Tis true, thro' Want of Learning, and of that Superior Genius which Men as Men lay claim to, she [Astell] was ignorant of the *Natural Inferiority* of our Sex, which our Masters lay down as a Self-Evident and Fundamental Truth. She saw nothing in the Reason of Things, to make this either a Principle or a Conclusion, but much to the contrary; it being Sedition at least, if not Treason to assert it in this Reign.
>
> For if by the Natural Superiority of their Sex, they mean that *every* Man is by Nature superior to *every* Woman, which is the obvious meaning, and that which must be stuck to if they would speak Sense, it wou'd be a Sin in *any* Woman to have Dominion over *any* Man, and the greatest Queen ought not to command but to obey her Footman, because no Municipal Laws can supersede or change the Law of Nature; so that if the Dominion of the Men be such, the *Salique Law,*[2] as unjust as *English Men* have ever thought it, ought to take place over all the Earth, and the most glorious Reigns in the *English, Danish, Castilian*, and other Annals, were wicked Violations of the Law of Nature!
>
> If they mean that *some* Men are superior to *some* Women this is no great Discovery; had they turn'd the Tables they might have seen that *some* Women are Superior to *some* Men. Or had they been pleased to remember their Oaths of Allegiance and Supremacy, they might have known that *One* Woman is superior to *All* the Men in these Nations, or else they have sworn to very little purpose.[3] And it must not be suppos'd, that their Reason and Religion wou'd suffer them to take Oaths, contrary to the Laws of Nature and Reason of things.[4]

We can symbolize the different interpretations Astell offers of Nicholls' claim that men are superior to women: He either meant that every man

[2]The Salique law was the common law of France which prohibited the crown be passed on to female heirs.

[3]In 1706, England was ruled by Queen Anne.

[4]Mary Astell, *Reflections upon Marriage*, 1706 Preface, iii–iv, and Mary Astell, *Political Writings*, ed. Patricia Springborg, Cambridge University Press, 1996, 9–10.

is superior to every woman, i.e.,

$$\forall x(Mx \rightarrow \forall y(Wy \rightarrow Sxy))$$

or that some men are superior to some women,

$$\exists x(Mx \land \exists y(Wy \land Sxy)).$$

The latter is true, but so is

$$\exists y(Wy \land \exists x(Mx \land Syx)).$$

(some women are superior to some men), so that would be "no great discovery." In fact, since the Queen is superior to all her subjects, it's even true that some woman is superior to every man, i.e.,

$$\exists y(Wy \land \forall x(Mx \rightarrow Syx)).$$

But this is incompatible with the "obvious meaning" of Nicholls' claim, i.e., the first reading. So what Nicholls claims amounts to treason against the Queen!

Practice exercises

A. Using this symbolization key:

domain: all animals
Ax: $\_\_\_\_\__x$ is an alligator
Mx: $\_\_\_\_\__x$ is a monkey
Rx: $\_\_\_\_\__x$ is a reptile
Zx: $\_\_\_\_\__x$ lives at the zoo
Lx,y: $\_\_\_\_\__x$ loves $\_\_\_\_\__y$
a: Amos
b: Bouncer
c: Cleo

symbolize each of the following sentences in FOL:

1. If Cleo loves Bouncer, then Bouncer is a monkey.
   $Lc,b \rightarrow Mb$
2. If both Bouncer and Cleo are alligators, then Amos loves them both.
   $(Ab \land Ac) \rightarrow (La,b \land La,c)$

3. Cleo loves a reptile.
   $\exists x(Rx \land Lc,x)$
   Comment: this English expression is ambiguous; in some contexts, it can be read as a generic, along the lines of 'Cleo loves reptiles'. (Compare 'I do love a good pint'.)
4. Bouncer loves all the monkeys that live at the zoo.
   $\forall x((Mx \land Zx) \rightarrow Lb,x)$
5. All the monkeys that Amos loves love him back.
   $\forall x((Mx \land La,x) \rightarrow Lx,a)$
6. Every monkey that Cleo loves is also loved by Amos.
   $\forall x((Mx \land Lc,x) \rightarrow La,x)$
7. There is a monkey that loves Bouncer, but sadly Bouncer does not reciprocate this love.
   $\exists x(Mx \land Lx,b \land \neg Lb,x)$

B. Using the following symbolization key:

domain: all animals
Dx: $\_\_\_\_\__x$ is a dog
Sx: $\_\_\_\_\__x$ likes samurai movies
Lx,y: $\_\_\_\_\__x$ is larger than $\_\_\_\_\__y$
r: Rave
h: Shane
d: Daisy

symbolize the following sentences in FOL:

1. Rave is a dog who likes samurai movies.
   $Dr \land Sr$
2. Rave, Shane, and Daisy are all dogs.
   $Dr \land Dh \land Dd$
3. Shane is larger than Rave, and Daisy is larger than Shane.
   $Lh,r \land Ld,h$
4. All dogs like samurai movies.
   $\forall x(Dx \rightarrow Sx)$
5. Only dogs like samurai movies.
   $\forall x(Sx \rightarrow Dx)$
   Comment: the FOL sentence just written does not require that anyone likes samurai movies. The English sentence might suggest that at least some dogs *do* like samurai movies?
6. There is a dog that is larger than Shane.
   $\exists x(Dx \land Lx,h)$

7. If there is a dog larger than Daisy, then there is a dog larger than Shane.
   $\exists x(Dx \wedge Lxd) \rightarrow \exists x(Dx \wedge Lx,h)$
8. No animal that likes samurai movies is larger than Shane.
   $\forall x(Sx \rightarrow \neg Lx,h)$
9. No dog is larger than Daisy.
   $\forall x(Dx \rightarrow \neg Lx,d)$
10. Any animal that dislikes samurai movies is larger than Rave.
    $\forall x(\neg Sx \rightarrow Lx,r)$
    Comment: this is very poor, though! For 'dislikes' does not mean the same as 'does not like'.
11. There is an animal that is between Rave and Shane in size.
    $\exists x((Lb,x \wedge Lx,h) \vee (Lh,x \wedge Lx,r))$
12. There is no dog that is between Rave and Shane in size.
    $\forall x\big(Dx \rightarrow \neg\big[(Lb,x \wedge Lx,h) \vee (Lh,x \wedge Lx,r)\big]\big)$
13. No dog is larger than itself.
    $\forall x(Dx \rightarrow \neg Lx,x)$
14. Every dog is larger than some dog.
    $\forall x(Dx \rightarrow \exists y(Dy \wedge Lx,y))$
    Comment: the English sentence is potentially ambiguous here. I have resolved the ambiguity by assuming it should be paraphrased by 'for every dog, there is a dog smaller than it'.
15. There is an animal that is smaller than every dog.
    $\exists x\forall y(Dy \rightarrow Ly,x)$
16. If there is an animal that is larger than any dog, then that animal does not like samurai movies.
    $\forall x(\forall y(Dy \rightarrow Lx,y) \rightarrow \neg Sx)$
    Comment: I have assumed that 'larger than any dog' here means 'larger than every dog'.

C. Using the symbolization key given, translate each English-language sentence into FOL.

domain: candies
Cx: $\underline{\quad\quad}_x$ has chocolate in it.
Mx: $\underline{\quad\quad}_x$ has marzipan in it.
Sx: $\underline{\quad\quad}_x$ has sugar in it.
Tx: Boris has tried $\underline{\quad\quad}_x$.
Bx,y: $\underline{\quad\quad}_x$ is better than $\underline{\quad\quad}_y$.

1. Boris has never tried any candy.
2. Marzipan is always made with sugar.

3. Some candy is sugar-free.
4. The very best candy is chocolate.
5. No candy is better than itself.
6. Boris has never tried sugar-free chocolate.
7. Boris has tried marzipan and chocolate, but never together.
8. Any candy with chocolate is better than any candy without it.
9. Any candy with chocolate and marzipan is better than any candy that lacks both.

D. Using the following symbolization key:

domain: people and dishes at a potluck
Rx: $\_\_\_\_\_\__x$ has run out.
Tx: $\_\_\_\_\_\__x$ is on the table.
Fx: $\_\_\_\_\_\__x$ is food.
Px: $\_\_\_\_\_\__x$ is a person.
Lx,y: $\_\_\_\_\_\__x$ likes $\_\_\_\_\_\__y$.
e: Eli
f: Francesca
g: the guacamole

symbolize the following English sentences in FOL:

1. All the food is on the table.
   $\forall x(Fx \rightarrow Tx)$
2. If the guacamole has not run out, then it is on the table.
   $\neg Rg \rightarrow Tg$
3. Everyone likes the guacamole.
   $\forall x(Px \rightarrow Lx,g)$
4. If anyone likes the guacamole, then Eli does.
   $\exists x(Px \wedge Lx,g) \rightarrow Le,g$
5. Francesca only likes the dishes that have run out.
   $\forall x\big[(Lf,x \wedge Fx) \rightarrow Rx\big]$
6. Francesca likes no one, and no one likes Francesca.
   $\forall x\big[Px \rightarrow (\neg Lf,x \wedge \neg Lx,f)\big]$
7. Eli likes anyone who likes the guacamole.
   $\forall x((Px \wedge Lx,g) \rightarrow Le,x)$
8. Eli likes anyone who likes the people that he likes.
   $\forall x\big[\big(Px \wedge \forall y[(Py \wedge Le,y) \rightarrow Lx,y]\big) \rightarrow Le,x\big]$
9. If there is a person on the table already, then all of the food must have run out.
   $\exists x(Px \wedge Tx) \rightarrow \forall x(Fx \rightarrow Rx)$

E. Using the following symbolization key:

domain: people
Dx: $\_\_\_\_\__x$ dances ballet.
Fx: $\_\_\_\_\__x$ is female.
Mx: $\_\_\_\_\__x$ is male.
Cx,y: $\_\_\_\_\__x$ is a child of $\_\_\_\_\__y$.
Sx,y: $\_\_\_\_\__x$ is a sibling of $\_\_\_\_\__y$.
e: Elmer
j: Jane
p: Patrick

symbolize the following sentences in FOL:

1. All of Patrick's children are ballet dancers.
   $\forall x(Cx,p \rightarrow Dx)$
2. Jane is Patrick's daughter.
   $Cj,p \wedge Fj$
3. Patrick has a daughter.
   $\exists x(Cx,p \wedge Fx)$
4. Jane is an only child.
   $\neg\exists x Sx,j$
5. All of Patrick's sons dance ballet.
   $\forall x\big[(Cx,p \wedge Mx) \rightarrow Dx\big]$
6. Patrick has no sons.
   $\neg\exists x(Cx,p \wedge Mx)$
7. Jane is Elmer's niece.
   $\exists x(Sx,e \wedge Cj,x \wedge Fj)$
8. Patrick is Elmer's brother.
   $Sp,e \wedge Mp$
9. Patrick's brothers have no children.
   $\forall x\big[(Sp,x \wedge Mx) \rightarrow \neg\exists y\, Cy,x\big]$
10. Jane is an aunt.
    $Fj \wedge \exists x(Sx,j \wedge \exists y Cy,x)$
11. Everyone who dances ballet has a brother who also dances ballet.
    $\forall x\big[Dx \rightarrow \exists y(My \wedge Sy,x \wedge Dy)\big]$
12. Every woman who dances ballet is the child of someone who dances ballet.
    $\forall x\big[(Fx \wedge Dx) \rightarrow \exists y(Cx,y \wedge Dy)\big]$

CHAPTER 30

Ambiguity

In chapter 16 we discussed the fact that sentences of English can be ambiguous, and pointed out that sentences of TFL are not. One important application of this fact is that the structural ambiguity of English sentences can often, and usefully, be straightened out using different symbolizations. One common source of ambiguity is *scope ambiguity*, where the English sentence does not make it clear which logical word is supposed to be in the scope of which other. Multiple interpretations are possible. In FOL, every connective and quantifier has a well-determined scope, and so whether or not one of them occurs in the scope of another in a given sentence of FOL is always determined.

For instance, consider the English idiom,

1. Everything that glitters is not gold.

If we think of this sentence as of the form 'every F is not G' where Fx symbolizes '$\underline{\quad\quad}_x$ glitters' and Gx is '$\underline{\quad\quad}_x$ is not gold', we would symbolize it as:

1. $\forall x(Fx \rightarrow \neg Gx)$,

in other words, we symbolize it the same way as we would 'Nothing that glitters is gold'. But the idiom does not mean that! It means that one should not assume that just because something glitters, it is gold; not everything that appears valuable is in fact valuable. To capture the actual meaning of the idiom, we would have to symbolize it instead as we would 'Not everything that glitters is gold', i.e., in the following way:

1. $\neg\forall x(Fx \rightarrow Gx)$

Compare the first of these with the previous symbolization: again we see that the difference in the two meanings of the ambiguous sentence lies in whether the '$\neg$' is in the scope of the '$\forall$' (in the first symbolization) or '$\forall$' is in the scope of '$\neg$' (in the second).

Of course we can alternatively symbolize the two readings using existential quantifiers as well:

1. $\neg\exists x(Fx \wedge Gx)$
2. $\exists x(Fx \wedge \neg Gx)$

In chapter s:SymbolisingSimpleFOL we discussed how to symbolize sentences involving 'only'. Consider the sentence:

2. Only young cats are playful.

According to our schema, we would symbolize it this way:

1. $\forall x(Px \rightarrow (Yx \wedge Cx))$

The meaning of this sentence of FOL is something like, 'If an animal is playful, it is a young cat'. (Assuming that the domain is animals, of course.) This is probably not what's intended in uttering sentence 2, however. It's more likely that we want to say that old cats are not playful. In other words, what we mean to say is that if something is a cat and playful, it must be young. This would be symbolized as:

1. $\forall x((Cx \wedge Px) \rightarrow Yx)$

There is even a third reading! Suppose we're talking about young animals and their characteristics. And suppose you wanted to say that of all the young animals, only the cats are playful. You could symbolize this reading as:

1. $\forall x((Yx \wedge Px) \rightarrow Cx)$

Each of the last two readings can be made salient in English by placing the stress appropriately. For instance, to suggest the last reading, you would say 'Only young *cats* are playful', and to get the other reading you would say 'Only *young* cats are playful'. The very first reading can be indicate by stressing both 'young' and 'cats': 'Only *young cats* are playful' (but not old cats, or dogs of any age).

In chapter 29 we discussed the importance of the order of quantifiers. This is relevant here because, in English, the order of quantifiers is sometimes not completely determined. When both universal ('all') and existential ('some', 'a') quantifiers are involved, this can result in scope ambiguities. Consider:

3. Everyone went to see a movie.

This sentence is ambiguous. In one interpretation, it means that there is a single movie that everyone went to see. In the other, it means that everyone went to see some movie or other, but not necessarily the same one. The two readings can be symbolized, respectively, by

1. $\exists x(Mx \land \forall y(Py \to Sy, x))$
2. $\forall y(Py \to \exists x(Mx \land Sy, x))$

We assume here that the domain contains (at least) people and movies, and the symbolization key,

Py: $\_\_\_\_\__y$ is a person,
Mx: $\_\_\_\_\__x$ is a movie
Sy, x: $\_\_\_\_\__y$ went to see $\_\_\_\_\__x$.

In the first reading, we say that the existential quantifier has *wide scope* (and its scope contains the universal quantifier, which has *narrow scope*), and the other way round in the second.

Practice exercises

A. Each of the following sentences is ambiguous. Provide a symbolization key for each, and symbolize all readings.

1. Noone likes a quitter.
2. CSI found only red hair at the scene.

PART VI

FOL-Semantics

CHAPTER 31

Sets and Relations

We have discussed the syntax of the language of first-order language and, using our intuitive understanding of the vocabulary of the language, proposed a guideline for symbolizing English sentences in the language of FOL. We now turn to the semantics, that is, we now want to spell out truth-conditions for sentences of the language. In TFL the basic building blocks of the language were atomic sentences that were declared true or false by a valuation. In FOL the basic building blocks are names and predicates, which cannot be true or false as they are not sentences. But if their *semantic value* is not a truth value, what is it?

31.1 Sets

In symbolizing English sentences in FOL we were required to specify a domain, that is, the collection of objects the quantifiers range over. For example, an adequate domain for symbolizing the sentence

- All philosophers are fashionable

might consist of people, e.g., the collection of all living people in the world. From now on we call a collection of objects (people, things, numbers,...) a SET and we will take domains to be sets. The domain we just alluded to would thus be the set consisting of all living people.

Any collection of objects form a set and the objects forming the set are the MEMBERS or ELEMENTS of the set. For example, we can form the set consisting of Poppy and Naomi, which we convey by writing

$$\{\text{Poppy, Naomi}\}.$$

The order in which we list the objects is irrelevant, that is, by writing

$$\{\text{Naomi, Poppy}\}$$

we denote the (same) set consisting of Poppy and Naomi. More generally, sets are identical if and only if they have the same members. So, in particular, as mentioned above we obtain:

$$\{\text{Poppy, Naomi}\} = \{\text{Naomi, Poppy}\}.$$

When specifying a domain we often deal with many objects, and it would be cumbersome and often impossible to list them all in order to specify the corresponding set: we cannot list all the living people in the world...

In such a case we use an alternative method for denoting sets: we single out all objects that have a specific property.[1] For example the set consisting of all living people will be denoted by:

$$\{o \mid o \text{ is a living person}\}.$$

Similarly, the set consisting of all students of the UoB is denoted by

$$\{o \mid o \text{ is a student of the UoB}\},$$

and the set consisting of all cats and dogs by

$$\{o \mid o \text{ is a cat or } o \text{ is a dog}\}.$$

Given this alternate way of denoting sets it is no longer immediate whether a given object is a member of a set. To convey that a given object o is a member of a set S we shall write

$$o \in S.$$

Recall that members of a set are also called the elements of the set. This may be helpful for remembering the meaning of $\in$. For example, since Tom is a cat we have

$$\text{Tom} \in \{o \mid o \text{ is a cat or } o \text{ is a dog}\}.$$

[1]One needs to spell this idea out with some care for otherwise Russell's paradox ensues.

The latter just tells us that Tom is a member (element) of the set consisting of all cats and dogs.

It is important to remember that sets are identical if and only if they have the same members. It doesn't matter which property was used to single out the elements of a set. This means that

$$\{o \mid o \text{ is an animal with a heart}\}$$

and

$$\{o \mid o \text{ is an animal with kidneys}\}$$

denote the same set, as—at least according to the traditional example—the animals that have a heart coincides with the animals that have kidneys, i.e.,

$$\{o \mid o \text{ is an animal with a heart}\} = \{o \mid o \text{ is an animal with kidneys}\}$$

This does not *mean* that the two properties are identical, but that any possible difference between the two properties gets lost once we move to sets.

As just discussed, sets are collections of objects, but weirdly there is also a set that has no members, i.e., a set with no elements. This set is called the EMPTY SET and denoted by $\emptyset$. There are many different ways to specify the empty set. For example,

$$\{o \mid o \neq o\} = \emptyset.$$

31.2 Relations

Suppose we have specified as our domain Dom the set of all animals, i.e.,

$$\mathsf{Dom} = \{o \mid o \text{ is an animal}\}.$$

How are we to understand a predicate like '$______x$is a cat'? The predicate will pick out those animals from the domain that are cats, that is, the set of all cats. As we will see, the sentence 'Tom is a cat' when symbolized in FOL will be true iff the object picked out by the name 'Tom', that is, Tom is a member of the set

$$\{o \mid o \in \mathsf{Dom} \text{ and } o \text{ is a cat}\}.$$

However, '$______x$is a cat' is a one-place predicate, what about a two-place predicate like '$______x$is bigger than$______y$'? Presumably the predicate applies to Tom and Jerry, as Tom is bigger than Jerry. Importantly, however it is no longer sufficient to merely consider the objects the predicate applies to, but we also need to specify their order: '$______x$is bigger than$______y$' applies to Tom and Jerry *in that order*; it does not apply to Tom and Jerry in the order Jerry, Tom—Jerry is not bigger than Tom.

Summing up, two place predicates such as '$______x$is bigger than$______y$' or '$______x$loves$______y$' apply to ordered pairs, that is, pairs of object in a specific order. For example, '$______x$is bigger than$______y$' applies to the ordered pair ⟨Tom, Jerry⟩. The ORDERED PAIR ⟨Tom, Jerry⟩ is different to the set {Tom, Jerry}, but also different to the ordered pair ⟨Jerry, Tom⟩. ⟨Tom, Jerry⟩ tells us that Tom is the first object and Jerry is the second object in contrast ⟨Jerry, Tom⟩ where it is the other way around and the set {Tom, Jerry} for which the order does not matter (recall that {Tom, Jerry} = {Jerry, Tom}).

Now, that we know what an ordered pairs are we can form sets of ordered pairs and such thats, that is, sets that only have ordered pairs as their elements are called BINARY RELATIONS. While, as discussed the one-place predicate like 'is a cat' picks out the set of all cats in the domain, the two place predicate picks out a binary relation. More precisely, it picks out the set

$$\{\langle o_1, o_2\rangle \mid o_1 \text{ is bigger than } o_2\}.$$

where we always assume that both o_1 and o_2 members of the domain. To recap: one-place predicates pick out sets and two-place predicates pick out binary relations. For example, the set

$$\{\langle \text{Oslo, Norway}\rangle, \langle \text{Stockholm, Sweden}\rangle, \langle \text{Helsinki, Finland}\rangle, \langle \text{Copenhague, Denmark}\rangle\}$$

is a binary relation picked out by the two-place predicate '$______x$is a capital of$______y$' with a domain, which, e.g., is a set that has Scandinavian cities and countries as its elements.

In Part 25 we learned that there are not only two-place predicates but also three-place predicates, four-place predicates and, indeed, at least in principle predicates for any number of argument places. What does a three place predicate pick out? It picks out a TERNARY RELATION. While a binary relation is a set of ordered pairs, a ternary relation is a set of ordered triples. An ordered triple such as ⟨Pooh, Eeyore, The trampoline⟩ consists of the three-objects Pooh, Eeyore, The trampoline *in that particular order*. The three-place predicate

'_____$_x$bounces with_____$_y$on_____$_z$' picks out the following ternary relation:

$$\{\langle o_1, o_2, o_3\rangle \mid o_1 \text{ bounces with } o_2 \text{ on } o_3\}.$$

So intuitively the sentence

- ▹ Pooh bounces with Eeyore on the trampoline

is true if and only if

$$\langle \text{Pooh, Eeyore, The trampoline}\rangle \in \{\langle o_1, o_2, o_3\rangle | o_1 \text{ bounces with } o_2 \text{ on } o_3\}.$$

In the same way that a three-place predicate picks out a ternary relation, an n-place predicate picks out an *n*-ARY RELATION, that is, a set of n-tuples $\langle o_1, o_2, \ldots, o_n\rangle$.

31.3 Models

We defined a VALUATION in TFL as any assignment of truth and falsity to atomic sentences. In FOL, we are going to define an MODEL. Sentences of FOL will be true or false *on a model*.

A model provides

- ▹ specifies a non-empty set as the domain, Dom;
- ▹ specifies an interpretation of the FOL-names and FOL-predicates.
  - assigns exactly one object of the domain to every name (specifies the referent of the name);
  - assigns a subset of the domain to any one-placed predicate.
  - assigns a binary relation on the domain to every two-placed predicate.
  - assigns a ternary relation on the domain to every three-placed preciate.
  - ...

For a model, $\mathcal{M}$, we denote its assignment to η as $[\![\eta]\!]_{\mathcal{M}}$. We will often drop the explicit reference to $\mathcal{M}$ and simply describe a model by providing a domain and $[\![\eta]\!]$ for each name and predicate.

A model thus provides us with all the necessary information for interpreting the predicates and names, as well as, information about the range of the quantifiers. This will be enough to determine the truth value of all sentences of FOL, which we will turn to in 32

Importantly, we are not guaranteed that the denotation of an n-place predicate is nonempty, i.e., it is possible that $\emptyset$ is assigned to a predicate. In general, a convenient way of specifying a model is to supplement a given symbolization key (cf. Part V) with the relevant model of the names and predicates.

Suppose we have just symbolized the sentence

▹ Pooh bounces with Eeyore on the trampoline

using the symbolization key;

domain: Animals and trampolines
p: Pooh
e: Eeyore
t: The trampoline
$Bxyz$: $\_\_\_\_\__x$ bounces with $\_\_\_\_\__y$ on $\_\_\_\_\__z$

Taking the symbolization key as a guideline we can construct an appropriate model as follows:

Dom: $\{o \mid o$ is an animal or o is trampolines$\}$
$[\![p]\!]$: *Pooh*
$[\![e]\!]$: *Eeyore*
$[\![t]\!]$: *The trampoline*
$[\![B]\!]$: $\{\langle o_1, o_2, o_3\rangle \mid o_1$ bounces with o_2 on $o_3\}$

Symbolization keys and models have similar components. Indeed we will sometimes be lazy and simply assume the model to be suitably specified by the symbolization key. However, symbolization keys and models are very different things: a model of the vocabulary of FOL inolves specifying objects, sets and relations while in the symblization key we specify which FOL-symbols are to stand for the English names and predicates of a sentence, e.g, which FOL-name is to stand for the name ‘Pooh’. In the model we are saying that the FOL-name ‘p’ picks out Pooh (the bear) or, in some different model, a different object of the domain.

We may then ask whether the FOL-sentence $Bpet$ is true in the model so specified. The answer is that it is true iff, indeed, Pooh bounces with Eeyore on the trampoline, that is, iff

$$\langle \text{Pooh, Eeyore, The trampoline} \rangle \in \{\langle o_1, o_2, o_3 \rangle \mid o_1 \text{ bounces with } o_2 \text{ on } o_3\}.$$

The model we have given above is arguably the *intended model*, that is, the model that fits the symbolization key and our understanding of the English sentence. But there are many alternative models for the FOL-sentence $Bpet$, the only thing which we require is that the model specifies a domain, that the FOL-name ‘p’, ‘e’, and ‘t’ pick out objects from the domain, and that the three-place B is assigned a ternary relation.

Here are two other models that could be provided.

$$\begin{aligned} \text{Dom}&: \{\textit{Alice}, \textit{Bob}, \textit{Carly}\}\\ [\![p]\!]&: \textit{Alice}\\ [\![e]\!]&: \textit{Bob}\\ [\![t]\!]&: \textit{Carly}\\ [\![B]\!]&: \{\langle \textit{Alice}, \textit{Bob}, \textit{Carly} \rangle\} \end{aligned}$$

or

$$\begin{aligned} \text{Dom}&: \{1, 2, 3, 4\}\\ [\![p]\!]&: 1\\ [\![e]\!]&: 2\\ [\![t]\!]&: 1\\ [\![B]\!]&: \{\langle 1, 2, 3 \rangle, \langle 3, 2, 1 \rangle\} \end{aligned}$$

As in the case TFL: once we move to the language of FOL we forget about the particular English sentences and their meaning.

It is often convenient to present a model *diagrammatically*.

Suppose we want to consider just a single two-place predicate, Rxy and we wish to consider the following model:

$$\begin{aligned} \text{Dom}&: \{1, 2, 3, 4\}\\ [\![R]\!]&: \{\langle 1, 1 \rangle, \langle 1, 2 \rangle, \langle 2, 3 \rangle, \langle 3, 4 \rangle, \langle 4, 1 \rangle, \langle 1, 3 \rangle\} \end{aligned}$$

Then we can equivalently offer

$$\text{Dom}: \{x \mid x \text{ is a number in the diagram}\}$$

$[\![R]\!]$: $\{\langle x, y\rangle \mid$ there is an arrow from x to y in the diagram$\}$

along with the diagram:

$$\begin{tikzcd}
1 \arrow[loop left] \arrow[r] \arrow[dr] & 2 \arrow[d] \\
4 \arrow[u] & 3 \arrow[l]
\end{tikzcd}$$

It must be abundantly clear what the domain is, what the names refer to, and which objects satisfy the predicates.

CHAPTER 32

Truth in FOL

We know what models are. Since, among other things, they tell us how we are to understand FOL-predicates and FOL-names, they will provide us with an account of the truth of atomic sentences. We already hinted at this account in the previous chapter. However, we must also present a detailed account of what it is for an arbitrary FOL sentence to be true or false in a model.

But we defined what a sentence was by first specifying what a formula is. Formulas like Ex aren't the sorts of things that are true or false on a model. Only *sentences* are true or false. But if we provide extra information we can determine the truth of Ex: we need to specify what object x stands for. This is done by using a variable assignment:

> A VARIABLE ASSIGNMENT assigns to each variable an object of the domain.

For example, if our domain is the set of characters in Winnie the Pooh, a variable assignment α might be:

	x	y	z	x_1	...
α :	*Pooh*	*Eeyore*	*Pooh*	*Tigger*	...

We often talk and compare different variable assignments so it's convenient to have names for them and we use Greek letters α, β, γ for this purpose.

Now consider a model where the domain Dom is the set of natural number and the predicate E picks out the even numbers:

$$\begin{aligned}\mathsf{Dom} &= \{o \mid o \text{ is a natural number}\}\\ [\![E]\!] &= \{o \mid o \text{ is an even number}\}\end{aligned}$$

Then Ex will be true in this model under a variable assignment α that assigns, say, 4 to x—we write $\alpha(x) = 4$—but false under a variable assignment β that assigns 3 to x, i.e., $\beta(x) = 3$.

We know from §26 that there are three kinds of formulas in FOL:

- atomic formulas
- formulas whose main logical operator is a sentential connective
- formulas whose main logical operator is a quantifier

We need to explain truth for all three kinds of formula.

We will provide a completely general explanation in this section. However, to try to keep the explanation comprehensible, we will, at several points, use the following model:

$$\begin{aligned}\mathsf{Dom}&: \{o \mid o \text{ is a person born before before 2000}\textsc{ce}\})\\ [\![a]\!]&: \textit{Aristotle}\\ [\![b]\!]&: \textit{Beyoncé}\\ [\![P]\!]&: \{o \mid o \text{ is a philosopher}\}\\ [\![R]\!]&: \{\langle o_1, o_2\rangle \mid o_1 \text{ was born before } o_2\}\end{aligned}$$

This will be our *go-to example* in what follows.

32.1 Atomic formulas

Atomic formulas are things like Px, Pb or Rax.

To see whether an atomic sentence like 'Pb' is true in a model, we need to check whether the object that is picked out by the name 'b' is a member of the set the model assigns to 'P'. Looking at our *go-to example* this means that 'Pb' is true in the model iff

$$\textit{Beyoncé} \in \{o \mid o \text{ is a philosopher}\}.$$

Since Beyoncé is not a philosopher, i.e.,

$$\textit{Beyoncé} \notin \{o \mid o \text{ is a philosopher}\}$$

'Pb' is false in that model.

What about Px? This reads something like 'they are a philosopher'. The question is who 'they' refers to, or in the logic terms: who x is. This depends on a variable assignment. A variable assignment needs to give an object in our domain for the variable x. For example, a variable assignment α assign Beyoncé to x, i.e., $\alpha(x)$=Beyoncé and since Beyoncé is not in the model of 'P' 'Px' is not true in the model under the variable assignment α. However, since Aristotle is a philosopher 'Px would be true in the model under a variable assignment β, where

$$\beta(x) = \text{Aristotle}.$$

A variable assignment doesn't need to specify one of the objects that are named, it can give us anyone in our domain, e.g. Queen Elizabeth II. Under the variable assignment which assigns x Queen Elizabeth II, 'Px' is false: Queen Elizabeth II is not a philosopher.

Likewise, on this model, Rab is true iff the object named by a was born before the object named by b. Well, Aristotle was born before Beyoncé. So Rab is true. Equally, Raa is false: Aristotle was not born before Aristotle. How about Rax? Well what does our variable assignment specify for x? If we have a variable assignment where x is Queen Elizabeth II, then Rax is true: Aristotle was born before Queen Elizabeth II.

Dealing with atomic formulas, then, is very intuitive. The only thing to be attentive to is the difference between names or variables:

A TERM is either a name or a variable.

- If t is a name, then $[\![t]\!]^\alpha$ is just $[\![t]\!]$, the object of the domain assigned to the FOL-name t by the model.
- If t is a variable, then $[\![t]\!]^\alpha$ the object of the domain assigned to the variable t by the variable assignment; that is, it is $\alpha(t)$.

With this final bit of terminology out of way we can state the truth conditions for atomic formuals. Let R be an n-place predicate and $t_1, t_2 \ldots t_n$ terms. Then

$Rt_1t_2 \ldots t_n$ is true in a model under a variable assignment α **iff**

$$\langle [\![t_1]\!]^\alpha, \ldots, [\![t_n]\!]^\alpha \rangle \in [\![R]\!]$$

From this definition of the truth of atomic formulas we can tell that an assignment only plays an important role if one of the terms is a variable. If all terms are name then an atomic formula is true in a model under a variable assignment **iff** it is true in the model under **every** variable assignment. We will see that that is the case for sentences more generally: the choice of the variable assignment is only relevant to determine whether a formulas with free variables is true in a model (under the chosen variable assignment). For this reason we sometimes do not mention the variable assignment when we talk about the truth of sentences. However, as we shall see the variable assignment will still have a role to play when we talk about the truth and falsity of quantified sentencesd.

32.2 Sentential connectives

We saw in §26 that FOL formulas can be built up from simpler ones using the truth-functional connectives that were familiar from TFL. The rules governing these truth-functional connectives are *exactly* the same as they were when we considered TFL. Here they are:

$\neg X$ is true in a model under a variable assignment **iff** X is false in that model under that variable assignment

$(X \wedge Y)$ is true in a model under a variable assignment **iff** both X and Y is true in that model under that variable assignment

$(X \vee Y)$ is true in a model under a variable assignment **iff** either X is true or Y is true in that model under that variable assignment

$(X \rightarrow Y)$ is true in a model under a variable assignment **iff** either X is false or Y is true in that model under that variable assignment

This is just another presentation of the truth rules we gave for the connectives in TFL; it just does so in a slightly different way. Some examples will probably help to illustrate the idea. On our go-to model:

- 'Pa' is true

- '$Rab \wedge Pb$' is false because, although 'Rab' is true, 'Pb' is false
- '$\neg Pa$' is false
- '$Pa \wedge \neg(Pb \wedge Rab)$' is true, because '$Pa$' is true and '$Pb$' is false, so '$Pb \wedge Rab$' is false, thus '$\neg(Pb \wedge Rab)$' is also true.

Make sure you understand these examples.

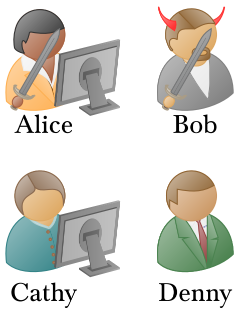
We can also carry variable assignments around with us. Consider a variable assignment which assigns David Hume to x. Then

- 'Px' is true under this variable assignment: David Hume was a philosopher
- 'Bxa' is false under this variable assignment: David Hume was born after Aristotle
- '$Px \rightarrow Bxa$' is false under this variable assignment: 'Px' is true and 'Bxa' is false, so by our rule for $\rightarrow$, '$Px \rightarrow Bxa$' is false.

32.3 When the main logical operator is a quantifier

The exciting innovation in FOL, though, is the use of *quantifiers*.

Consider the following model:



Dom: $\{Alice, Bob, Cathy, Denny\}$

$[\![H]\!]$: $\{x \mid x \text{ has horns}\} = \{Bob\}$

$[\![S]\!]$: $\{x \mid x \text{ is carrying a sword}\} = \{Alice, Bob\}$

$[\![C]\!]$: $\{x \mid x \text{ has a computer}\} = \{Alice, Cathy\}$

Is '$\exists x Sx$' true in this model under a variable assignment? Intuitively this should be true iff there is an object in the domain which has a sword and since $[\![S]\!] = \{Alice, Bob\}$ should indeed come out as true. We can directly spell out these intuitive truth conditions by using the notion of an v-variant of the initial variable assignment:

Before we give a general definition of whether a formula is true or false *under a variable assignment* in a give model, we need to introduce one further useful bit of terminology.

> Let β be a variable assignment and v a variable. Then the variable assignment $\beta[v:o]$ is a variable assignment that assigns the same objects to the variables than β except possibly to the variable v to which it assigns o. $\beta[v:o]$ is called an v-VARIANT of β.

Notice that v is metavariable for variables of FOL, so that it can be done for any variable. For example:

	x	y	z	x_1	...
α:	*Alice*	*Cathy*	*Cathy*	*Bob*	...
$\alpha[y:Bob]$:	*Alice*	*Bob*	*Cathy*	*Bob*	...

Variants of variable assignments will be important for stating the truth conditions of quantified formulas.

We can then say:

> '$\exists x Sx$' true in this model under the variable assignment α **iff** there exists an object o of the domain, i.e. $o \in \mathsf{Dom}$, such that 'Sx' is true in the model under the variable assignment $\alpha[x:o]$.

Then, to show that '$\exists x Sx$' true in the model under an arbitrary variable assignment it suffices to find an object in the domain, say, *Alice* such that 'Sx' is true under the variable assignment $\alpha[x:Alice]$ (we can also denote this variable assignment by $x \mapsto$ Alice). Under this variable assignment, Sx is true: Since Alice does have a sword, i.e., $Alice \in [\![S]\!]$, Sx is indeed true under this assignment. This guaranteed the truth of '$\exists x Sx$': There is a choice of an object in our domain for x under which Sx is true.

What about $\forall x Sx$? Intuitively, this should be true iff every object in the domain which has a sword. Appealing to the x-variant of out initial variable assignment we can again spell out the truth conditions in a straightforward:

> '$\forall x Sx$' true in this model under the variable assignment α **iff** for every object o of the domain 'Sx' is true in the model under the variable assignment $\alpha[x:o]$.

Since *Denny* $\notin [\![S]\!] = \{Alice, Bob\}$ '$\forall x Sx$' to comes out false in this model: 'Sx' is not true in the model under the variable assignment $\alpha[x{:}Denny]$.

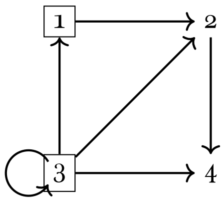
What about $\forall x(Sx \rightarrow (Hx \vee Cx))$? To check whether it is true we need to check whether '$Sx \rightarrow (Hx \vee Cx)$ is true for every choice of object that is assigned to x:

x	$(Sx$	$\rightarrow$	$(Hx$	$\vee$	$Cx))$
Alice	T	T	F	T	T
Bob	T	T	T	T	F
Cathy	F	T	F	T	T
Denny	F	T	F	F	F

So '$Sx \rightarrow (Hx \vee Cx)$' is true in the model under every assignment $\alpha[x{:}o]$ for every object $o \in$ Dom. And so $\forall x(Sx \rightarrow (Hx \vee Cx))$ is true is true in the model (under every variable assignment).

We have to tread more carefully once we start having multiple quantifiers. Let's walk through some cases.

Consider a new model:



Dom: $\{1,2,3,4\}$
$[\![R]\!]$: $\{\langle o_1, o_2\rangle |$There is an arrow from o_1 to o_2 in the diagram$\}$.
$[\![S]\!]$: $\{o \mid$ There is a square around o in the diagram$\}$.

Is $\exists x \forall y Rxy$ true? Intuitively, the sentence is true iff there is an object in the domain from which there is an arrow to all other arrows. Now, let's work through the example step by step and see whether we get the intuitive outcome.

1. '$\exists x \forall y Rxy$' is true in the model under an (arbitrary) variable assignment α iff there is a number $n \in \{1,2,3,4\}$ such '$\forall y Rxy$' is true in the model under the variable assignment $\alpha[x{:}n]$.

    ▹ We choose n to be the number 3 (with some foresight)

2. '$\forall y Rxy$' is true in the model under the variable assignment $\alpha[x:3]$ iff Rxy is true in the model under the variable assignment $\alpha[x:3][y:m]$ for all $m \in \{1,2,3,4\}$.

   ▹ $\alpha[x:3][y:m]$ looks very complicated; it is a y-variant of an x-variant... But all it says is that the variable y gets assigned the number m, i.e., $y \mapsto m$ and the variable x gets assigned the number 3, i.e., $x \mapsto 3$.

3. 'Rxy' is true in the model under the variable assignment $\alpha[x:3][y:m]$ for all $m \in \{1,2,3,4\}$ iff there is an arrow from 3 to every number in the set $\{1,2,3,4\}$. In that case we have that 'Rxy' is true in the model under the variable assignment $\alpha[x:3][y:1]$, the variable assignment $\alpha[x:3][y:2]$, the variable assignment $\alpha[x:3][y:3]$ and also $\alpha[x:3][y:4]$.

4. Since there is an arrow from 3 to every number we can conclude that '$\exists x \forall y Rxy$' is true in the model

What about '$\exists x \exists y(Rxy \wedge Ryx)$'? To show that it is true we will want to choose an object o that we can assign to x such that '$\exists y(Rxy \wedge Ryx)$' is true uunder the assignment $\alpha[x:o]$. Let's consider $x \mapsto 3$ (again I'm using my foresight of what will come to choose carefully). Now is '$\exists y(Rxy \wedge Ryx)$' true under the variable assignment $x \mapsto 3$ (i.e., $\alpha[x:3]$? We need to find y-variant of this assignment which chooses an object such that '$Rxy \wedge Ryx$' is true. Consider $y \mapsto 3$. We now have a variable assignment $\alpha[x:3][y:3]$. x and y are different variables but there's nothing stopping them being assigned the same object. And we can then consider whether $Rxy \wedge Ryx$ is true under this model. Well, 'Rxy' is true: 3 does have an arrow to 3. And 'Ryx' is also true: 3 does have an arrow to 3. So by our clause for $\wedge$, '$Rxy \wedge Ryx$' is true under the variable assignment $\alpha[x:3][y:3]$. And so '$\exists y(Rxy \wedge Ryx)$' is true under the variable assignment $\alpha[x:3]$. And so $\exists x \exists y(Rxy \wedge Ryx)$ is true in this model.

One more example: $\forall x(Sx \to \exists y Rxy)$? To check this is true we will need to check for every number n whether '$Sx \to \exists y Rxy$' is true under the variable assignment $\alpha[x:n]$ for each n in the domain.

x	Sx	$\exists y Rxy$	$Sx \to \exists y Rxy$
1	T	?	?
2	F		T
3	T	?	?
4	F		T

So we need to check whether $\exists y Rxy$ is true under the variable assignments $x \mapsto 1$ and $x \mapsto 3$. In the two other cases the conditional will be trivially true

Consider a variable assignment where $x \mapsto 1$. We can find an object for y such that 'Rxy' is true under that variable assignment, namely, $y \mapsto 2$. Since there is an arrow from 1 to 2, 'Rxy' is true in the variable assignment $\alpha[x:1][y:2]$. Thus '$\exists y Rxy$' is true under the variable assignment $\alpha[x:1]$. For $x \mapsto 3$ we can also find an assignment for y such that 'Rxy' is true (which one?) and thus '$\exists y Rxy$' is also true under the variable assignment $\alpha[x:3]$.

	Sx	$\exists y Rxy$	$Sx \rightarrow \exists y Rxy$
$x \mapsto 1$	T	T	T
$x \mapsto 2$	F	?	T
$x \mapsto 3$	T	T	T
$x \mapsto 4$	F	?	T

So $\forall x(Sx \rightarrow \exists y Rxy)$ is true. Informally we might say this as: for every number that has a square around it has an arrow going out of it.

One final example: $\forall x \forall y Rxy$. To check this we will need to consider all choices for x and all choices for y and check Rxy is true on all of them. There are 16 such choices. But we won't have to go through them all: it'll be false. Consider $x \mapsto 1$ and $y \mapsto 4$. Rxy is false under this variable assignment: there is no arrow from 1 to 4. Thus $\forall y Rxy$ is false on the variable assignment $\alpha[x:1]$. And so $\forall x \forall y Rxy$ is false in the model under the variable assignment α. However, α itself played no role in the evaluation of $\forall x \forall y Rxy$. All that mattered was the objects that were assigned to x and y, and for x and y we considered the x-variants/y-variants. So $\forall x \forall y Rxy$ is not only true under the variable assignment α in this model but true under all variable assignments. This goes back to what we said at the end of Section 32.1: the particular choice of an variable assignment is only relevant if we are evaluating the truth of formulas with free variables. $\forall x \forall y Rxy$ is a sentence, as it has no free variables.

More generally, if we have a quantifier like $\exists x$ or $\forall x$ we ignore whatever our original variable assignment told us about x but consider suitable x-variants. So when all our variables are bound by quantifiers, all the original components of our variable assignment are ignored. To summarise: Sentences are simply true or false in models, variable assignments don't matter.

Now for our formal definition:

$\forall v X$ is true in the model under a variable assignment α
iff X is true in the model under the variable assignment $\alpha[v:o]$ for every $o \in \mathsf{Dom}$.

$\exists v X$ is true in the model under a variable assignment α
iff X is true in the model under the variable assignment $\alpha[v:o]$ for some $o \in \mathsf{Dom}$.

Practice exercises

A. Consider the following model:

Dom: $\{Corwin, Benedict\}$
$[\![A]\!]$: $\{Corwin, Benedict\}$
$[\![B]\!]$: $\{Benedict\}$
$[\![N]\!]$: $\emptyset$
$[\![c]\!]$: $Corwin$

Determine whether each of the following sentences is true or false in that model:

1. Bc False
2. $Ac \leftrightarrow \neg Nc$ True
3. $Nc \rightarrow (Ac \lor Bc)$ True
4. $\forall x\, Ax$ True
5. $\forall x \neg Bx$ False
6. $\exists x(Ax \land Bx)$ True
7. $\exists x(Ax \rightarrow Nx)$ False
8. $\forall x(Nx \lor \neg Nx)$ True
9. $\exists x\, Bx \rightarrow \forall x\, Ax$ True

B. Consider the following model:

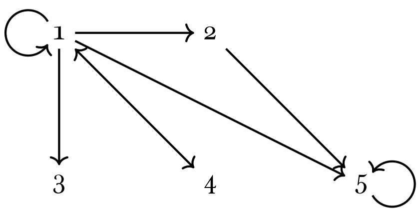
Dom: $\{Lemmy, Courtney, Eddy\}$
$[\![G]\!]$: $\{Lemmy, Courtney, Eddy\}$
$[\![H]\!]$: $\{Courtney\}$
$[\![M]\!]$: $\{Lemmy, Eddy\}$
$[\![c]\!]$: $Courtney$
$[\![e]\!]$: $Eddy$

Determine whether each of the following sentences is true or false in that model:

1. Hc True
2. He False
3. $Mc \lor Me$ True
4. $Gc \lor \neg Gc$ True
5. $Mc \to Gc$ True
6. $\exists x\, Hx$ True
7. $\forall x\, Hx$ False
8. $\exists x \neg Mx$ True
9. $\exists x(Hx \land Gx)$ True
10. $\exists x(Mx \land Gx)$ True
11. $\forall x(Hx \lor Mx)$ True
12. $\exists x\, Hx \land \exists x\, Mx$ True
13. $\forall x(Hx \leftrightarrow \neg Mx)$ True
14. $\exists x\, Gx \land \exists x \neg Gx$ False
15. $\forall x \exists y(Gx \land Hy)$ True

C. Following the diagram conventions introduced at the end of §23, consider the following model:

Dom: $\{1, 2, 3, 4, 5\}$
$[\![R]\!]$: $\{\langle o_1, o_2\rangle \mid$ there is an arrow from o_1 to o_2 in the diagram$\}$



Determine whether each of the following sentences is true or false in that model:

1. $\exists x\, Rxx$ True
2. $\forall x\, Rxx$ False
3. $\exists x \forall y\, Rxy$ True
4. $\exists x \forall y\, Ryx$ False
5. $\forall x \forall y \forall z((Rxy \land Ryz) \to Rxz)$ False
6. $\forall x \forall y \forall z((Rxy \land Rxz) \to Ryz)$ False
7. $\exists x \forall y \neg Rxy$ True
8. $\forall x(\exists y\, Rxy \to \exists y\, Ryx)$ True
9. $\exists x \exists y(\neg x = y \land Rxy \land Ryx)$ True

10. $\exists x \forall y(Rxy \leftrightarrow x = y)$ True
11. $\exists x \forall y(Ryx \leftrightarrow x = y)$ False
12. $\exists x \exists y(\neg x = y \land Rxy \land \forall z(Rz, x \leftrightarrow y = z))$ True

CHAPTER 33

Semantic concepts

Offering a precise definition of truth in FOL was more than a little fiddly, but now that we are done, we can define various central logical notions. These will look very similar to the definitions we offered for TFL. However, remember that they concern *models*, rather than valuations.

$$X_1, X_2, \ldots, X_n \therefore Z$$

is VALID iff there is no model in which all of $X_1, X_2, \ldots, X_n$ are true and in which Z is false.

The other logical notions also have corresponding definitions in FOL:

- An FOL sentence X is a LOGICAL TRUTH iff X is true in every model.
- X is a CONTRADICTION iff X is false in every model.
- Two FOL sentences X and Y are LOGICALLY EQUIVALENT iff they are true in exactly the same models as each other.
- The FOL sentences $X_1, X_2, \ldots, X_n$ are JOINTLY LOGICALLY CONSISTENT iff there is some model in which all of the sentences are true. They are JOINTLY LOGICALLY INCONSISTENT iff there is no such model.

CHAPTER 34

Using models

34.1 Logical truths and contradictions

Suppose we want to show that '$\exists x Axx \rightarrow Bd$' is *not* a logical truth. This requires showing that the sentence is not true in every model; i.e., that it is false in some model. If we can provide just one model in which the sentence is false, then we will have shown that the sentence is not a logical truth.

In order for '$\exists x Axx \rightarrow Bd$' to be false, the antecedent ($\exists x Axx$) must be true, and the consequent (Bd) must be false. To construct such a model, we start by specifying a domain. Keeping the domain small makes it easier to specify what the predicates will be true of, so we will start with a domain that has just one member. For concreteness, let's say it is the city of Paris.

Dom: $\{Paris\}$

The name d must refer to something in the domain, so we have no option but:

$[\![d]\!]$: *Paris*

Recall that we want $\exists x Axx$ to be true, so we want all members of the domain to be paired with themselves in the extension of A. We can just offer:

$[\![A]\!]$: $\{o \mid o = o\}$

Now Add is true, so it is surely true that $\exists x Axx$. Next, we want Bd to be false, so the referent of d must not be in the extension of B. We might simply offer:

$[\![B]\!]$: $\{o \mid o \text{ is in Germany}\}$

Since Paris is not in Germany the set $\{o \mid o \text{ is in Germany}\}$ does no have any elements that are in the domain of the model. This means that we have a model where '$\exists x Axx$' is true, but where 'Bd' is false. So there is a model where $\exists x Axx \rightarrow Bd$ is false. So $\exists x Axx \rightarrow Bd$ is not a logical truth.

We can just as easily show that '$\exists x Axx \rightarrow Bd$' is not a contradiction. We need only specify a model in which '$\exists x Axx \rightarrow Bd$' is true; i.e., a model in which either '$\exists x Axx$' is false or 'Bd' is true. Here is one:

Dom: $\{\textit{Paris}\}$
$[\![d]\!]$: *Paris*
$[\![A]\!]$: $\{o \mid o = o\}$
$[\![B]\!]$: $\{o \mid o \text{ is in France}\}$

This shows that there is a model where '$\exists x Axx \rightarrow Bd$' is true. So '$\exists x Axx \rightarrow Bd$' is not a contradiction.

34.2 Logical equivalence

Suppose we want to show that $\forall x Sx$ and $\exists x Sx$ are not logically equivalent. We need to construct a model in which the two sentences have different truth values; we want one of them to be true and the other to be false. We start by specifying a domain. Again, we make the domain small so that we can specify extensions easily. In this case, we will need at least two objects. (If we chose a domain with only one member, the two sentences would end up with the same truth value. In order to see why, try constructing some partial models with one-member domains.) For concreteness, let's take:

Dom: $\{$*Ornette Coleman, Miles Davis*$\}$

We can make $\exists x Sx$ true by including something in the extension of S, and we can make $\forall x Sx$ false by leaving something out of the extension of S. For concreteness we will offer:

$[\![S]\!]$: $\{o \mid o \text{ plays saxophone}\}$

Now '$\exists x Sx$' is true, because Sx is true of *Ornette Coleman.* Slightly more precisely, extend our model by allowing c to name *Ornette Coleman.* 'Sc' is true in this extended model, so '$\exists x Sx$' was true in the original model. Similarly, '$\forall x Sx$' is false, because *Miles Davis* $\notin [\![S]\!]$. Slightly more precisely, extend our model by allowing d to name *Miles Davis*. Then 'Sd' is false in this extended model, so '$\forall x Sx$' was false in the original model. We have provided a counter-model to the claim that '$\forall x Sx$' and '$\exists x Sx$' are logically equivalent.

> To show that X is not a logical truth, it suffices to find a model where X is false.
> To show that X is not a contradiction, it suffices to find a model where X is true.
> To show that X and Y are not logically equivalent, it suffices to find a model where one is true and the other is false.

34.3 Validity, logical consequence and consistency

To test for validity, logical consequence, or consistency, we typically need to produce models that determine the truth value of several sentences simultaneously.

Consider the following argument in FOL:

$$\exists x(Gx \rightarrow Ga) \therefore \exists x Gx \rightarrow Ga$$

To show that this is invalid, we must make the premise true and the conclusion false. The conclusion is a conditional, so to make it false, the antecedent must be true and the consequent must be false. Clearly, our domain must contain two objects. Let's try:

Dom: $\{$*Karl Marx, Ludwig von Mises*$\}$
$[\![G]\!]$: $\{o \mid o \text{ hated communism}\}$ (=$\{$*Ludwig von Mises*$\}$)
$[\![a]\!]$: *Karl Marx*

Given that Marx wrote *The Communist Manifesto*, Ga is plainly false in this model. But von Mises famously hated communism, so $\exists x Gx$ is true in this model. Hence '$\exists x Gx \rightarrow Ga$' is false, as required.

Does this model make the premise true? Yes it does! Note that $Ga \rightarrow Ga$ is true. (Indeed, it is a logical truth.) But then certainly '$\exists x(Gx \rightarrow Ga)$' is true, so the premise is true, and the conclusion is false, in this model. The argument is therefore invalid.

In passing, note that we have also shown that $\exists x Gx \rightarrow Ga$ is *not* a logical consequence of $\exists x(Gx \rightarrow Ga)$. Equally, we have shown that the sentences $\exists x(Gx \rightarrow Ga)$ and $\neg(\exists x Gx \rightarrow Ga)$ are jointly consistent.

Let's consider a second example. Consider:

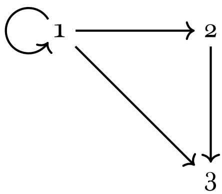
$$\forall x \exists y Lxy \therefore \exists y \forall x Lxy$$

Again, we want to show that this is invalid. To do this, we must make the premises true and the conclusion false. Here is a suggestion:

Dom: $\{o | o$ is a UK citizen currently in a civil partnership with another UK citizen
$[\![L]\!]$: $\{\langle o_1, o_2 \rangle \mid o_1$ is in a civil partnership with $o_2\}$

The premise is clearly true on this model. Anyone in the domain is a UK citizen in a civil partnership with some other UK citizen. That other citizen will also, then, be in the domain. So for everyone in the domain, there will be someone (else) in the domain with whom they are in a civil partnership. Hence '$\forall x \exists y Lxy$' is true. However, the conclusion is clearly false, for that would require that there is some single person who is in a civil partnership with everyone in the domain, and there is no such person, so the argument is invalid. We observe immediately that the sentences '$\forall x \exists y Lxy$ and $\neg \exists y \forall x Lxy$' are jointly consistent and that '$\exists y \forall x Lxy$, is not a logical consequence of '$\forall x \exists y Lxy$'.

For our third example, we'll mix things up a bit. In §31.3, we described how we can present some models using diagrams. For example:



Using the conventions employed in §31.3, the domain of this model is the first three positive whole numbers, and Rxy is true of x and y just in case there is an arrow from x to y in our diagram. Here are some sentences that the model makes true:

- $\forall x \exists y Ryx$
- $\exists x \forall y Rxy$ — witness 1
- $\exists x \forall y (Ryx \leftrightarrow x = y)$ — witness 1
- $\exists x \exists y \exists z ((\neg y = z \land Rxy) \land Rzx)$ — witness 2
- $\exists x \forall y \neg Rxy$ — witness 3
- $\exists x (\exists y Ryx \land \neg \exists y Rxy)$ — witness 3

This immediately shows that all of the preceding six sentences are jointly consistent. We can use this observation to generate *invalid* arguments, e.g.:

$$\forall x \exists y Ryx, \exists x \forall y Rxy \therefore \forall x \exists y Rxy$$
$$\exists x \forall y Rxy, \exists x \forall y \neg Rxy \therefore \neg \exists x \exists y \exists z (\neg y = z \land (Rxy \land Rzx))$$

and many more besides.

> To show that $X_1, X_2, \ldots, X_n \therefore Z$ is invalid, it suffices to find a model where all of $X_1, X_2, \ldots, X_n$ are true and where Z is false. That same model will show that $X_1, X_2, \ldots, X_n$ do not entail Z. It will also show that $X_1, X_2, \ldots, X_n, \neg Z$ are jointly consistent.

When you provide a model to refute a claim—to logical truth, say, or to entailment—this is sometimes called providing a *counter-model* (or providing a *counter-model*).

Practice exercises

A. Show that each of the following is neither a validity nor a contradiction:

1. $Da \land Db$

   The sentence is true in this model:

   Dom: Stan
   $[\![D]\!]$: Stan
   $[\![a]\!]$: Stan
   $[\![b]\!]$: Stan

   And it is false in this model:

   Dom: Stan
   $D(x)$:
   $[\![a]\!]$: Stan
   $[\![b]\!]$: Stan

2. $\exists x\, Tx, h$

   The sentence is true in this model:

   Dom: Stan
   $T(x,y)$: $\langle$Stan, Stan$\rangle$
   $[\![h]\!]$: Stan

And it is false in this model:

Dom: Stan
$T(x,y)$:
$[\![h]\!]$: Stan

3. $Pm \wedge \neg\forall x\, Px$

The sentence is true in this model:

Dom: Stan, Ollie
$P(x)$: Stan
$[\![m]\!]$: Stan

And it is false in this model:

Dom: Stan
$P(x)$:
$[\![m]\!]$: Stan

4. $\forall z\, Jz \leftrightarrow \exists y\, Jy$
5. $\forall x(Wx,m,n \vee \exists yLx,y)$
6. $\exists x(Gx \rightarrow \forall y\, My)$
7. $\exists x(x = h \wedge x = i)$

B. Show that the following pairs of sentences are not logically equivalent.

1. Ja, Ka

Making the first sentence true and the second false:

Dom: 0
$J(x)$: 0
$K(x)$:
$[\![a]\!]$: 0

2. $\exists x\, Jx$, Jm

Making the first sentence true and the second false:

Dom: 0, 1
$J(x)$: 0
$[\![m]\!]$: 1

3. $\forall x\, Rx,x$, $\exists x\, Rx,x$

Making the first sentence false and the second true:

Dom: 0, 1
$R(x,y)$: $\langle 0,0 \rangle$

4. $\exists x\, Px \rightarrow Qc$, $\exists x(Px \rightarrow Qc)$

Making the first sentence false and the second true:

Dom: $0, 1$
$P(x)$: 0
$Q(x)$:
$[\![c]\!]$: 0

5. $\forall x(Px \rightarrow \neg Qx)$, $\exists x(Px \land \neg Qx)$
Making the first sentence true and the second false:

Dom: 0
$P(x)$:
$Q(x)$:

6. $\exists x(Px \land Qx)$, $\exists x(Px \rightarrow Qx)$
Making the first sentence false and the second true:

Dom: 0
$P(x)$:
$Q(x)$: 0

7. $\forall x(Px \rightarrow Qx)$, $\forall x(Px \land Qx)$
Making the first sentence true and the second false:

Dom: 0
$P(x)$:
$Q(x)$: 0

8. $\forall x \exists y\, Rx,y$, $\exists x \forall y\, Rx,y$
Making the first sentence true and the second false:

Dom: $0, 1$
$R(x,y)$: $\langle 0, 1\rangle$, $\langle 1, 0\rangle$

9. $\forall x \exists y\, Rx,y$, $\forall x \exists y\, Ry,x$
Making the first sentence false and the second true:

Dom: $0, 1$
$R(x,y)$: $\langle 0, 0\rangle$, $\langle 0, 1\rangle$

C. Show that the following sentences are jointly satisfiable:

1. $Ma, \neg Na, Pa, \neg Qa$
2. $Le,e, Le,g, \neg Lg,e, \neg Lg,g$
3. $\neg(Ma \land \exists x\, Ax), Ma \lor Fa, \forall x(Fx \rightarrow Ax)$
4. $Ma \lor Mb, Ma \rightarrow \forall x \neg Mx$
5. $\forall y\, Gy, \forall x(Gx \rightarrow Hx), \exists y \neg Iy$
6. $\exists x(Bx \lor Ax), \forall x \neg Cx, \forall x\big[(Ax \land Bx) \rightarrow Cx\big]$
7. $\exists x\, Xx, \exists x\, Yx, \forall x(Xx \leftrightarrow \neg Yx)$
8. $\forall x(Px \lor Qx), \exists x \neg(Qx \land Px)$

9. $\exists z(Nz \wedge Oz,z), \forall x \forall y(Ox,y \rightarrow Oy,x)$
10. $\neg\exists x \forall y\, Rx,y, \forall x \exists y\, Rx,y$
11. $\neg Ra,a,\ \forall x(x = a \vee Rx,a)$
    The sentences are both true in this model:

    Dom: Harry, Sally
    $R(x,y)$: ⟨Sally, Harry⟩
    $[\![a]\!]$: Harry

12. $\forall x \forall y \forall z[(x = y \vee y = z) \vee x = z],\ \exists x \exists y\ \neg x = y$
    There are no predicates or constants, so we only need to give a domain. Any domain with 2 elements will do.
13. $\exists x \exists y((Zx \wedge Zy) \wedge x = y),\ \neg Zd,\ d = e$

D. Show that the following arguments are invalid:

1. $\forall x(Ax \rightarrow Bx) \therefore \exists x\, Bx$
2. $\forall x(Rx \rightarrow Dx), \forall x(Rx \rightarrow Fx) \therefore \exists x(Dx \wedge Fx)$
3. $\exists x(Px \rightarrow Qx) \therefore \exists x\, Px$
4. $Na \wedge Nb \wedge Nc \therefore \forall x\, Nx$
5. $Rd,e, \exists x\, Rxd \therefore Re,d$
6. $\exists x(Ex \wedge Fx), \exists x\, Fx \rightarrow \exists x\, Gx \therefore \exists x(Ex \wedge Gx)$
7. $\forall x\, Ox,c, \forall x\, Oc,x \therefore \forall x\, Ox,x$
8. $\exists x(Jx \wedge Kx), \exists x \neg Kx, \exists x \neg Jx \therefore \exists x(\neg Jx \wedge \neg Kx)$
9. $La,b \rightarrow \forall x\, Lx,b, \exists x\, Lx,b \therefore Lb,b$
10. $\forall x(Dx \rightarrow \exists y\, Ty,x) \therefore \exists y \exists z\ \neg y = z$

CHAPTER 35

Reasoning about all models

35.1 Logical truths and contradictions

We can show that a sentence is *not* a logical truth just by providing one carefully specified model: a model in which the sentence is false. To show that something is a logical truth, on the other hand, it would not be enough to construct ten, one hundred, or even a thousand models in which the sentence is true. A sentence is only a logical truth if it is true in *every* model, and there are infinitely many models. We need to reason about all of them, and we cannot do this by dealing with them one by one!

Sometimes, we can reason about all models fairly easily. For example, we can offer a relatively simple argument that '$Raa \rightarrow Raa$' is a logical truth:

> Any relevant model will give 'Raa' a truth value. If 'Raa' is true in a model, then '$Raa \rightarrow Raa$' is true in that model. If Raa is false in a model, then $Raa \rightarrow Raa$ is true in that model. These are the only alternatives. So '$Raa \leftrightarrow Raa$' is true in every model. Therefore, it is a logical truth.

This argument is valid, of course, and its conclusion is true. However, it is not an argument in FOL. Rather, it is an argument in English *about* FOL: it is an argument in the metalanguage.

Note another feature of the argument. Since the sentence in question contained no quantifiers, we did not need to think about how to interpret a and R; the point was just that, however we interpreted them, Raa would have some truth value or other. (We could ultimately have given the same argument concerning TFL sentences.)

Here is another bit of reasoning. Consider the sentence $\forall x(Rxx \to Rxx)$. Again, it should obviously be a logical truth, but to say precisely why is quite a challenge. We cannot say that $Rxx \to Rxx$ is true in every model, since $Rxx \to Rxx$ is not even a *sentence* of FOL (remember that x is a variable, not a name). So we have to be a bit cleverer.

> Consider some arbitrary model. Consider some arbitrary member of the domain, which, for convenience, we will call *obbie*, and suppose we extend our original model by adding a new name, c, to name *obbie*. Then either Rcc will be true or it will be false. If Rcc is true, then $Rcc \leftrightarrow Rcc$ is true. If Rcc is false, then $Rcc \to Rcc$ will be true. So either way, $Rcc \to Rcc$ is true. Since there was nothing special about *obbie*—we might have chosen any object—we see that no matter how we extend our original model by allowing c to name some new object, $Rcc \to Rcc$ will be true in the new model. So $\forall x(Rxx \to Rxx)$ was true in the original model. But we chose our model arbitrarily, so $\forall x(Rxx \to Rxx)$ is true in every model. It is therefore a logical truth.

This is quite longwinded, but, as things stand, there is no alternative. In order to show that a sentence is a logical truth, we must reason about *all* models.

35.2 Other cases

Similar points hold of other cases too. Thus, we must reason about all models if we want to show:

- that a sentence is a contradiction; for this requires that it is false in *every* model.
- that two sentences are logically equivalent; for this requires that they have the same truth value in *every* model.

- that some sentences are jointly inconsistent; for this requires that there is no model in which all of those sentences are true together; i.e. that, in *every* model, at least one of those sentences is false.
- that an argument is valid; for this requires that the conclusion is true in *every* model where the premises are true.
- that some sentences entail another sentence.

The problem is that, with the tools available to you so far, reasoning about all models is a serious challenge! Let's take just one more example. Here is an argument which is obviously valid:

$$\forall x(Hx \wedge Jx) \therefore \forall xHx$$

After all, if everything is both H and J, then everything is H. But we can only show that the argument is valid by considering what must be true in every model in which the premise is true. To show this, we would have to reason as follows:

> Consider an arbitrary model in which the premise $\forall x(Hx \wedge Jx)$ is true. It follows that, however we expand the model with a new name, for example c, $Hc \wedge Jc$ will be true in this new model. Hc will, then, also be true in this new model. But since this held for *any* way of expanding the model, it must be that $\forall xHx$ is true in the old model. We've assumed nothing about the model except that it was one in which $\forall x(Hx \wedge Jx)$ is true, so any model in which $\forall x(Hx \wedge Jx)$ is true is one in which $\forall xHx$ is true. The argument is valid!

Even for a simple argument like this one, the reasoning is somewhat complicated. For longer arguments, the reasoning can be extremely torturous.

The following table summarises whether a single (counter-)model suffices, or whether we must reason about all models.

	Yes	**No**
logical truth?	all models	one counter-model
contradiction?	all models	one counter-model
equivalent?	all models	one counter-model
consistent?	one model	all models
valid?	all models	one counter-model
entailment?	all models	one counter-model

This might usefully be compared with the table at the end of §12. The key difference resides in the fact that TFL concerns truth tables, whereas FOL concerns models. This difference is deeply important, since each truth-table only ever has finitely many lines, so that a complete truth table is a relatively tractable object. By contrast, there are infinitely many models for any given sentence(s), so that reasoning about all models can be a deeply tricky business.

PART VII

Natural deduction for FOL

CHAPTER 36

Basic rules for FOL

FOL makes use of all of the connectives of TFL. So proofs in FOL will use all of the basic and derived rules from Part IV. We will also use the proof-theoretic notions (particularly, the symbol '$\vdash$') introduced there. However, we will also need some new basic rules to govern the quantifiers.

36.1 Universal elimination

Consider:

1. Everyone is happy

∴ Therefore: Catrin is happy.

This is a valid argument. Generally, then, from the claim that everything is F, you can infer that any particular thing is F. You name it; it's F. So the following should be fine:

1	$\forall x Fx$	
2	Fa	$\forall$E, 1

We obtained line 2 by dropping the universal quantifier and replacing 'x' with 'a'.

This isn't restricted to simple properties. Consider the following argument:

1. Every cat is sleeping.

∴ Therefore: If Fluffy is a cat, then she is sleeping.

This is a valid argument. And it will be allowed by our rule ∀E:

1	$\forall x(Cx \to Sx)$	
2	$Cf \to Sf$	∀E, 1

Note here that we have to replace two instances of 'x' with our name: 'f' or Fluffy. Indeed it would have been fine to do with any name. We can even do it with names we already have. Consider the following:

1. Pavel owes everyone money.

∴ Therefore: Pavel owes himself money.

We we symbolise this as:

1. $\forall x Opx$

∴ Therefore: Opp

This is valid: the premise says *everything* in the domain owes money to Pavel; and Pavel is something in the domain. So it implies that Pavel owes money to himself.

This argument is also directly allowed by our rule ∀E:

1	$\forall x Opx$	
2	Opp	∀E, 1

We can now give our general rule: whenever you have a sentence $\forall x X(\ldots x \ldots x \ldots)$, for example $\forall x Fx$, $\forall x(Cx \to Sx)$, $\forall x Opx$; one can conclude that we have the sentence which is obtained by stripping of the quantifier and replacing every free occurrence of the variable by a name, be it $a, b, c \ldots$. So we could derive Fa, $Cf \to Sf$ or Opp.

Here is the formal specification of the universal elimination rule (∀E):

m	$\forall x X(\ldots x \ldots x \ldots)$	
	$X(\ldots c \ldots c \ldots)$	∀E, m

The point is that you can obtain any *substitution instance* of a universally quantified formula: **replace every instance of the quantified variable with any name you like**. (Of course, you need to replace every free occurrence of x in X by the same name.)

I should emphasize that (as with every elimination rule) you can only apply the $\forall$E rule when the universal quantifier is the main logical operator. Thus the following is outright banned:

1	$\forall x Bx \rightarrow Bk$	
2	$Bb \rightarrow Bk$	naughtily attempting to invoke $\forall$E, 1

This is illegitimate, since '$\forall x$' is not the main logical operator in line 1. (If you need a reminder as to why this sort of inference should be banned, reread §??.)

36.2 Existential introduction

The following argument is valid:

1. Catrin is happy

∴ Therefore: Someone is happy.

This is the idea of our existential introduction rule: from the claim that some particular thing is F, you can infer that something is F:

1	Fa	
2	$\exists x Fx$	$\exists$I, 1

We obtained line 2 by replacing the name 'a' with the variable 'x' and adding $\exists x$ in front of the sentence. This will be permissible by our rule of $\exists$I.

This isn't restricted to simple properties.

1. Bob is a money and knows sign language.

∴ Therefore: There is a monkey who knows sign language.

1	$Mb \wedge Sb$	
2	$\exists x(Mx \wedge Sx)$	$\exists$I, 1

Or even

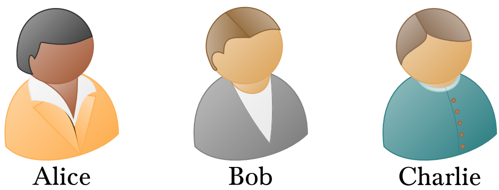
1. Catrin is friends with someone who is friends with everyone.

∴ Therefore: Someone is friends with someone who is friends with everyone.

1	$\exists x(Fcx \wedge \forall y Fxy)$	
2	$\exists z \exists x(Fzx \wedge \forall y Fxy)$	$\exists$I, 1

We replaced the name, 'c' with the variable 'z' and added $\exists z$ at the beginning of the sentence.

This rule will now allow us to carefully work through our brain teaser from §19.2

> Three people are standing in a row looking at eachother.
>
> 
>
>
> Alice is happy. Charlie is not happy. Is there someone who is happy who is looking at someone who is not happy?

Answer: Yes.

We can formalise this argument as:

$$Lab, Lbc, Ha, \neg Hc \;\therefore\; \exists x \exists y(Hx \wedge (Lxy \wedge \neg Hx)).$$

And we can show it is valid. In §19.2 we wrote this in a pseudo-formal style.

1	Bob is either happy or he’s not happy	
2		Suppose Bob is happy
3		Then happy Bob is looking at unhappy Charlie
4		So someone who is happy is looking at someone who is not happy.
5		Suppose Bob is unhappy
6		Then happy Alice is looking at unhappy Bob
7		So someone who is happy is looking at someone who is not happy.
8	Therefore, someone who is happy is looking at someone who is not happy.	

We can now fill out the details of this to see that it’s a formal proof:

1	Lab	
2	Lbc	
3	Ha	
4	$\neg Hc$	
5	$Hb \lor \neg Hb$	LEM
6	$\quad Hb$	
7	$\quad Lbc \land \neg Hc$	$\land$I, 2, 6
8	$\quad Hb \land (Lbc \land \neg Hc)$	$\land$I, 7, 6
9	$\quad \exists y(Hb \land (Lby \land \neg Hy))$	$\exists$I, 8
10	$\quad \exists x \exists y(Hx \land (Lxy \land \neg Hx))$	$\exists$I, 9
11	$\quad \neg Hb$	
12	$\quad Lab \land \neg Hb$	$\land$I, 11, 1
13	$\quad Ha \land (Lab \land \neg Hb)$	$\land$I, 3, 12
14	$\quad \exists y(Ha \land (Lay \land \neg Hy))$	$\exists$I, 13
15	$\quad \exists x \exists y(Hx \land (Lxy \land \neg Hx))$	$\exists$I, 14
16	$\exists x \exists y(Hx \land Lxy \land \neg Hx)$	$\lor$E, 5, 6–10, 11–15

Consider the following example:

1. Narcissus loves himeself.

∴ Therefore: There is someone who loves Narcissus.

This is a valid argument. The formalised version, which will be allowed by our rule is:

1	Lnn	
2	$\exists x Lxn$	$\exists$I, 1

This shows us that we do not have to replace *all* instances of the name with the variable. Though of course we can if we wish: we could also deduce there is someone who loves himself.

To give our rule in general we need to introduce some new notation for this ability to replace just some of our instances of the name: If X is a sentence containing the name c, we can emphasize this by writing '$X(\ldots c \ldots c \ldots)$'. **We will write '$X(\ldots x \ldots c \ldots)$' to indicate any formula obtained by replacing *some or all* of the instances of the name c with the variable x.** Armed with this, our introduction rule is:

m	$X(\ldots c \ldots c \ldots)$	
	$\exists x X(\ldots x \ldots c \ldots)$	$\exists$I, m

x must not occur in $X(\ldots c \ldots c \ldots)$

All the cases we've seen in this section follow this rule.

You might have noticed the additional constraint that's added to the rule. It is part of the rule; so if you are asked to write the rule $\exists$I you must include this constraint. However, you will not need to worry about it in practice. It's simply there to guarantee that applications of the rule yield sentences of FOL. If the rule were not there we would be allowed to argue as follows:

1	$\exists x Lnx$	
2	$\exists x \exists x Lxx$	naughtily attempting to invoke $\exists$I, 1

But this expression on line 2 contains clashing variables. It will not count as a sentence of FOL.

36.3 Empty domains

The following proof combines our two new rules for quantifiers:

1	$\forall x Fx$	
2	Fa	$\forall$E, 1
3	$\exists x Fx$	$\exists$I, 2

Could this be a bad proof? If anything exists at all, then certainly we can infer that something is F, from the fact that everything is F. But what if *nothing* exists at all? Then it is surely vacuously true that everything is F; however, it does not following that something is F, for there is nothing to *be* F. So if we claim that, as a matter of logic alone, '$\exists x F x$' follows from '$\forall x F x$', then we are claiming that, as a matter of *logic alone*, there is something rather than nothing. This might strike us as a bit odd.

Actually, we are already committed to this oddity. In §25, we stipulated that domains in FOL must have at least one member. We then defined a logical truth (of FOL) as a sentence which is true in every model. Since '$\exists x(Ax \vee \neg Ax)$' will be true in every model, this *also* had the effect of stipulating that it is a matter of logic that there is something rather than nothing.

Since it is far from clear that logic should tell us that there must be something rather than nothing, we might well be cheating a bit here.

If we refuse to cheat, though, then we pay a high cost. Here are three things that we want to hold on to:

- $\forall x F x \vdash F a$: after all, that was $\forall$E.
- $F a \vdash \exists x F x$: after all, that was $\exists$I.
- the ability to copy-and-paste proofs together: after all, reasoning works by putting lots of little steps together into rather big chains.

If we get what we want on all three counts, then we have to countenance that $\forall x F x \vdash \exists x F x$. So, if we get what we want on all three counts, the proof system alone tells us that there is something rather than nothing. And if we refuse to accept that, then we have to surrender one of the three things that we want to hold on to!

Before we start thinking about which to surrender, we might want to ask how *much* of a cheat this is. Granted, it may make it harder to engage in theological debates about why there is something rather than nothing. But the rest of the time, we will get along just fine. So maybe we should just regard our proof system (and FOL, more generally) as having a very slightly limited purview. If we ever want to allow for the possibility of *nothing*, then we will have to cast around for a more complicated proof system. But for as long as we are content to ignore that possibility, our proof system is perfectly in order. (As, similarly, is the stipulation that every domain must contain at least one object.)

36.4 Universal introduction

Suppose you had shown of each particular thing that it is F (and that there are no other things to consider). Then you would be justified in claiming that everything is F. This would motivate the following proof rule. If you had established each and every single substitution instance of '$\forall x Fx$', then you can infer '$\forall x Fx$'.

Unfortunately, that rule would be utterly unusable. To establish each and every single substitution instance would require proving 'Fa', 'Fb', ..., 'Fj_2', ..., 'Fr_{79002}', ..., and so on. Indeed, since there are infinitely many names in FOL, this process would never come to an end. So we could never apply that rule. We need to be a bit more cunning in coming up with our rule for introducing universal quantification.

Our cunning thought will be inspired by considering:

$$\forall x Fx \therefore \forall y Fy$$

This argument should *obviously* be valid. After all, alphabetical variation ought to be a matter of taste, and of no logical consequence. But how might our proof system reflect this? Suppose we begin a proof thus:

1	$\forall x Fx$	
2	Fa	$\forall$E, 1

We have proved 'Fa'. And, of course, nothing stops us from using the same justification to prove 'Fb', 'Fc', ..., 'Fj_2', ..., '$Fr_{79002}, \ldots$, and so on until we run out of space, time, or patience. But reflecting on this, we see that there is a way to prove Fc, for any name c. And if we can do it for *any* thing, we should surely be able to say that 'F' is true of *everything*. This therefore justifies us in inferring '$\forall y Fy$', thus:

1	$\forall x Fx$	
2	Fa	$\forall$E, 1
3	$\forall y Fy$	$\forall$I, 2

The crucial thought here is that 'a' was just some *arbitrary* name. There was nothing special about it—we might have chosen any other name—and still the proof would be fine. And this crucial thought motivates the universal introduction rule ($\forall$I):

$$
\begin{array}{r|ll}
m & X(\ldots c \ldots c \ldots) & \\
 & \forall x X(\ldots x \ldots x \ldots) & \forall\text{I}, m
\end{array}
$$

c must not occur in any undischarged assumption or premise.
x must not occur in $X(\ldots c \ldots c \ldots)$

It is important to appreciate that to apply the rule correctly **we must replace every occurrence of the name c by the variable x**. A crucial aspect of this rule, though, is bound up in the first constraint. This constraint ensures that we are always reasoning at a sufficiently general level (the second constraint guarantees that the variable x is not already bound by a different quantifier in X which would lead to unintended results.)

To see the constraint in action, consider this terrible argument:

> Everyone loves Kylie Minogue; therefore everyone loves themselves.

We might symbolize this obviously invalid inference pattern as:

$$\forall x Lxk \therefore \forall x Lxx$$

Now, suppose we tried to offer a proof that vindicates this argument:

$$
\begin{array}{r|ll}
1 & \forall x Lxk & \\
\hline
2 & Lkk & \forall\text{E}, 1 \\
3 & \forall x Lxx & \text{naughtily attempting to invoke } \forall\text{I}, 2
\end{array}
$$

This is not allowed, because 'k' occurred already in an undischarged assumption, namely, on line 1. The crucial point is that, if we have made any assumptions about the object we are working with, then we are not reasoning generally enough to license ∀I.

Although the name may not occur in any *undischarged* assumption, it may occur as a discharged assumption. That is, it may occur in a subproof that we have already closed. For example:

$$
\begin{array}{r|l l}
1 & \quad \begin{array}{|l} Gd \\ \hline \end{array} & \\
2 & \quad \begin{array}{|l} Gd \end{array} & \text{R, 1} \\
3 & Gd \to Gd & \to\text{I, 1–2} \\
4 & \forall z(Gz \to Gz) & \forall\text{I, 3}
\end{array}
$$

This tells us that '$\forall z(Gz \to Gz)$' is a *theorem*. And that is as it should be.

36.5 Existential elimination

Suppose we know that *something* is F. The problem is that simply knowing this does not tell us which thing is F. So it would seem that from '$\exists x Fx$' we cannot immediately conclude 'Fa', 'Fe_{23}', or any other substitution instance of the sentence. What can we do?

Suppose we know that something is F, and that everything which is F is G. In (almost) natural English, we might reason thus:

> Since something is F, there is some particular thing which is an F. We do not know anything about it, other than that it's an F, but for convenience, let's call it 'obbie'. So: obbie is F. Since everything which is F is G, it follows that obbie is G. But since obbie is G, it follows that something is G. And nothing depended on which object, exactly, obbie was. So, something is G.

We might try to capture this reasoning pattern in a proof as follows:

$$
\begin{array}{r|l l}
1 & \exists x Fx & \\
2 & \forall x(Fx \to Gx) & \\ \hline
3 & \quad \begin{array}{|l} Fo \\ \hline \end{array} & \\
4 & \quad \begin{array}{|l} Fo \to Go \end{array} & \forall\text{E, 2} \\
5 & \quad \begin{array}{|l} Go \end{array} & \to\text{E, 4, 3} \\
6 & \quad \begin{array}{|l} \exists x Gx \end{array} & \exists\text{I, 5} \\
7 & \exists x Gx & \exists\text{E, 1, 3–6}
\end{array}
$$

Breaking this down: we started by writing down our assumptions. At line 3, we made an additional assumption: 'Fo'. This was just a substitution instance of '$\exists xFx$'. On this assumption, we established '$\exists xGx$'. Note that we had made no *special* assumptions about the object named by 'o'; we had *only* assumed that it satisfies 'Fx'. So nothing depends upon which object it is. And line 1 told us that *something* satisfies 'Fx', so our reasoning pattern was perfectly general. We can discharge the specific assumption 'Fo', and simply infer '$\exists xGx$' on its own.

Putting this together, we obtain the existential elimination rule ($\exists$E):

$$
\begin{array}{l|ll}
m & \exists x\mathscr{A}(\ldots x \ldots x \ldots) & \\
i & \quad \mid \underline{\mathscr{A}(\ldots c \ldots c \ldots)} & \\
 & \quad \mid \quad \vdots & \\
j & \quad \mid \mathscr{B} & \\
 & \mathscr{B} & \exists\text{E}, m, i\text{–}j
\end{array}
$$

c must be new to the proof: it does not occur in any line $< i$;
c must not occur in $\mathscr{B}$.

As with universal introduction, the constraints are extremely important. To see why, consider the following terrible argument:

> Tim Button is a lecturer. There is someone who is not a lecturer. So Tim Button is both a lecturer and not a lecturer.

We might symbolize this obviously invalid inference pattern as follows:

$$Lb, \exists x\neg Lx \therefore Lb \wedge \neg Lb$$

Now, suppose we tried to offer a proof that vindicates this argument:

$$
\begin{array}{l|ll}
1 & Lb & \\
2 & \underline{\exists x\neg Lx} & \\
3 & \quad \mid \underline{\neg Lb} & \\
4 & \quad \mid Lb \wedge \neg Lb & \wedge\text{I}, 1, 3 \\
5 & Lb \wedge \neg Lb & \text{naughtily attempting to invoke } \exists\text{E}, 2, 3\text{–}4
\end{array}
$$

The last line of the proof is not allowed. The name that we used in our substitution instance for '$\exists x \neg Lx$' on line 3, namely 'b', occurs in line 4. The following proof would be no better:

1	Lb	
2	$\exists x \neg Lx$	
3	$\quad \neg Lb$	
4	$\quad Lb \wedge \neg Lb$	$\wedge$I, 1, 3
5	$\quad \exists x(Lx \wedge \neg Lx)$	$\exists$I, 4
6	$\exists x(Lx \wedge \neg Lx)$	naughtily attempting to invoke $\exists$E , 2, 3–5

The last line of the proof would still not be allowed. For the name that we used in our substitution instance for '$\exists x \neg Lx$', namely 'b', occurs in an undischarged assumption, namely line 1.

The moral of the story is this. *If you want to squeeze information out of an existential quantifier, choose a new name for your substitution instance.* That way, you can guarantee that you meet all the constraints on the rule for $\exists$E.

Practice exercises

A. The following two 'proofs' are *incorrect*. Explain why both are incorrect. Also, provide models which would invalidate the fallacious argument forms the 'proofs' enshrine:

1	$\forall x Rxx$	
2	Raa	$\forall$E, 1
3	$\forall y Ray$	$\forall$I, 2
4	$\forall x \forall y Rxy$	$\forall$I, 3

1	$\forall x \exists y Rxy$	
2	$\exists y Ray$	$\forall$E, 1
3	$\quad Raa$	
4	$\quad \exists x Rxx$	$\exists$I, 3
5	$\exists x Rxx$	$\exists$E, 2, 3–4

B. The following three proofs are missing their citations (rule and line numbers). Add them, to turn them into bona fide proofs.

$$
\begin{array}{r|l}
1 & \forall x \exists y (Rxy \vee Ryx) \\
2 & \forall x \neg Rmx \\
\hline
3 & \exists y (Rmy \vee Rym) \\
4 & \quad \begin{array}{|l} Rma \vee Ram \\ \hline \end{array} \\
5 & \quad \begin{array}{|l} \neg Rma \end{array} \\
6 & \quad \begin{array}{|l} Ram \end{array} \\
7 & \quad \begin{array}{|l} \exists x Rxm \end{array} \\
8 & \exists x Rxm
\end{array}
$$

$$
\begin{array}{r|l}
1 & \forall x (\exists y Lxy \to \forall z Lzx) \\
2 & Lab \\
\hline
3 & \exists y Lay \to \forall z Lza \\
4 & \exists y Lay \\
5 & \forall z Lza \\
6 & Lca \\
7 & \exists y Lcy \to \forall z Lzc \\
8 & \exists y Lcy \\
9 & \forall z Lzc \\
10 & Lcc \\
11 & \forall x Lxx
\end{array}
$$

$$
\begin{array}{r|l}
1 & \forall x (Jx \to Kx) \\
2 & \exists x \forall y Lxy \\
3 & \forall x Jx \\
\hline
4 & \quad \begin{array}{|l} \forall y Lay \\ \hline \end{array} \\
5 & \quad \begin{array}{|l} Laa \end{array} \\
6 & \quad \begin{array}{|l} Ja \end{array} \\
7 & \quad \begin{array}{|l} Ja \to Ka \end{array} \\
8 & \quad \begin{array}{|l} Ka \end{array} \\
9 & \quad \begin{array}{|l} Ka \wedge Laa \end{array} \\
10 & \quad \begin{array}{|l} \exists x (Kx \wedge Lxx) \end{array} \\
11 & \exists x (Kx \wedge Lxx)
\end{array}
$$

C. In §?? problem A, we considered fifteen syllogistic figures of Aristotelian logic. Provide proofs for each of the argument forms. NB: You will find it *much* easier if you symbolize (for example) 'No F is G' as '$\forall x(Fx \to \neg Gx)$'.

D. Aristotle and his successors identified other syllogistic forms which depended upon 'existential import'. Symbolize each of the following argument forms in FOL and offer proofs.

- **Barbari**. Something is H. All G are F. All H are G. So: Some H is F
- **Celaront**. Something is H. No G are F. All H are G. So: Some H is not F
- **Cesaro**. Something is H. No F are G. All H are G. So: Some H is not F.
- **Camestros**. Something is H. All F are G. No H are G. So: Some H is not F.
- **Felapton**. Something is G. No G are F. All G are H. So: Some H is not F.
- **Darapti**. Something is G. All G are F. All G are H. So: Some H is F.
- **Calemos**. Something is H. All F are G. No G are H. So: Some H is not F.
- **Fesapo**. Something is G. No F is G. All G are H. So: Some H is not F.
- **Bamalip**. Something is F. All F are G. All G are H. So: Some H are F.

E. Provide a proof of each claim.

1. $\vdash \forall x Fx \lor \neg\forall x Fx$
2. $\vdash \forall z(Pz \lor \neg Pz)$
3. $\forall x(Ax \to Bx), \exists x Ax \vdash \exists x Bx$
4. $\forall x(Mx \leftrightarrow Nx), Ma \land \exists x Rxa \vdash \exists x Nx$
5. $\forall x\forall y Gxy \vdash \exists x Gxx$
6. $\vdash \forall x Rxx \to \exists x\exists y Rxy$
7. $\vdash \forall y\exists x(Qy \to Qx)$
8. $Na \to \forall x(Mx \leftrightarrow Ma), Ma, \neg Mb \vdash \neg Na$
9. $\forall x\forall y(Gxy \to Gyx) \vdash \forall x\forall y(Gxy \leftrightarrow Gyx)$
10. $\forall x(\neg Mx \lor Ljx), \forall x(Bx \to Ljx), \forall x(Mx \lor Bx) \vdash \forall x Ljx$

F. Write a symbolization key for the following argument, symbolize it, and prove it:

> There is someone who likes everyone who likes everyone that she likes. Therefore, there is someone who likes herself.

G. Show that each pair of sentences is provably equivalent.

1. $\forall x(Ax \to \neg Bx)$, $\neg\exists x(Ax \land Bx)$
2. $\forall x(\neg Ax \to Bd)$, $\forall x Ax \lor Bd$
3. $\exists x Px \to Qc$, $\forall x(Px \to Qc)$

H. For each of the following pairs of sentences: If they are provably equivalent, give proofs to show this. If they are not, construct a model to show that they are not logically equivalent.

1. $\forall x Px \rightarrow Qc, \forall x(Px \rightarrow Qc)$
2. $\forall x\forall y\forall z Bxyz, \forall x Bxxx$
3. $\forall x\forall y Dxy, \forall y\forall x Dxy$
4. $\exists x\forall y Dxy, \forall y\exists x Dxy$
5. $\forall x(Rca \leftrightarrow Rxa), Rca \leftrightarrow \forall x Rxa$

I. For each of the following arguments: If it is valid in FOL, give a proof. If it is invalid, construct a model to show that it is invalid.

1. $\exists y\forall x Rxy \therefore \forall x\exists y Rxy$
2. $\forall x\exists y Rxy \therefore \exists y\forall x Rxy$
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. $\exists x Mx \rightarrow \exists x Nx, \neg\exists x Nx \therefore \forall x\neg Mx$

CHAPTER 37

Conversion of quantifiers

In this section, we will add some additional rules to the basic rules of the previous section. These govern the interaction of quantifiers and negation.

In §25, we noted that $\neg\exists x X$ is logically equivalent to $\forall x \neg X$. We will add some rules to our proof system that govern this. In particular, we add:

m	$\forall x \neg X$	
	$\neg\exists x X$	CQ, m

and

m	$\neg\exists x X$	
	$\forall x \neg X$	CQ, m

Equally, we add:

m	$\exists x \neg X$	
	$\neg \forall x X$	CQ, m

and

m	$\neg \forall x X$	
	$\exists x \neg X$	CQ, m

Practice exercises

A. Show in each case that the sentences are provably inconsistent:

1. $Sa \rightarrow Tm, Tm \rightarrow Sa, Tm \wedge \neg Sa$
2. $\neg \exists x Rxa, \forall x \forall y Ryx$
3. $\neg \exists x \exists y Lxy, Laa$
4. $\forall x(Px \rightarrow Qx), \forall z(Pz \rightarrow Rz), \forall y Py, \neg Qa \wedge \neg Rb$

B. Show that each pair of sentences is provably equivalent:

1. $\forall x(Ax \rightarrow \neg Bx), \neg \exists x(Ax \wedge Bx)$
2. $\forall x(\neg Ax \rightarrow Bd), \forall x Ax \vee Bd$

C. In §??, we considered what happens when we move quantifiers 'across' various logical operators. Show that each pair of sentences is provably equivalent:

1. $\forall x(Fx \wedge Ga), \forall x Fx \wedge Ga$
2. $\exists x(Fx \vee Ga), \exists x Fx \vee Ga$
3. $\forall x(Ga \rightarrow Fx), Ga \rightarrow \forall x Fx$
4. $\forall x(Fx \rightarrow Ga), \exists x Fx \rightarrow Ga$
5. $\exists x(Ga \rightarrow Fx), Ga \rightarrow \exists x Fx$
6. $\exists x(Fx \rightarrow Ga), \forall x Fx \rightarrow Ga$

NB: the variable 'x' does not occur in 'Ga'.

When all the quantifiers occur at the beginning of a sentence, that sentence is said to be in *prenex normal form*. These equivalences are sometimes called *prenexing rules*, since they give us a means for putting any sentence into prenex normal form.

CHAPTER 38

Derived rules

As in the case of TFL, we first introduced some rules for FOL as basic (in §36), and then added some further rules for conversion of quantifiers (in §37). In fact, the CQ rules should be regarded as *derived* rules, for they can be derived from the *basic* rules of §36. (The point here is as in §22.) Here is a justification for the first CQ rule:

1	$\forall x \neg Px$	
2	$\quad \exists x Px$	
3	$\quad\quad Pb$	
4	$\quad\quad\quad Da \vee \neg Da$	
5	$\quad\quad\quad \neg Pb$	$\forall$E, 1
6	$\quad\quad\quad Pb$	R, 3
7	$\quad\quad \neg(Da \vee \neg Da)$	$\neg$I, 4–6
8	$\quad \neg(Da \vee \neg Da)$	$\exists$E, 2, 3–7
9	$\quad Da \vee \neg Da$	LEM
10	$\neg \exists x Px$	$\neg$I, 2–9

Here is a justification of the third CQ rule:

1	$\exists x \neg Px$	
2	$\quad \forall x Px$	
3	$\quad\quad \neg Pb$	
4	$\quad\quad\quad Da \lor \neg Da$	
5	$\quad\quad\quad Pb$	$\forall$E, 2
6	$\quad\quad\quad \neg Pb$	R, 3
7	$\quad\quad \neg(Da \lor \neg Da)$	$\neg$I, 4–6
8	$\quad \neg(Da \lor \neg Da)$	$\exists$E, 1, 3–7
9	$\quad Da \lor \neg Da$	LEM
10	$\neg \forall x Px$	$\neg$I, 2–10

This explains why the CQ rules can be treated as derived. Similar justifications can be offered for the other two CQ rules.

Practice exercises

A. Offer proofs which justify the addition of the second and fourth CQ rules as derived rules.

CHAPTER 39

Proof-theoretic and semantic concepts

We have used two different turnstiles in this book. This:

$$X_1, X_2, \ldots, X_n \vdash Y$$

means that there is some proof which starts with assumptions $X_1, X_2, \ldots, X_n$ and ends with Y (and no undischarged assumptions other than $X_1, X_2, \ldots, X_n$). This is a *proof-theoretic notion.*

By contrast, this:

$$X_1, X_2, \ldots, X_n \vDash Y$$

means that there is no valuation (or model) which makes all of $X_1, X_2, \ldots, X_n$ true and makes Y false. This concerns assignments of truth and falsity to sentences. It is a *semantic notion.*

It cannot be emphasized enough that these are different notions. But we can emphasize it a bit more: *They are different notions.*

Once you have internalised this point, continue reading.

Although our semantic and proof-theoretic notions are different, there is a deep connection between them. To explain this connection,we will start by considering the relationship between logical truths and theorems.

To show that a sentence is a theorem, you need only produce a proof. Granted, it may be hard to produce a twenty line proof, but it is not so hard to check each line of the proof and confirm that it is legitimate; and if each line of the proof individually is legitimate, then the whole proof is legitimate. Showing that a sentence is a logical truth, though, requires reasoning about all possible models. Given a choice between showing that a sentence is a theorem and showing that it is a logical truth, it would be easier to show that it is a theorem.

Contrawise, to show that a sentence is *not* a theorem is hard. We would need to reason about all (possible) proofs. That is very difficult. However, to show that a sentence is not a logical truth, you need only construct a model in which the sentence is false. Granted, it may be hard to come up with the model; but once you have done so, it is relatively straightforward to check what truth value it assigns to a sentence. Given a choice between showing that a sentence is not a theorem and showing that it is not a logical truth, it would be easier to show that it is not a logical truth.

Fortunately, *a sentence is a theorem if and only if it is a logical truth.* As a result, if we provide a proof of X on no assumptions, and thus show that X is a theorem, we can legitimately infer that X is a logical truth; i.e., $\vDash X$. Similarly, if we construct a model in which X is false and thus show that it is not a logical truth, it follows that X is not a theorem.

More generally, we have the following powerful result:

$$X_1, X_2, \ldots, X_n \vdash Y \textbf{ iff } X_1, X_2, \ldots, X_n \vDash Y$$

This shows that, whilst provability and entailment are *different* notions, they are extensionally equivalent. As such:

- An argument is *valid* iff *the conclusion can be proved from the premises.*
- Two sentences are *logically equivalent* iff they are *provably equivalent.*
- Sentences are *provably consistent* iff they are *not provably inconsistent.*

For this reason, you can pick and choose when to think in terms of proofs and when to think in terms of valuations/models, doing whichever is easier for a given task. The table on the next page summarises which is (usually) easier.

It is intuitive that provability and semantic entailment should agree. But—let us repeat this—do not be fooled by the similarity of the symbols '$\vDash$' and '$\vdash$'. These two symbols have very different meanings. The fact that provability and semantic entailment agree is not an easy result to come by.

In fact, demonstrating that provability and semantic entailment agree is, very decisively, the point at which introductory logic becomes intermediate logic.

	Yes	**No**
Is X a **logical truth**?	give a proof which shows $\vdash X$	give a model in which X is false
Is X a **contradiction**?	give a proof which shows $\vdash \neg X$	give a model in which X is true
Are X and Y **equivalent**?	give two proofs, one for $X \vdash Y$ and one for $Y \vdash X$	give a model in which X and Y have different truth values
Are $X_1, X_2, \ldots, X_n$ **jointly consistent**?	give a model in which all of $X_1, X_2, \ldots, X_n$ are true	prove a contradiction from assumptions $X_1, X_2, \ldots, X_n$
Is $X_1, X_2, \ldots, X_n \therefore C$ **valid**?	give a proof with assumptions $X_1, X_2, \ldots, X_n$ and concluding with C	give a model in which each of $X_1, X_2, \ldots, X_n$ is true and C is false

Appendices

APPENDIX A

Symbolic notation

1.1 Alternative nomenclature

Truth-functional logic. TFL goes by other names. Sometimes it is called *sentential logic*, because it deals fundamentally with sentences. Sometimes it is called *propositional logic*, on the idea that it deals fundamentally with propositions. We have stuck with *truth-functional logic*, to emphasize the fact that it deals only with assignments of truth and falsity to sentences, and that its connectives are all truth-functional.

First-order logic. FOL goes by other names. Sometimes it is called *predicate logic*, because it allows us to apply predicates to objects. Sometimes it is called *quantified logic*, because it makes use of quantifiers.

Formulas. Some texts call formulas *well-formed formulas*. Since 'well-formed formula' is such a long and cumbersome phrase, they then abbreviate this as *wff*. This is both barbarous and unnecessary (such texts do not countenance 'ill-formed formulas'). We have stuck with 'formula'.

In §4.3, we defined *sentences* of TFL. These are also sometimes called 'formulas' (or 'well-formed formulas') since in TFL, unlike FOL, there is no distinction between a formula and a sentence.

Valuations. Some texts call valuations *truth-assignments*, or *truth-value assignments*.

Expressive adequacy. Some texts describe TFL as *truth-functionally complete*, rather than expressively adequate.

n-place predicates. We have chosen to call predicates 'one-place', 'two-place', 'three-place', etc. Other texts respectively call them 'monadic', 'dyadic', 'triadic', etc. Still other texts call them 'unary', 'binary', 'ternary', etc.

Names. In FOL, we have used 'a', 'b', 'c', for names. Some texts call these 'constants'. Other texts do not mark any difference between names and variables in the syntax. Those texts focus simply on whether the symbol occurs *bound* or *unbound*.

Domains. Some texts describe a domain as a 'domain of discourse', or a 'universe of discourse'.

1.2 Alternative symbols

In the history of formal logic, different symbols have been used at different times and by different authors. Often, authors were forced to use notation that their printers could typeset.

This appendix presents some common symbols, so that you can recognize them if you encounter them in an article or in another book.

Negation. Two commonly used symbols are the *hoe*, '$\neg$', and the *swung dash* or *tilda*, '$\sim$.' There are some issues typing '$\neg$' on a keyboard, and '$\sim$' is perfectly acceptable for you to use. In some more advanced formal systems it is necessary to distinguish between two kinds of negation; the distinction is sometimes represented by using both '$\neg$' and '$\sim$'. Older texts sometimes indicate negation by a line over the formula being negated, e.g., $\overline{A \wedge B}$. Some texts use '$x \neq y$' to abbreviate '$\neg\, x = y$'.

Disjunction. The symbol '$\vee$' is typically used to symbolize inclusive disjunction. One etymology is from the Latin word 'vel', meaning 'or'.

Conjunction. The two symbols commonly used for conjuction are *wedge*, '$\land$', and *ampersand*, '&'. The ampersand is a decorative form of the Latin word 'et', which means 'and'. (Its etymology still lingers in certain fonts, particularly in italic fonts; thus an italic ampersand might appear as '*&*'.) We have chosen to use it to allow for easier typing on a keyboard during these online-heavy times. However there are some substantial reservations about this choice. This symbol is commonly used in natural English writing (e.g. 'Smith & Sons'), and so even though it is a natural choice, many logicians use a different symbol to avoid confusion between the object and metalanguage: as a symbol in a formal system, the ampersand is not the English word '&'. The most common choice now is '$\land$', which is a counterpart to the symbol used for disjunction. Sometimes a single dot, '$\bullet$', is used. In some older texts, there is no symbol for conjunction at all; 'A and B' is simply written 'AB'.

Material Conditional. There are two common symbols for the material conditional: the *arrow*, '$\rightarrow$', and the *hook*, '$\supset$'.

Material Biconditional. The *double-headed arrow*, '$\leftrightarrow$', is used in systems that use the arrow to represent the material conditional. Systems that use the hook for the conditional typically use the *triple bar*, '$\equiv$', for the biconditional.

Quantifiers. The universal quantifier is typically symbolized as a rotated 'A', and the existential quantifier as a rotated, 'E'. In some texts, there is no separate symbol for the universal quantifier. Instead, the variable is just written in parentheses in front of the formula that it binds. For example, they might write '$(x)Px$' where we would write '$\forall x\, Px$'.

These alternative typographies are summarised below:

negation	$\neg$, $\sim$
conjunction	$\land$, &, $\bullet$
disjunction	$\lor$
conditional	$\rightarrow$, $\supset$
biconditional	$\leftrightarrow$, $\equiv$
universal quantifier	$\forall x$, (x)

APPENDIX B

Quick reference

2.1 Sentences of TFL

Definition of being a sentence of TFL:

1. Every atomic sentence is a sentence.
    - $A, B, C \ldots, W$, or with subscripts $A_1, B_3, A_{100}, J_{375}$
2. If X is a sentence, then $\neg X$ is a sentence.
3. If X and Y are sentences, then $(X \wedge Y)$ is a sentence.
4. If X and Y are sentences, then $(X \vee Y)$ is a sentence.
5. If X and Y are sentences, then $(X \rightarrow Y)$ is a sentence.
6. If X and Y are sentences, then $(X \leftrightarrow Y)$ is a sentence.
7. Nothing else is a sentence.

2.2 Truth Rules for Connectives in TFL

$$\neg: \begin{array}{l} \text{T} \rightsquigarrow \text{F} \\ \text{F} \rightsquigarrow \text{T} \end{array} \qquad \wedge: \begin{array}{l} \text{T, T} \rightsquigarrow \text{T} \\ \text{T, F} \rightsquigarrow \text{F} \\ \text{F, T} \rightsquigarrow \text{F} \\ \text{F, F} \rightsquigarrow \text{F} \end{array} \qquad \vee: \begin{array}{l} \text{T, T} \rightsquigarrow \text{T} \\ \text{T, F} \rightsquigarrow \text{T} \\ \text{F, T} \rightsquigarrow \text{T} \\ \text{F, F} \rightsquigarrow \text{F} \end{array}$$

$$\rightarrow: \quad \begin{array}{l} \mathrm{T, T} \rightsquigarrow \mathrm{T} \\ \mathrm{T, F} \rightsquigarrow \mathrm{F} \\ \mathrm{F, T} \rightsquigarrow \mathrm{T} \\ \mathrm{F, F} \rightsquigarrow \mathrm{T} \end{array}$$

2.3 Symbolization

Rough Meaning of the TFL Connectives

symbol	name	rough meaning
$\neg$	negation	'It is not the case that...'
$\wedge$	conjunction	'... and ...'
$\vee$	disjunction	'... or ...'
$\rightarrow$	conditional	'If ... then ...'
$\leftrightarrow$	biconditional	'... if and only if ...'

Sentential Connectives

It is not the case that P	$\neg P$
P or Q	$(P \vee Q)$
P and Q	$(P \wedge Q)$
If P, then Q	$(P \rightarrow Q)$
P if and only if Q	$(P \leftrightarrow Q)$

Further symbolisation help:

Neither P nor Q	$\neg(P \vee Q)$ or $(\neg P \wedge \neg Q)$
P but Q	$(P \wedge Q)$
P unless Q	$(P \vee Q)$
P only if Q	$(P \rightarrow Q)$

Predicates

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

Only c is G	$\forall x(Gx \rightarrow x = c)$ or perhaps $\leftrightarrow$.
Everything besides c is G	$\forall x(\neg x = c \rightarrow Gx)$
The F is G	$\exists x(Fx \wedge \forall y(Fy \rightarrow x = y) \wedge Gx)$
It is not the case that the F is G	$\neg\exists x(Fx \wedge \forall y(Fy \rightarrow x = y) \wedge Gx)$
The F is non-G	$\exists x(Fx \wedge \forall y(Fy \rightarrow x = y) \wedge \neg Gx)$

2.4 Using identity to symbolize quantities

There are at least _____ Fs.

one: $\exists x Fx$

two: $\exists x_1 \exists x_2 (Fx_1 \wedge Fx_2 \wedge \neg x_1 = x_2)$

three: $\exists x_1 \exists x_2 \exists x_3 (Fx_1 \wedge Fx_2 \wedge Fx_3 \wedge \neg x_1 = x_2 \wedge \neg x_1 = x_3 \wedge \neg x_2 = x_3)$

four: $\exists x_1 \exists x_2 \exists x_3 \exists x_4 (Fx_1 \wedge Fx_2 \wedge Fx_3 \wedge Fx_4 \wedge$
$\neg x_1 = x_2 \wedge \neg x_1 = x_3 \wedge \neg x_1 = x_4 \wedge \neg x_2 = x_3 \wedge \neg x_2 = x_4 \wedge \neg x_3 = x_4)$

n: $\exists x_1 \ldots \exists x_n (Fx_1 \wedge \ldots \wedge Fx_n \wedge \neg x_1 = x_2 \wedge \ldots \wedge \neg x_{n-1} = x_n)$

There are at most _____ Fs.

One way to say 'there are at most n Fs' is to put a negation sign in front of the symbolization for 'there are at least $n + 1$ Fs'. Equivalently, we can offer:

one: $\forall x_1 \forall x_2 \big[(Fx_1 \wedge Fx_2) \rightarrow x_1 = x_2\big]$

two: $\forall x_1 \forall x_2 \forall x_3 \big[(Fx_1 \wedge Fx_2 \wedge Fx_3) \rightarrow (x_1 = x_2 \vee x_1 = x_3 \vee x_2 = x_3)\big]$

three: $\forall x_1 \forall x_2 \forall x_3 \forall x_4 \big[(Fx_1 \wedge Fx_2 \wedge Fx_3 \wedge Fx_4) \rightarrow$
$(x_1 = x_2 \vee x_1 = x_3 \vee x_1 = x_4 \vee x_2 = x_3 \vee x_2 = x_4 \vee x_3 = x_4)\big]$

n: $\forall x_1 \ldots \forall x_{n+1} \big[(Fx_1 \wedge \ldots \wedge Fx_{n+1}) \rightarrow (x_1 = x_2 \vee \ldots \vee x_n = x_{n+1})\big]$

There are exactly _____ Fs.

One way to say 'there are exactly n Fs' is to conjoin two of the symbolizations above and say 'there are at least n Fs and there are at most n Fs.' The following equivalent formulae are shorter:

zero: $\forall x \neg Fx$

one: $\exists x \big[Fx \wedge \forall y(Fy \rightarrow x = y)\big]$

two: $\exists x_1 \exists x_2 \big[Fx_1 \wedge Fx_2 \wedge \neg x_1 = x_2 \wedge \forall y\big(Fy \rightarrow (y = x_1 \vee y = x_2)\big)\big]$

three: $\exists x_1 \exists x_2 \exists x_3 \big[Fx_1 \wedge Fx_2 \wedge Fx_3 \wedge \neg x_1 = x_2 \wedge \neg x_1 = x_3 \wedge \neg x_2 = x_3 \wedge$
$\forall y\big(Fy \rightarrow (y = x_1 \vee y = x_2 \vee y = x_3)\big)\big]$

$$n\colon\ \exists x_1 \ldots \exists x_n \big[F x_1 \wedge \ldots \wedge F x_n \wedge \neg x_1 = x_2 \wedge \ldots \wedge \neg x_{n-1} = x_n \wedge \\ \forall y \big(F y \rightarrow (y = x_1 \vee \ldots \vee y = x_n) \big) \big]$$

2.5 Basic deduction rules for TFL

Conjunction

$$\begin{array}{r|ll}
m & X & \\
n & Y & \\
 & X \wedge Y & \wedge\text{I}, m, n
\end{array}$$

$$\begin{array}{r|ll}
m & X \wedge Y & \\
 & X & \wedge\text{E}, m
\end{array}$$

$$\begin{array}{r|ll}
m & X \wedge Y & \\
 & Y & \wedge\text{E}, m
\end{array}$$

Disjunction

$$\begin{array}{r|ll}
m & X & \\
 & X \vee Y & \vee\text{I}, m
\end{array}$$

$$\begin{array}{r|ll}
m & X & \\
 & Y \vee X & \vee\text{I}, m
\end{array}$$

$$\begin{array}{r|ll}
m & X \vee Y & \\
i & \quad \vert\ \underline{X} & \\
 & \quad \vert\ \vdots & \\
j & \quad \vert\ Z & \\
k & \quad \vert\ \underline{Y} & \\
 & \quad \vert\ \vdots & \\
l & \quad \vert\ Z & \\
 & Z & \vee\text{E}, m, i\text{–}j, k\text{–}l
\end{array}$$

Conditional

$$\begin{array}{r|ll}
m & \quad \vert\ \underline{X} & \\
 & \quad \vert\ \vdots & \\
n & \quad \vert\ Y & \\
 & X \rightarrow Y & \rightarrow\text{I}, m\text{–}n
\end{array}$$

$$\begin{array}{r|ll}
m & X \rightarrow Y & \\
n & X & \\
 & Y & \rightarrow\text{E}, m, n
\end{array}$$

Negation

$$\begin{array}{r|ll}
m & \quad \vert\ \underline{X} & \\
 & \quad \vert\ \vdots & \\
k & \quad \vert\ Y \wedge \neg Y & \\
 & \neg X & \neg\text{I}, m\text{–}k
\end{array}$$

$$\begin{array}{r|ll}
m & \quad \vert\ \underline{\neg X} & \\
 & \quad \vert\ \vdots & \\
k & \quad \vert\ Y \wedge \neg Y & \\
 & X & \neg\text{E}, m\text{–}k
\end{array}$$

Reiteration

$$
\begin{array}{r|ll}
m & X & \\
 & X & \text{R}, m
\end{array}
$$

Law of Excluded Middle

$$
\begin{array}{|ll}
X \lor \neg X & \text{LEM}
\end{array}
$$

2.6 Derived rules for TFL

Disjunctive syllogism

$$\begin{array}{r|ll} m & X \vee Y & \\ n & \neg X & \\ & Y & \text{DS}, m, n \end{array}$$

$$\begin{array}{r|ll} m & X \vee Y & \\ n & \neg Y & \\ & X & \text{DS}, m, n \end{array}$$

Modus Tollens

$$\begin{array}{r|ll} m & X \rightarrow Y & \\ n & \neg Y & \\ & \neg X & \text{MT}, m, n \end{array}$$

Double-negation elimination

$$\begin{array}{r|ll} m & \neg\neg X & \\ & X & \text{DNE}, m \end{array}$$

Explosion

$$\begin{array}{r|ll} m & X & \\ n & \neg X & \\ & Y & \text{Explosion}, m, n \end{array}$$

De Morgan Rules

$$\begin{array}{r|ll} m & \neg(X \vee Y) & \\ & \neg X \wedge \neg Y & \text{DeM}, m \end{array}$$

$$\begin{array}{r|ll} m & \neg X \wedge \neg Y & \\ & \neg(X \vee Y) & \text{DeM}, m \end{array}$$

$$\begin{array}{r|ll} m & \neg(X \wedge Y) & \\ & \neg X \vee \neg Y & \text{DeM}, m \end{array}$$

$$\begin{array}{r|ll} m & \neg X \vee \neg Y & \\ & \neg(X \wedge Y) & \text{DeM}, m \end{array}$$

2.7 Basic deduction rules for FOL

Universal elimination

$$\begin{array}{l|ll} m & \forall x X(\ldots x \ldots x \ldots) & \\ & X(\ldots c \ldots c \ldots) & \forall\mathrm{E},\ m \end{array}$$

Universal introduction

$$\begin{array}{l|ll} m & X(\ldots c \ldots c \ldots) & \\ & \forall x X(\ldots x \ldots x \ldots) & \forall\mathrm{I},\ m \end{array}$$

c must not occur in any undischarged assumption

x must not occur in $X(\ldots c \ldots c \ldots)$

Existential introduction

$$\begin{array}{l|ll} m & X(\ldots c \ldots c \ldots) & \\ & \exists x X(\ldots x \ldots c \ldots) & \exists\mathrm{I},\ m \end{array}$$

x must not occur in $X(\ldots c \ldots c \ldots)$

Existential elimination

$$\begin{array}{l|l|ll} m & \exists x X(\ldots x \ldots x \ldots) & & \\ i & & \underline{X(\ldots c \ldots c \ldots)} & \\ & & \vdots & \\ j & & Y & \\ & Y & & \exists\mathrm{E},\ m,\ i\text{–}j \end{array}$$

c must not occur in
any undischarged assumption,
in $\exists x X(\ldots x \ldots x \ldots)$,
or in Y

Identity introduction

$$\begin{array}{|ll} c = c & =\mathrm{I} \end{array}$$

Identity elimination

$$\begin{array}{l|ll} m & a = b & \\ n & X(\ldots a \ldots a \ldots) & \\ & X(\ldots b \ldots a \ldots) & =\mathrm{E},\ m,\ n \end{array}$$

$$\begin{array}{l|ll} m & a = b & \\ n & X(\ldots b \ldots b \ldots) & \\ & X(\ldots a \ldots b \ldots) & =\mathrm{E},\ m,\ n \end{array}$$

2.8 Derived rules for FOL

$$\begin{array}{l|ll} m & \forall x \neg X & \\ & \neg \exists x X & \text{CQ}, m \end{array}$$

$$\begin{array}{l|ll} m & \exists x \neg X & \\ & \neg \forall x X & \text{CQ}, m \end{array}$$

$$\begin{array}{l|ll} m & \neg \exists x X & \\ & \forall x \neg X & \text{CQ}, m \end{array}$$

$$\begin{array}{l|ll} m & \neg \forall x X & \\ & \exists x \neg X & \text{CQ}, m \end{array}$$

In the Introduction to his volume *Symbolic Logic*, Charles Lutwidge Dodson advised: "When you come to any passage you don't understand, *read it again*: if you *still* don't understand it, *read it again*: if you fail, even after *three* readings, very likely your brain is getting a little tired. In that case, put the book away, and take to other occupations, and next day, when you come to it fresh, you will very likely find that it is *quite* easy."

The same might be said for this volume, although readers are forgiven if they take a break for snacks after *two* readings.