forall$x$

BRISTOL (Accessible)

For PHIL10032 Logic and Critical Thinking
University of Bristol

In-progress version for 2026-27
Last updated August 17, 2026. Please check for updates.

Maintainer for this version:
Catrin Campbell-Moore catrin.campbell-moore@bristol.ac.uk

Links:

- Most recent version of textbook for 2026–27
- Accessibility formatting of most recent version of textbook for 2026–27
  Designed with students in dyslexia in mind.
  Adopted from Richard Zach (see discussion).
- Other editions
- Information

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.

forall$x$:Bristol is licensed under a Creative Commons Attribution-ShareAlike 4.0 license. You are free to copy and

redistribute the material in any medium or format, and remix, transform, and build upon the material for any purpose, even commercially, under the following terms:

- You must give appropriate credit, provide a link to the license, and indicate if changes were made. You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use.
- If you remix, transform, or build upon the material, you must distribute your contributions under the same license as the original.

Preface

The book is divided into the following parts. Part I introduces the topic and notions of logic in an informal way, without introducing a formal language yet. Part 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). Part VII–VIII 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 alterations for the Bristol course. The resulting text is licensed under a Creative Commons Attribution–ShareAlike 4.0 license.

Contents

Preface v

I Arguments 1

1 Arguments 2
2 The scope of logic 25

II Truth-functional logic 54

3 A Prolegomenon to TFL 55
4 The connectives 65
5 Validity 81
6 Use and mention 101
7 Complex sentences of TFL 113
8 Truth Tables 138

9 Validity and the method of truth tables 154
10 Other logical notions 162
11 Truth table shortcuts 174
12 Partial truth tables 184

III Symbolizations in TFL 190

13 First steps towards Symbolization 191
14 Symbolization 195
15 On Truth-functional connectives 219
16 Expanding our Symbolisation Guide 227
17 Ambiguity 261

IV Natural deduction for TFL 273

18 The very idea of natural deduction 274
19 The First Basic Rules 286
20 The remaining Basic Rules for TFL 304
21 Proofs and Validity 330
22 Derived rules for TFL 342
23 Derivations of the derived rules 357
24 Soundness and completeness 366
25 Proof-theoretic concepts 380

V First-order logic 386

26 Building blocks of FOL 387
27 Sentences of FOL 441

VI First Order Logic Semantics 457

28 Truth in a model 458
29 Truth in a model, more carefully 496
30 Logical notions 519

VII First-order logic symbolisation 530

31 Common quantifier phrases 531
32 Ambiguity 549

VIII Natural deduction for FOL 562

33 Basic rules for FOL 563
34 Conversion of quantifiers 600
35 Derived rules 604
36 Proof-theoretic and semantic concepts 607

Appendices 614

A Answers to practice exercises 614
B Symbolic notation 702
C Quick reference 708

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:

- Premise 1
- Premise 2

  ...
- Premise *n*

∴ Therefore: Conclusion

For example

- 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.

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:

– If I have free will, then there is some event that I could have caused to go differently.
– 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:

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

'Since' is also a very common and useful indicator. Consider

- Since the first Brexit referendum was made under false pretences, the voters deserve a further say on any final deal agreed with Brussels.

This both gives us a premise:

- The first Brexit referendum was made under false pretences.

and a conclusion:

- The voters deserve a further say on any final deal agreed with Brussels.

> Since *X*, *Y*
>
> indicates that *X* is a premise, and *Y* is the conclusion.

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.

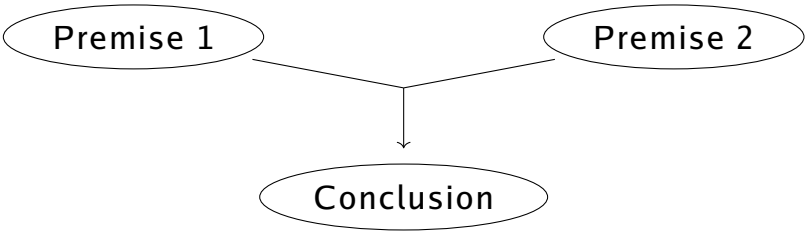
Sometimes, though, people aren't giving arguments but are simply presenting facts or stating their opinion. For example, the following do not contain arguments, they're not trying to convince us of anything.

- 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.

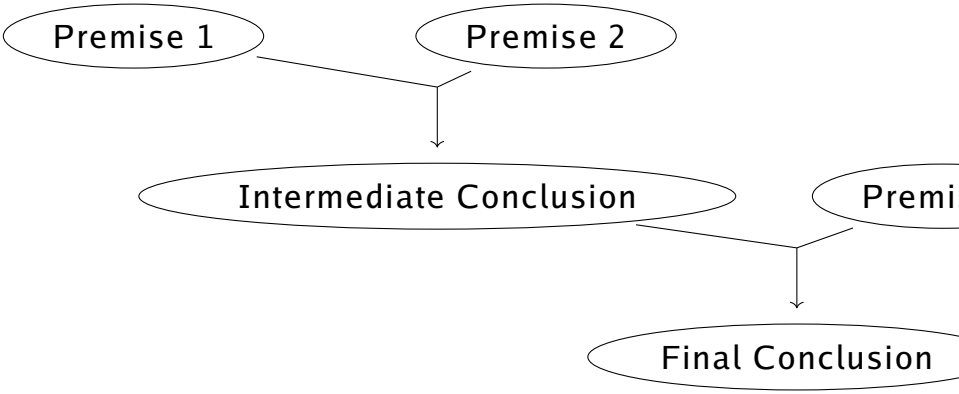
1.3 Intermediate Conclusions

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'.

Definition

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:

- 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 (Chapter 1)

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.

Some of the answers to the practice questions are in chapter A, but it is important to first try yourself. The answers are provided on a best–endeavours basis. There will be gaps and mistakes. Questions with answers are indicated with stars. Mistakes can be reported to catrin.campbell–moore@bristol.ac.uk.

So here's the first exercise.

I.

a) Are arguments always presented in our standard form?
b) Do conclusions always come after the premises in arguments in texts?
c) Might premises and conclusions be clauses within sentences?
d) Can questions be premises?

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

a) It is sunny. So I should take my sunglasses. (∗)
b) It must have been sunny. I did wear my sunglasses, after all. (∗)
c) 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!
d) 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.
e) 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.
f) 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.

g) 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**

h) 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.

i) 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

III. 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.

a) Answer:

– 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.

b) It is sunny. So I should take my sunglasses. (∗)

c) It must have been sunny. I did wear my sunglasses, after all. (∗)

d) 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! (∗)

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

f) 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.

g) 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. (∗)

Chapter 2

The scope of logic

2.1 Consequence and validity

In chapter 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:

– The butler or the gardener did it.
– 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:

- If the driver did it, the maid didn't do it.
- The maid didn't do it.
- ∴ Therefore: The driver did it.

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", perhaps it was the gardener. 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:

– The crocodile or the kangaroo stole the cake.
– 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.

Definition

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.

- If the driver did it, the maid didn't do it.
- 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:

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

- Priya is either an ophthalmologist or a dentist.
- Priya isn't a dentist.

∴ Therefore: Priya is an ophthalmologist.

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

- *a* is either an *F* or a *G*.
- *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,

– Mei is either a mathematician or a botanist.
– 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:

– Priya is either an ophthalmologist or a dentist.
– 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:

– *a* is either an *F* or a *G*.
– *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":

– Mei is either a mathematician or a botanist.
– 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:

– Oranges are either fruit or musical instruments.
– 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:

– London is in England.
– 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

- Priya is either an ophthalmologist or a dentist.
- 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

- 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:

– 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:

- 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:

– 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 *F*s have been *G*s, we infer all *F*s are *G*s. 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

– Priya is an ophthalmologist.
∴ Therefore: Priya is an eye doctor.

– 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 (Chapter 2)

I.

a) What kind of things are valid or invalid?
b) When is an argument said to be valid?
c) When is an argument said to be sound?

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

x. – Every good zoo has a giraffe.
– It is a zoo.
∴ Therefore: It has a giraffe.
Invalid. It is a zoo, but not a good one. (And has a giraffe.)

a) – Everyone in group 1 handed in their homework.
– Jenny is in group 1.
∴ Therefore: Jenny handed in her homework.

(∗)

b) – If she won the lottery then she is rich.
– She is rich.

∴ Therefore: She won the lottery.

(∗)

c) – Most people are scared of spiders.
∴ Therefore: Oscar is scared of spiders.

(∗)

d) – She is a donkey.
∴ Therefore: She does not talk.

e) – Socrates is a man.
– All men are carrots.
∴ Therefore: Socrates is a carrot. (∗)

f) – Abe Lincoln was either born in Illinois or he was president.
– Abe Lincoln was not president.
∴ Therefore: Abe Lincoln was born in Illinois. (∗)

g) – If I pull the trigger, Abe Lincoln will die.
– I do not pull the trigger.
∴ Therefore: Abe Lincoln will not die. (∗)

h) – Abe Lincoln was either from France or from Luxembourg.
– Abe Lincoln was not from Luxembourg.
∴ Therefore: Abe Lincoln was from France.

(∗)

i) – If the world were to end today, then I would not need to get up tomorrow morning.
– I will need to get up tomorrow morning.
∴ Therefore: The world will not end today. (∗)

j) – Joe is now 19 years old.
– Joe is now 87 years old.
∴ Therefore: Bob is now 20 years old. (∗)

III. Could there be:

a) A valid argument that has one false premise and one true premise? (∗)
b) A valid argument that has only false premises? (∗)
c) A valid argument with only false premises and a false conclusion? (∗)
d) An invalid argument that can be made valid by the addition of a new premise? (∗)
e) A valid argument that can be made invalid by the addition of a new premise? (∗)

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

Part II

Truth-functional logic

Chapter 3

A Prole-gomenon 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 chapter 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 chapter 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:

- It is raining outside.
- If it is raining outside, then Jenny is miserable.

∴ Therefore: Jenny is miserable.

and another argument:

- Jenny is an anarcho-syndicalist.

- 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:

- A
- 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:

- Jenny is miserable.
- If it is raining outside, then Jenny is miserable.

∴ Therefore: It is raining outside.

The form of this argument is:

– *B*
– 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

– *A* or *B*
– not–*A*
∴ Therefore: *B*

as well as the form

– not–(*A* and *B*)
– *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 chapter 26.

3.1 Vocabulary of TFL

The vocabulary of TFL is very simple. There are three parts to the vocabulary: atomic sentences, connectives, and brackets.

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 chapter 26 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.

Connectives

The second category of the vocabulary of TFL is the connectives. They can be used to form new sentences out of old ones.

symbol	what it is called	rough meaning
¬	negation	'It is not the case that...'
∧	conjunction	'... and ...',
∨	disjunction	'... or ...'
→	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 '¬','∧','∨' and '→'

combine other sentences. For example, the atomic sentences 'R' and 'M' combine with '$\wedge$' to form the sentence '$R \wedge M$'. If think of 'R' and 'M' as representing the English sentences

R: It is raining outside
M: Jenny is miserable

'$R \wedge M$' can be read as:

- It is raining and Jenny is miserable.

Similarly, on this understanding we get the following readings:

$\neg R$	It is not the case that it is raining outside. (It is not raining outside).
$R \wedge M$	It is raining outside and Jenny is miserable.
$R \vee M$	It is raining outside or Jenny is miserable.
$R \rightarrow M$	If it is raining outside, then Jenny is miserable.

Brackets

Brackets are the remaining component of the vocabulary of TFL. They will be used to provide the grammar. They are vital when we start looking at complex sentences of TFL. For example '$A \wedge (B \vee C)$'. But, to start to understand the connectives we will ignore this for now and just focus on one connective applied to atomic sentences.

Chapter 4

The connectives

Whilst this provides a gloss on our new symbols of TFL, their role in TFL in terms of their contribution to the truth conditions of a sentence is explicitly specified. The important feature of truth functional logic is that the truth value of a sentence such as '$A \wedge B$' is determined entirely by the truth value of 'A' and the truth value of 'B'. The truth value of $\neg A$ is determined entirely by the truth value of A.

4.1 Negation

The '¬'–connective is called negation. It should roughly be understood as 'it is not the case' or, perhaps, simply 'not'.

If think of 'A' as representing the English sentence 'Alice is happy':

A: Alice is happy

Then '$\neg A$' will be used to symbolise the English sentence:

1. It is not the case that Alice is happy

Is this sentence true or false? It depends on whether Alice is happy. She either is or isn't happy, that is 'Alice is happy' might be true or false:

A	
T	Alice is happy
F	Alice is not happy

In which cases is '$\neg A$' true? Just when 'A' is false:

A	$\neg A$
T	F
F	T

The negation symbol of TFL, $\neg$, is defined to have this affect on the truth value of the sentence. Whether $\neg A$ is then a good symbolisation of a given sentence then depends on whether it is appropriately governed by this truth table.

4.2 Conjunction

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

If think of 'A' and 'B' as representing the English sentences

A: Alice is happy
B: Beth is happy

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

2. Alice is happy **and** Beth is happy.

Is it true or false? It depends on whether Alice is happy and whether Beth is happy. Here are the four possibilities:

	A	B	
v_1	T	T	Alice and Beth are both happy
v_2	T	F	Alice is happy but Beth is not
v_3	F	T	Alice is not happy but Beth is
v_4	F	F	Neither Alice nor Beth are happy

In which cases is '$A \wedge B$' true? Only v_1. That is, only if 'A' is true and 'B' is true. This is:

A	B	$A \wedge B$
T	T	T
T	F	F
F	T	F
F	F	F

4.3 Disjunction

If think of 'G' and 'F' as representing the English sentences

G: She understands German
F: She understands French

'$G \vee F$' can be read as:

3. She understands German or she understands French.

Or, more concisely:

4. She understands German or French.

Again whether this is true will depend on G and F. Here are the four possibilities:

	G	F	
v_1	T	T	She understands German and French
v_2	T	F	She understands German but not French
v_3	F	T	She understands French but not German
v_4	F	F	She understands neither German nor French

Clearly 'She understands German or French' is false in v_4 where she understands neither; and

true in v_2 and v_3 where she understands one of these languages:

	G	F	$G \vee F$
v_1	T	T	
v_2	T	F	T
v_3	F	T	T
v_4	F	F	F

But what about v_1, where she understands both? 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 understand German or French (or both).
6. She can understand German or French (but not both).

Sometimes English is ambiguous between the two readings.

In logic there can be no ambiguity. We stipulate that $\vee$ stands for the **inclusive or**, giving it the truth value true in v_1:

	G	F	$G \vee F$
v_1	T	T	T
v_2	T	F	T
v_3	F	T	T
v_4	F	F	F

That is '$G \vee F$' is true if and only if 'G' is true or 'F' is true (or both). The "or both" will be dropped throughout, and we will understand the English 'or' to be given an inclusive reading unless explicitly discussed.

If one wishes to capture the exclusive use of English, this will be perfectly possible in TFL, but for that, we'll have to use a more complex sentence. We will essentially make explicit the whispered "but not both". This will end up being symbolised as:

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

She can understand German or French, and it's not the case that she can understand both German and French.

4.4 Conditional

The '→'–connective is called the conditional–connective and it is meant to be related to our understanding of **if..., then...** sentences.

Consider

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

These are very different. We might have atomic sentences:

B: She is drinking beer.
E: She is over eighteen.

We will then represent sentence 7 as '$B \to E$' and sentence 8 as '$E \to B$'.

Since it matters which order they come in, we is valuable to give a name to indicate which part goes where:

Definition

In $X \to Y$, X is called the **antecedent**, and Y is called the **consequent**.

In what cases is '$B \to E$' true or false? Here are the four cases:

	B	E	
v_1	T	T	She is drinking beer and is over eighteen
v_2	T	F	She is drinking beer and is not over eighteen
v_3	F	T	She is not drinking beer and she is over eightee
v_4	F	F	She is not drinking beer and she is not over eig

In v_2 it is false, and in v_1 it is true.

	B	E	$B \to E$
v_1	T	T	T
v_2	T	F	F
v_3	F	T	
v_4	F	F	

What about v_3 and v_4? We stipulate that it is true in v_3 and v_4: if the antecedent of the conditional is false, the conditional is true.

	B	E	$B \to E$
v_1	T	T	T
v_2	T	F	F
v_3	F	T	T
v_4	F	F	T

To understand the rationale for this, let us think about when a bartender would get into trouble (clearly the conditional 7 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.

For our purposes $\rightarrow$–connective will be understood in terms of truth rules given in this section. That is, we define '$\rightarrow$' by these truth 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

That is, $X \rightarrow Y$ is true if and only if it is not the case that X is true and Y is false; or, equivalently, if X is false or Y is true.

The conditional which is defined by this truth rule is sometimes marked by calling it the **material conditional**. It is the conditional that we will use in the unit.

This material conditional says nothing about what the world **would be like** were the antecedent to be true. For the material conditional, if the antecedent doesn't apply, in this case, she is not drinking beer, then the conditional is simply trivially true independently of what would happen were she to be drinking a beer.

There is a rich philosophical and logical investigation into conditionals, but for this unit, we will stick with the material conditional. The material conditional is the best that can be done to capture the conditional in a truth–functional way, where the truth can be determined simply from the truth values of its constituent parts. See https://plato.stanford.edu/entries/conditionals and section 15.2 for further discussion

As we highlighted, the order matters. So lets consider '$E \to B$': if she is over eighteen then she is drinking beer. How does it differ from $B \to E$? We can apply this rule, but we have to be careful which order to apply it in. To help make sure we apply it in the right order, it might be helpful to

copy down the values for E and B before applying the rule:

	B	E	$B \to E$	$E \to B$	
v_1	T	T	T	T	T
v_2	T	F	F	F	T
v_3	F	T	T	T	F
v_4	F	F	T	F	T

and we can then apply the rule. We will write the answer underneath the '$\to$'. As we have extra truth values hanging around in the table, we have highlighted the column of truth values which give the truth value for the sentence '$E \to B$'.

	B	E	$B \to E$	$E \to B$	
v_1	T	T	T	T	T	T
v_2	T	F	F	F	T	T
v_3	F	T	T	T	F	F
v_4	F	F	T	F	T	T

'$E \to B$' is false when she is over eighteen and not drinking beer. But otherwise it is true.

4.5 Truth rules

These are the rules associated with the connectives. They are motivated by the English uses, but are stipulated to be governed by these truth rules.

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 $X \rightarrow Y$ is:
T	T	$\rightsquigarrow$	T
T	F	$\rightsquigarrow$	F
F	T	$\rightsquigarrow$	T
F	F	$\rightsquigarrow$	T

If you want to capture an exclusive disjunction, you can use a more complex sentence of TFL. In fact, anything where the truth values of

the sentence depends entirely on the truth of the component sentences is something that can be captured in truth-functional logic, possibly by a more complex sentence. Non-truth-functional notions, however, go beyond the resources of TFL. See chapter 15.

Practice Exercises (Chapter 4)

I. Write down (from memory) the truth rules for each of the connectives.

II.

a) Suppose X is true, what can you say about $X \vee Y$?
b) Suppose X is false, what can you say about $X \rightarrow Y$?

Chapter 5

Validity

5.1 Validity

Example

We are given a cards on the table which has a number on one side and a colour on the other side. You know a rule: ‘If the card is blue, it has an even number on it’. You can see the card is red. What can you conclude from this? You try to reason as follows:

- If the card is blue, it has an even number on it.
- The card is not blue.

$\therefore$ Therefore: The card does not have an even number.

Is this a good pattern of reasoning? Is it a valid argument?

We will symbolise this argument as:

$$B \rightarrow E,\ \neg B\ \therefore\ \neg E$$

An argument is valid if the premises guarantee the conclusion. There is no way for the premises to be true and the conclusion false. Whether the sentences in this argument are true or false depend on whether B is true or false and whether E is true or false. Here are the four possibilities:

	B	E	
v_1	T	T	The card is blue and is even
v_2	T	F	The card is blue and is not even
v_3	F	T	The card is not blue and is even
v_4	F	F	The card is not blue and is not even

They are called “valuations”.

Definition

A **valuation** is any assignment of truth values (true or false) to the atomic sentences.

What is special about the logic we are introducing here, called Truth Functional Logic, is that to know the truth value of any sentence, one just needs to specify the truth value of the atomic sentences, which is exactly what is given by a valuation.

Now, to check validity, we will evaluate whether the premises and conclusions are true or false in each of these valuations.

			premise 1	premise 2	conclusion	
	B	E	$B \rightarrow E$	$\neg B$	$\neg E$	
v_1	T	T	T	F	F	
v_2	T	F	F	F	T	
v_3	F	T	**T**	**T**	**F**	⇜ counter
v_4	F	F	T	T	T	

We look through these lines and see if we can find one of them where the premises are all true and the conclusion is false. We can: v_3!

The card might be red but still have an even number on it. Nothing in the information you are given (i.e., the premises) rules out the possibility. Your drawing the conclusion the the card was odd was a mistake of reasoning.

5.2 Definition of Validity

Definition

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 $X_1, \ldots, X_n$ are all true relative to v, but Y is false relative to v.

To check validity with truth tables, draw a truth table for all the sentences simultaneously and look for a row like:

Valuation	Premise 1	Premise 2	⋯	Conclusion
v_1				
★ v_2	T	T	T	F
v_3				
v_4				

Invalid

If there is no such row, then it is **valid**.

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

We will also say:

Definition

Y is a **logical consequence** of $X_1, \ldots, X_n$ if and only if there is no valuation on which $X_1, \ldots, X_n$ are true and Y false.

That is, Y is a logical consequence of $X_1, \ldots, X_n$ if and only if $X_1, \ldots, X_n \therefore Y$ is valid.

Example

Consider this argument:

– The cook or the gardener did it.
– The cook did not do it.
∴ Therefore: The gardener did it.

We will follow the same approach:

1. Symbolise the premises and conclusion.
2. List every valuation and calculate the truth value of the premises and conclusion on each valuation.
3. Determine whether there is one possibility where the premises are all true and the conclusion is false.

Let:

C: The cook did it.
G: The gardener did it.

So the argument is:

$$C \vee G,\ \neg C \therefore G$$

	C	G	premise 1 $C \vee G$	premise 2 $\neg C$	conclusion G
v_1	T	T	T	F	T
v_2	T	F	T	F	F
v_3	F	T	**T**	**T**	**T**
v_4	F	F	F	T	F

For this argument, there is no valuation where both the premises are true and the conclusion is false. It is only v_3 where the premises are both true, and on v_3 the conclusion is also true. So the argument is valid.

More valuations

Consider this argument:

- If Alex went to the party, then Beth went to the party.
- If Beth went to the party, then Chris went to the party.

– Alex went to the party.
∴ Therefore: Chris went to the party.

Using the symbolisation key:

A: Alex went to the party.
B: Beth went to the party.
C: Chris went to the party.

The argument is symbolised as:

$$A \rightarrow B,\ B \rightarrow C,\ A \therefore C$$

Now we need to again list all the valuations and determine the truth of each of the premises and the conclusion on each. Now we have three atomic sentences to deal with. There are eight valuations:

	A	B	C
v_1	T	T	T
v_2	T	T	F
v_3	T	F	T
v_4	T	F	F
v_5	F	T	T
v_6	F	T	F
v_7	F	F	T
v_8	F	F	F

We can now use our truth rules to determine the truth of each sentence on each valuation. To aid this calculation, we first copy across the truth values for A, B and C to make sure we apply the truth rule for $\rightarrow$ to the right values in the right order. We then highlight the column for the main connective, which provides the truth value for the overall sentence at each valuation, indicating that the others were just calculation columns.

				premise 1		premise 2		premise 3	conclu
	A	B	C	$A \rightarrow$	B	$B \rightarrow$	C	A	C
v_1	T	T	T	T **T**	T	T **T**	T	**T**	**T**
v_2	T	T	F	T T	T	T F	F	T	F
v_3	T	F	T	T F	F	F T	T	T	T
v_4	T	F	F	T F	F	F T	F	T	F
v_5	F	T	T	F T	T	T T	T	F	T
v_6	F	T	F	F T	T	T F	F	F	F
v_7	F	F	T	F T	F	F T	T	F	T
v_8	F	F	F	F T	F	F T	F	F	F

We check whether there is a valuation where the premises are all true and the conclusion is false. v_1 is the only valuation where the premises are all true, and on this valuation, the conclusion is also true. It is therefore a valid argument.

5.3 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		A	B
T	→	T	T
	→	T	F
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		A	B	C
T	T	→	T	T	T
		→	T	T	F
T	F	→	T	F	T
		→	T	F	F
F	T	→	F	T	T
		→	F	T	F
F	F	→	F	F	T
		→	F	F	F

What if we add a further atomic sentence, say P_{23}? We want to consider the valuations for atomic sentences A, B, C, P_{23}. For each valuation

for A, B, C there is a version of it where P_{23} is true, and another where P_{23} is false. We thus duplicate each line of the table so far and add these two possible truth values for P_{23} in the two variants.

A	B	C		A	B	C	P_{23}
T	T	T	→	T	T	T	T
			→	T	T	T	F
T	T	F	→	T	T	F	T
			→	T	T	F	F
T	F	T	→	T	F	T	T
			→	T	F	T	F
T	F	F	→	T	F	F	T
			→	T	F	F	F
F	T	T	→	F	T	T	T
			→	F	T	T	F
F	T	F	→	F	T	F	T
			→	F	T	F	F
F	F	T	→	F	F	T	T
			→	F	F	T	F
F	F	F	→	F	F	F	T
			→	F	F	F	F

The list of all possible valuations gets very big very fast. It doubles for each new atomic sentence.

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 part 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.

5.4 Careful application of the definition

Check all the valuations

Consider this argument:

- If Ava studies, then Ava passes.
- Ava passes.

∴ Therefore: Ava studies.

Using the symbolisation key:

S: Ava studies.
P: Ava passes.

This argument is symbolised as:

$$S \to P,\ P \therefore S$$

	S	P	premise 1 $S \to P$	premise 2 P	conclusion S
v_1	T	T	**T**	**T**	**T**
v_2	T	F	F	F	T
v_3	F	T	**T**	**T**	**F**
v_4	F	F	T	F	F

This argument is invalid. There is a valuation where the premises are true and the conclusion is false: v_3. It does not matter that there is another valuation where the premises and the conclusion are all true, namely v_1.

Compare it to:

	A	B	C	premise 1 $A \to B$	premise 2 $B \to C$	premise 3 $A \vee B$	conclus C
v_1	T	T	T	**T**	**T**	**T**	**T**
v_2	T	T	F	T	F	T	F
v_3	T	F	T	F	T	T	T
v_4	T	F	F	F	T	T	F
v_5	F	T	T	**T**	**T**	**T**	**T**
v_6	F	T	F	T	F	T	F
v_7	F	F	T	T	T	F	T
v_8	F	F	F	T	T	F	F

Again we have two valuations where the premises are all true, v_1 and v_5, but here the conclusion is true on both of them.

Inconsistent premises

Consider:

- Jenny is happy.
- Jenny is not happy.

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

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

We can draw out a truth table:

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

This is a valid argument! There is no valuation where the premises are all true, and thus, no valuation where the premises are true and the conclusion is false.

Recall, validity is only a judgement about the connection between the premises and the conclusion. It does not imply that the conclusion is true, only that it follows from the premises. If the premises have flaws, then a valid argument isn't of much help, except to help us identify that they have flaws. This is a common pattern of

reasoning: identify some proposals, and see where they lead. If what they lead to is implausible, and your argument was valid, there must be a mistake with the premises.

5.5 Beyond Logical Validity

Some valuations might not correspond to conceptual possibilities. Consider this argument:

– Jonas is a puppy.
∴ Therefore: Jonas is dog.

			premise	conclusion
	P	*D*	*P*	*D*
v_1	T	T	T	F
v_2	T	F	T	T
v_3	F	T	F	F
v_4	F	F	F	T

So, in v_2, the premise is true and the conclusion is false. Therefore, the argument is (logically) invalid.

v_2, however, is not a conceptually possible valuation. Given what 'puppy' and 'dog' mean, v_2 is not possible. Sometimes the notion of 'conceptual validity' is introduced, which considers not every valuation but only those which are compatible with the meanings of the terms at stake. In this unit, we will instead say it is invalid but encourage you to include an additional implicit premise: $P \rightarrow D$.

Chapter 6

Use and mention

We have now talked **about** sentences of TFL and we are just about to do so in much more detail. We should pause to explain an important, and very general, point.

6.1 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.

6.2 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:

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

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.

6.3 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 chapter 7, when we presented the recursive definition of a sentence of TFL. We used uppercase script letters to do this, namely:

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

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 X is a sentence, then $\neg X$ 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:

> '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 '¬*X*' is a sentence.

But this is no good: '¬*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 '¬' 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 '¬*X*' should simply be read **directly** in terms of rules for concatenation. So, **officially**, the metalanguage expression '¬*X*' simply abbreviates:

> the result of placing the symbol '¬' in front of the sentence *X*

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

6.4 Quotation conventions for arguments

One of our main purposes for using TFL is to study arguments. 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 '$\therefore$' is simply to indicate which sentences are the premises and which is the conclusion.

Strictly, the symbol '$\therefore$' 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 (Chapter 6)

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

a) ¬, ∧, ∨ and → are the basic logical connectives of the language of TFL. **(∗)**

b) In a well-formed sentence for every opening bracket '(' there must be a closing bracket). (∗)

c) The sentence snow is white is true if and only if snow is white. (∗)

d) Kevin is a lecturer, whereas Kevin names Kevin and the latter is an expression of the language (and not a lecturer). (∗)

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

a) $X \wedge Y$ (∗)

b) $Y \rightarrow Z$ (∗)

c) $\neg(Y \vee Z)$ (∗)

III. Consider the sentence

▹ Sharky is so-called because of his teeth.

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

Chapter 7

Complex sentences of TFL

7.1 Forming complex sentences

We have introduced our new connectives and specified how the truth of a sentence like $\neg A$, $A \wedge B$, $A \vee B$ and $A \rightarrow B$ depends on the truth of A and the truth of B.

These cover many arguments, but some involve sentences with more complex logical form,

with multiple connectives. For example, they might involve the following sentences.

1. It is not the case that both Alice and Beth are happy.
2. Alice or Beth is happy, and not both.
3. Either Alice and Beth are both happy, or Charlie is happy.
4. If Alice is happy, then Beth is not happy.

In order to use our logic to evaluate them, we will have to symbolise them in our logical system, TFL. We thus need to consider what complex sentences of TFL look like and then how to symbolise these sentences.

For example:

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

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

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

$$A \rightarrow \neg B$$

We gave truth-rules for each of our connectives telling us how the truth value of a

sentence depends on the truth-values of its component parts. This can be iterated to apply to more complex sentences. This will allow us to determine the truth value of complex sentences such as '$(A \vee B) \wedge \neg(A \wedge B)$' from knowing the truth value of A and of B. To do this, it needs to be clear which order to apply the truth rules and this is the value of brackets.

Brackets play a similar role in mathematics to tell you which order to apply your operations.

5. $9 + (3 \times 4)$ $(= 9 + 12 = 21)$
6. $(9 + 3) \times 4$ $(= 12 \times 4 = 48)$

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, brackets must be included. 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.

This will specify what counts as a sentence, and given a sentence, will help us determine in which order it was constructed and thus which order to apply our truth rules.

We will describe a procedure for constructing sentences of TFL:

Definition

Sentences of TFL are given by the following rules:

1. The **atomic sentences** of TFL are the capital letters

   $$A, B, C, \ldots W,$$

   Plus these subscripted by a number, $A_1, A_2, C_{37}, F_{102}, \ldots$

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

3. If X and Y are sentences of TFL, then $(X \wedge Y)$ is a sentence of TFL.

4. If X and Y are sentences of TFL, then $(X \vee Y)$ is a sentence of TFL.

5. If X and Y are sentences of TFL, then $(X \rightarrow Y)$ is a sentence of TFL.

6. Nothing else is a sentence of TFL.

Definitions like this are called **inductive definitions**. They 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.

This tells us that

$$\neg A$$

is a sentence: A is a sentence by our base clause, then adding a negation in front of it obtains a new sentence. Further negations could still be added, obtaining

$$\neg\neg A$$

It also tells us that

$$(A \wedge B)$$

is a sentence: A is a sentence by our base clause, and so is B, then putting a '$\wedge$' between them and adding brackets on the outside we get a further sentence.

We adopt a useful convention which allows us to write $A \wedge B$ instead of the more cumbersome $(A \wedge B)$; and $A \rightarrow \neg B$ instead of $(A \rightarrow \neg B)$:

- One is allowed to drop the outermost brackets of a sentence.

Thus we allow ourselves to write $A \wedge B$ instead of the strict 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!

An alternative way of thinking about it: the brackets should always be there, but we are allowed to write them in invisible ink.[1] We will sometimes display this invisible ink in gray.

$$(A \wedge B)$$

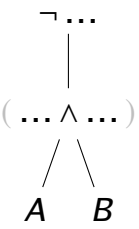
When the sentence is written by itself, it can be dropped, but when it is included in another sentence, those brackets now need to be written.

So when we write

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

those brackets must no longer be invisible.

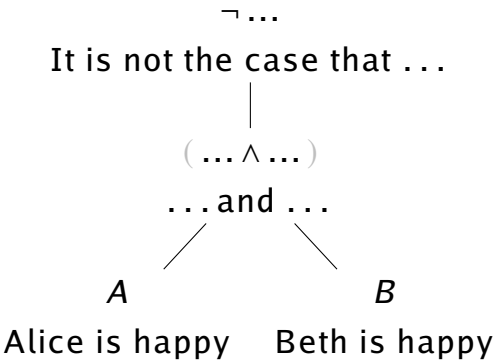
[1] Thanks to Kevin Blackwell for this metaphor.



We might indicate the construction of this sentence visually through the use of braces:

$$\neg \underbrace{(A \wedge B)}$$

This determining of the order of construction is important when knowing how to read it and how to apply the truth rules. We can link to English again:



We will use the sentence $\neg(A \wedge B)$ to be the symbolisation of the English sentence:

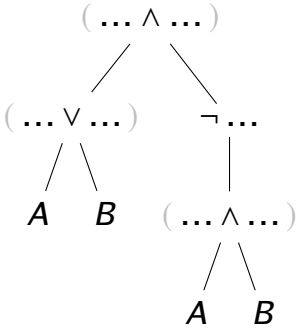
7. It is not the case that both Alice and Beth are happy.

Another example

Now consider

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

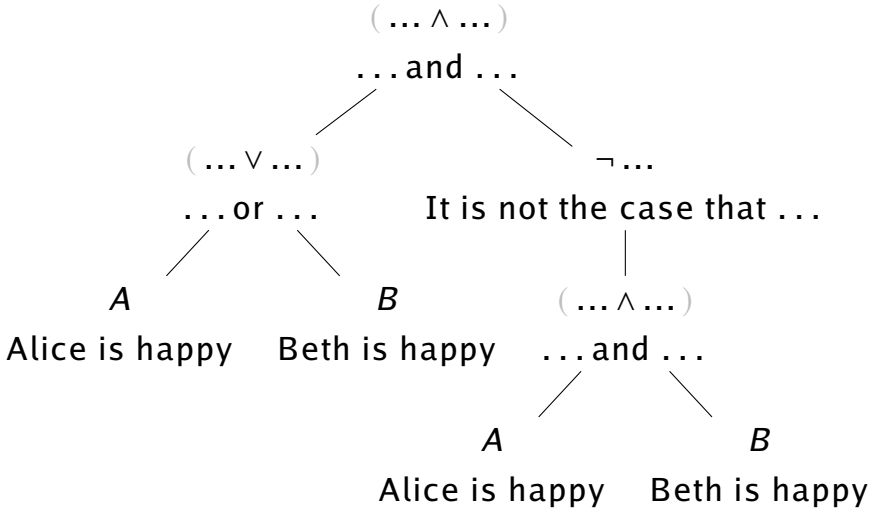
This sentence is constructed in stages:



Displayed with braces as:

$$\underbrace{(A \vee B)} \wedge \underbrace{\neg \underbrace{(A \wedge B)}}$$

It can be read as:



This will be an appropriate symbolisation for:

8. Alice or Beth are happy, and it is not the case that both Alice and Beth are happy.

Or, more concisely put:

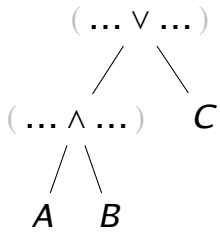
9. Alice or Beth are happy, and not both.

Another example

Consider

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

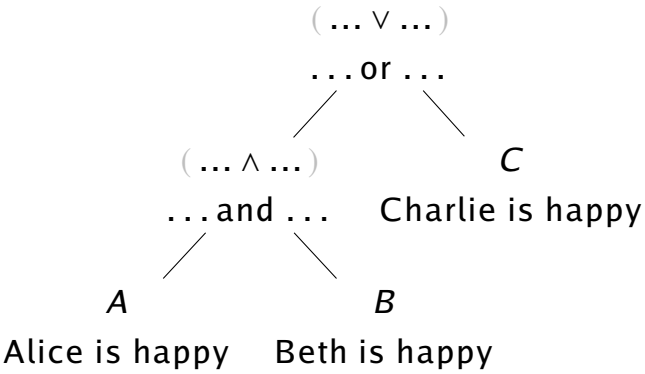
It is formed as:



We can also indicate with braces

$$\underbrace{(A \wedge B)} \vee C$$

We can link it with English:



When giving English sentences we have to give some mechanisms to indicate which order to read it. Commas, or pauses when speaking, or words like 'both' holding components of the sentence together. This is a sentence which might naturally be symbolised as $(A \wedge B) \vee C$:

10. Alice and Beth are both happy, or Charlie is happy

English, however, can be ambiguous. Indeed that is part of the value of offering such symbolisations: no ambiguity is possible. Consider:

11. Alice and Beth or Charlie is happy

Or, more plainly,

12. Alice is happy and Beth is happy or Charlie is happy

Here it is not at all clear whether it should be symbolised as $(A \wedge B) \vee C$ or as $A \wedge (B \vee C)$. Moreover, this isn't a purely stylistic choice: they

mean different things and are true in different circumstances. They are not logically equivalent. The sentence is simply ambiguous. In logic, there is no ambiguity possible. You will not be allowed to write $A \wedge B \vee C$. Instead, you have to either choose $(A \wedge B) \vee C$ or $A \wedge (B \vee C)$.

7.2 Defining Main Connective and Scope

Given any sentence of TFL, there is a unique way to read the sentence. In each sentence there is a singe main connective.

Definition

The **main connective** of a sentence is the last connective that was introduced in the construction of the sentence.

In $(A \wedge B) \vee C$, the main connective is $\vee$. In $A \wedge (B \vee C)$, the main connective is $\wedge$.

Another important notion is that of scope.

To add to our notion of main connective, we introduce the notion of **scope**.

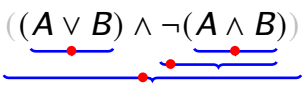
Definition

The **scope** of an occurrence of a connective in a sentence is the sub-sentence for which that connective is the main connective.

For example, consider

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

We can add braces to indicate how to read it. We also indicate the scope of each connective with the red dot on the brace.



We can more explicitly list the scope of each connective:

- The scope of the central $\wedge$ is the whole sentence:
  $(A \vee B) \wedge \neg(A \wedge B)$.
  The main connective of the whole sentence is $\wedge$:

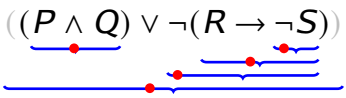
$(A \vee B) \wedge \neg(A \wedge B)$

- The scope of $\vee$ is $(A \vee B)$.
  The main connective of $(A \vee B)$ is $\vee$.
- The scope of $\neg$ is $\neg(A \wedge B)$.
  The main connective of $\neg(A \wedge B)$ is $\neg$.
- The scope of the right-hand $\wedge$ is $(A \wedge B)$.
  The main connective of $(A \wedge B)$ is $\wedge$.

For another example, consider:

$$(P \wedge Q) \vee \neg(R \rightarrow \neg S)$$

We can visually indicate scopes and main connectives as:



We can more explicitly list the scope of each connective:

- The scope of $\vee$ is the whole sentence, $((P \wedge Q) \vee \neg(R \rightarrow \neg S))$.
  The main connective of this whole sentence is $\vee$. We can more explicitly list the scope of each connective:
- The scope of $\wedge$ is $(P \wedge Q)$.
  The main connective of $(P \wedge Q)$ is $\wedge$.
- The scope of $\neg$ is $\neg S$.
  The main connective of $\neg S$ is $\neg$.
- The scope of $\rightarrow$ is $(R \rightarrow \neg S)$.
  The main connective of $(R \rightarrow \neg S)$ is $\rightarrow$.
- The scope of the outer $\neg$ is $\neg(R \rightarrow \neg S)$.
  The main connective of $\neg(R \rightarrow \neg S)$ is the outer $\neg$.

7.3 Negations

No brackets

When we were combining two sentences with one of our connectives, we are told to add brackets. Here were three of our rules for what counts as a sentence of TFL:

3.–5. If X and Y are sentences then so are $(X \wedge Y)$, $(X \vee Y)$ and $(X \rightarrow Y)$.

But the rule for negation did not include brackets:

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

Using these rules we write:

$$A \rightarrow \neg B$$

which is to be read as:

$$(\underbrace{A} \rightarrow \underbrace{\neg B})$$

You might be tempted to write:

$$(A \rightarrow (\neg B)) \qquad \text{⇜ wrong!}$$

That would be wrong. Negations do not introduce brackets.

Identifying scope with negations

An important example is

$$\neg A \vee B$$

The main connective of the sentence is $\vee$. It is to be understood as

$$\underbrace{\neg A}_{} \vee \underbrace{B}_{}$$

We can see this because it can be obtained from dropping the outer brackets from the sentence

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

To instead have $\neg$ as the main connective, one must include brackets:

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

These brackets cannot be dropped because they are inner brackets.

One way you might want to think about it: $\neg$ binds more strongly than the other connectives. You should do it first, unless there's brackets telling you explicitly to first do something else.

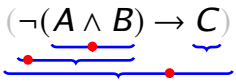
Similarly, in

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

The main connective is the $\rightarrow$.

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

The scope of the $\neg$ is $\neg(A \wedge B)$ rather than the whole sentence.



7.4 More bracketing rules

Binding strength

We give another rule for brackets.

In arithmetic, there is a convention telling you in which order to apply operations. You learn that you have to do a $\times$ calculation before the $+$, unless there are brackets instructing you otherwise. You learn to treat $2 + 3 \times 4$ as $2 + (3 \times 4)$.

We can describe this as that $\times$ binds more strongly than $+$. We apply an exactly analogous rule in logic: $\wedge$ and $\vee$ bind more strongly than $\rightarrow$. You need to calculate their truth value before you calculate that with the $\rightarrow$.

So

$$A \wedge B \rightarrow C$$

is to be understood as shorthand for $(A \wedge B) \rightarrow C$ or, even more carefully: $((A \wedge B) \rightarrow C)$. The brackets are allowed to be dropped, writing simply $A \wedge B \rightarrow C$.

In $A \wedge B \rightarrow C$, the scope of $\wedge$ is narrower than that of $\rightarrow$. In determining the truth value, one first works out the truth value of $A \wedge B$.

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

Similarly

$$A \wedge B \rightarrow C \vee D$$

is shorthand for $(A \wedge B) \rightarrow (C \vee D)$.

$$A \wedge B \rightarrow C \vee D$$

There is no rule that tells us which of $\wedge$ or $\vee$ is stronger. $A \wedge B \vee C$ is dangerously ambiguous and must not be written. You have to include brackets, giving either $(A \wedge B) \vee C$ or $A \wedge (B \vee C)$.

Brackets in strings of conjunctions or disjunctions

There is one final bracketing rule which we will use, although we will not properly talk about it until later: Brackets in strings of $\wedge$ may be dropped. When you read a sentence with a string of $\wedge$ and no brackets, one can read it as an abbreviation for the sentence adding brackets back towards the left.

$P \wedge Q \wedge R$ abbreviates $(P \wedge Q) \wedge R$. The reason that this is a sensible rule, is that it doesn't actually matter which order we compute this. $(P \wedge Q) \wedge R$ is logically equivalent to $P \wedge (Q \wedge R)$. (We will introduce the notion of logical equivalence in section 10.3).

Similarly for $\vee$: $P \vee Q \vee R$ abbreviates $(P \vee Q) \vee R$.

There is however no rule for mixing $\wedge$ and $\vee$. $P \wedge Q \vee R$ is not allowed to be written.

There is also no analogous rule for $\rightarrow$: You must not write $P \rightarrow Q \rightarrow R$.

Summarising our bracketing rules

Definition

- Outer brackets may be dropped.
  $P \wedge Q$ is an abbreviation of $(P \wedge Q)$.
- $\wedge$ and $\vee$ bind more strongly than $\rightarrow$.
  $P \wedge Q \rightarrow R \vee S$ is an abbreviation of $(P \wedge Q) \rightarrow (R \vee S)$
- Brackets in strings of $\wedge$ or $\vee$ may be dropped.
  A sentence with such brackets dropped is an abbreviation of one where they are added to the left.
  $P \wedge Q \wedge R$ is an abbreviation of $(P \wedge Q) \wedge R$.
  $P \vee Q \vee R$ is an abbreviation of $(P \vee Q) \vee R$.

Practice Exercises (Chapter 7)

I. Add brackets to form a sentence of TFL.

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

a) Where the '$\rightarrow$' is the main connective

b) Where the '$\vee$' is the main connective.
c) Where the left '$\neg$' is the main connective.
d) Can it be done so that the right '$\neg$' is the main connective?

II. 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, include braces indicating the scope of each connective.

a) (A) (*)
b) $J_{374} \vee \neg J_{374}$ (*)
c) $\neg\neg\neg\neg F$ (*)
d) $\neg \wedge S$ (*)
e) $(G \wedge \neg G)$ (*)
f) $(A \rightarrow (A \wedge \neg F)) \vee (D \rightarrow E)$ (*)
g) $((A \rightarrow S) \rightarrow W) \wedge (J \vee A)$ (*)
h) $(F \rightarrow \neg D \rightarrow J) \vee (C \wedge D)$ (*)

III.

a) Does $\neg A \vee B$ have the form $X \vee Y$?
b) Does $\neg(A \vee B)$ have the form $X \vee Y$?

c) Does $(A \wedge B) \vee C$ have the form $X \vee Y$?
d) Does $(A \wedge B) \vee C$ have the form $X \wedge Y$?

IV.

a) 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. (*)

b) Are there any sentences of TFL that contain no atomic sentences? Explain your answer. (*)

c) Are there sentences of TFL that don't have a main connective? Explain your answer. (*)

Chapter 8

Truth Tables

In section 4.5 we gave truth rules for the connectives of TFL. Now we will more carefully talk about applying them to complex sentences.

We will make use of the truth rules along with our instructions about reading a sentence and which order it was formed.

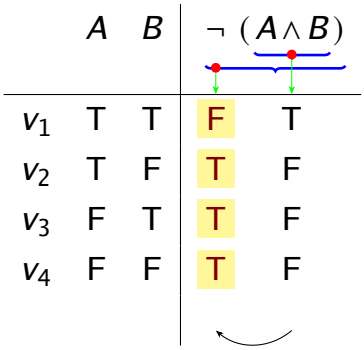
Whether $A \wedge B$ is true depends on whether Alice is happy and whether Beth is happy. There are four possible combinations. There are four valuations. Recall that a valuation is any assignment of true and false to each atomic sentence.

	A	B	$A \wedge B$
v_1	T	T	T
v_2	T	F	F
v_3	F	T	F
v_4	F	F	F

What about $\neg(A \wedge B)$? This will make use of our truth rule for its main connective, $\neg$. The only difference now is that it is applied not to an atomic sentence but something that is more complicated. Recall the truth rule for $\neg$ describing whether $\neg X$ is true or false on a valuation depending on whether X is true or false on that valuation.

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

So suppose we are in a case where neither Alice nor Beth are happy (v_4). Then $(A \wedge B)$ is false. So, according to our rule, $\neg(A \wedge B)$ is true. To keep track of this reasoning in the four possibilities, we write the truth value of the sentence underneath its main connective, in this case, $\neg$. That is, in v_1, $\neg(A \wedge B)$ is false, and in v_2, v_3 and v_4 it is true.



We write the truth value for each sub–sentence underneath its main connective, as indicated by the red dot. We highlight the column giving the truth value for the overall sentence to indicate that it is the final answer, and the other column was just a calculation column.

8.1 Definitions

Recall the definition:

Definition

A **valuation** is any assignment of truth values (true or false) to the atomic sentences.

We now say what a complete truth table is:

Definition

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.

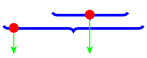
For example, a complete truth table is:

	A	B	$\neg(A \wedge B)$
v_1	T	T	F
v_2	T	F	T
v_3	F	T	T
v_4	F	F	T

Which one can also display with calculation stages, as we did above, where one then needs to highlight the column which gives the final answer:

	A	B	$\neg$	$(A \wedge B)$
v_1	T	T	F	T
v_2	T	F	T	F
v_3	F	T	T	F
v_4	F	F	T	F

We will now turn to giving further worked examples of determining complete truth tables. Before doing so, we recall the truth rules introduced in section 4.5:



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 $X \rightarrow Y$ is:
T	T	$\rightsquigarrow$	T
T	F	$\rightsquigarrow$	F
F	T	$\rightsquigarrow$	T
F	F	$\rightsquigarrow$	T

To determine the complete truth table for an arbitrary sentence of TFL, one just needs to apply these truth rules in the right order, which is

determined by considering the formation order of the sentence.

8.2 Complete Truth Tables

Consider any sentence of TFL, for example

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

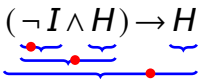
Our task is to draw a complete truth table for this sentence.

The first task is to identify the atomic sentences that appear in it, in this case I and H, and list all the valuations, as described in section 5.3

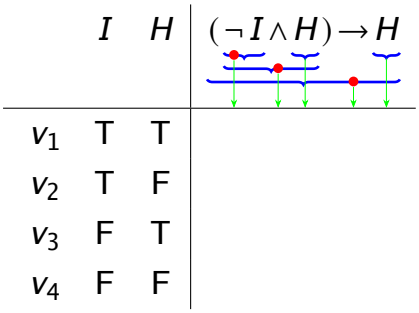
Valuation	I	H	$(\neg I \wedge H) \rightarrow H$
v_1	T	T	
v_2	T	F	
v_3	F	T	
v_4	F	F	

We want to find out the truth values of '$(\neg I \wedge H) \rightarrow H$' on each of the four valuations.

The next step is to identify the formation structure of this sentence, which we do by drawing braces:

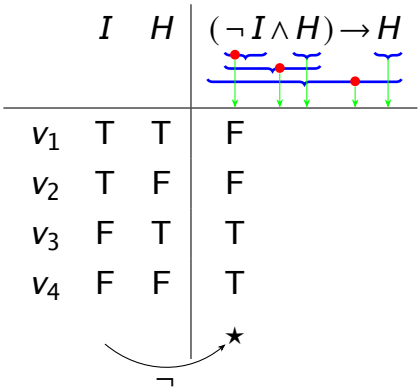


We will apply the different truth rules in turn, and write the answer underneath the main connective of each sub-sentence

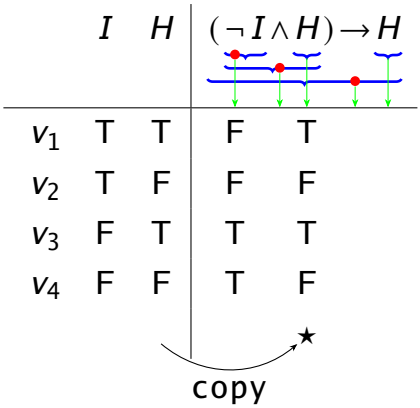


We cannot work out the truth of the overall sentence until we have determined the truth of its components.

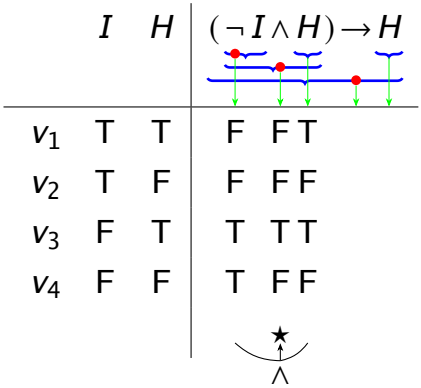
We first calculate $\neg I$, which is to simply apply the rule for $\neg$ from the truth value for I, which is specified directly by the valuation.



Now to determine the truth value for $\neg I \wedge H$, we take our truth value for $\neg I$, which we have calculated, along with the truth value for H, and apply our rule for $\wedge$. For convenience, we first copy the value for H:

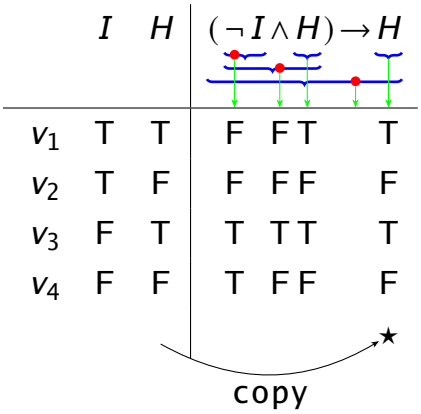


Then we can apply the rule for $\wedge$, applied to our two columns so far:

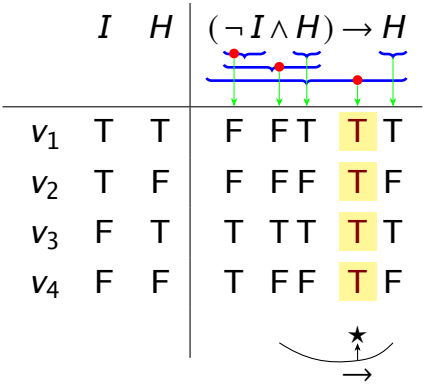


We can then start looking at the truth value for our overall sentence, which will combine the truth value for $(\neg I \wedge H)$ we have just calculated with the

truth value for H, using the rule for $\rightarrow$. To make sure we apply the rule for $\rightarrow$ the right way around, we will copy across the truth value for H:



Finally, we can apply our rule for $\rightarrow$.



We highlight this final column as it indicates our final answer: it displays the truth value of the sentence on each valuation.

Now, many of these columns were calculation columns, and the answer could be given without including them, unless the question on the exam explicitly asks them to be given. That is, we could give the complete truth table as:

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

8.3 Complete truth tables for multiple sentences simultaneously

Often we will want to consider multiple sentences simultaneously.

Consider the argument:

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

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

We thus need to do complete truth tables for the three sentences simultaneously.

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

We will work in stages, applying the appropriate truth rule as we go to determine the truth value of the sentence on each 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	F	F	T	T
v_2 :	T	F	T	T	T	F	T	T	F
v_3 :	F	T	F	T	T	T	F	T	T
v_4 :	F	F	T	F	F	T	T	F	F

Again, many of these are calculation stages, and an answer without calculation would be:

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

Practice Exercises (Chapter 8)

I. Complete truth tables for each of the following:

a) $A \to A$ (∗)

b) $C \to \neg C$ (∗)

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

d) $(A \to B) \vee (B \to A)$ (∗)

e) $(A \wedge B) \rightarrow (B \vee A)$ (∗)

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

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

h) $\big((A \wedge B) \wedge \neg(A \wedge B)\big) \wedge C$ (∗)

i) $((A \wedge B) \wedge C) \rightarrow B$ (∗)

j) $\neg\big((C \vee A) \vee B\big)$ (∗)

k) $(A \rightarrow B) \rightarrow (\neg B \rightarrow \neg A)$ (∗)

l) $(C \rightarrow (D \vee E)) \wedge \neg C$

m) $\neg(G \wedge (B \wedge H)) \rightarrow (G \vee (B \vee H))$

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

a) $(D \wedge \neg D) \rightarrow G$ (∗)

b) $(\neg P \vee \neg M) \rightarrow M$

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

d) $((D \wedge R) \rightarrow I) \rightarrow \neg(D \vee R)$ (∗)

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

Definition

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:

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

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

The complete truth table for this argument is:

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

The only valuation on which both '$\neg L \rightarrow (J \vee L)$' and '$\neg J$' are true is v_3, and on v_3, '$\neg L \rightarrow J$' 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	⋯	Conclusion
v_1				
★ v_2	T	T	T	F
v_3				
v_4				

Invalid

If there is no such row, then it is **valid**.

We will also say:

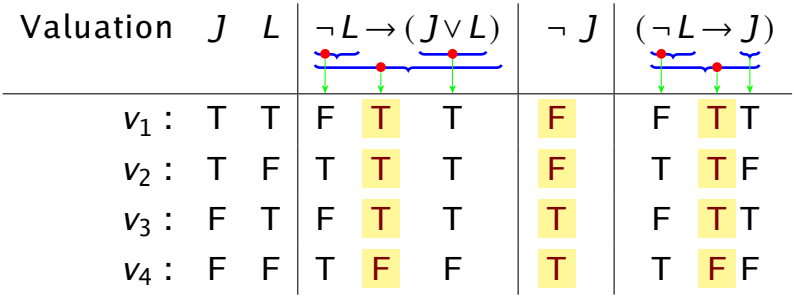
Definition

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.

When doing this in practice, your truth table might include the calculation columns, but these are ignored for the purposes of determining validity.



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

Example

$D \rightarrow V,\ \neg D \ \therefore\ \neg V$

			Premise 1			Premise 2		Conclusion	
Valuation	D	V	$D \rightarrow$		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: $D \rightarrow V$, $D \therefore V$

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

- If I have a pet, then I have a cat or a dog.
- I do not have a cat.

$\therefore$ Therefore: If I have a pet, then I have a dog.

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

	P	C	D	premise 1 $P \to (C \vee D)$			prem 2 $\neg$	C	conclusion $P \to$	D
v_1	T	T	T	T	**T**	T	F		T	T	T
v_2	T	T	F	T	**T**	T	F		T	F	F
v_3	T	F	T	T	**T**	T	**T**		T	**T**	T
v_4	T	F	F	T	F	F	**T**		T	F	F
v_5	F	T	T	F	**T**	T	F		F	T	T
v_6	F	T	F	F	**T**	T	F		F	T	F
v_7	F	F	T	F	**T**	T	**T**		F	**T**	T
v_8	F	F	F	F	**T**	F	**T**		F	**T**	F

There is no valuation where both premises are true and the conclusion is false. So the argument is valid.

Example

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

	A	B	C	premise 1 $A \vee B$	premise 2 $\neg (B \wedge \neg C)$			conclusion $\neg C \vee A$		
v_1	T	T	T	**T**	**T**	T	F F	F	**T**	T
v_2	T	T	F	T	F	T	T T	T	T	T
v_3	T	F	T	**T**	**T**	F	F F	F	**T**	T
v_4	T	F	F	**T**	**T**	F	F T	T	**T**	T
v_5	F	T	T	**T**	**T**	T	F F	F	**F**	F
v_6	F	T	F	T	F	T	T T	T	T	F
v_7	F	F	T	F	T	F	F F	F	F	F
v_8	F	F	F	F	T	F	F T	T	T	F

This argument is invalid. In v_5, both premises are true and the conclusion is false.

Inconsistent premises

- Jenny is happy.
- 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:

Definition

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

Definition

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.

Definition

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

Definition

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

I

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:

Definition

X and *Y* are **logically equivalent** iff they are true on exactly the same valuations.

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	F	F	F
T	F	F	T	F	F	T
F	T	F	T	T	F	F
F	F	T	F	T	T	T

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 (Chapter 10)

I. Revisit your answers to section 8.3. Determine which sentences were tautologies, which were contradictions, and which were neither tautologies nor contradictions.

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

a) $A \rightarrow A$, $\neg A \rightarrow \neg A$, $A \wedge A$, $A \vee A$

b) $A \vee B$, $A \rightarrow C$, $B \rightarrow C$
c) $B \wedge (C \vee A)$, $A \rightarrow B$, $\neg(B \vee C)$
d) $A \leftrightarrow (B \vee C)$, $C \rightarrow \neg A$, $A \rightarrow \neg B$

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

a) $\neg B \wedge B$
b) $\neg D \vee D$
c) $(A \wedge B) \vee (B \wedge A)$
d) $\neg(A \rightarrow (B \rightarrow A))$
e) $A \leftrightarrow (A \rightarrow (B \wedge \neg B))$
f) $((A \wedge B) \leftrightarrow B) \rightarrow (A \rightarrow B)$

IV. 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.”

a) A and $\neg A$

b) $A \wedge \neg A$ and $\neg B \leftrightarrow B$
c) $((A \vee B) \vee C)$ and $(A \vee (B \vee C))$
d) $A \vee (B \wedge C)$ and $(A \vee B) \wedge (A \vee C)$
e) $(A \wedge (A \vee B)) \rightarrow B$ and $A \rightarrow B$

V. 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."

a) $A \rightarrow A$ and $A \leftrightarrow A$
b) $\neg(A \rightarrow B)$ and $\neg A \rightarrow \neg B$
c) $A \vee B$ and $\neg A \rightarrow B$
d) $(A \rightarrow B) \rightarrow C$ and $A \rightarrow (B \rightarrow C)$
e) $A \leftrightarrow (B \leftrightarrow C)$ and $A \wedge (B \wedge C)$

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

a) $A \wedge \neg B$, $\neg(A \rightarrow B)$, $B \rightarrow A$
b) $A \vee B$, $A \rightarrow \neg A$, $B \rightarrow \neg B$
c) $\neg(\neg A \vee B)$, $A \rightarrow \neg C$, $A \rightarrow (B \rightarrow C)$

d) $A \to B$, $A \wedge \neg B$

e) $A \to (B \to C)$, $(A \to B) \to C$, $A \to C$

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

a) $\neg B$, $A \to B$, A

b) $\neg(A \vee B)$, $A \leftrightarrow B$, $B \to A$

c) $A \vee B$, $\neg B$, $\neg B \to \neg A$

d) $A \leftrightarrow B$, $\neg B \vee \neg A$, $A \to B$

e) $(A \vee B) \vee C$, $\neg A \vee \neg B$, $\neg C \vee \neg B$

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

a) Suppose X is true and Y is true
Is $X \vee Y$ true. $(*)$

b) Suppose X is false
Is $X \vee Y$ true? $(*)$

c) Suppose X is false
Is $X \to Y$ true. $(*)$

d) Suppose that X, Y and Z are jointly inconsistent.
What can you say about $X \wedge Y \wedge Y$? (∗)

e) Suppose X is a tautology and $X \therefore Y$ is a valid argument.
What can you say about Y?

f) Suppose that $(X \wedge Y) \rightarrow Z$ is not a tautology.
What can you say about whether $X, Y \therefore Z$ is valid? (∗)

g) Suppose that X is a contradiction.
What can you say about whether $X, Y \vDash Z$? (∗)

h) Suppose that X and Y are logically equivalent.
What can you say about $(X \vee Y)$? (∗)

i) Suppose that X and Y are **not** logically equivalent.
What can you say about $X \vee Y$? (∗)

IX. Are the following statements true? Why?

▹ if Y is a logical consequence of X, then $X \rightarrow Y$ is a logical truth. (∗)

- ▹ if $X \rightarrow Y$ is a logical truth, then Y is a logical consequence of X. (∗)
- ▹ if $X \rightarrow Y \wedge \neg Y$ is a logical truth, then X is a logical contradiction. (∗)
- ▹ if $X \rightarrow Y \wedge \neg Y$ is true relative to a valuation, then X is a logical contradiction. (∗)
- ▹ if $X \vee \neg X \rightarrow Y$ is a logical truth, then Y is a tautology. (∗)
- ▹ if $X \vee \neg X \rightarrow Y$ is true relative to a valuation, then Y is a tautology. (∗)

X. 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 ×
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	$((P \to Q) \to P) \to P$
T	T	× T
T	F	× T
F	T	T F T
F	F	T F T

So '$((P \to Q) \to P) \to P$' 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:

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

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	$\neg L \rightarrow (J \vee L)$	$\neg L$	J
T	T	×	×	T
T	F	×	×	T
F	T	?	?	F
F	F	?	?	F

where the '×' marks indicate entries that we are not going to bother filling in (since the line is already known not to be bad), and the question-marks indicate that we still need to keep investigating.

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

J	L	$\neg L \rightarrow (J \vee L)$	$\neg L$	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	$\neg L \rightarrow (J \vee L)$	$\neg L$	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	$A \vee B$	$\neg\ (A \wedge C)$		$\neg\ (B \wedge \neg D)$		$(\neg C \vee D)$	
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	TT	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 (Chapter 11)

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

a) $\neg B \wedge B$ $(*)$

b) $\neg D \vee D$ $(*)$

c) $(A \wedge B) \vee (B \wedge A)$ $(*)$

d) $\neg(A \rightarrow (B \rightarrow A))$ $(*)$

e) $A \leftrightarrow (A \rightarrow (B \wedge \neg B))$ $(*)$

f) $\neg(A \wedge B) \leftrightarrow A$ $(*)$

g) $A \rightarrow (B \vee C)$ $(*)$

h) $(A \wedge \neg A) \rightarrow (B \vee C)$ $(*)$

i) $(B \wedge D) \leftrightarrow (A \leftrightarrow (A \vee C))$ $(*)$

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-l
contradiction?	complete truth table	one-l
equivalent?	complete truth table	one-l
consistent?	one-line partial truth table	compl
valid?	complete truth table	one-l
logical consequence?	complete truth table	one-l

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

- It is raining outside.

– If it is raining outside, then Jenny is miserable.
∴ Therefore: Jenny is miserable.

We determined that the argument was of the form:

– *A*
– If *A* then *B*
∴ Therefore: *B*

We will 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

- It is raining outside.
- It is not raining outside or Jenny is miserable.

∴ Therefore: Jenny is miserable.

The form of the argument seems to be the following:

- A
- 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

Symbolization

14.1 Atomic sentences

Symbolisations will be relative to a symbolisation key. For example:

> *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 in chapter 4 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 . . . $_X$, then . . . $_Y$'

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

Our task when given a sentence: can you find something to fill out the ... to understand it in one of these ways. By symbolising it with one of our connectives, we are specifying that the truth value of the overall sentence should be governed by the truth rule for that connective. So the truth value of the overall sentence should depend on the truth values of the sentences you pick to fill out ... in the way governed by that connective. If it does not, then you need to think about a different paraphrasing or simply use an atomic sentence.

By symbolising it with one of these connectives, the truth conditions for the sentence is determined in accordance with the truth rule. Symbolising it this way is appropriate when these are appropriate truth conditions; but these standard English paraphrasing is good guidance for when this is the case.

Typically one needs to apply this paraphrasing trick a number of times to result in the symbolisation for the sentence. Let's go

through an example to see how to apply the strategy. Consider:

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

Our first task: identify its main logical form. For this sentence we can find something to fill out ... where the sentence is paraphrased as

> It is not the case that . . .

In particular:

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

'You won't get both soup and salad' is true if and only if 'you will get both soup and salad' is false. It is governed by the truth rule for ¬ depending on the truth of 'you will get both soup and salad'.

We will keep track of the stages of our symbolisation in a tree form, which bears a close relationship to the formation tree we've seen already.

You won't get both soup and salad.
It is not the case that . . .
¬ . . .
|
You will get both soup and salad
?

This should be read as: we take our initial sentence 'You won't get both soup and salad.' and we paraphrase it as 'it is not the case that . . . ', with what fills out the ... being placed underneath. We then know that we'll give a symbolisation of the overall sentence as ¬ ..., where however we symbolise 'You will get both soup and salad' fills out the ....

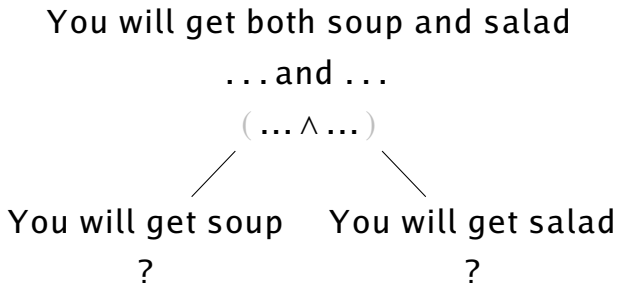
So, how to symbolise 'You will get both soup and salad'? We can paraphrase it as the form

. . . and ...

In particular:

you will get soup and you will get salad

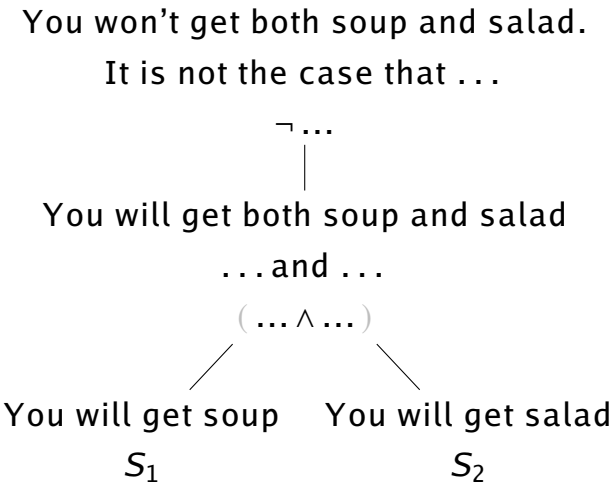
With the sentence ‘You will get both soup and salad’ being true if and only if both ‘you will get soup’ and ‘you will get salad’ are true. It is appropriately governed by $\wedge$.



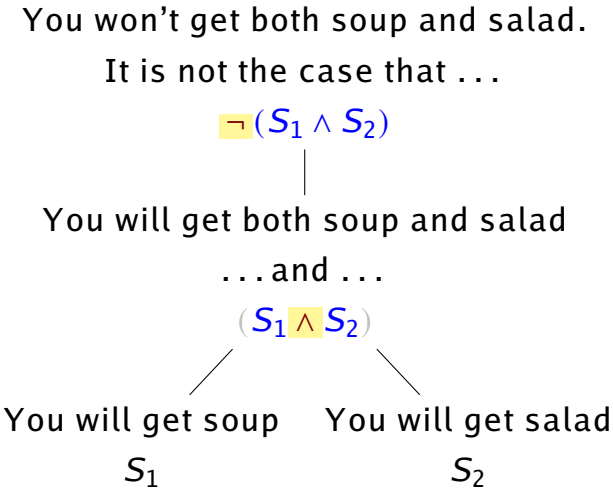
Both ‘You will get soup’ and ‘You will get salad’ cannot be paraphrased in any of our standard forms so we have to use an atomic TFL–sentence. As we haven’t already introduced any atomic TFL–sentences, we need to introduce new ones. We can chose from *A*–*W* or with numerical subscripts. Let’s choose.

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

We can then build our symbolisation tree:



Now to obtain the symbolisation of our overall sentence we just need to work back up the tree, replacing . . . with the determined symbolisation.



Here's a description of this strategy:

1. Consider whether the sentence can be paraphrased as sentences in one of our standard forms, and governed by the associated truth rule.

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

   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.

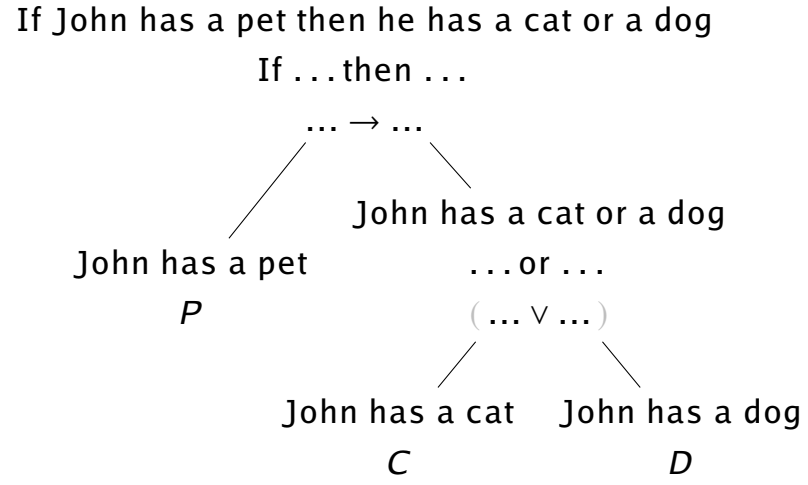
3. Check the symbolization key to see if we already have already chosen an atomic TFL-sentence to symbolise that English sentence.

   - If so, replace the English sentence by that atomic TFL-sentence.
   - Otherwise, choose a new atomic

Consider:

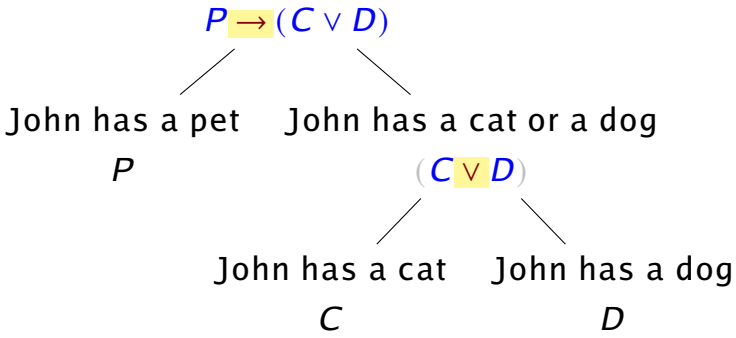
2. If John has a pet, then he has a cat or a dog.

We apply this strategy to break the sentence down:



Which, when we put back together as:

If John has a pet then he has a cat or a dog



So we obtain our symbolisation:

$$P \rightarrow (C \vee D)$$

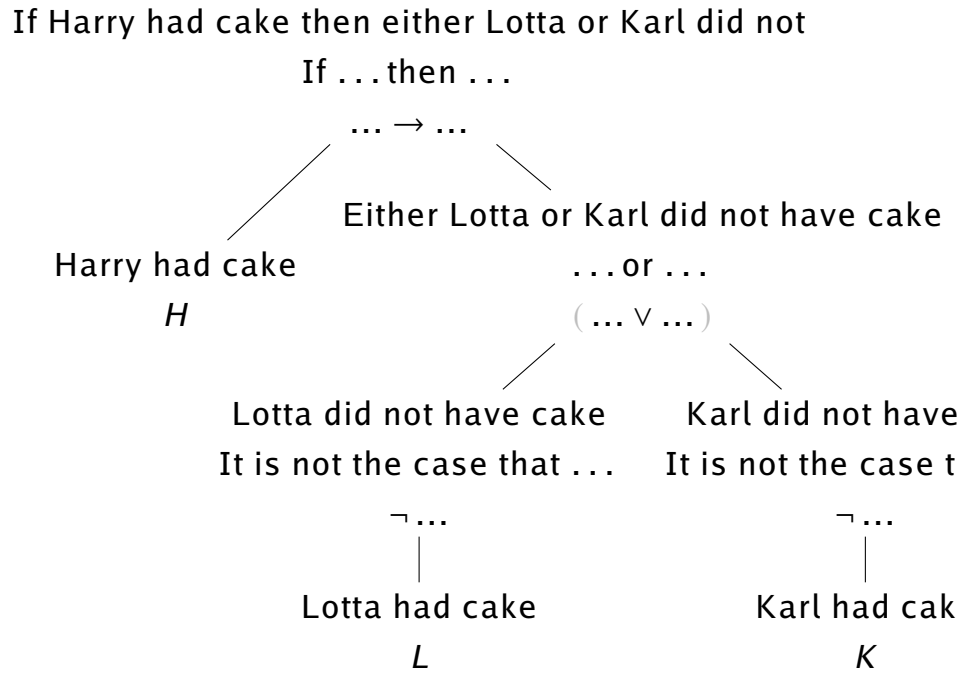
with a symbolisation key:

P: John has a pet
C: John has a cat
D: John has a dog

Consider

3. If Harry had cake then either Lotta or Karl did not.

We can break the sentence down as follows:



obtaining a symbolisation

$$H \rightarrow (\neg L \vee \neg K)$$

with a symbolisation key:

H: Harry had cake
L: Lotta had cake
K: Karl had cake

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):

- Rishi Sunak will make Great Britain great again or he will loose the election.
- 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!).

14.4 Does the symbolisation give the right truth–conditions?

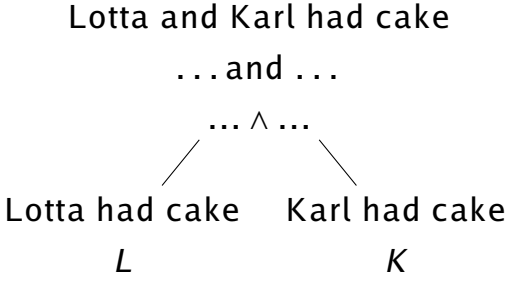
We have already had to apply some analysis to determine how to split a sentence up.

Consider:

4. Lotta and Karl had cake

We will expand this out to:

Lotta had cake and Karl had cake



We judge that it is correct that 'Lotta and Karl had cake' is determined in accordance the rule for $\wedge$:

Lotta had cake L	Karl had cake K	Lotta and Karl had cake $L \wedge K$
T	T	T
T	F	F
F	T	F
F	F	F

This was appropriate. But is it always appropriate?

Consider

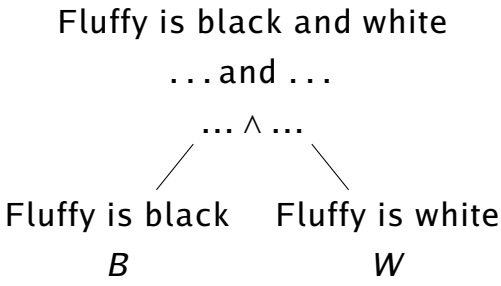
5. Fluffy is black and white.

This has ‘and’ in it, but it is not appropriately paraphrased as:

6. Fluffy is black and Fluffy is white

We might try to symbolise it as

$$B \wedge W$$



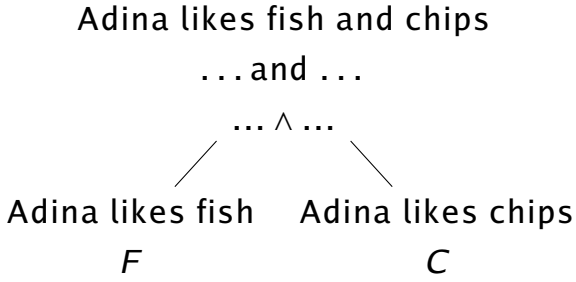
This clearly does not capture the intended truth conditions of ‘Fluffy is black and white’.

Fluffy is black B	Fluffy is white W	Fluffy is black and white $B \wedge W$
T	T	T
T	F	F
F	T	F
F	F	F

Similarly, consider:

7. Adina likes fish and chips.

One might try to break it down as:



which is governed by:

Adina likes fish	Adina likes chips	Adina likes fish and ch
F	C	$F \wedge C$
T	T	T
T	F	F
F	T	F
F	F	F

but the problem is that probably the author of this sentence meant that she likes eating fish–and–chips, rather than giving two separate facts, that she likes fish and that she likes chips.

Sometimes it is a matter of judgement, though. What about:

8. Priya likes nuts and cheese.

What does the author of the sentence mean?

9. Priya likes nuts and cheese (together)
10. Priya likes nuts and Priya likes cheese

Only the latter is to be sy If we symbolise it as $N \wedge C$ it will be given by the truth conditions:

Priya likes nuts N	Priya likes cheese C	Priya likes nuts and $N \wedge C$
T	T	T
T	F	F
F	T	F
F	F	F

When symbolizing sentences, make sure the paraphrasing matches the truth conditions

14.5 How much paraphrasing is appropriate?

Consider the following sentence:

11. The information is irretrievable.

Even though the sentence does not explicitly include 'not', it does have 'ir–' which indicates a negation. We will paraphrase it as

12. It is not the case that
    the information is retrievable

The information is irretrievable
It is not the case that . . .
|
The information is retrievable
R

So we symbolise it as '¬*R*' with

R: The information is retrievable.

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. In these cases, we should expect to paraphrase it as 'it is not the case that . . .' and use ¬ in the symbolisation.

Consider:

13. That action is good.
14. That action is bad.

Let's symbolise sentence 13 with an atomic sentence, giving a symbolisation key with:

G: That action is good.

Should we paraphrase sentence 14 as

15. It is not the case that that action is good

symbolising it as ¬*G*.

There are two reasons we will not do this.

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 steals some food. In other words, there may be things that are neither good nor bad. However, in TFL if '*G*' is false, then '¬*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 14 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 sentence 13 and sentence 14

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:

- Stealing is good or donating for charity is good.
- Stealing 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 sentence 14 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, 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 sentence 14 is the negation of sentence 13 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!

Chapter 15

On Truth-functional connectives

Often English sentences come with different ways of joining sentences than our standard ‘and’, ‘or’, ‘not’ or ‘if’.

Consider these sentences:

1. Neither Alice nor Beth will come to the party.
2. Alice won’t come to the party unless Beth comes.

3. Alice will come to the party only if Beth will too.

These are all sentences whose truth values depend on whether Alice and Beth come to the party. The highlighted connectives are components of English which join together two simpler sentences so that the truth value of the overall sentence just depends on the truth values of the simpler sentences which are combined.

15.1 Truth-functionality

Let's introduce an important idea.

Definition

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**.

15.2 Non truth-functional connectives

In plenty of languages, for example 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:

1. 2 + 2 = 4

2. 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 → was pretty bad at capturing some uses of 'if. . . , then. . .' in English. It is especially bad at capturing use of **subjunctive conditionals** of English. The problem is that a subjunctive conditional is not truth functional. Consider the two sentences:

3. If Mitt Romney had won the 2012 election, then he would have been the 45th President of the USA.
4. 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 3 is true; sentence 4 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 in TFL for capturing conditionals. TFL just doesn't have the required resources as the subjunctive conditional is not truth-functional.

15.3 Truth-functional completeness

There is an important result about TFL, or truth functional logic: not only are its connectives truth functional, in fact, it can capture all truth-functional connectives, possibly by use of a more complex sentence of TFL.

Consider the following:

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

One can come up with a sentence of TFL which has this truth table, we already gave it as the exclusive disjunction:

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

This is a general fact about TFL and the reason it is called "Truth-functional logic", it is the logic of truth functions, the logic of all truth-functional connectives. Whatever truth table you give me, I can find a sentence of TFL which has that truth table. It can encode any possible truth functional connective.

Practice Exercises (Chapter 15)

I.

a) Is 'I know that' truth functional? (∗)

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

a) (∗)

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

b) (∗)

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

c) (∗)

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

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

a) Can you explain why? (∗)
b) Can you show for each of these combinations that there is a sentence of TFL with that column describing its truth.
c) Can you show there's always a formula just using ¬ and ∧ with that column describing its truth.

Chapter 16

Expanding our Symbolisation Guide

We have our standard symbolisation guide:

- 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 . . . $_X$, then . . . $_Y$'
  it can be symbolised as $X \rightarrow Y$.

But often sentences of English aren't neatly given in one of these forms. 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. Beth won't come to the party unless Alice comes.
6. Neither Alice nor Beth will come to the party.

7. Alice will come to the party only if Beth will too.
8. Alice will come to the party if Beth will too.
9. 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 chapter 15), possibly by a more complex sentence.

We will expand our symbolisation guide to give instructions about how to treat a variety of English sentences whose logical form isn't presented in a nicely canonical way. The completed symbolisation guide can be found in section 3.4.

16.1 Ways of expressing conjunction

The symbolization guideline for conjunction we introduced Section section 14.2 was:.

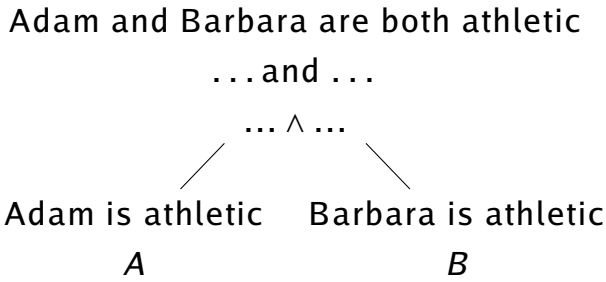
If a sentence can be paraphrased as '. . . and . . . '
it can be symbolised as $(X \wedge Y)$.

1. Adam and Barbara are both athletic.

We will appropriately paraphrase it as

2. Adam is athletic and Barbara is athletic.

In our symbolisation tree format:



Giving a symbolisation

$$A \wedge B$$

with a symbolisation key:

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

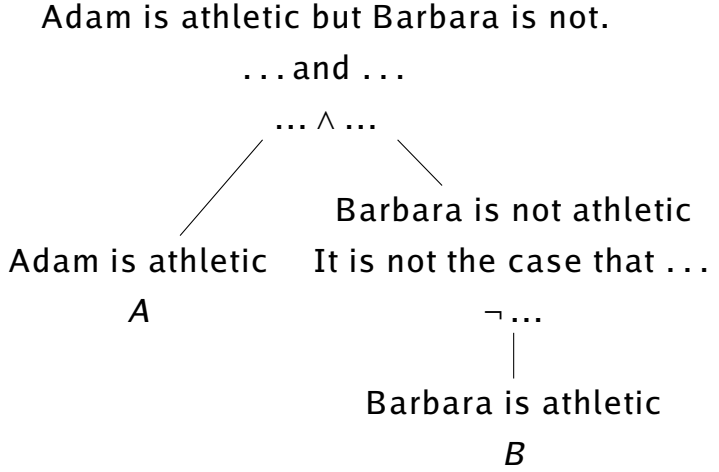
This appropriately captures the truth conditions of the original sentence. It is true just if Adam is athletic and Barbara is athletic, i.e., if both *A* and *B* are true.

Now consider:

3. Adam is athletic but Barbara is not.

The word ‘but’ sets up a contrast between the first part of the sentence and the second part. Nevertheless, it still tells us two things: that Adam is athletic, and that Barbara is not athletic.

We will treat it as just another way of expressing a conjunction ‘and’:



So we give

$$A \wedge \neg B$$

There is a difference in tone between these two sentences, but no difference in truth conditions.

4. Adam is athletic but Barbara is not.
5. Adam is athletic and Barbara is not athletic.

They are both true just when Adam is athletic and Barbara is not athletic:

A	B	$A \wedge \neg B$
T	T	T F F
T	F	T T T
F	T	F F F
F	F	F F T

We have lost some nuance in this symbolisation, but TFL does not (and cannot) preserve that nuance. The word ‘but’ often carries an additional suggestion of contrast or unexpectedness between the two clauses. This is not part of the literal truth-conditions of the sentence, but arises from what is called a conversational implicature: something suggested or conveyed by the way a sentence is used in conversation, rather than by its strict logical content. If there was no surprise in the difference between the facts, then we would still say that the sentence combined with ‘but’ is infelicitous or misleading but is nonetheless true.

A similar point can be made about ‘although’:

6. Although Adam is athletic, Barbara is not.
7. Barbara is not athletic, although Adam is.

This raises the same sort of issue. The word 'although' indicates a contrast, but that is not something TFL can represent. As far as truth conditions go, the sentence again is true just if Adam is athletic and Barbara is not athletic. We will symbolise these sentences as $A \wedge \neg B$.

We can add these to our toolbox for symbolisation:

If a sentence can be paraphrased as

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

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

16.2 Exclusive Disjunction

In section 14.2, we proposed that:

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

Let's consider some examples again

8. Fatima will play videogames, or she will watch movies.
9. 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 section 4.3 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 9 to be true, if either Fatima or Omar play videogames, or both play videogames. In contrast, on the exclusive reading we take sentence 9 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 9 the inclusive reading is available and using the above symbolization key it can be symbolized as $F \vee O$. For Sentence 8 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 8 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 8 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 '$\neg(F \wedge M)$'.

Sometimes in English we use 'either. . . or. . .' in a way which indicates that it should be read exclusively, consider, for example:

10. Either Fatima will play videogames, or she will watch movies.
11. Either Fatima will play videogames or Omar will play videogames.

However, this is not always the case. Consider

12. If either Fatima or Omar play videogames, then the house will remain a mess.

Here it's not so clear that an exclusive disjunction is being indicated. Often 'either' is introduced in the English to play something like the role of brackets. It helps indicate what is being held together.

We suggest to always capture 'or' or 'either. . . or. . . ' with a simple inclusive reading by symbolising it as '$\vee$', however, bear in mind that when evaluating validity of arguments that the exclusive reading might be needed, or encouraged by adding an implicit premise.

If one wishes to symbolize the exclusive 'or' in TFL that can be done. It can be symbolised as:

- 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 11 should be symbolized by the sentence

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

This has the truth table

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

So it is true just when exactly one of Fatima or Oma play video games. It captures an exclusive disjunction.

> If a sentence can be paraphrased in English as '**either ... or ... (and not both)**', it can be symbolised as $(X \vee Y) \wedge \neg(X \wedge Y)$.

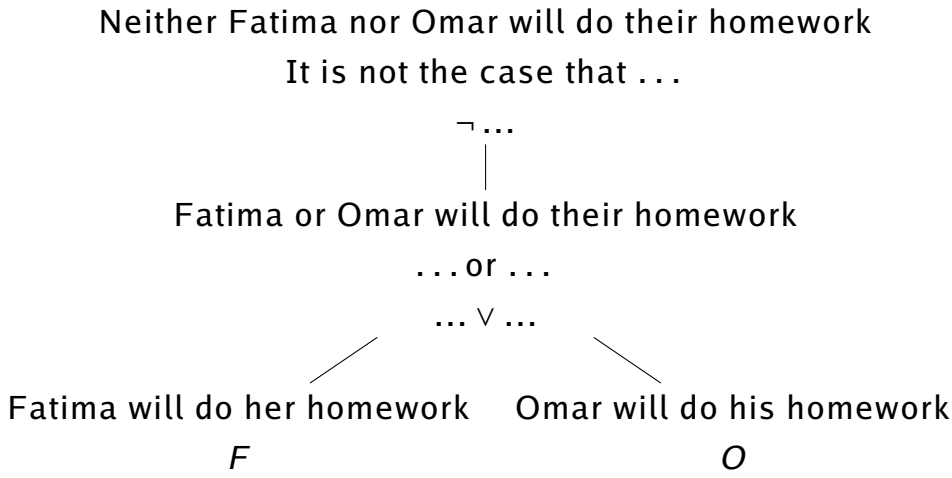
16.3 Neither–nor

Consider these sentences, which are all paraphrases of one another. Their main (in fact, only) English connective is '**neither... nor...**'.

13. Neither Fatima nor Omar will do their homework.

'**neither... nor...**' is also a truth-functional English connective, whose truth depends just on the truth of the components. But it is not directly captured with one of our TFL connectives. Instead we need to use a more complex sentence of TFL.

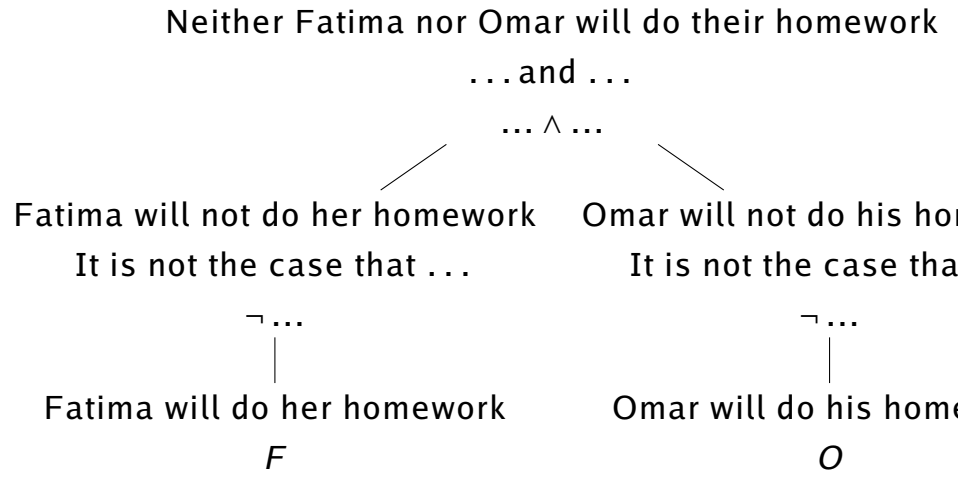
Here is a way to paraphrase it and break it down into our standard connectives:



Giving the symbolisation:

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

There is a completely different but also entirely adequate way to paraphrase it and break it down:



Giving the symbolisation

$$\neg F \wedge \neg O$$

We can check that these are logically equivalent and match the intended reading of 'neither–nor'.

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

Whilst in general we do **not** replace one symbolisation with another simply because they are logically equivalent, for 'neither–nor' this alternative symbolization is allowed.

Symbolisation in TFL aims to capture the **structure** of an English sentence, not merely its truth conditions. Two sentences can be logically equivalent yet represent different logical constructions in English, so normally we do not treat them as interchangeable.

Whilst in general you are not allowed to substitute in any logically equivalent sentence, for neither–nor this alternative way of capturing the logical form is permitted.

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

16.4 Conditionals

Let us again recall the symbolization guideline that we introduced in Section 14.2:

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

Now consider the sentences

14. If the number can be divided by 4, then it is even.
15. If it can be divided by 4, it is even.
16. It is an even number if it 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 15 and 16 are just a rephrasing of sentence 14. So we will again symbolise them as $(P \rightarrow F)$.

Now consider

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

sentence 17 is also a conditional. But, intuitively, while sentence 14–sentence 15 are true sentence 16 is false, as, e.g., 6 is an even number but cannot be divided by 4. This suggest that sentence 16 needs to be symbolized differently than sentence 14–sentence 15. Indeed, 'only if' inverses the order of the conditional: whereas sentence 14–sentence 15 are symbolized as $(P \rightarrow F)$, sentence 16 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 sentence 17 fits with our intuitive judgement.

In fact, sentence 17 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 sentence 14.

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 $(X \to Y)$ such that X assumes the position of $\ldots_X$ and Y that of $\ldots_Y$.

Following this guidance, we symbolise all the following as ‘$A \to B$’:

18. If Alice is on shift then Beth is on shift.
19. Beth is on shift if Alice is.
20. Alice is on shift only if Beth is.

At this point, a word of warning about the connective ‘$\to$’ 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 '→' and '**if..., then...**' as we discussed in section 15.2.

16.5 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:

21. Laika is a dog only if she is a mammal
22. Laika is a dog if she is a mammal
23. 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 21, for reasons discussed above, can be symbolized by '$D \rightarrow M$'. in contrast Sentence 22 can be paraphrased as, 'If Laika is a mammal then Laika is a dog'. So it can be symbolized by '$M \rightarrow D$'.

Sentence 23 says something stronger than both sentence 21 and sentence 22. 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 21 and 22. 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 ‘$(X \to Y) \wedge (Y \to X)$’ 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 $((X \to Y) \wedge (Y \to X))$.

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.

16.6 Unless

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

24. Unless you wear a jacket, you will catch a cold.
25. You will catch a cold unless you wear a jacket.

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

26. If you do not wear a jacket, then you will catch a cold.
27. If you do not catch a cold, then you wore a jacket.
28. 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.

16.7 Summary

Symbolisation Guide

Paraphrased as	Symbolised as
[.5em] It is not the case that ...	$\neg X$
... and ...	$X \wedge Y$
Both ... and ...	$X \wedge Y$
Although ..., ...	$X \wedge Y$
... but ...	$X \wedge Y$
... or ...	$X \vee Y$
Either ... or ... (and not both)	$(X \vee Y) \wedge \neg(X \wedge Y)$
Neither ... nor ...	$\neg(X \vee Y)$ or $\neg X \wedge \neg Y$
If $\ldots_X$, then $\ldots_Y$	$X \rightarrow Y$
$\ldots_Y$ if $\ldots_X$	$X \rightarrow Y$
$\ldots_X$ only if $\ldots_Y$	$X \rightarrow Y$
$\ldots_X$ if and only if $\ldots_Y$	$(X \rightarrow Y) \wedge (Y \rightarrow X)$ **(can be shortened**
Unless $\ldots_X$, $\ldots_Y$	$X \vee Y$ or $\neg X \rightarrow Y$ or $\neg$

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 discussed some examples in chapter 15.

Practice Exercises (Chapter 16)

I. 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.

a) Those creatures are not men in suits. (∗)
b) Those creatures are men in suits, or they are not. (∗)
c) Those creatures are either gorillas or chimpanzees. (∗)

d) Those creatures are neither gorillas nor chimpanzees. (∗)
e) If those creatures are chimpanzees, then they are neither gorillas nor men in suits. (∗)
f) Unless those creatures are men in suits, they are either chimpanzees or they are gorillas. (∗)

II. 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.

a) Either Mister Ace or Mister Edge was murdered. (∗)
b) If Mister Ace was murdered, then the cook did it. (∗)
c) If Mister Edge was murdered, then the cook did not do it. (∗)

d) Either the butler did it, or the Duchess is lying. (∗)
e) The cook did it only if the Duchess is lying. (∗)
f) If the murder weapon was a frying pan, then the culprit must have been the cook. (∗)
g) If the murder weapon was not a frying pan, then the culprit was either the cook or the butler. (∗)
h) Mister Ace was murdered if and only if Mister Edge was not murdered. (∗)
i) The Duchess is lying, unless it was Mister Edge who was murdered. (∗)
j) If Mister Ace was murdered, he was done in with a frying pan. (∗)
k) Since the cook did it, the butler did not. (∗)
l) Of course the Duchess is lying! (∗)

III. 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.

a) Ava and Harrison are both electricians. (∗)
b) If Ava is a firefighter, then she is satisfied with her career. (∗)
c) Ava is a firefighter, unless she is an electrician. (∗)
d) Harrison is an unsatisfied electrician. (∗)
e) Neither Ava nor Harrison is an electrician. (∗)
f) Both Ava and Harrison are electricians, but neither of them find it satisfying. (∗)
g) Harrison is satisfied only if he is a firefighter. (∗)
h) If Ava is not an electrician, then neither is Harrison, but if she is, then he is too. (∗)
i) Ava is satisfied with her career if and only if Harrison is not satisfied with his. (∗)
j) If Harrison is both an electrician and a firefighter, then he must be satisfied with his work. (∗)
k) It cannot be that Harrison is both an electrician and a firefighter. (∗)

l) Harrison and Ava are both firefighters if and only if neither of them is an electrician. (∗)

IV. Give a symbolization key and symbolize the following English sentences in TFL. (∗)

a) Alice and Bob are both spies. (∗)
b) If either Alice or Bob is a spy, then the code has been broken. (∗)
c) If neither Alice nor Bob is a spy, then the code remains unbroken. (∗)
d) The German embassy will be in an uproar, unless someone has broken the code. (∗)
e) Either the code has been broken or it has not, but the German embassy will be in an uproar regardless. (∗)
f) Either Alice or Bob is a spy, but not both. (∗)

V. Give a symbolization key and symbolize the following English sentences in TFL. (∗)

a) If there is food to be found in the pridelands, then Rafiki will talk about squashed bananas. (∗)

b) Rafiki will talk about squashed bananas unless Simba is alive. (∗)
c) Rafiki will either talk about squashed bananas or he won't, but there is food to be found in the pridelands regardless. (∗)
d) Scar will remain as king if and only if there is food to be found in the pridelands. (∗)
e) If Simba is alive, then Scar will not remain as king. (∗)

VI. 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?

a) 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. (∗)
b) 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. (∗)

c) 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. (∗)

VII. 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 15)

a) It is going to rain soon. I know because my leg is hurting, and my leg hurts if it's going to rain.

b) **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.

c) **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 17

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:

– 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:

- One morning I shot an elephant in my pajamas.
- 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:

- $A \vee (G \vee H)$

The interpretation which got him executed, on the other hand, can be symbolized by:

– $(A \vee B) \vee (G \vee H)$

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 Ărefsent:scamb1 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 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 '$(S \wedge D) \vee T$' and that of sentence 4 is '$S \wedge (D \vee T)$'. In the first, $\wedge$is in the scope of $\vee$, and in the second $\vee$is in the scope of $\wedge$.

17.1 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 chapter 16, 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.

Practice Exercises (Chapter 17)

I. The following sentences are ambiguous. Give symbolization keys for each and symbolize the different readings. Show that they are not logically equivalent.

a) Haskell is a birder and enjoys watching cranes.
b) The zoo has lions or tigers and bears.
c) The flower is not red or fragrant.

Part IV

Natural deduction for TFL

Chapter 18

The very idea of natural deduction

Way back in chapter 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 $A \rightarrow (B \wedge C)$ 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, $X \rightarrow Y \therefore Y$ is valid: there are no valuations where X and $X \rightarrow Y$ are true but Y is false. So if our premises $A \rightarrow (B \wedge C)$ and A are true, then $B \wedge C$ must also be true: that's just Modus Ponens. So now from our supposition of $A \rightarrow (B \wedge C)$ 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 $A \rightarrow (B \wedge C)$ and A we can conclude B by walking someone through these two steps. This will be enough to show that $A \rightarrow (B \wedge C)$, $A \therefore B$ 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	$A \rightarrow (B \wedge C)$	
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
4	Beth came to the party

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.

18.1 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.

- Alice, or Betty, or Carys, or Dan, or Ella stole the teacher's pen.
- It wasn't Alice.
- It wasn't Betty,
- It wasn't Carys,
- 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,

- Alice, or Betty, or Carys, . . . , or Uli or Volker stole the teacher's pen.
- It wasn't Alice.
- It wasn't Betty,
- ⋮
- 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 26) 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

[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.

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.

Chapter 19

The First Basic Rules

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 **??**.

19.1 Reiteration

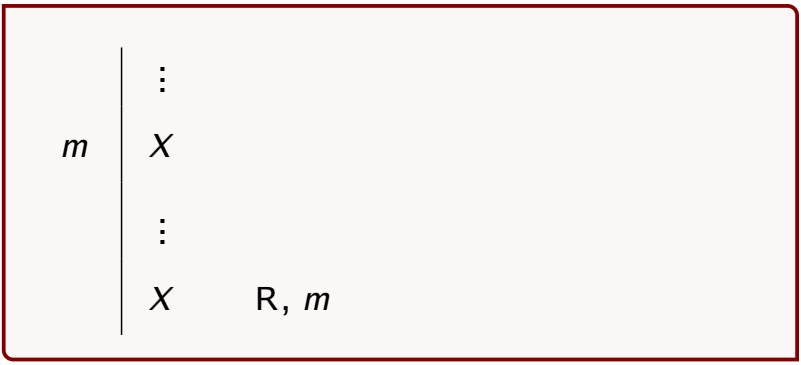
The very first rule is so breathtakingly obvious that it is surprising we bother with it at all.

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:

4	$A \wedge B$	
⋮	⋮	
10	$A \wedge B$	R, 4

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:



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 chapter 6). 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.

19.2 Modus Ponens

Consider the following argument:

> If Jane is smart then she is fast. Jane is smart. $\therefore$ 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 sometimes called Modus Ponens.

In our proof system, we choose a list of rules organised by the connectives. Each connective will have an introduction rule and an elimination rule. So instead of using the name Modus Ponens, we use the name "Conditional Elimination" or "Arrow Elimination", writing it as "→–Elimination", or, more concisely, →E. It is an elimination rule because we start with something including the connective → and we derive something without the connective, that is we have **eliminated** the → connective.

In a simple use of this rule, we might just use it to derive F from the premises $S \to F$ and S:

1	$S \to F$	
2	S	
3	F	→E, 1, 2

This would then be a natural deduction proof that $S \rightarrow F$, $S \therefore F$ 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 "→E 1,2" to say that we obtained line 3 by use of this rule → 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	
	⋮	
8	$S \to F$	some rule
	⋮	
15	S	another rule
	⋮	
23	F	$\to$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	
	⋮	
8	$S \to F$	
	⋮	
23	F	$\to$E, 8, 1

We write our general rule as:

	$\vdots$	
m	$X \to Y$	
	$\vdots$	
n	X	
	$\vdots$	
	Y	$\to$E, m, n

We can apply it to any sentences of TFL, X and Y. For example,

1	$(A \vee B) \to \neg F$	
2	$(A \vee B)$	
3	$\neg F$	$\to$E, 1, 2

In this, X is $(A \vee B)$, and Y is $\neg F$.

The two lines needed to apply this can come in either order.

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.

19.3 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:

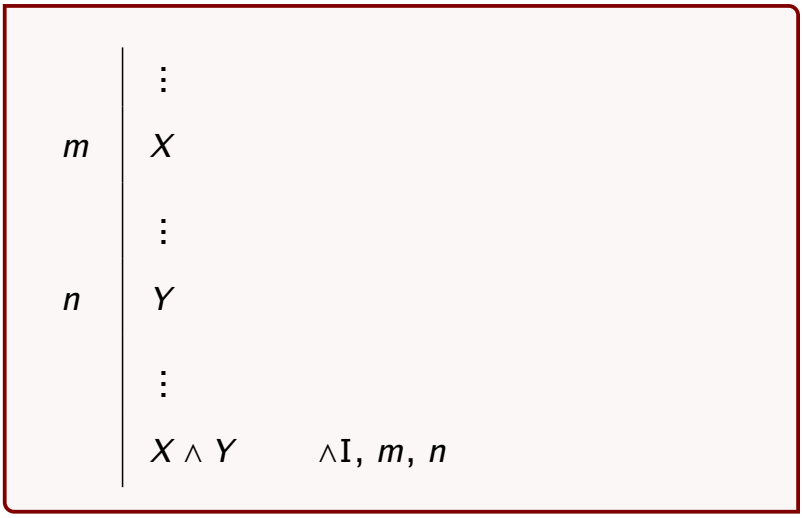
> *A*: Alice came to the party
> *B*: Beth came to the party

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:

8	A	
15	B	
23	$A \wedge B$	$\wedge$I, 8, 15

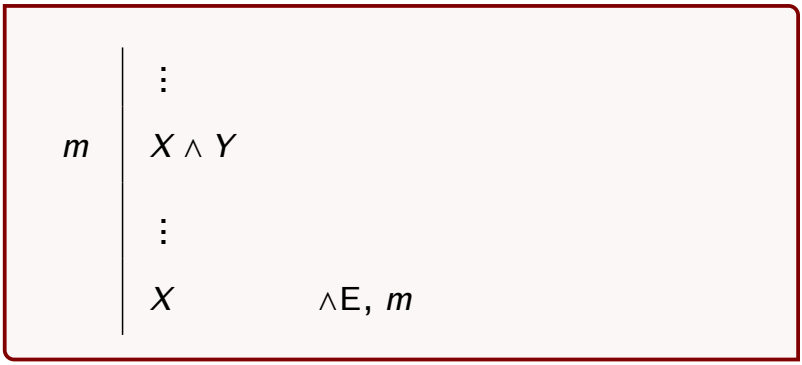
Note that every line of our proof must either be an assumption, or must be justified by some rule. We cite '$\wedge$I 8, 15' here to indicate that the line is obtained by the rule of conjunction introduction ($\wedge$I) applied to lines 8 and 15. More generally, here is our conjunction introduction rule:



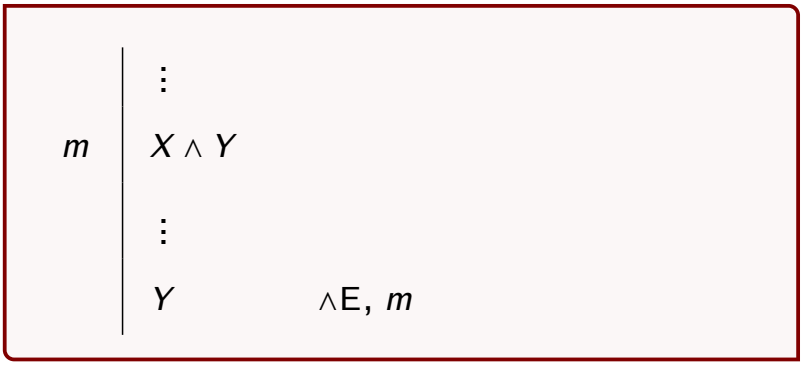
The rule is called 'conjunction **introduction**' because it introduces the symbol '$\wedge$' into our proof where it may have been absent.

19.4 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):



and equally:



The point is simply that, when you have a conjunction on some line of a proof, you can obtain either of the conjuncts by ∧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 '$C \vee (D \wedge E)$'!)

19.5 Constructing proofs

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.

Examples

Consider

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

We know our proof will look something like

1	$A \to (B \wedge C)$	
2	$A \wedge \neg D$	
	?	
n	$C \wedge \neg D$	?

And our task is to fill out the '?'.

Our aim is to get to $C \wedge \neg D$. The main connective of this is $\wedge$. So we look at the introduction rule for $\wedge$. So it would be great if we'd be able to get both to C and to $\neg D$ then we could use the $\wedge$I rule to get to our conclusion.

1	$A \to (B \wedge C)$	
2	$A \wedge \neg D$	
	?	
$n-2$	C	?
$n-1$	$\neg D$	?
n	$C \wedge \neg D$	$\wedge$I, $n-2$, $n-1$

Let's look at our premises and see how they can be used to get to both $\neg D$ and to C.

$\neg D$ is actually easy: we have a premise $A \wedge \neg D$, so the $\wedge$E rule immediately gets us $\neg D$.

1	$A \rightarrow (B \wedge C)$	
2	$A \wedge \neg D$	
	?	
$n-2$	C	?
$n-1$	$\neg D$	$\wedge$E, 2
n	$C \wedge \neg D$	$\wedge$I, $n-2$, $n-1$

But what about C? It's in premise 1, but the elimination rule for the main connective is

m	$X \rightarrow Y$	
n	X	
	Y	$\rightarrow$E, m, n

We can only use premise 1 if we also have A. Can we get A? Yes! By $\wedge$E from premise. That'll allow us to close our gap and complete our proof:

1	$A \rightarrow (B \wedge C)$	
2	$A \wedge \neg D$	
3	A	$\wedge$E, 2
4	$B \wedge C$	$\rightarrow$E, 1, 3
5	C	$\wedge$E, 4
6	$\neg D$	$\wedge$E, 2
7	$C \wedge \neg D$	$\wedge$I, 5, 6

Practice Exercises (Chapter 19)

I. 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	$(*)$
4	R	$(*)$
5	$A \vee E$	$(*)$

II. Give a proof corresponding to each of the following arguments:

1. $A \wedge (B \vee C) \therefore (B \vee C) \wedge A$ $(*)$
2. $A \rightarrow B,\ A \rightarrow C,\ A \wedge D \therefore B \wedge C$

Chapter 20

The remaining Basic Rules for TFL

To introduce a conditional we need to introduce another kind of thing: a sub-proof.

20.1 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	∧I, 1, 2
4	Thus, **if** Beth came to the party, both Alice and Beth came.	→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	$\quad B$	

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:

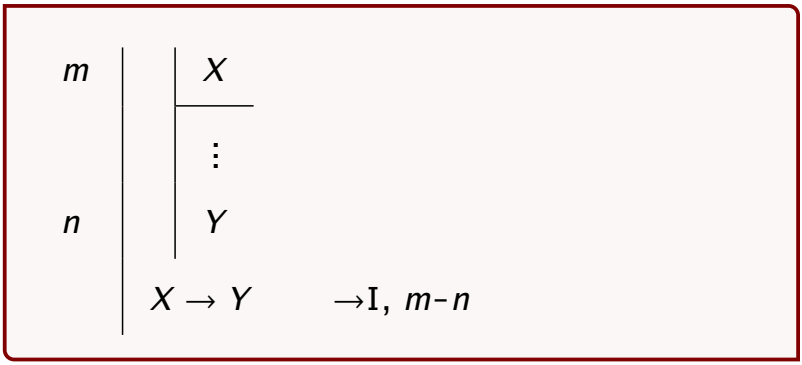
1	A	
2	$\quad B$	
3	$\quad A \wedge B$	$\wedge$I, 1, 2

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 '$B \rightarrow (A \wedge B)$':

1	A	
2	$\quad B$	
3	$\quad A \wedge B$	$\wedge$I, 1, 2
4	$B \rightarrow (A \wedge B)$	$\rightarrow$I, 2–3

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:



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:

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

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:

1	$P \to Q$	
2	$Q \to R$	
3	$\quad P$	
4	$\quad Q$	→E, 1, 3
5	$\quad R$	→E, 2, 4
6	$P \to R$	→I, 3-5

The subproof also doesn't need to start immediately. For example:

1	$(P \to Q) \wedge O$	
2	$Q \to R$	
3	$P \to Q$	$\wedge$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

20.2 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.

Consider this proof:

1	A	
2	$\quad B$	
3	$\quad B$	R, 2
4	$B \rightarrow B$	$\rightarrow$I, 2–3

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:

1	A	
2	$\quad B$	
3	$\quad B$	R, 2
4	$B \rightarrow B$	$\rightarrow$I, 2–3
5	B	naughty attempt to invoke $\rightarrow$E, 4, 3

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 →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 →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,

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,

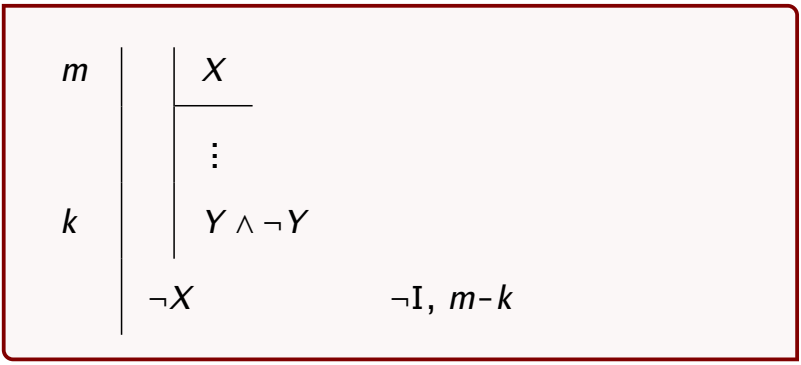
you must assume the antecedent of the conditional in a subproof.

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.

20.3 Negation

Negation Introduction

If assuming something leads you to a contradiction, then the assumption must be wrong. This thought motivates the following rule:



It does not matter whether in the subproof you first derive Y and then derive $\neg Y$ or whether you

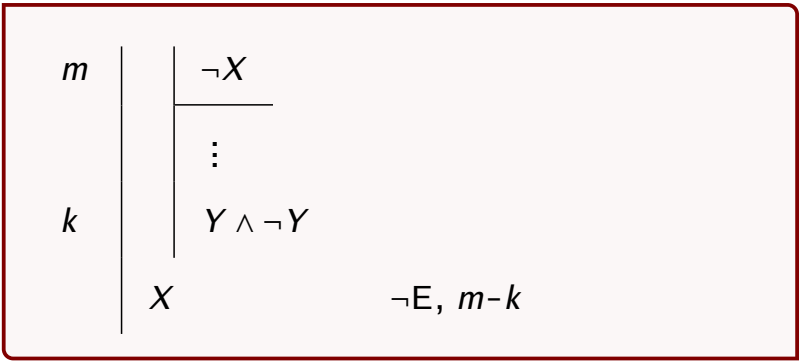
first derive $\neg Y$ and the derive Y. It is only important that in the subproof you derive both Y and $\neg Y$.

To see this in practice, and interacting with negation, consider this proof:

1	D	
2	$\neg D$	
3	$D \wedge \neg D$	$\wedge$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.



Formally, the rule is very similar to ¬I, but they are different.[1] 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.[2]

[1] 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.

[2] 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

To see this rule in action, here is an example proof using it:

1	$\neg\neg D$	
2	$\quad\neg D$	
3	$\quad\neg D \wedge \neg\neg D$	$\wedge$I, 1, 2
4	D	$\neg$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	$\quad\neg B$	
4	$\quad A \wedge \neg A$	$\wedge$I, 1, 2
5	B	$\neg$E, 3–4

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.

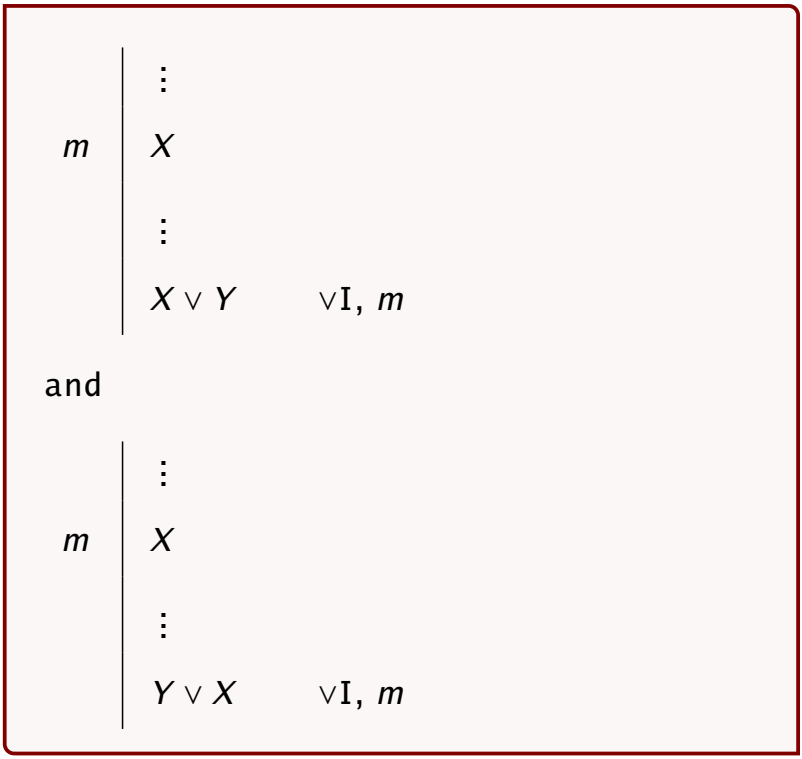
20.4 Disjunction

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):



Notice that Y can be **any** sentence whatsoever, so the following is a perfectly acceptable proof:

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

Using a truth table to show this would have taken 128 lines.

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 (∨E) rule:

m	$X \vee Y$	
i	$\quad X$	
	$\quad \vdots$	
j	$\quad Z$	
k	$\quad Y$	
	$\quad \vdots$	
l	$\quad Z$	
	Z	$\vee$E, m, i–j, k–l

This is obviously a bit clunkier to write down than our previous rules, but the point is fairly simple. Suppose we have some disjunction, $X \vee Y$. 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:

$$(P \wedge Q) \vee (P \wedge R) \therefore P$$

A proof corresponding to this argument is:

1	$(P \wedge Q) \vee (P \wedge R)$	
2	$\quad P \wedge Q$	
3	$\quad P$	$\wedge$E, 2
4	$\quad P \wedge R$	
5	$\quad P$	$\wedge$E, 4
6	P	$\vee$E, 1, 2–3, 4–5

20.5 Biconditional

We strictly speaking have not got the biconditional '$\leftrightarrow$' in our language of TFL. We have, however, let you use it in symbolizations. We said in section 16.5 that $X \leftrightarrow Y$ is an abbreviation for $(X \rightarrow Y) \wedge (Y \rightarrow X)$.

We then have rules in our natural deduction system that let you exchange $X \leftrightarrow Y$ for its explicit form.

	$\vdots$	
m	$X \leftrightarrow Y$	
	$\vdots$	
	$(X \rightarrow Y) \wedge (Y \leftrightarrow X)$	$\leftrightarrow$E, m

	$\vdots$	
m	$(X \rightarrow Y) \wedge (Y \rightarrow X)$	
	$\vdots$	
	$X \leftrightarrow Y$	$\leftrightarrow$I, m

Practice Exercises (Chapter 20)

I. 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	$S \rightarrow R$
3	S
4	R
5	$R \vee E$

1	$A \rightarrow D$	
2	$A \wedge B$	$(*)$
3	A	$(*)$
4	D	$(*)$
5	$D \vee E$	$(*)$
6	$(A \wedge B) \rightarrow (D \vee E)$	$(*)$

II. For each of the following arguments, give a proof using natural deduction.

a) $J \rightarrow \neg J \therefore \neg J$ $(*)$

b) $Q \rightarrow (Q \wedge \neg Q) \therefore \neg Q$ $(*)$

c) $A \rightarrow (B \rightarrow C) \therefore (A \wedge B) \rightarrow C$ $(*)$

d) $(C \wedge D) \vee E \therefore E \vee D$ $(*)$

e) $(Z \wedge K) \vee (K \wedge M), K \rightarrow D \therefore D$ $(*)$

f) $P \wedge (Q \vee R), P \rightarrow \neg R \therefore Q \vee E$ $(*)$

g) $\neg(P \rightarrow Q) \therefore \neg Q$ $(*)$

h) $\neg(P \rightarrow Q) \therefore P$ $(*)$

Chapter 21

Proofs and Validity

21.1 Proof Strategies

> These are all of the basic rules for the proof system for TFL. For ease of reference, they're listed again in section 3.7.

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 $\rightarrow$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 a proof by contradiction. 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 ¬E.

Law of Excluded Middle. There's one more convenient approach, closely related to proof by contradiction: if you are allowed to use derived rules, then the rule LEM is sometimes very helpful section 22.1. It will never be necessary, instead a proof by contradiction would do the job instead, but it can be very convenient.

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. . .

Example

Consider $C \therefore B \to ((A \to B) \wedge C)$ How to come up with a proof of this?

1. First step is always to just start organising the proof:

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

2. We'll use introduction rule for the main connective in the conclusion: $\to$Introduction

1	C	
2	B	
	?	
$n-1$	$(A \to B) \wedge C$	?
n	$B \to ((A \to B) \wedge C)$	$\to$I, 2-$n-1$

3. To fill out this, we'll use introduction rule for $\wedge$. We already have C, it was our premise, so we just need to show $A \to B$.

1	C	
2	$\quad B$	
	$\quad ?$	
$n-2$	$\quad A \to B$	?
$n-1$	$\quad (A \to B) \wedge C$	$\wedge$I, $n-2$, 1
n	$B \to ((A \to B) \wedge C)$	$\to$I, 2-$n-1$

4. How to get $A \to B$? Introduction rule for $\to$

1	C	
2	$\quad B$	
3	$\quad\quad A$	
	$\quad\quad ?$	
$n-3$	$\quad\quad B$	?
$n-2$	$\quad A \to B$	$\to$I, 3-$n-3$
$n-1$	$\quad (A \to B) \wedge C$	$\wedge$I, $n-2$, 1
n	$B \to ((A \to B) \wedge C)$	$\to$I, 2-$n-1$

5. How can we get to B from what we have? Well we actually already have it! We have alredy assumed B, on line 2 (and that is still available to us, as we're within that vertical line still). So we can just use Reiteration.

1	C	
2	$\quad B$	
3	$\quad\quad A$	
	$\quad\quad$ no more '?' to fill – we're done	
$n-3$	$\quad\quad B$	R, 2
$n-2$	$\quad A \to B$	$\to$I,
$n-1$	$\quad (A \to B) \wedge C$	$\wedge$I,
n	$B \to ((A \to B) \wedge C)$	$\to$I,

6. Just need to fill out the numbers properly now and we have our completed proof:

1	C	
2	$\quad B$	
3	$\quad\quad A$	
4	$\quad\quad B$	R, 2
5	$\quad A \to B$	$\to$I, 3–4
6	$\quad (A \to B) \wedge C$	$\wedge$I, 5, 1
7	$B \to ((A \to B) \wedge C)$	$\to$I, 2–6

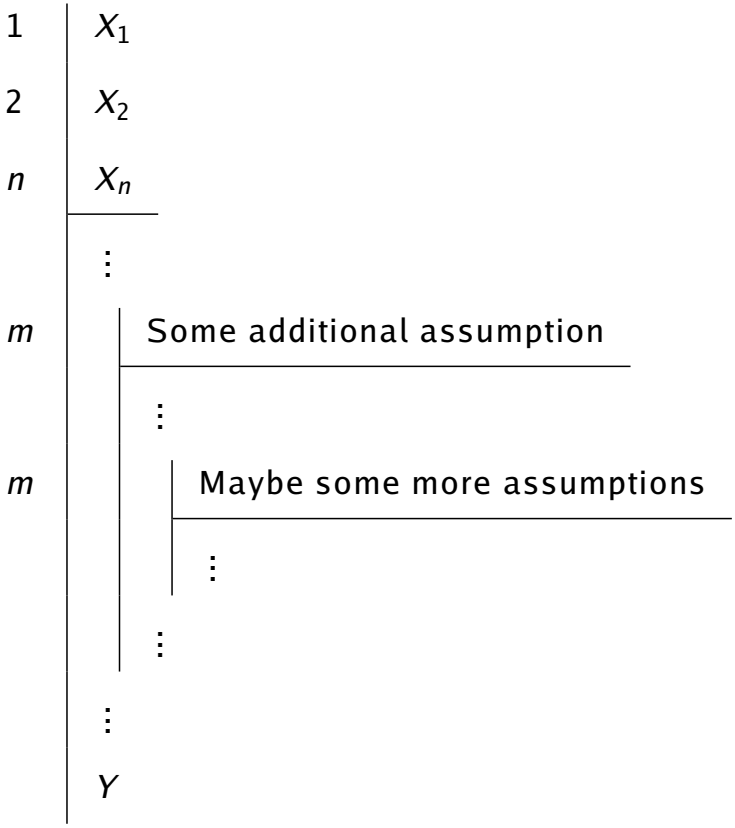
21.2 Proofs and Validity

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:

1	X_1
2	X_2
n	X_n
	⋮
	Y

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:



But any subproofs need to have been closed by the time we get to Y. So, for example if we gave

1	A	
2	$\quad B$	
3	$\quad B$	R, 2

this does not count as a proof corresponding to the argument $A \therefore B$. But

1	A	
2	$\quad B$	
3	$\quad B$	R, 2
4	$B \to B$	$\to$I, 2–3

counts as a proof corresponding to $A \therefore B \to B$ 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 $X_1 \dots X_n \therefore Y$, then this argument $X_1 \dots X_n \therefore Y$ 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!

Chapter 22

Derived rules for TFL

We have introduced all 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 chapter 23 we will see that they are not strictly speaking **necessary**.)

22.1 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 $X \vee \neg X$. In section 10.2 we saw that $A \vee \neg A$ is a tautology: it is true on all valuations. More generally, any sentence of the form $X \vee \neg X$ is a tautology. The rule **Law of Excluded Middle** encodes this fact: it simply says that you are always allowed to write $X \vee \neg X$:

⋮	
$X \vee \neg X$	LEM

As always, X can be whatever you want, e.g. $(A \wedge (B \rightarrow C))$. Then the rule tells us, e.g. that you can write $(A \wedge (B \rightarrow C)) \vee \neg(A \wedge (B \rightarrow C))$ on any line of the proof.

Consider the following brain teaser:[1]

> Three people are standing in a row looking at each other.

Alice

Bob

Charlie

[1] Originally by Hector Levesque.

> 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]

[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 section 33.2.

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 w
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 w
8		Therefore, someone who is happy is looking at someo

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 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

22.2 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:

m	$X \vee Y$	
n	$\neg X$	
	Y	DS, m, n

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.

22.3 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	$X \to Y$	
n	$\neg Y$	
	$\neg X$	MT, m, n

As usual, the premises may occur in either order.

22.4 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.

22.5 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*

22.6 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 (Chapter 22)

I. 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	$N \vee C$	

II. Give a proof for each of these arguments:

1. $\neg F \rightarrow G, F \rightarrow H \therefore G \vee H$ $(*)$
2. $E \vee F, F \vee G, \neg F \therefore E \wedge G$
3. $M \vee (N \rightarrow M) \therefore \neg M \rightarrow \neg N$
4. $(M \vee N) \wedge (O \vee P), N \rightarrow P, \neg P \therefore M \wedge O$
5. $A \wedge B) \vee (A \wedge C), \neg(A \wedge D), D \vee M \therefore M$

Chapter 23

Derivations of the 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 chapter 22 are not strictly speaking necessary, but can be derived from the basic rules of chapter 19.

23.1 Derivation of Disjunctive syllogism

Suppose that you are in a proof, and you have something of this form:

m	$X \vee Y$
n	$\neg X$

You now want, on line k, to prove Y. You can do this with the rule of DS, introduced in chapter 22, but equally well, you can do this with the **basic** rules of chapter 19:

m	$X \vee Y$	
n	$\neg X$	
k	$\quad X$	
$k+1$	$\quad\quad \neg Y$	
$k+2$	$\quad\quad X \wedge \neg X$	$\wedge$I, k, n
$k+3$	$\quad Y$	$\neg$E, $k+1$–$k+2$
$k+4$	$\quad Y$	
$k+5$	$\quad Y$	R, $k+4$
$k+6$	Y	$\vee$E, m, k–$k+3$, $k+4$–$k+5$

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 chapter 19.

23.2 Derivation of Modus Tollens

Suppose in the course of you proof you already have $X \to Y$, 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 chapter 22. But you could also do this with the **basic** rules of chapter 19:

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 chapter 19.

23.3 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 chapter 19.

23.4 Derivation of Explosion

Here is a demonstration of how we could derive explosion using the basic rules:

m	X	
n	$\neg X$	
$n+1$	$\quad \neg Y$	
$n+2$	$\quad X \wedge \neg X$	$\wedge$I, m, n
	Y	$\neg$E, $n+1$–$n+2$

23.5 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	$\quad X$	
$k+1$	$\quad\quad Y$	
$k+2$	$\quad\quad X \wedge Y$	$\wedge$I, k, $k+1$
$k+3$	$\quad\quad (X \wedge Y) \wedge \neg(X \wedge Y)$	$\wedge$I, $k+2$, m
$k+4$	$\quad \neg Y$	$\neg$I, $k+1$–$k+3$
$k+5$	$\quad \neg X \vee \neg Y$	$\vee$I, $k+4$
$k+6$	$\quad \neg X$	
$k+7$	$\quad \neg X \vee \neg Y$	$\vee$I, $k+6$
$k+8$	$\neg X \vee \neg Y$	$\vee$E, j, k–$k+5$, $k+6$–k

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 (Chapter 23)

I. Provide proof schemes that justify the addition of the third and fourth De Morgan rules as derived rules.

II. The proofs you offered in response to the practice exercises of chapter 22–25 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 24

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.

Logical Consequence

For this chapter we will make use of a further symbols:

We use the symbol $\models$ 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:

$$X_1, X_2, \ldots, X_n \models Y$$

The symbol '$\models$' is known as **the double–turnstile**, since it looks like a turnstile with two horizontal beams.

Let us be clear be clear. '$\models$' 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 chapter 6). So the metalanguage sentence:

- $P, P \rightarrow Q \models Q$

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 '$\models$'. Indeed, we can even consider the limiting case where there is no sentence (this should ring familiar from **??**):

$$\models Y$$

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 '⊢' is called the **single turnstile**. We want to emphasize that this is not the double turnstile symbol ('⊨') that we introduced for 'logical consequence'. The single turnstile, '⊢', concerns the existence of proofs; the double turnstile, '⊨', concerns the existence of valuations (or interpretations, when used for FOL). **They are very different notions.**

24.1 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
$$X_1, X_2, \ldots, X_n \models Y$$

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 $\models$ 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 (as in chapter 7). 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 (as in chapter 7). 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$_{\models}$. For instance the proof $X, Y \vdash X \wedge Y$ corresponds to a truth table that shows $X, Y \models X \wedge Y$ 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$_{\models}$ proof into an invalid$_{\models}$ 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 $X \models Y$. 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.

24.2 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 table 24.1

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 table 24.2:

Practice Exercises (Chapter 24)

I. Use either a derivation or a truth table for each of the following.

a) Show that $A \rightarrow (((B \wedge C) \vee D) \rightarrow A)$ is a tautology.
b) Show that $A \rightarrow (A \rightarrow B)$ is not a tautology
c) Show that the sentence $A \rightarrow \neg A$ is not a contradiction.
d) Show that the sentence $A \leftrightarrow \neg A$ is a contradiction.
e) Show that the sentence $\neg(W \rightarrow (J \vee J))$ is contingent

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	An argument whose truth table has no lines where there are all Ts under main	An argument where one can

f) Show that the sentence $\neg(X \vee (Y \vee Z)) \vee (X \vee (Y \vee Z))$ is not contingent
g) Show that the sentence $B \rightarrow \neg S$ is equivalent to the sentence $\neg\neg B \rightarrow \neg S$
h) Show that the sentence $\neg(X \vee O)$ is not equivalent to the sentence $X \wedge O$
i) Show that the sentences $\neg(A \vee B)$, C, $C \rightarrow A$ are jointly inconsistent.
j) Show that the sentences $\neg(A \vee B)$, $\neg B$, $B \rightarrow A$ are jointly consistent
k) Show that $\neg(A \vee (B \vee C)) \therefore \neg C$ is valid.
l) Show that $\neg(A \wedge (B \vee C)) \therefore \neg C$ is invalid.

II. Use either a derivation or a truth table for each of the following.

a) Show that $A \rightarrow (B \rightarrow A)$ is a tautology
b) Show that $\neg(((N \leftrightarrow Q) \vee Q) \vee N)$ is not a tautology
c) Show that $Z \vee (\neg Z \leftrightarrow Z)$ is contingent
d) show that $(L \leftrightarrow ((N \rightarrow N) \rightarrow L)) \vee H$ is not contingent
e) Show that $(A \leftrightarrow A) \wedge (B \wedge \neg B)$ is a contradiction

f) Show that $(B \leftrightarrow (C \vee B))$ is not a contradiction.

g) Show that $((\neg X \leftrightarrow X) \vee X)$ is equivalent to X

h) Show that $F \wedge (K \wedge R)$ is not equivalent to $(F \leftrightarrow (K \leftrightarrow R))$

i) Show that the sentences $\neg(W \rightarrow W)$, $(W \leftrightarrow W) \wedge W$, $E \vee (W \rightarrow \neg(E \wedge W))$ are inconsistent.

j) Show that the sentences $\neg R \vee C$, $(C \wedge R) \rightarrow \neg R$, $(\neg(R \vee R) \rightarrow R)$ are consistent.

k) Show that $\neg\neg(C \leftrightarrow \neg C), ((G \vee C) \vee G) \therefore ((G \rightarrow C) \wedge G)$ is valid.

l) Show that $\neg\neg L, (C \rightarrow \neg L) \rightarrow C) \therefore \neg C$ is invalid.

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 contradic–tion	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

Chapter 25

Proof-theoretic concepts

Armed with our '$\vdash$' symbol, we can introduce some new terminology.

Definition

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	$A \wedge \neg A$	
2	A	$\wedge$E, 1
3	$\neg A$	$\wedge$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:

Definition

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:

Definition

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 (Chapter 25)

I. Show that each of the following sentences is a theorem:

a) $O \to O$
b) $N \vee \neg N$

c) $J \leftrightarrow (J \vee (L \wedge \neg L))$
d) $((A \rightarrow B) \rightarrow A) \rightarrow A$

II. Provide proofs to show each of the following:

a) $C \rightarrow (E \wedge G), \neg C \rightarrow G \vdash G$
b) $M \wedge (\neg N \rightarrow \neg M) \vdash (N \wedge M) \vee \neg M$
c) $(Z \wedge K) \leftrightarrow (Y \wedge M), D \wedge (D \rightarrow M) \vdash Y \rightarrow Z$
d) $(W \vee X) \vee (Y \vee Z), X \rightarrow Y, \neg Z \vdash W \vee Y$

III. Show that each of the following pairs of sentences are provably equivalent:

a) $R \leftrightarrow E$, $E \leftrightarrow R$
b) G, $\neg\neg\neg\neg G$
c) $T \rightarrow S$, $\neg S \rightarrow \neg T$
d) $U \rightarrow I$, $\neg(U \wedge \neg I)$
e) $\neg(C \rightarrow D)$, $C \wedge \neg D$
f) $\neg G \leftrightarrow H$, $\neg(G \leftrightarrow H)$

IV. If you know that $X \vdash Y$, what can you say about $(X \wedge Z) \vdash Y$? What about $(X \vee Z) \vdash Y$? Explain your answers.

V. 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 26

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:

– Alice is a logician.
– All logicians wear funny hats.
∴ Therefore: Alice wears a funny hat.

We can symbolise this in TFL (follow the strategy as in section 14.2). 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.
> *H*: Alice wears a funny hat.
> *A*: All logicians wear funny hats.

This TFL argument is invalid. But the original English argument seems valid. The premises guaranted the colcusion independent of the subject matter, that is, independently of whether they are talking about Alice, logicians, funny hats or loving Manchester United. They do seem to be valid in virtue of the argument form. The problem lies with TFL. 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.

We will here give a sneak-preview by saying how the first argument will be symbolised in FOL to give you some ideas of how FOL works.

You might feel the pull to symbolise premise 2 as

$$L \to H.$$

With the given symbolisation key, this is to say

1. if Alice is a logician, she wears a funny hat.

That’s enough to get to the conclusion, and it’s a consequence of premise 2, but it’s not what premise 2 says. It’s not an adequate symbolisation of “All logicians wear funny hats”.

Instead you might try: well, lets just let L stand for 'logician' and H stand for 'wears a funny hat'. Then we can just use $L \to H$. The problem is that these aren't declarative sentences. They're not the sort of thing to be true or false. They only form declarative sentences once the person they're to be applied to is given, such as Alice.

In First Order Logic, we can capture the idea that whoever is being referred to, if they are a logician, then they wear a funny hat. Or "If someone is a logician, they wear a funny hat." In this sentence, 'they' isn't referring to any particular individual, but instead its reference is varying. In First Order Logic, we will have variables such as x, y, z which play this role. We then want to say of all people, if they are a logician then they wear a funny hat.

- for all x, if x is a logician, then x wears a funny hat.

We will then end up symbolising it as

$$\forall x(Lx \to Hx)$$

where L and H are not atomic sentences, but are instead predicates. They express properties of individuals: being a logician, or wearing a funny hat.

Saying that they apply to Alice, we will introduce a name in FOL 'a' which will stand for 'Alice'. La will then stand for 'Alice is a logician' and Ha stands for 'Alice wears a funny hat'.

In First Order Logic, then we can capture the argument as:

$$La,\ \forall x(Lx \rightarrow Hx)\ \therefore\ Ha$$

We can also capture arguments which talk about 'some' or 'exists'.

- Every logician wears a funny hat.
- There is a logician.

$\therefore$ Therefore: Somone wears a funny hat.

Which we will can end up symbolising as

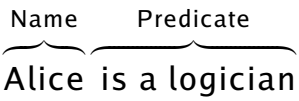
$$\forall x(Lx \rightarrow Hx),\ \exists xLx\ \therefore\ \exists xHx$$

26.1 Names and Predicates

Consider

2. 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.



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 ... for names (except x, y, z which are used for variables as we will later see), and upper case letters like A, B, C, ... 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: $\_\_\_\_\_{}_x$ is a logician

(We will say more about the "x" subscript later.) and we can then symbolise 'Alice is a logician' as

$$La.$$

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: _____$_x$ 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.

3. Rocky is strong
4. Joe Biden is a Democrat
5. Michael Palin is a member of Monty Python

In each of these cases the relevant symbolisation key would then be:

3 *a*: Rocky

Px: _____$_x$ is strong

4 *a*: Joe Biden
Px: _____$_x$ is a Democrat

5 *a*: Michael Palin
Px: _____$_x$ is a member of Monty Python

Names don't have to name people, for example we can also symbolise

6. 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

7. Buses are red.

You might think that this has the same form and symbolise it as *La* with the symbolisation key:

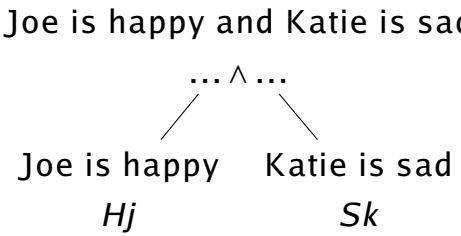
a: Buses
Lx: _____$_x$ is red

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.

26.2 Names, predicates and connectives

In FOL we will also make use of all of the tools from TFL. We can symbolise

8. Joe is happy and Katie is sad



as

$$Hj \wedge Sk$$

with the symbolisation key:

Hx: $\_\_\_\_\_x$ is happy
Sx: $\_\_\_\_\_x$ is sad
j: Joe
k: Katie

To symbolise

9. 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

10. If Joe is happy, then Katie is too

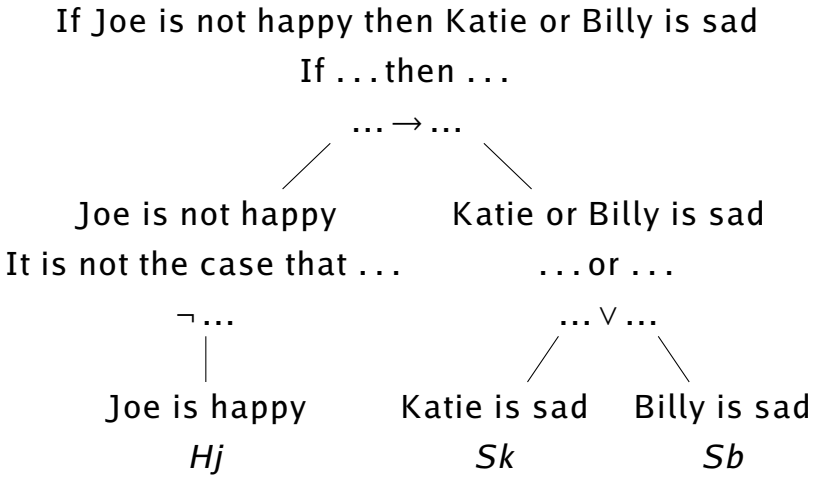
we observe that it can be naturally paraphrased as 'If Joe is happy then Katie is happy' and thus

symbolised as

$$Hj \rightarrow Hk$$

We can also symbolise more complex sentences, for example:

11. If Joe is not happy then Katie or Billy is sad.



so giving the symbolisation:

$$Hj \rightarrow (Sk \vee Sb)$$

To symbolise:

12. 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.

However, just like with TFL, we have to be careful. To symbolise:

13. Adina is a skilled surgeon

Should one simply symbolise it as

$$Ka$$

using

Kx: $\_\_\_\_\_x$ is a skilled surgeon
a: Adina

Or should one split it into

$$La \wedge Sa$$

using

Lx: $\_\_\_\_\_x$ is skilled
Sx: $\_\_\_\_\_x$ is a surgeon
a: Adina

This symbolisation would naturally lead to a symbolisation of this argument which would make it valid:

> Adina is a skilled surgeon. She also plays tennis. Therefore, Adina is a skilled tennis player.

by symbolising it as

$$La \wedge Sa,\ Ta \ \therefore\ La \wedge Ta$$

26.3 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:

14. Karl loves Imre.
15. Imre loves himself.
16. Karl loves Imre, but not vice versa.
17. Karl is loved by Imre.

We will start with the following symbolisation key:

domain: people
i: Imre
k: Karl
Lxy: _____$_x$ loves _____$_y$

Sentence 14 will now be symbolized by *Lki*.

Sentence 15 can be paraphrased as 'Imre loves Imre'. It can now be symbolized by *Lii*.

Sentence 16 is a conjunction. We might paraphrase it as 'Karl loves Imre, and Imre does

not love Karl'. It can now be symbolized by $Lki \wedge \neg Lik$.

Sentence 17 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

18. 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.

19. 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 $\_\_\_\_\_x_4$.

26.4 Universal Quantifier

The universal quantifier is

$$\forall$$

We read it out loud as “for all”.

Consider

20. All logicians wear funny hats.

The sentence can be paraphrased as:

> If someone is a logician, then they wear a funny hat.

Here the pronoun ‘they’ is not referring to a particular person, Harry or Katie, but instead its reference can vary, and we want to say of all people, if they are a logician, then they wear a funny hat. This is made explicit in our logic with quantifiers and variables. We use a variable, such as x, or y or z instead of the English pronoun ‘they’.

21. For all people, x: if x is a logician, then x wears a funny hat.

And we symbolise it as

$$\forall x(Lx \rightarrow Hx)$$

with our symbolisation key:

Lx: $\_\_\_\_\_{}_x$ is a logician.
Hx: $\_\_\_\_\_{}_x$ wears a funny hat.

We also want to use our logic to talk not only about people but also other kinds of objects, such as coins.

22. Everything in my pocket is a coin.

If something is in my pocket, then it is a coin.

Here 'it' is not referring to a particular object, instead the object that it refers to can vary. This sentence is saying of every thing, if it is in my pocket then it is a coin.

For all things, if it is in my pocket, then it is a coin.

We make this explicit in our logic with variables and quantifiers. A variable such as *x*, or *y* or *z* is used not to refer to a particular object but to have a varying reference.

For all *x*, if *x* is in my pocket, then *x* is a coin.

We symbolise this as

$$\forall x(Px \to Cx)$$

$\forall$ is read as "for all", and is also used to naturally paraphrase those sentences of English which use "every".

$\forall x$ ... is read as:

- For all x: ...

We also use it for:

- For everything, x: ...

Restricted universal quantifiers

The examples we have seen so far have a common pattern, and it is a typical one which we see with the universal quantifier. Their pattern might be given in one of the following (equivalent) ways:

23. All the things like **this** are like **that**.
24. Everything like **this** is like **that**.
25. If something is like **this** then it is like **that**.

We will always consider this to be given by:

For all x: if x is like **this**, then x is like **that**.

So the common pattern is that when you see $\forall$ you often see it paired with $\rightarrow$.

When we say

26. Every dog is energetic

So we're expecting to use $\forall$. And then very likely we'll be using this form with $\rightarrow$. So we expect to be able to paraphrase it as:

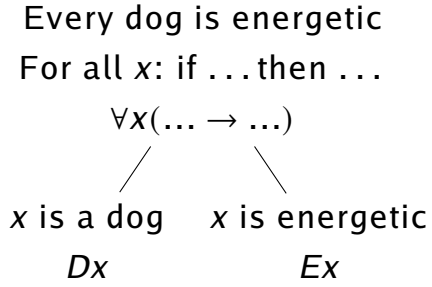
For all x: if . . . then . . .

We just need to think about English ways of filling out these '. . .'. In this case: 'x is a dog' and 'x is energetic' seem to do the job. This is an appropriate paraphrase of the original sentence.

For all x: if x is a dog then x is energetic.

So to symbolise it, we use $\forall x(\ldots \rightarrow \ldots)$, filling the gaps with how we will symbolise 'x is a dog' (Dx)

and 'x is energetic' (Ex). In our symbolisation tree presentation:



Resulting in:

$$\forall x(Dx \rightarrow Ex).$$

Here is the general instruction:

> 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
>
> $$\forall x(Fx \rightarrow Gx).$$

This will be used also for more complicated sentences. Consider:

27. Every enthusiastic student is not tired.

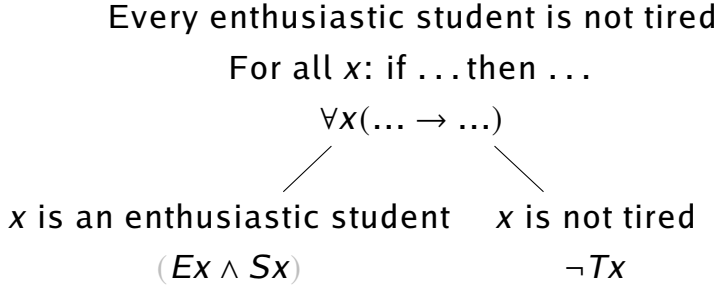
We see that it's main form is saying something about everyone. So we're expecting to use ∀. ∀*x* is understood as everything is like blah. What do we want to say of everything? We want to say, if it is like **this** and **that**. So we need to decide what to use to play the role of '**this**' and '**that**'. Take this sentence and try to find things filling out ..., things about *x*, which results in a natural paraphrasing of the original sentence.

> For all *x*: if . . . , then . . . .

In this case:

> For all *x*: if *x* is an enthusiastic student, then *x* is not tired.

So when we're symbolising it we apply our symbolisation strategy and break it up in stages.



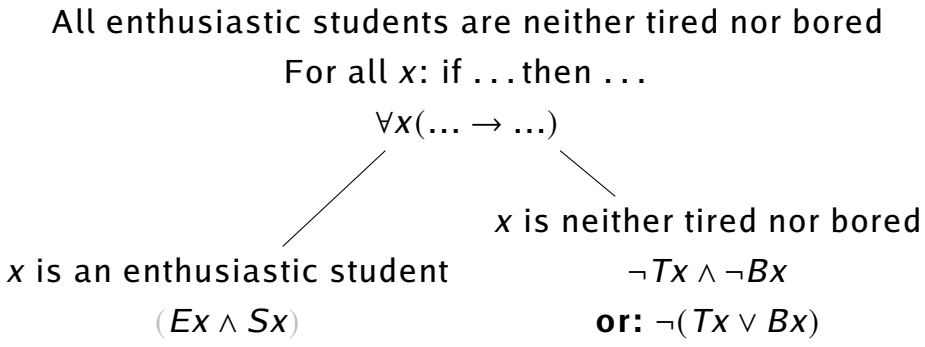
so provide the symbolisation

$$\forall x((Ex \wedge Sx) \rightarrow \neg Tx)$$

Recalling our bracketing rules, we can also drop some brackets:

$$\forall x(Ex \wedge Sx \rightarrow \neg Tx)$$

This can also apply when things get complicated. We make use of all of our symbolisation tricks we learned from TFL. Consider a sentence involving 'neither':



Giving us

$$\forall x((Ex \wedge Sx) \rightarrow (\neg Tx \wedge \neg Bx))$$

Recalling our bracketing rules, we can also drop some brackets:

$$\forall x(Ex \wedge Sx \rightarrow \neg Tx \wedge \neg Bx)$$

26.5 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 “something” or “there exists”.

$\exists x$ ... is read as:

- There exists at least one x: ...

Or, shorter:

- There exists x: ...

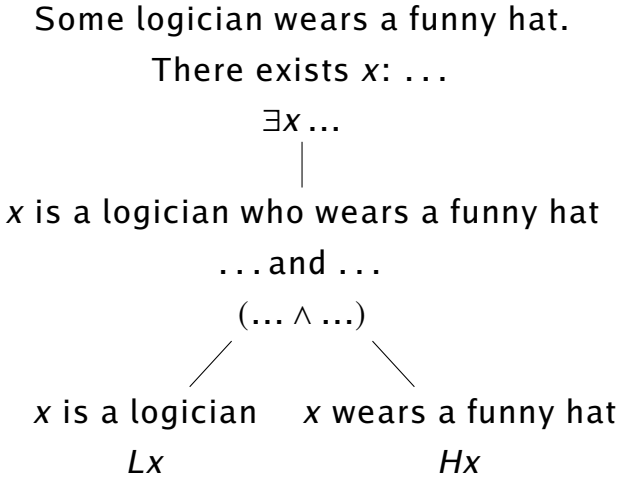
We also use it for:

- For something, x: ...

28. Some logician wears a funny hat.

We paraphrase it as

29. There exists at least one x: x is a logician who wears a funny hat.



giving

$$\exists x(Lx \wedge Hx)$$

When using the existential quantifier, $\exists$, we typically find it with $\wedge$, unlike the universal quantifier, $\forall$, which typically comes paired with $\rightarrow$. Once we do FOL semantics we can see why this is right. For now, we simply motivate it with paraphrasing.

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

$$\exists x(Fx \wedge Gx).$$

30. There's a coin in my pocket.

There's a coin in my pocket.
There exists x: ... and ...

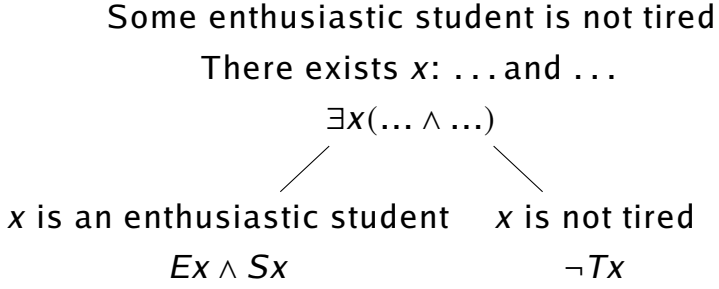
$$\exists x(\ldots \wedge \ldots)$$

x is a coin — Cx
x is in my pocket — Px

giving

$$\exists x(Cx \wedge Px)$$

31. Some enthusiastic student is not tired.



giving

$$\exists x((Ex \wedge Sx) \wedge \neg Tx)$$

or

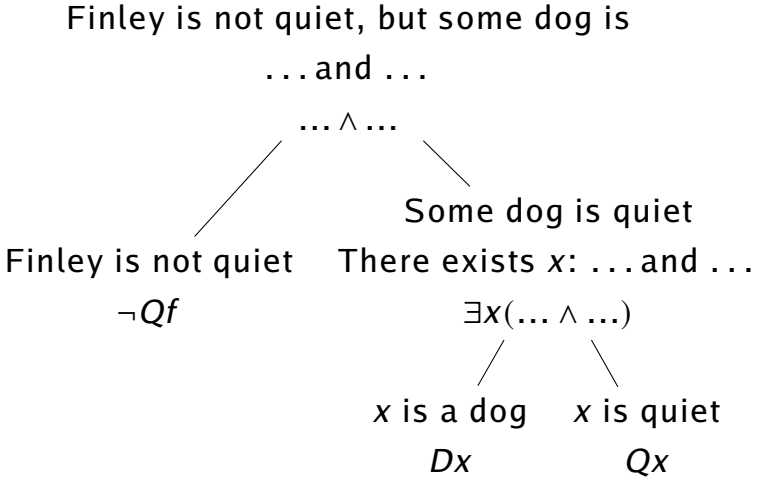
$$\exists x(Ex \wedge Sx \wedge \neg Tx)$$

26.6 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.

32. Finley is not quiet, but some dog is.

This sentence has the form '. . . but . . .', which we treat as '. . . and . . .'. We then symbolise each of those components separately, then put them back together to provide our symbolisation for this sentence.



So we symbolise this sentence as

$$\neg Qf \wedge \exists x(Dx \wedge Qx)$$

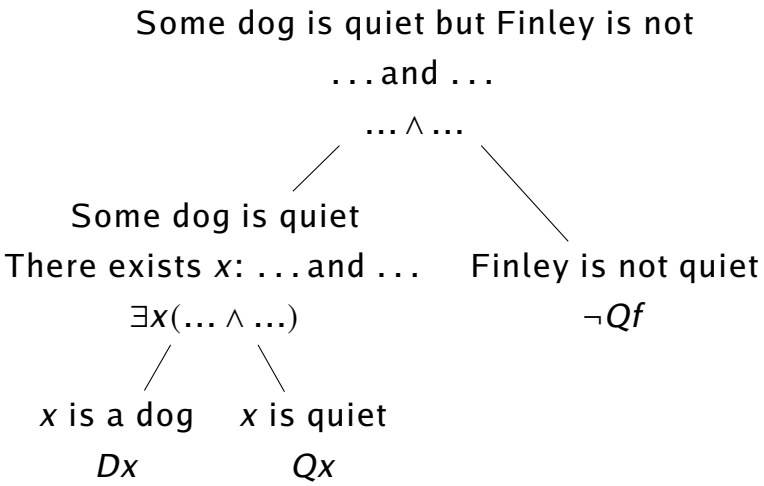
giving the symbolisation key

Dx: _____$_x$ is a dog
Qx: _____$_x$ is quiet

f: Finley

Very similarly, consider:

33. Some dog is quiet but Finley is not



So we result in:

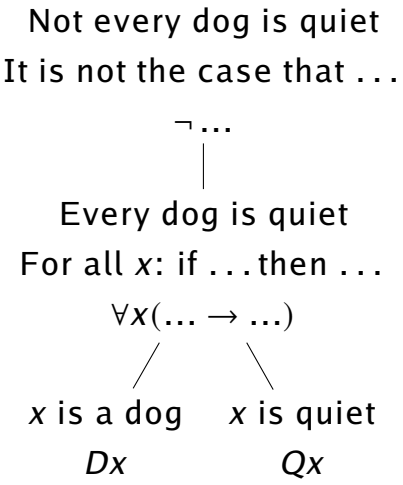
$$\exists x(Dx \wedge Qx) \wedge \neg Qf$$

Here the $\exists x$ is at the beginning of the sentence, but it is not the main logical operator. The main logical operator is $\wedge$.

Consider:

34. Not every dog is quiet

We work as follows:



So we symbolise this sentence as

$$\neg\forall x(Dx \rightarrow Qx)$$

26.7 Quantifiers and many-placed predicates

Consider

35. All vets like Snowball.

We will symbolise this with $\forall$.

$$\forall x \ldots .$$

Again, we are expecting to see its form $\forall x(\ldots \rightarrow \ldots)$. So we need to think of ways of filling out '. . .' with parts of English which we can then work on symbolising.

All vets like Snowball
For all x: if . . . then . . .
$\forall x(\ldots \rightarrow \ldots)$

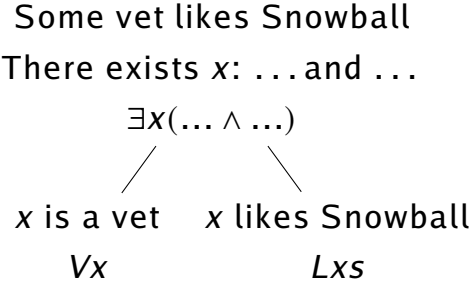
x is a vet	x likes Snowball
Vx	Lxs

So giving

$$\forall x(Vx \rightarrow Lxs)$$

Consider

36. Some vet likes Snowball.



So we give:

$$\exists x(Vx \wedge Lxs)$$

Now consider

37. Charlie likes every animal.

This sentence doesn't start with 'every' or 'all'. But since it includes 'every', we might think that it will be paraphrased as something like 'every **blah** is a **blah**'. Can we come up with something of English which is a natural paraphrasing of this sentence.

> For everything, if it is like **this** then it is like **that**

Here's a suggestion:

For everything, if it is an animal, then Charlie likes it.
For all x, if x is an animal then
Charlie likes x.

So we give:

Charlie likes every animal.
For all x: if . . . then . . .
$\forall x(\ldots \rightarrow \ldots)$

x is an animal	Charlie likes x
Ax	Lcx

$$\forall x(Ax \rightarrow Lcx)$$

26.8 Combining quantifiers

First Order Logic is particularly powerful once we start combining quantifiers.

38. Some vet is such that every animal likes them.

We have both ‘exists’ and ‘forall’ doing work here. But which is the main one. Can we paraphrase it as “exists x:. . . ” or as “forall x:. . . ”. ‘Some’ comes first, it’s perhaps the first one you should try. Can we paraphrase it as:

> There exists *x*: . . . and . . .

Yes:

> There exists *x*: *x* is a vet and
> every animal likes *x*.

We write it out in a symbolisation tree to help keeping track, and then work on symbolising each component.

Some vet is such that every animal likes them.

There exists *x*: . . . and . . .

$\exists x(\ldots \wedge \ldots)$

x is a vet — *Vx*

every animal likes *x* — ?

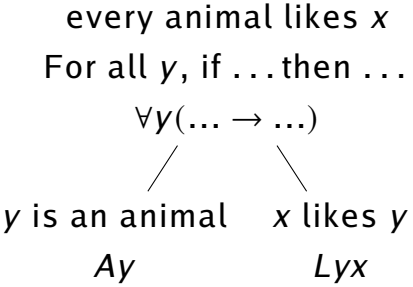
'*x* is a vet' can be symbolised as *Vx*. How should we symbolise 'every animal likes *x*'? We can paraphrase it as:

> For all things: if it is an animal, then it likes *x*.

Now, we have already got variable *x* in play – it's the vet – so we should choose a different variable, such as *y*.

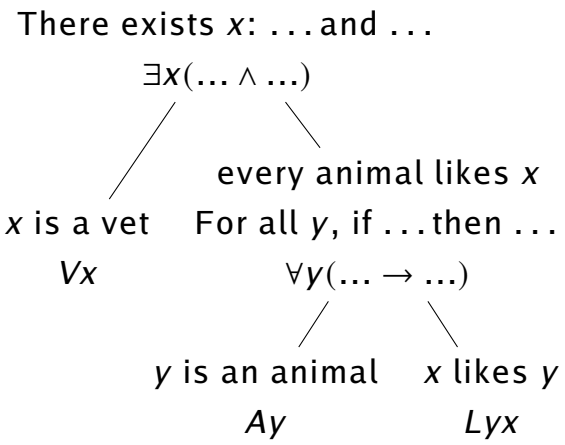
> For all *y*: if *y* is an animal then *y* likes *x*.

In our symbolisation tree:



Putting these together:

Some vet is such that every animal likes them.



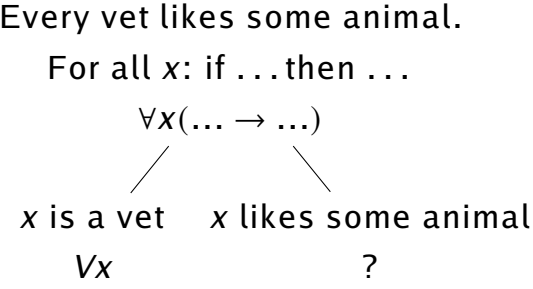
So we give:

$$\exists x(Vx \wedge \forall y(Ay \rightarrow Lyx))$$

Consider

39. Every vet likes some animal.

Again we can break it down:



And our task which we should now focus on is how to symbolise

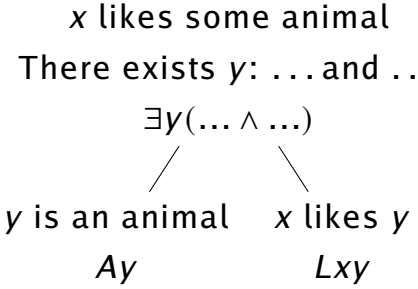
> *x* likes some animal.

We can see it has an existential form: "there exists something". We can paraphrase it as:

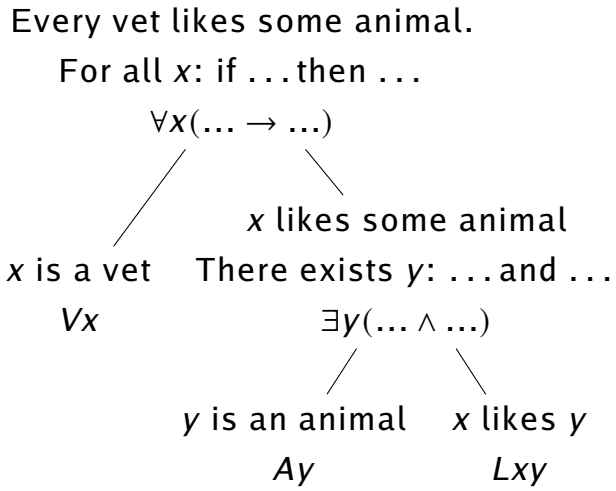
> There exists something which is an animal and which *x* likes.

But again, *x* is our vet; so we need to use a different variable, lets use *y*.

> There exists *y*: *y* is an animal and *x* likes *y*.



So piecing them together we get



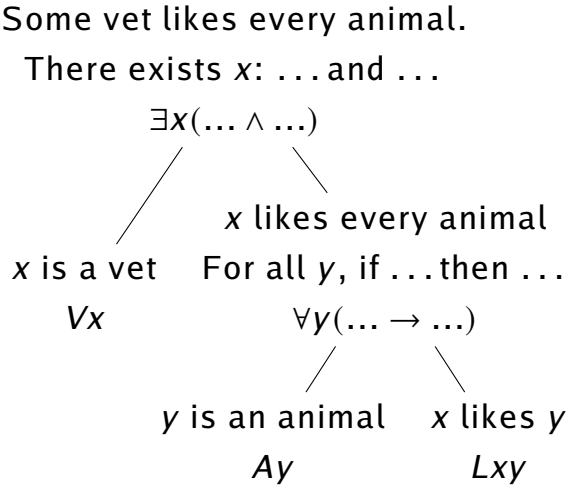
So we give:

$$\forall x(Vx \rightarrow \exists y(Ay \wedge Lxy))$$

Consider:

40. Some vet likes every animal.

Again we can manage to break it down step by step:

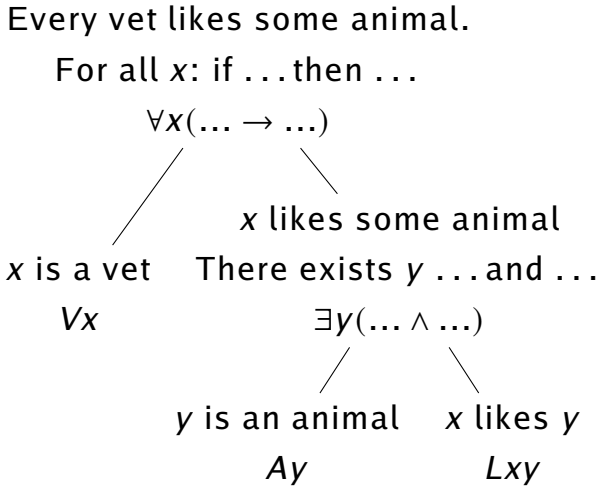


So we give:

$$\exists x(Vx \wedge \forall y(Ay \rightarrow Lxy))$$

Consider:

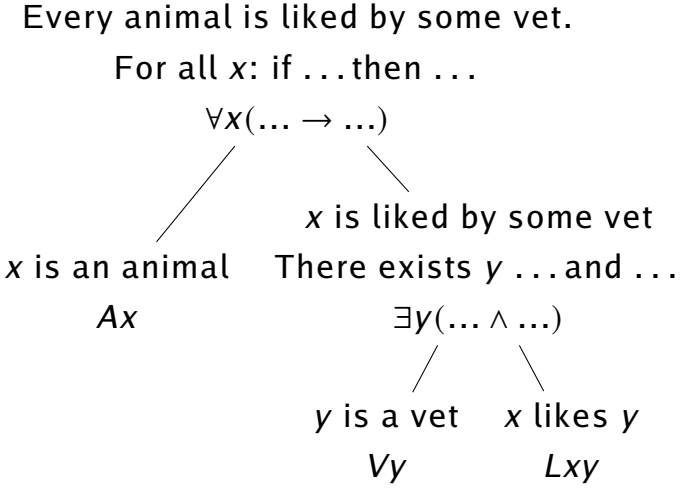
41. Every vet likes some animal.



So we give:

$$\forall x(Vx \rightarrow \exists y(Ay \wedge Lxy))$$

42. Every animal is liked by some vet.

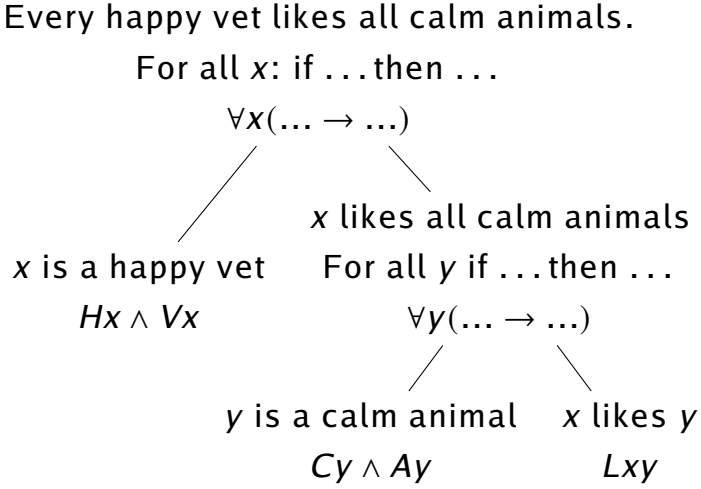


Giving:

$$\forall x(Ax \rightarrow \exists y(Vy \wedge Lxy))$$

These strategies can be applied with more complicated case

43. Every happy vet likes all calm animals.



Giving:

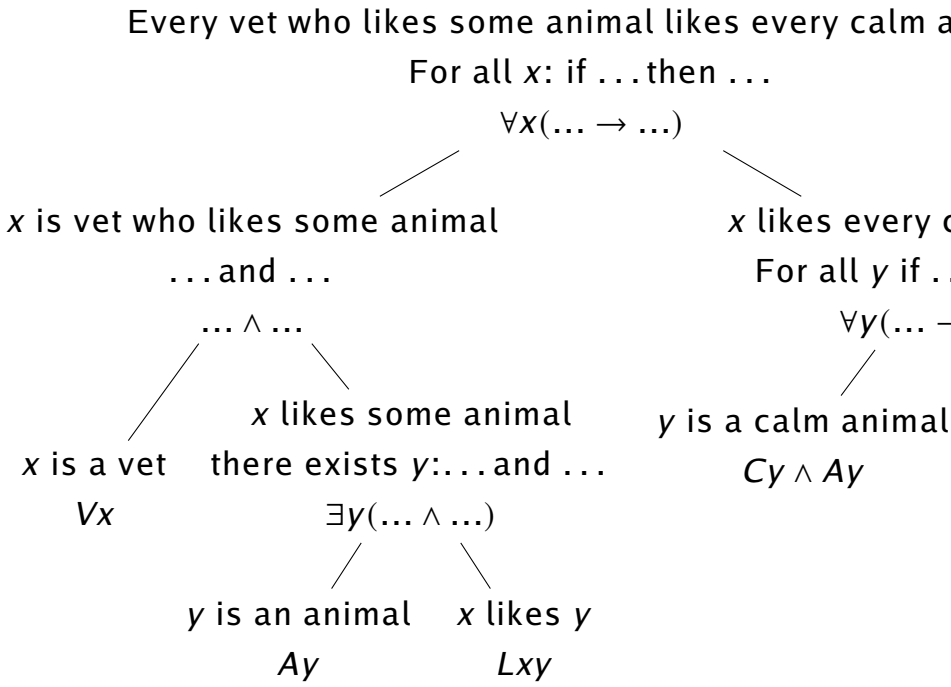
$$\forall x((Hx \wedge Vx) \rightarrow \forall y((Cy \wedge Ay) \rightarrow Lxy))$$

With our rules for dropping brackets fully utilised we could just give:

$$\forall x(Hx \wedge Vx \rightarrow \forall y(Cy \wedge Ay \rightarrow Lxy))$$

The remaining brackets cannot be dropped.

44. Every vet who likes some animal which is not calm likes every calm animal.



giving:

$$\forall x(Vx \wedge \exists y(Ay \wedge Lxy) \rightarrow \forall y(Cy \wedge Ay \rightarrow Lxy))$$

When evaluating this, the y that appears on the left hand side, in the scope of the $\exists y$ quantifier, is evaluated entirely separately from the y in the scope of the $\forall y$ quantifier.

These strategies of breaking down the sentence and working on symbolising the components separately using the symbolisation key can take you very far.

Let us now return to more carefully see the syntax of the language of FOL. How to read a sentence? Where do brackets go?

Practice Exercises (Chapter 26)

I. Using this symbolization key:

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:

a) If Cleo loves Bouncer, then Bouncer is a monkey.

(∗)

b) If both Bouncer and Cleo are alligators, then Amos loves them both.

(∗)

c) Cleo loves a reptile.

(∗)

d) Bouncer loves all the monkeys that live at the zoo.

(∗)

e) All the monkeys that Amos loves love him back.

(∗)

f) Every monkey that Cleo loves is also loved by Amos.

(∗)

g) There is a monkey that loves Bouncer, but sadly Bouncer does not reciprocate this love.

(∗)

II. Using the following symbolization key:

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:

a) All the food is on the table.

(*)

b) If the guacamole has not run out, then it is on the table.

(*)

c) Everyone likes the guacamole.

(*)

d) If anyone likes the guacamole, then Eli does.

(*)

e) Francesca only likes the dishes that have run out.

(*)

f) Eli likes anyone who likes the guacamole.

(*)

g) Eli likes anyone who likes the people that he likes.

(∗)

h) If there is a person on the table already, then all of the food must have run out.

(∗)

III. Using the following symbolization key:

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:

a) All of Patrick's children are ballet dancers.

(∗)

b) Jane is Patrick's daughter.

(∗)

c) Patrick has a daughter.

(∗)

d) All of Patrick's sons dance ballet.

(∗)

e) Jane is Elmer's niece.

(∗)

f) Patrick is Elmer's brother.

(∗)

g) Jane is an aunt.

(∗)

h) Everyone who dances ballet has a brother who also dances ballet.

(∗)

i) Every woman who dances ballet is the child of someone who dances ballet.

(∗)

Chapter 27

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. This is important to, for example, know precisely what is allowed in writing expressions of FOL and where brackets are to be added or not.

27.1 Vocabulary of FOL

We'll start by summarising, a bit more formally, the vocabulary of FOL:

Definition

The vocabulary of FOL is as follows:

Predicates $A, B, C, \ldots, W$,
or with subscripts, as needed: $A_1, W_2, A_{25}, J_{375}, \ldots$.[a]

Names $a, b, c, \ldots, s, t$,
or with subscripts $a_1, b_{224}, h_7, m_{32}, \ldots$

Variables x, y, z,
or with subscripts $x_1, y_1, z_1, x_2, \ldots$.
u, v, w may also be used.

Connectives $\neg, \wedge, \vee, \rightarrow$

Brackets (,)

Quantifiers $\forall, \exists$

[a] Each predicate will have a number of **places** associated with it. We should thus really introduce:

Zero-Place Predicates = Atomic sentences of TFL
$A^0, B^0, \ldots, W^0$, with subscripts, as needed: $A^0_1, W^0_2, A^0_{25}, J^0_{375}, \ldots$

One-Place Predicates $A^1, B^1, \ldots, W^1$, with subscripts, as needed: $A^1_1, W^1_2, A^1_{25}, J^1_{375}, \ldots$

Two-Place Predicates $A^2, B^2, \ldots, W^2$, with subscripts, as needed: $A^2_1, W^2_2, A^2_{25}, J^2_{375}, \ldots$

Three-Place Predicates $A^3, B^3, \ldots, W^3$, with subscripts, as needed: $A^3_1, W^3_2, A^3_{25}, J^3_{375}, \ldots$

etc. We drop the superscripts for ease.

Definition

An **expression of FOL** is any string of symbols of FOL. Take any of the symbols of FOL and write them down, in any order, and you have an expression.

27.2 Formulas

In chapter 7, 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.

During our symbolisation strategy we have stages like symbolising "x is happy" or "x likes someone". These are not the sorts of things that are true or false outright. They are not sentences of English. To determine if they are true or false, not only do you need to know what the world is like (be given a model), you also need a specification of who x is.

These are symbolised in FOL as Hx or $\exists yLxy$, and the variable x is unbound. To know whether it

is true, you need to know who x is. These are not **sentences** of FOL, but they are well-formed **formulas**. Formulas are the sorts of things that form sentences when you put quantifiers in front of them. Sentences are special kinds of formulas, those which have no so-called free-variables. No additional specification of who x is is needed in $\forall x\, Hx$. It is a sentence.

What counts as a formula of FOL is given by an inductive definition similar to that which we gave for TFL. 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**.[1]

[1] In FOL variables and names are both called (singular) terms. Notice that in the above definition P and t are

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.

---

metavariables ranging over names of predicates and names of terms respectively.

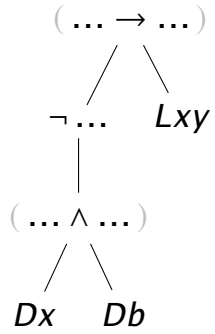
Definition

A **formula of FOL** is defined inductively as follows:

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 in TFL, we can construct more complex formulas from simple ones. Like in TFL, we will write outer brackets in invisible ink, which here is

displayed in gray. Such outer brackets can be dropped when the formula is written by itself but must be included once it is embedded in a further formula.



$$\neg(Dx \wedge Db) \rightarrow Lxy$$

We can also display the construction of formulas by with braces, where we also indicate the main connective of the braced part of the sentence with a red dot.

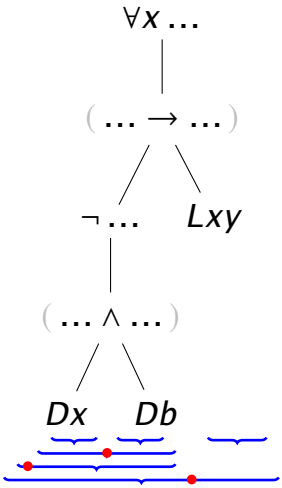
$$(\neg(Dx \wedge Db) \rightarrow Lxy)$$

This is exactly as in the case of TFL, the only difference is that the smallest components of the formula have more structure to them: they are predicates applied to names or variables rather than simply the single atomic sentences that we had in TFL.

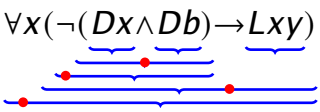
FOL also allows us to add quantifiers with variables in the front of a formula we have. We can thus construct

$$\forall x(\neg(Dx \wedge Db) \rightarrow Lxy)$$

in this process:



Or indicating this with scope braces:



Note that the quantifier phrase, $\forall x$, does not itself add any new brackets. Brackets are only introduced with the binary connectives: $\wedge, \vee, \rightarrow$.

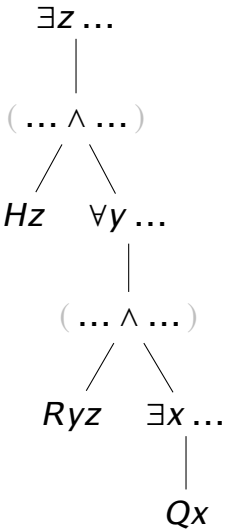
If our sentence only includes quantifiers and negations and none of these binary connectives, then there are no brackets:

Which is also usefully indicated with scope brackets:

$$\neg\forall x\neg Hx$$

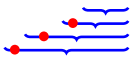
This is a sentence symbolising ‘Not everything is not happy’.

We can also add existential quantifiers, and mix up the variables in use, for example:

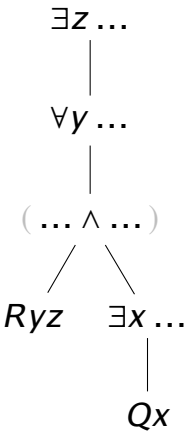


Or indicating this with scope braces:

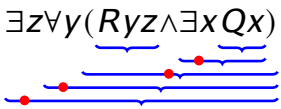
$$\exists z(Hz \wedge \forall y(Ryz \wedge \exists x Qx))$$



If no binary connectives are used, the quantifiers just get stacked on top of one another without no brackets.



Or indicating this with scope braces:



The notions of scope and main logical operators that were given in chapter 7 equally applies to FOL but now the main logical operator might be a quantifier. These were:

Definition

The **main logical operator** of a formula is the last logical operator ($\neg$, $\wedge$, $\vee$, $\rightarrow$, $\exists v$ or $\forall v$) that was introduced in the construction of the formula.

The **scope** of an occurrence of a logical operator in a sentence is the sub–formula for which that operator is the main logical operator.

$$\overbrace{\neg\,\exists z\big(\underbrace{\exists y(\,\overbrace{\forall x Rxy}^{\text{scope of }\forall x}\,\wedge Ryz)}_{\text{scope of }\exists y}\wedge\,\underbrace{\forall y Rzy}_{\text{scope of }\forall y}\,\big)}^{\text{scope of }\exists z}$$

Since there is no interaction between the scope of $\exists y$ and $\forall y$ the same variable y can be used in both places without confusion or danger, the evaluation is done entirely separately.

27.3 Sentences

Consider this formula, which we used to symbolise the English sentence ‘some vet likes

every animal'.

$$\exists x(Vx \wedge \forall y(Ay \rightarrow Lxy))$$

Indicating the scope of the quantifiers we get:

$$\overbrace{\exists x\,(Vx \wedge \underbrace{\forall y\,(Ay \rightarrow Lxy)}_{\text{scope of } \forall y})}^{\text{scope of } \exists x}$$

Looking at this, we can see that every occurrence of a variable occurs within the scope of a quantifier which binds that variable:

$$\overbrace{\exists x\,(V\,x \wedge \underbrace{\forall y\,(A\,y \rightarrow L\,x\,y)}_{\text{scope of } \forall y})}^{\text{scope of } \exists x}$$

Compare that to this formula:

$$\forall y(Ay \rightarrow Lxy)$$

This symbolises something of pseudo-English: 'Every animal is liked by x'. But who does 'x' refer to? There is an occurrence of a variable which is not in the scope of a quantifier which binds it.

Let's see some more examples:

$$Hx \wedge \neg Px \tag{27.1}$$

$$\exists y(Lyx \wedge \neg Lxy) \tag{27.2}$$

$$\underbrace{\forall x Hx}_{\text{scope of } \forall x} \rightarrow \exists y Rxy \tag{27.3}$$

In each case, there are some variable occurrences which are not in the scope of a quantifier binding them: the ones highlighted in yellow. We will call this being free.

> An occurrence of a variable v is **bound** if it is within the scope of either $\forall v$ or $\exists v$.
>
> An occurrence of a variable v is **free** if it is not within the scope of either $\forall v$ or $\exists v$.

Consider this example and note that the highlighted occurrence of the variable occurs free. Even though there is a quantifier over x in the sentence, it is not within the scope of that quantifier.

$$\underbrace{\forall x Hx}_{\text{scope of } \forall x} \rightarrow \exists y Rxy \tag{27.4}$$

If a formula has a free variable, the formula doesn't get a fixed truth value in a model. To know whether it is true, you need to specify who x is. This won't be fixed by a model because it is important for the semantics that who x refers to can vary. What this means is that it won't get a truth value on the model without the additional information given by a variable assignment specifying who x is.

Finally we can say what a sentence of FOL is

Definition

A **sentence** of FOL is any formula of FOL that contains no free variables.

It is sentences which are given truth values on a model. For formulas which still have free variables, they are only given truth values on a model with additional information specified by a variable assignment.

27.4 Bracketing conventions

We will adopt the same notational conventions governing brackets that we did for TFL (see

section 7.4): we may omit the outermost brackets of a formula.

Practice Exercises (Chapter 27)

I. Identify which variables are bound and which are free.

a) $\exists x Lxy \wedge \forall y Lyx$
b) $\forall x Ax \wedge Bx$
c) $\forall x(Ax \wedge Bx) \wedge \forall y(Cx \wedge Dy)$
d) $\forall x \exists y(Rxy \rightarrow (Jz \wedge Kx)) \vee Ryx$
e) $\forall x_1(Mx_2 \rightarrow Lx_2x_1) \wedge \exists x_2 Lx_3x_2$

Part VI

First Order Logic Semantics

Chapter 28

Truth in a model

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. This will be done with more care in chapter 29 but here we will introduce the ideas.

In TFL, the notion of 'interpretation' was spelled out by a valuation, which assigns a truth value to each atomic sentence. A valuation

provides enough information to determine the truth values of all the sentences of TFL, by applying the rules associated with each of the connectives and working up the formation tree.

What plays the equivalent role for FOL? What is the information that one needs to provide in order to determine the truth value of each of the sentences of FOL? The answer will be called a **model**.

Suppose we have symbolised

1. Alice is happy

as *Ha* with a symbolisation key

Px: _____$_x$ is a philosopher
Hx: _____$_x$ is happy
a: Alice

To evaluate whether it is true, one needs to know who Alice is, and if she is happy.

If we have symbolised

2. All philosophers are happy

as $\forall x(Px \rightarrow Hx)$ To evaluate whether it is true, one needs to work out who the philosophers are and ask each of them whether they are happy.

If we have symbolised

3. Some philosopher likes Alice.

as $\exists x(Px \wedge Lxa)$ To evaluate whether it is true, one needs to work out who Alice is, who the philosophers are, and whether one of them likes Alice.

To evaluate $\forall xSx$ we need to know what all the objects are, and whether they have property S.

The information needed to determine these matters is the following:

- A domain. A collection of objects.
- A specification of who/what (in the domain) does each name refer to. Who is 'Alice'?

- A specification of who/what (in the domain) does each predicate apply to. Who is happy? Who likes whom? Who has property S.

When we give a model, we specify who/what is in the domain, who/what each name refers to, and who/what each predicate applies to.

This is an example of a model:

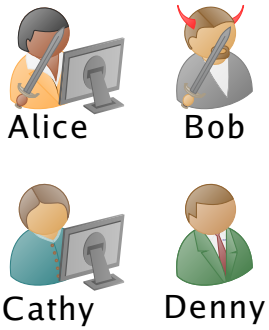
Dom: {Alice, Bob, Cathy, Denny}
$[\![H]\!]$: {Bob}
$[\![S]\!]$: {Alice, Bob}
$[\![C]\!]$: {Alice, Cathy}
$[\![a]\!]$: Alice
$[\![b]\!]$: Bob

This is the **intended model** when the symbolisation key is:

Domain: People in the below picture.
Hx: $\_\_\_\_\_x$ has horns
Sx: $\_\_\_\_\_x$ has a sword
Cx: $\_\_\_\_\_x$ is looking at a computer

a: Alice
b: Bob

with the associated picture:



The model specifies a collection of people (or more generally objects) as the domain, and says which predicates are true of each of them.

28.1 Introduction to sets

A collection of objects is called a **set**. The domain in a model is a set. The extensions of the predicates assigned by a model are also sets, they are subsets of the domain.

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

[1] One needs to spell this idea out with some care for otherwise Russell's paradox ensues.

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

$$\text{Tom} \in \{o \mid o \text{ is a cat or } o \text{ is a dog}\}.$$

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.,

$$\{o|o \text{ is an animal with a heart}\} = \{o|o \text{ is an animal with kidne}$$

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 $\varnothing$. There are many different ways to specify the empty set. For example,

$$\{o \mid o \neq o\} = \varnothing.$$

28.2 Truth of simple sentences

A model $\mathscr{M}$ provides

- specifies a non-empty set as the domain.
- 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 . . . to other predicates. [We'll talk about this in section 28.8.]

We denote the domain of a model as $\text{Dom}_{\mathscr{M}}$; its assignment of an object to a name by, for example, $[\![a]\!]_{\mathscr{M}}$; and its assignment of a subset to a predicate by, for example, $[\![H]\!]_{\mathscr{M}}$.

We typically simply drop the subscript to $\mathscr{M}$ when it is clear which model we are referring to.

So, here is a model, $\mathscr{M}$:

Dom: {Alice, Bob, Cathy, Denny}
$[\![H]\!]$: {Bob}
$[\![S]\!]$: {Alice, Bob}
$[\![C]\!]$: {Alice, Cathy}
$[\![a]\!]$: Alice
$[\![b]\!]$: Bob

Is 'Ca' true in this model? To answer this, we simply ask if the object the model assigns to 'a', in this case the person Alice, is a member of the set that the model assigns to 'C', in this case {Alice, Cathy}. That is, 'Ca' is true in $\mathcal{M}$ iff

$$[\![a]\!]_{\mathcal{M}} \in [\![C]\!]_{\mathcal{M}}.$$

On this model, that is the question of whether Alice $\in$ {Alice, Cathy}. Since she is in this set, it is true.

Is Ha true in this model? It is true in $\mathcal{M}$ iff

$$[\![a]\!]_{\mathcal{M}} \in [\![H]\!]_{\mathcal{M}}.$$

That is, iff Alice $\in$ {Bob} Since Alice $\notin$ {Bob}, Ha is false in $\mathcal{M}$.

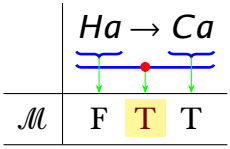
What about '$Ha \rightarrow Ca$'? Just like in TFL, the truth of this conditional is determined by the truth

of its constituent parts according to the truth rule for '$\rightarrow$'.

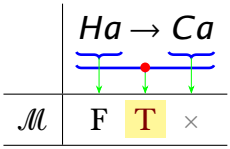
Recall the truth rule for $\rightarrow$:

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

So we can see:

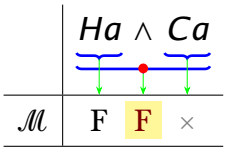


Observe that the truth value of *Ca* is irrelevant here: once you know that *Ha* is false, you can write T. This observation is important.



A closely related observation is that if Ha is false, then you don't need to consult the truth value of Ca to determine the truth value for $Ha \wedge Ca$.

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



28.3 Universal quantifier

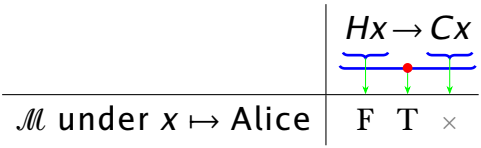
What about $\forall x(Hx \rightarrow Cx)$? This would symbolise

4. Everyone who has horns is looking at a computer.

$$\forall x\,(Hx \rightarrow Cx)$$

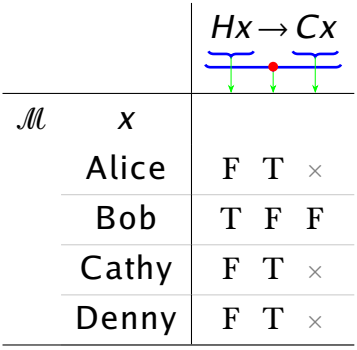
The idea we've applied so far is to determine the truth value of complex sentences by applying the rules for each connective to its constituent parts. The idea then is that we apply the rule for $\forall x$ to the value calculated for $(Hx \to Cx)$. But $Hx \to Cx$ doesn't get a fixed truth value on the model because the model doesn't specify who 'x' is, and indeed that's important. It's important that the reference of x can vary.

To evaluate $Hx \to Cx$, then we will need an additional component: a variable assignment. This will specify the referent of the variables. Here we only care about the variable x, so we just need to specify who x is, for example, $x \mapsto$ Alice. We can then evaluate $Hx \to Cx$ on this variable assignment.



Now, to evaluate $\forall x(Hx \to Cx)$ we need to consider not only this variable assignment, but all the possible assignments of objects to the

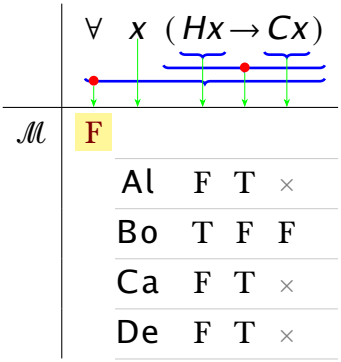
variable x. In this model, there are four:



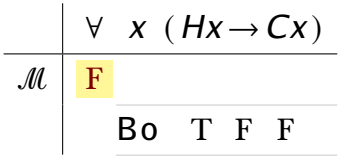
When $x \mapsto$Bob, makes $Hx \to Cx$ false. This means $\forall x(Hx \to Cx)$ is false in $\mathscr{M}$. There is some assignment of an object in the domain to the variable x on which $Hx \to Cx$ is false.

We need to consider different variable assignments at the point of evaluating a quantifier. In the quantified $\forall x$, we will list the options of assignments of objects to x underneath the x, and then write the truth value of $\forall x(Hx \to Cx)$

underneath the quantifier:



Since we only need to consider the variable assignment $x \mapsto$ Bo to show that it is false, we could just give:



But if the model makes $\forall xX$ true, then you can only see that by checking every instance, so need to include all the variable assignments.

	$\forall\ x\ X$
$\mathscr{M}$	T
	o_1 T
	⋮ T
	o_n T

In fact, domains may be infinite and then there isn't even in principle a presentation in terms of a table like this. But it can help you get the idea.

Here is the clause for $\forall$:

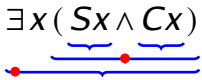
- $\forall xX$ is true in $\mathscr{M}$ iff for every $o \in \text{Dom}_{\mathscr{M}}$, X is true in $\mathscr{M}$ under the variable assignment $x \mapsto o$.

Once we have multiple variables interacting, we'll have to give a more general clause.

28.4 Existential quantifier

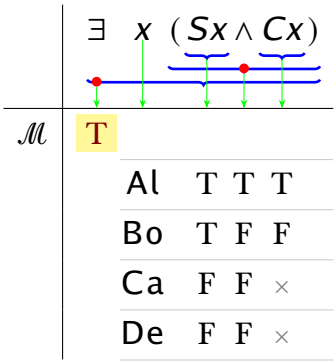
What about $\exists x(Sx \wedge Cx)$? This would symbolise

5. Some sword carrier is looking at a computer.

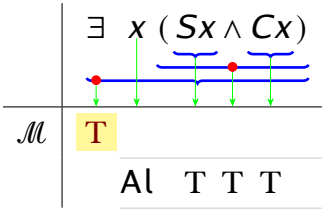


Just like for the universal quantifier, we need to consider possible assignments of objects from the domain to the variable x. Here is the clause for $\exists$:

- $\exists xX$ is true in $\mathscr{M}$ iff for at least one $o \in \mathrm{Dom}_{\mathscr{M}}$, X is true in $\mathscr{M}$ under the variable assignment $x \mapsto o$.



Since we only need to consider the variable assignment $x \mapsto$ Al we could just give:



But if the model makes $\exists xX$ false, then you can only see that by checking every instance, so need to include all the variable assignments.

	$\exists\ x\ X$
$\mathscr{M}$	F
	o_1 F
	⋮ F
	o_n F

Again, domains may be infinite and then there isn't even in principle a presentation in terms of a table like this. But it can help you get the idea.

28.5 When the main connective is something from TFL

Note that we can still then apply all the connectives. For example:

	$\neg\ \forall\ x\ (Hx \rightarrow Cx)$
$\mathscr{M}$	T F
	Bo T F F

or

	$Hb\ \wedge\ \neg \forall\ x\ (Hx \rightarrow Cx)$
$\mathscr{M}$	T T T F
	Bo T F F

We can also combine multiple sentences which both include variables.

If the sentence has the main connective $\wedge$, it truth value is simply defined using the rule for $\wedge$.

Consider

$$\forall x Hx \wedge \forall x(Hx \rightarrow Cx)$$

This would symbolise:

6. Not everyone has horns but everyone with horns is looking at a computer.

Is it true? Note that its main connective is $\wedge$:

$$\underbrace{\forall x Hx}\ \wedge\ \underbrace{\forall x(Hx \to Cx)}$$

The truth of the sentence overall just depends on the truth of $\forall x Hx$ and that of $\forall x(Hx \to Cx)$, and then they are combined using the truth rule for $\wedge$.

Actually, it doesn't matter what the truth value of $\forall x Hx$ since we already know that $\forall x(Hx \to Cx)$ is false:

	$\forall x\, Hx \wedge \forall\ x\ (Hx \to Cx)$
$\mathscr{M}$	$\times$ F F
	Bo T F F

Consider $\neg\forall x Hx \vee \neg\forall x\neg Hx$. We need to consider who x is when evaluating $\forall x Hx$ and, separately, when evaluating $\forall x\neg Hx$. There is no interaction between these, they are done separately then the results joined according to the

TFL connective rules.

	$\neg\forall x\, Hx \vee \neg\forall x\, \neg Hx$
$\mathscr{M}$	F T F F T Al F BoF T

The value we choose for x when evaluating $\forall x Hx$ has no relationship to the value we choose for x when evaluating $\forall x \neg Hx$.

28.6 Further Examples

Consider

- $\mathscr{M}_1$:

  Dom: $\{1, 2, 3, 4\}$
  $[\![C]\!]$: $\{1, 2\}$
  $[\![S]\!]$: $\{1, 2, 3\}$

- $\mathscr{M}_2$:

  Dom: $\{1, 2\}$
  $[\![C]\!]$: $\{1\}$
  $[\![S]\!]$: $\{2\}$

We are asked to evaluate this sentence on these two models

$$\neg\exists x Cx \vee \forall y(Cy \vee Sy)$$

Its truth depends on the truth of the constituent sentences under variable assignments. All the relevant cases are the '?' in the below table.

	$\neg\exists x\, Cx \vee \forall y\,(Cy \vee Sy)$
$\mathcal{M}_1$	?? ??
	1 ? 1 ? ? ?
	2 ? 2 ? ? ?
	3 ? 3 ? ? ?
	4 ? 4 ? ? ?
$\mathcal{M}_2$	?? ??
	1 ? 1 ? ? ?
	2 ? 2 ? ? ?

Although one doesn't need to fill out all the values, once you have found one variable assignment where the sentence is true, you can stop and put 'T' for $\exists xX$.

	$\neg\exists x\, Cx \vee \forall y\,(Sy \vee Cy)$
$\mathscr{M}_1$	F T F F 1 T 4 F F F
$\mathscr{M}_2$	F T T T 1 T 1 F T T 2 T T ×

We can summarise that, there is no need to present it as a table:

- On $\mathscr{M}_1$ it is false: $\exists xCx$ is true so $\neg\exists xCx$ is false. Also $\forall y(Sy \vee Cy)$ is false ($y \mapsto 4$). So the disjunction is false.
- On $\mathscr{M}_2$ it is true because $\forall y(Sy \vee Cy)$ is true: 1 is C and 2 is S.

We have used $\exists x$ and then $\forall y$. In fact, there would be no confusion if we gave

$$\neg\exists xCx \vee \forall x(Cx \vee Sx)$$

when we evaluate this, we would evaluate $\neg\exists xCx$ separately from $\forall x(Cx \vee Sx)$, and then combine their truth values according to the rule for $\vee$.

There is no confusion – who x refers in evaluating $\exists x Cx$ has no bearing on who x refers to when evaluating $\forall x(Cx \vee Sx)$.

	$\neg \exists x\, Cx \vee \forall x\,(Sx \vee Cx)$
$\mathscr{M}_1$	F T F F 1 T 4 F F F
$\mathscr{M}_2$	F T T T 1 T 1 F T T 2 T T ×

Since models can be infinite there may not even theoretically be a table like this to calculate the truth, but it should give you the idea of how to go about reasoning about it.

28.7 Relations

Suppose we have specified as our domain Dom the set of all animals, i.e.,

$$\text{Dom} = \{o \mid o \text{ is an animal}\}.$$

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

$$\{o \mid o \in \text{Dom and } o \text{ is a cat}\}.$$

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

{⟨Oslo, Norway⟩, ⟨Stockholm, Sweden⟩, ⟨Helsinki, Finland⟩,

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 chapter 26 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

$\langle$Pooh, Eeyore, The trampoline$\rangle \in \{\langle o_1, o_2, o_3\rangle | o_1$ bounces wi

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 $\langle o_1, o_2, \ldots, o_n \rangle$.

28.8 Models

In section 28.2, we introduced the idea of a model but we did not include relations. Now we can extend it to give the proper definition of a **model**. Sentences of FOL will be true or false **in a model**.

Definition

A model $\mathscr{M}$:

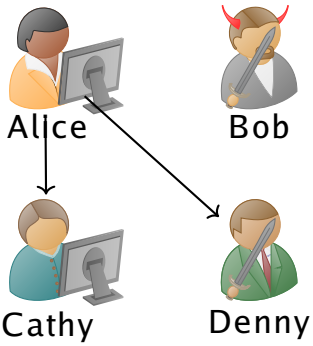
- specifies a non–empty set as the domain, Dom;
- specifies an interpretation of the FOL–names and FOL–predicates.
  - assigns exactly one object of the domain to every name (specifies the referent of the name);
  - assigns a subset of the domain to any one–placed predicate.
  - assigns a binary relation on the domain to every two–placed predicate.
  - assigns a ternary relation on the domain to every three–placed preciate.
  - ...

For a model, $\mathscr{M}$, we denote its assignment to η as $[\![\eta]\!]_{\mathscr{M}}$.

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.

Consider the following model:

Dom: {Alice, Bob, Cathy, Denny}
$[\![H]\!]$: {Bob}
$[\![S]\!]$: {Alice, Denny}
$[\![C]\!]$: {Alice, Cathy}
$[\![A]\!]$: {⟨Alice, Denny⟩, ⟨Alice, Cathy⟩}
$[\![a]\!]$: Alice
$[\![b]\!]$: Bob



Is '*Aab*' true? This is answered by looking at whether the sequence of the objects that the model assigns as denotations of '*a*' and '*b*' are in

the model's denotation of the relation '*A*', that is, whether

$$\langle [\![a]\!]_{\mathcal{M}}, [\![b]\!]_{\mathcal{M}} \rangle \in [\![A]\!]_{\mathcal{M}}.$$

In this model, this is the question of whether

$$\langle \text{Alice, Bob} \rangle \in \{\langle \text{Alice, Denny} \rangle, \langle \text{Alice, Cathy} \rangle\}?$$

Since
$\langle \text{Alice, Bob} \rangle \notin \{\langle \text{Alice, Denny} \rangle, \langle \text{Alice, Cathy} \rangle\}$,
'*Aab*' is false in $\mathcal{M}$.

Is $\exists x Aax$ true? This works just like what we have done so far, we have to consider variable assignments choosing each member of the domain to be the referent of the variable *x*. If there is at least one object in the domain we can choose for *x* which makes *Aax* true, then $\exists x Aax$ is true.

	∃ *x* *Aax*	
$\mathcal{M}$	?	
	Al	F
	Bo	F
	Ca	T
	De	T

So, $\exists x Aax$ is true. There is some choice of an object for x which makes Aax true: there someone for whom there is an arrow from Alice to them. For example, $x \mapsto$Cathy makes Aax true in $\mathscr{M}$; so $\exists x Aax$ is true in $\mathscr{M}$.

28.9 Interacting quantifiers

How about $\exists x(Cx \wedge \forall y(Sy \rightarrow Axy))$. This says that there is some computer watcher who has an arrow to all sword-bearers.

We need to consider the different possible variable assignments $x \mapsto o$ and evaluate whether $Cx \wedge \forall y(Sy \rightarrow Axy)$ is true under each of them. That's fine for Cx. And when Cx is false, that's enough to know that the conjunction is false. But we need to know then whether $\forall y(Sy \rightarrow Axy)$ is true under the variable assignment, say $x \mapsto$Alice.

	$\exists\ x\ (Cx \wedge \forall y(Sy \rightarrow Axy))$
$\mathscr{M}$	?
	Al T ??
	Bo F F×
	Ca T ??
	De F F×

To do this, we need to extend our variable assignment so it not only considers $x \mapsto$Alice but also considers a choice of who is y. A variable assignment in general assigns an object of the domain to every variable. So to evaluate $\forall y(Sy \rightarrow Axy)$ under the variable assignment $x \mapsto$Alice, we consider the variable assignments $x \mapsto$Alice; $y \mapsto o$, for each o in the domain. If $Sy \rightarrow Axy$ is true under each such, then $\forall y(Sy \rightarrow Axy)$ is true under $x \mapsto$Alice. But, $\forall y(Sy \rightarrow Axy)$ is false under $x \mapsto$Alice because under $x \mapsto$Alice and $y \mapsto$Bob, $Sy \rightarrow Axy$ is false.

	$\exists\ x\ (Cx \wedge \forall\ y\ (Sy \rightarrow Axy))$
$\mathscr{M}$	F
	Al T F F
	Al F T ×
	Bo T F F
	Ca F T ×
	De T T T
	Bo F F ×
	Ca T F F
	Al F T ×
	Bo T F F
	Ca F T ×
	De T F F
	De F F ×

28.10 Examples

Consider

- $\mathscr{M}_1$:

  Dom: $\{1, 2, 3, 4\}$
  $[\![C]\!]$: $\{1, 2\}$
  $[\![S]\!]$: $\{1, 2, 3\}$

$[\![R]\!]$: $\{\langle 1,3\rangle, \langle 2,4\rangle, \langle 3,1\rangle\}$

▹ $\mathscr{M}_2$:

Dom: $\{1,2\}$
$[\![C]\!]$: $\{1\}$
$[\![S]\!]$: $\{2\}$
$[\![R]\!]$: $\{\langle 1,1\rangle\}$

What about

$$\neg\forall x(Cx \to \neg\exists y(Cy \wedge Rxy)) \wedge \exists x\neg Cx$$

To calculate this, we need to reason through filling out enough of the '?' in the below table to

work out the answer.

	$\neg\forall x(Cx \rightarrow \neg\exists y(Cy \wedge Rxy)) \wedge \exists x \neg Cx$
$\mathcal{M}_1$	? ? ? ?
	1 ? ? ? ? 1 ? ? ? 1 ? ?
	2 ? ? ? 2 ? ?
	3 ? ? ? 3 ? ?
	4 ? ? ? 4 ? ?
	2 ? ? ? ?
	1 ? ? ?
	2 ? ? ?
	3 ? ? ?
	4 ? ? ?
	3 ? ? ? ?
	1 ? ? ?
	2 ? ? ?
	3 ? ? ?
	4 ? ? ?
	4 ? ? ? ?
	1 ? ? ?
	2 ? ? ?
	3 ? ? ?
	4 ? ? ?
$\mathcal{M}_2$	? ? ? ?
	1 ? ? ? ? 1 ? ? ? 1 ? ?
	2 ? ? ? 2 ? ?
	2 ? ? ? ?

This is a lot to fill out. But the guidance is that you take shortcuts: don't carry on if you already have enough information to determine the truth value. Or only consider those variable assignments which seem relevant. Once you've found one choice of x on which X is true, then $\exists xX$ is true; it doesn't matter what the other values are.

	$\neg\forall x(Cx\rightarrow\neg\exists y(Cy\wedge Rxy))\wedge\exists x\neg Cx$
$\mathscr{M}_1$	F T F T 1 T T T F 1 F T 1 T F F 2 F T 2 T F F 3 T F 3 F F T 4 T F 4 F F F 2 T T T F 1 T F F 2 T F F 3 F F F 4 F F T 3 F T × 4 F T ×
$\mathscr{M}_2$	T F T T 1 T F F T 1 F T 1 T T T 2 T F 2 F T ×

Chapter 29

Truth in a model, more carefully

Sentences of FOL are true or false in a model.

A model $\mathscr{M}$ 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.
  - ...

Definition

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.

For example $\mathcal{M}$ might be:

Dom: $\{1, 2, 3\}$
$[\![H]\!]$: $\{1\}$
$[\![S]\!]$: $\{1, 2\}$
$[\![C]\!]$: $\{2, 3\}$
$[\![A]\!]$: $\{\langle 1, 2\rangle, \langle 2, 2\rangle\}$
$[\![a]\!]$: 2
$[\![b]\!]$: 3

Notes:

- Domain must be non−empty.
- Domains may be infinite, e.g., Dom as $\mathbb{N}$, the natural numbers $\{0, 1, 2, \ldots\}$
- Predicates and relations may be empty, e.g., $[\![P]\!] = \varnothing$

- Multiple names may name the same object.

Since, among other things, models 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 in 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:

Definition

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 numbers and the predicate *E* picks out the even numbers:

$$\text{Dom} = \{o \mid o \text{ is a natural number}\}$$
$$[\![E]\!] = \{o \mid o \text{ is an even number}\}$$

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 chapter 27 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.

29.1 Atomic formulas

We already described the clause for atomic sentences, for example:

▹ Ca is true in $\mathcal{M}$ iff

$$[\![a]\!]_{\mathcal{M}} \in [\![C]\!]_{\mathcal{M}}.$$

▹ Lab is true in $\mathcal{M}$ iff

$$\langle [\![a]\!]_{\mathcal{M}}, [\![b]\!]_{\mathcal{M}} \rangle \in [\![L]\!]_{\mathcal{M}}.$$

But we also need to describe a clause for atomic formulas which contain variables.

- Cx is true in $\mathcal{M}$ under variable assignment α iff

$$\alpha(x) \in [\![C]\!]_{\mathcal{M}}$$

- Lxy is true in a model $\mathcal{M}$ under a variable assignment α iff

$$\langle \alpha(x), \alpha(y) \rangle \in [\![L]\!]_{\mathcal{M}}$$

- Lxb is true in a model $\mathcal{M}$ under a variable assignment α iff

$$\langle \alpha(x), [\![b]\!]_{\mathcal{M}} \rangle \in [\![L]\!]_{\mathcal{M}}$$

To write these out more precisely, we unify the terminology:

Definition

A **term** is either a name or a variable.

- If t is a name, then $[\![t]\!]^{\alpha}_{\mathcal{M}}$ is just $[\![t]\!]_{\mathcal{M}}$, the object of the domain assigned to the FOL-name t by $\mathcal{M}$.
- If t is a variable, then $[\![t]\!]^{\alpha}_{\mathcal{M}}$ 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.

Definition

For R an n–place predicate, t_1, t_2 ... t_n terms $Rt_1t_2 \ldots t_n$ is true in a model $\mathcal{M}$ under a variable assignment α **iff**

$$\langle [\![t_1]\!]^{\alpha}_{\mathcal{M}}, \ldots, [\![t_n]\!]^{\alpha}_{\mathcal{M}} \rangle \in [\![R]\!]_{\mathcal{M}}$$

29.2 Sentential connectives

We saw in chapter 27 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:

Definition

$\neg X$ is true in a model $\mathcal{M}$ under a variable assignment α **iff**
X is false in $\mathcal{M}$ under α.

$(X \wedge Y)$ is true in a model $\mathcal{M}$ under a variable assignment α **iff**
both X and Y are true in $\mathcal{M}$ under α.

$(X \vee Y)$ is true in a model $\mathcal{M}$ under a variable assignment α **iff**
either X is true or Y is true in $\mathcal{M}$ under α.

$(X \rightarrow Y)$ is true in a model $\mathcal{M}$ under a variable assignment α **iff**
either X is false or Y is true in $\mathcal{M}$ under α.

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

Dom: {*o* |
o is a person born before before 2000CE})

$[\![a]\!]$: **Aristotle**
$[\![b]\!]$: **Beyoncé**
$[\![P]\!]$: $\{o \mid o$ is a philosopher$\}$
$[\![R]\!]$: $\{\langle o_1, o_2\rangle \mid o_1$ was born before $o_2\}$

we have:

- '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

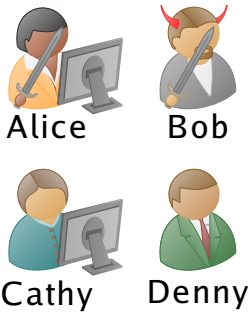
- 'Px' is true under this variable assignment: David Hume was a philosopher
- 'Bxa' is false under this variable assignment: David Hume was born after Aristotle

- '$Px \rightarrow Bxa$' is false under this variable assignment: 'Px' is true and 'Bxa' is false, so by our rule for $\rightarrow$, '$Px \rightarrow Bxa$' is false.

29.3 Quantifiers

The exciting innovation in FOL, though, is the use of **quantifiers**.

Consider the following model:



Is '$\exists xSx$' 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.

Definition

Let α be a variable assignment and v a variable.

$\alpha[v \mapsto o]$ is a variable assignment which assigns to the variables other than v the same objects as α does and it assigns o to v.

That is, $\alpha[v \mapsto o](v') = \begin{cases} \alpha(v') & \text{if } v' \neq v \\ o & \text{if } v' = v \end{cases}$

$\alpha[v \mapsto o]$ is called an **v-variant** of α.

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**	...
$\alpha[y \mapsto$ **Bob**$]$:	**Alice**	**Bob**	**Cathy**	**Bob**	...

Variants of variable assignments will be important for stating the truth conditions of quantified formulas. We will define

> '$\exists vX$' true in a model under the variable assignment α **iff** there exists an object *o* of the domain, i.e. $o \in$ Dom, such that '*X*' is true in the model under the variable assignment $\alpha[v \mapsto o]$.

To show that '$\exists xSx$' true in $\mathcal{M}$ under α we need to find an object in the domain, say, **Alice** such that '*Sx*' is true under the variable assignment $\alpha[x \mapsto$ **Alice**$]$. Under $\alpha[x \mapsto$ **Alice**$]$, *Sx* is in fact true: it assigns **Alice** to *x* and Alice does have a sword, i.e., **Alice** $\in [\![S]\!]$. This guaranteed the truth of '$\exists xSx$': There is a choice of an object in our domain for *x* under which *Sx* is true.

In this case, the other information in α is irrelevant as only the variable *x* is relevant for determining the truth of *Sx*. Whether $\exists xSx$ is true in $\mathcal{M}$ under α is independent of α and it can simply be said to be true or false in the model.

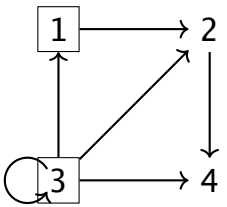
What about the universal quantifier?

> '$\forall vX$' true in a model under the variable assignment α **iff** for every object *o* of the domain, i.e. $o \in$ Dom, '*X*' is true in the model under the variable assignment $\alpha[v \mapsto o]$.

Since **Denny** $\notin [\![S]\!]$ = {**Alice, Bob**} '$\forall xSx$' to comes out false in this model: '*Sx*' is not true in the model under the variable assignment $\alpha[x \mapsto$ **Denny**$]$, or simply $x \mapsto$**Denny**, as the remaining information in the variable assignment α is irrelevant.

We have to tread more carefully once we start having multiple quantifiers. Let's walk through some cases.

Consider a new model:



Dom: $\{1, 2, 3, 4\}$
$[\![R]\!]$: $\{\langle o_1, o_2\rangle \mid$
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 objects. 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 $\alpha[x \mapsto n]$.

    ▸ We choose n to be the number 3 (with some foresight)

2. '$\forall y Rxy$' is true in the model under the variable assignment $\alpha[x \mapsto 3]$ iff Rxy is true in the model under the variable assignment $\alpha[x \mapsto 3][y \mapsto m]$ for all $m \in \{1, 2, 3, 4\}$.

- $\alpha[x \mapsto 3][y \mapsto m]$ 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 $\alpha[x \mapsto 3][y \mapsto m]$ 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 $\alpha[x \mapsto 3][y \mapsto 1]$, the variable assignment $\alpha[x \mapsto 3][y \mapsto 2]$, the variable assignment $\alpha[x \mapsto 3][y \mapsto 3]$ and also $\alpha[x \mapsto 3][y \mapsto 4]$.

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 $\alpha[x \mapsto o]$. 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., $\alpha[x \mapsto 3]$? 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 $\alpha[x \mapsto 3][y \mapsto 3]$. 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 $\alpha[x \mapsto 3][y \mapsto 3]$. And so '$\exists y(Rxy \wedge Ryx)$' is true under the variable assignment $\alpha[x \mapsto 3]$. And so $\exists x \exists y(Rxy \wedge Ryx)$ is true in this model.

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 underall 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 $\forall yRxy$ is false on the variable assignment $\alpha[x \mapsto 1]$. And so $\forall x \forall yRxy$ is false in the model under the variable assignment α. However, α itself played no role in the evaluation of $\forall x \forall yRxy$. 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 $\forall x \forall yRxy$ is not only true under the variable assignment α in this model but true under all variable assignments. The particular choice of an variable assignment is only relevant if we are evaluating the truth of formulas with free variables. $\forall x \forall yRxy$ 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:

Definition

$\forall v X$ is true in the model $\mathcal{M}$ under a variable assignment α
iff for every $o \in$ Dom, X is true in $\mathcal{M}$ under the variable assignment $\alpha[v \mapsto o]$.

$\exists v X$ is true in the model $\mathcal{M}$ under a variable assignment α
iff for some $o \in$ Dom, X is true in $\mathcal{M}$ under the variable assignment $\alpha[v \mapsto o]$.

We have now provided the full inductive definition of truth in a model under a variable assignment. This allows us to work up the formation tree of any sentence of FOL to detemrine its truth value. The special feature of sentences as opposed to formulas more generally was that there are no free variables. This means that information in a variable assignment α is irrelevant for the truth of the sentence and we can simply say that it is true or false on the model.

Practice Exercises (Chapter 29)

I. Consider the following model:

Dom: $\{1, 2\}$
$[\![A]\!]$: $\{1, 2\}$
$[\![B]\!]$: $\{2\}$
$[\![N]\!]$: $\varnothing$
$[\![c]\!]$: 1

Determine whether each of the following sentences is true or false in that model:

a) Bc (∗)
b) $Ac \leftrightarrow \neg Nc$ (∗)
c) $Nc \rightarrow (Ac \vee Bc)$ (∗)
d) $\forall x\, Ax$ (∗)
e) $\forall x \neg Bx$ (∗)
f) $\exists x(Ax \wedge Bx)$ (∗)
g) $\exists x(Ax \rightarrow Nx)$ (∗)
h) $\forall x(Nx \vee \neg Nx)$ (∗)
i) $\exists x\, Bx \rightarrow \forall x\, Ax$ (∗)

II. Consider the following model:

Dom: $\{$**Lemmy**, **Courtney**, **Eddy**$\}$
$[\![G]\!]$: $\{$**Lemmy**, **Courtney**, **Eddy**$\}$
$[\![H]\!]$: $\{$**Courtney**$\}$

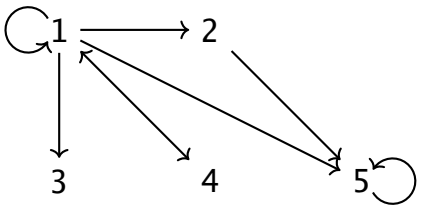
$[\![M]\!]$: {**Lemmy**, **Eddy**}
$[\![c]\!]$: **Courtney**
$[\![e]\!]$: **Eddy**

Determine whether each of the following sentences is true or false in that model:

a) Hc (∗)
b) He (∗)
c) $Mc \vee Me$ (∗)
d) $Gc \vee \neg Gc$ (∗)
e) $Mc \rightarrow Gc$ (∗)
f) $\exists x\, Hx$ (∗)
g) $\forall x\, Hx$ (∗)
h) $\exists x \neg Mx$ (∗)
i) $\exists x(Hx \wedge Gx)$ (∗)
j) $\exists x(Mx \wedge Gx)$ (∗)
k) $\forall x(Hx \vee Mx)$ (∗)
l) $\exists x\, Hx \wedge \exists x\, Mx$ (∗)
m) $\forall x(Hx \leftrightarrow \neg Mx)$ (∗)
n) $\exists x\, Gx \wedge \exists x \neg Gx$ (∗)
o) $\forall x \exists y(Gx \wedge Hy)$ (∗)

III. 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:

a) $\exists x\, Rxx$ (∗)
b) $\forall x\, Rxx$ (∗)
c) $\exists x \forall y\, Rxy$ (∗)
d) $\exists x \forall y\, Ryx$ (∗)
e) $\forall x \forall y \forall z((Rxy \wedge Ryz) \rightarrow Rxz)$ (∗)
f) $\forall x \forall y \forall z((Rxy \wedge Rxz) \rightarrow Ryz)$ (∗)
g) $\exists x \forall y \neg Rxy$ (∗)
h) $\forall x(\exists y\, Rxy \rightarrow \exists y\, Ryx)$ (∗)
i) $\exists x \exists y(\neg x = y \wedge Rxy \wedge Ryx)$ (∗)

j) $\exists x \forall y (Rxy \leftrightarrow x = y)$ $(*)$

k) $\exists x \forall y (Ryx \leftrightarrow x = y)$ $(*)$

l) $\exists x \exists y (\neg x = y \wedge Rxy \wedge \forall z (Rz, x \leftrightarrow y = z))$ $(*)$

Chapter 30

Logical notions

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.

30.1 Validity

Definition

A FOL–argument $X_1, \ldots, X_n \therefore Y$ is **valid** if and only if there is no model $\mathscr{M}$ such that all the premises, $X_1, \ldots, X_n$ are true in $\mathscr{M}$ but Y is false in $\mathscr{M}$.

Consider

$$\exists x(Px \wedge Hx) \ \therefore \ \forall x(Px \rightarrow Hx)$$

This argument is invalid.

We might be interested in this as a symbolisation of someone's argument:

– Some philosopher is happy
$\therefore$ Therefore: Every philosopher is happy

To show that the argument is invalid, we will find a model where the premise(s) are true and the conclusion is false.

Consider, for example:

Dom: $\{1, 2, 3\}$

$[\![P]\!]$: $\{1, 2\}$
$[\![H]\!]$: $\{1\}$

$\exists x(Px \wedge Hx)$ is true, after all under $x \mapsto 1$, $Px \wedge Hx$ is true. However, $\forall x(Px \to Hx)$ is false, after all, for $x \mapsto 2$, $Px \to Hx$ is false.

Here is another argument which is invalid:

$$\forall x(Px \to \exists y(By \wedge Lxy)) \ \therefore \ \exists y(By \wedge \forall x(Px \to Lxy))$$

Under an appropriate symbolisation key, this might be your symbolisation of

– Everyone likes some book
$\therefore$ Therefore: There is some book which everyone likes.

Consider

Dom: $\{1, 2, 3, 4\}$
$[\![P]\!]$: $\{1, 2\}$
$[\![B]\!]$: $\{3, 4\}$
$[\![L]\!]$: $\{\langle 1, 3\rangle, \langle 2, 4\rangle\}$

Every person likes some book (1 bears L to 3 and 2 bears L to 4), even though there is no particular book which everyone likes (for 3, 2 does not like it; for 4, 1 does not like it). This is a model on which the premise is true but the conclusion is false.

What about the converse, that is, the argument:

– There is some book which everyone likes.
∴ Therefore: Everyone likes some book

Symbolised as:

$$\exists y(By \wedge \forall x(Px \rightarrow Lxy)) \ \therefore \ \forall x(Px \rightarrow \exists y(By \wedge Lxy))$$

This argument is valid, but to show this, one needs to reason about every possible model. We need to demonstrate that there is no model where the premise is true and the conclusion is false. That on any model where the premise is true, the conclusion is also true.

In TFL, we simply listed all the valuations and checked whether there was such a counterexample valuation. This is not even in

principle possible in FOL as there are infinitely many models. Instead one needs to undertake semantic reasoning. Alternatively, one can give a proof in the natural deduction system which we will give in part VIII.

To show an argument is invalid, one needs to find a counterexample model: a model where the premises are all true and the conclusion is false.
To show an argument is valid, one needs to reason about all models. We will instead give a proof.

30.2 Other notions

The other logical notions also have corresponding definitions in FOL:

Definition

Y is a **logical consequence** of $X_1, \ldots, X_n$ if and only if $X_1, \ldots, X_n \therefore Y$ is valid. That is, iff there is no model where $X_1, \ldots, X_n$ are true and Y is false.

Definition

X and Y are **logically equivalent** iff they are true on exactly the same valuations.

An important example of logical equivalence is the close relationship between the two quantifiers:

- $\neg\exists x X$ and $\forall x \neg X$ are logically equivalent.
  - They are true in a model $\mathcal{M}$ under variable assignment α iff for every $o \in \text{Dom}_{\mathcal{M}}$, X is false in $\mathcal{M}$ under $\alpha[x \mapsto o]$.
- $\neg\forall x X$ and $\exists x \neg X$ are logically equivalent.
  - They are true in a model $\mathcal{M}$ under variable assignment α iff there is some $o \in \text{Dom}_{\mathcal{M}}$ such that X is false in $\mathcal{M}$ under $\alpha[x \mapsto o]$.

Definition

A sentence X of TFL is a **logical truth** if and only if it is true in every model.

"Tautology" is a name which was specific to TFL. The notion of being a "logical truth" is broader and also applies in FOL

Definition

X is a **contradiction** iff it is false in every model.

Definition

A collection of FOL–sentences $X_1, \ldots, X_n$ (with possibly $n = 1$) is **consistent** iff there is a model relative to which $X_1, \ldots, X_n$ are true. Otherwise the collection is **inconsistent**.

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 chapter 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.

Practice Exercises (Chapter 30)

I. Show that each of the following is neither a logical truth nor a contradiction:

a) $Da \wedge Db$ (∗)

b) $\exists x\, Tx, h$ (∗)

c) $Pm \wedge \neg \forall x\, Px$ (∗)

d) $\forall z\, Jz \leftrightarrow \exists y\, Jy$

e) $\forall x(Wx, m, n \vee \exists y Lx, y)$

f) $\exists x(Gx \to \forall y\, My)$

g) $\exists x(x = h \wedge x = i)$

II. Show that the following pairs of sentences are not logically equivalent.

a) Ja, Ka

(*)

b) $\exists x\, Jx$, Jm

(*)

c) $\forall x\, Rx, x$, $\exists x\, Rx, x$

(*)

d) $\exists x\, Px \to Qc$, $\exists x(Px \to Qc)$

(*)

e) $\forall x(Px \to \neg Qx)$, $\exists x(Px \wedge \neg Qx)$

(*)

f) $\exists x(Px \wedge Qx)$, $\exists x(Px \to Qx)$

(*)

g) $\forall x(Px \to Qx)$, $\forall x(Px \wedge Qx)$

(*)

h) $\forall x \exists y\, Rx, y$, $\exists x \forall y\, Rx, y$

(*)

i) $\forall x \exists y\, Rx, y$, $\forall x \exists y\, Ry, x$

(*)

III. Show that the following sentences are (jointly) consistent:

a) $Ma, \neg Na, Pa, \neg Qa$
b) $Le, e, Le, g, \neg Lg, e, \neg Lg, g$
c) $\neg(Ma \wedge \exists x\, Ax), Ma \vee Fa, \forall x(Fx \rightarrow Ax)$
d) $Ma \vee Mb, Ma \rightarrow \forall x \neg Mx$
e) $\forall y\, Gy, \forall x(Gx \rightarrow Hx), \exists y \neg Iy$
f) $\exists x(Bx \vee Ax), \forall x \neg Cx, \forall x\big((Ax \wedge Bx) \rightarrow Cx\big)$
g) $\exists x\, Xx, \exists x\, Yx, \forall x(Xx \leftrightarrow \neg Yx)$
h) $\forall x(Px \vee Qx), \exists x \neg(Qx \wedge Px)$
i) $\exists z(Nz \wedge Oz, z), \forall x \forall y(Ox, y \rightarrow Oy, x)$
j) $\neg \exists x \forall y\, Rx, y, \forall x \exists y\, Rx, y$
k) $\neg Ra, a, \forall x(x = a \vee Rx, a)$

(∗)

l) $\forall x \forall y \forall z((x = y \vee y = z) \vee x = z), \exists x \exists y\, \neg x = y$

(∗)

m) $\exists x \exists y((Zx \wedge Zy) \wedge x = y), \neg Zd,\ d = e$

IV. Show that the following arguments are invalid:

a) $\forall x(Ax \rightarrow Bx) \therefore \exists x\, Bx$
b) $\forall x(Rx \rightarrow Dx), \forall x(Rx \rightarrow Fx) \therefore \exists x(Dx \wedge Fx)$
c) $\exists x(Px \rightarrow Qx) \therefore \exists x\, Px$

d) $Na \wedge Nb \wedge Nc \therefore \forall x\, Nx$

e) $Rd, e, \exists x\, Rxd \therefore Re, d$

f) $\exists x(Ex \wedge Fx), \exists x\, Fx \rightarrow \exists x\, Gx \therefore \exists x(Ex \wedge Gx)$

g) $\forall x\, Ox, c, \forall x\, Oc, x \therefore \forall x\, Ox, x$

h) $\exists x(Jx \wedge Kx), \exists x \neg Kx, \exists x \neg Jx \therefore \exists x(\neg Jx \wedge \neg Kx)$

i) $La, b \rightarrow \forall x\, Lx, b, \exists x\, Lx, b \therefore Lb, b$

j) $\forall x(Dx \rightarrow \exists y\, Ty, x) \therefore \exists y \exists z\, \neg y = z$

Part VII

First-order logic symboli-sation

Chapter 31

Common quantifier phrases

We have already seem the key idea of FOL symbolisation in chapter 26. Our strategy for symbolising sentences extends the strategy that we used for TFL in section 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**
For all x, ...	$\forall x \ldots$
There exists some x 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.

What we will now do is discuss further symbolisation tricks to add to our symbolisation guide.

31.1 Restricted quantifiers

We have already discussed two of them:

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

$$\forall x(Fx \rightarrow Gx).$$

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

$$\exists x(Fx \wedge Gx).$$

So we symbolise

1. Every dog is energetic.

as

$$\forall x(Dx \to Ex)$$

and

2. Some dog is energetic.

as

$$\exists x(Dx \wedge Ex)$$

Notice that we needed to use a conditional with the universal quantifier, but we used a conjunction with the existential quantifier.

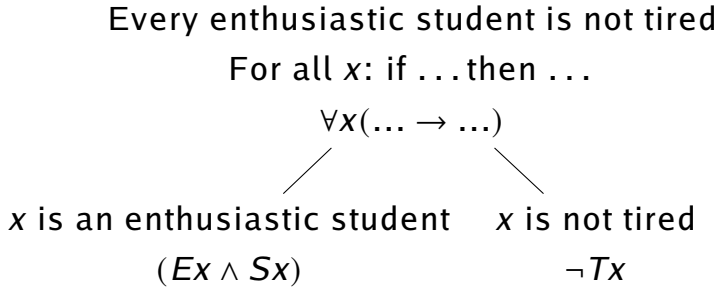
Suppose we had instead written $\exists x(Dx \to Ex)$. So long as there is a non-dog, o, in the model, then $Dx \to Ex$ is true under $x \mapsto o$. After all, a conditional with a false antecedent is always true. So it would be far too easy to make $\exists x(Dx \to Ex)$ true: it would be true in any model which contains a non-dog.

Similarly, suppose we had instead written $\forall x(Dx \wedge Ex)$. So long as there is a non-dog, o, in the model, then $Dx \wedge Ex$ is false under $x \mapsto o$. So it would be far too hard to make $\exists x(Dx \wedge Ex)$ true. It

would be false in any model which contains non–dogs.

Recall that this is a general strategy which can also be applied with more complex sentences

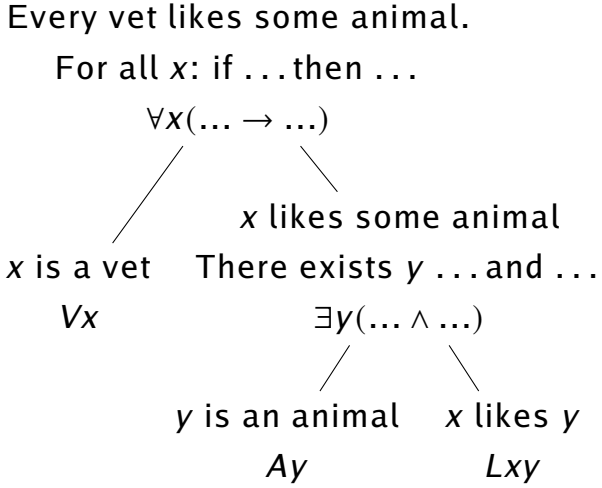
3. Every enthusiastic student is not tired.



$$\forall x((Ex \wedge Sx) \rightarrow \neg Tx)$$

or which can be iterated

4. Every vet likes some animal.



So we give:

$$\forall x(Vx \rightarrow \exists y(Ay \wedge Lxy))$$

These are just new tools which we add to our symbolisation toolbox.

Now, we will provide some additional symbolisation guidance.

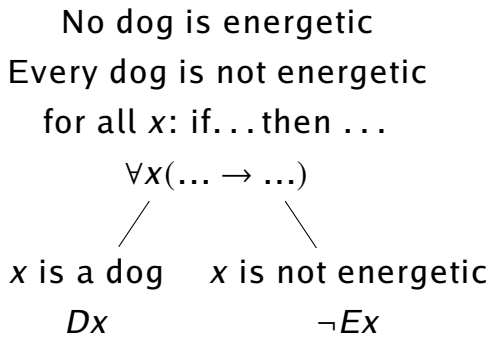
31.2 None

5. No dog is energetic

We have to paraphrase this sentence using the operators that we have. There are two different ways to do this:

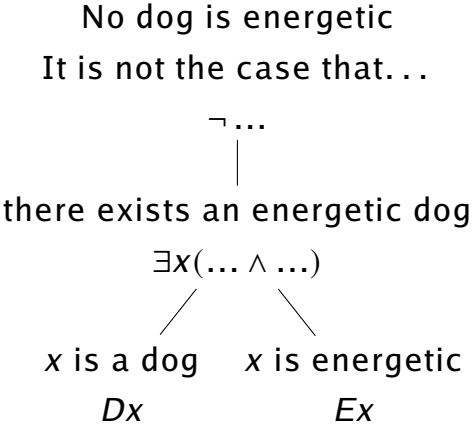
6. Every dog is not energetic.
7. It is not the case that there exists an energetic dog.

These different paraphrasing will give different symbolisations



Giving

$$\forall x(Dx \rightarrow \neg Ex)$$



Giving

$$\neg\exists x(Dx \wedge Ex)$$

These two sentences are logically equivalent and both appropriate symbolisations.

We give an example of semantic reasoning to show that they are logically equivalent.

Observe that in any model, $\mathcal{M}$, and any variable assignment, α,

$Dx \wedge Ex$ is true in $\mathcal{M}$ under α
iff both Dx and Ex are true in $\mathcal{M}$ under α
iff Dx is true in $\mathcal{M}$ under α and Ex is false in $\mathcal{M}$ under α
iff $Dx \rightarrow \neg Ex$ is false in $\mathcal{M}$ under α.

Then we can reason as follows:

$\neg\exists x(Dx \wedge Ex)$ is true in $\mathcal{M}$
iff there is no $o \in \text{Dom}_{\mathcal{M}}$
such that $Dx \wedge Ex$ is true in $\mathcal{M}$ under $x \mapsto o$
iff there is no $o \in \text{Dom}_{\mathcal{M}}$
such that $Dx \rightarrow \neg Ex$ is false in $\mathcal{M}$ under $x \mapsto o$
iff for every $o \in \text{Dom}_{\mathcal{M}}$, $Dx \rightarrow \neg Ex$ is true in $\mathcal{M}$ under x
iff $\forall x(Dx \rightarrow \neg Ex)$ is true in $\mathcal{M}$.

If a sentence can be paraphrased in English as

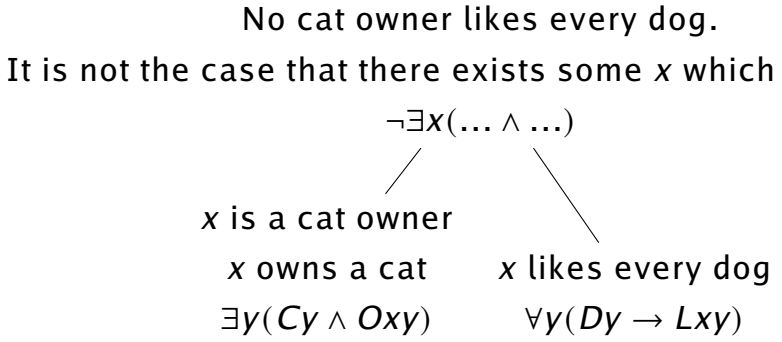
'no Fs are Gs',

it can be symbolised as

$\neg\exists x(Fx \wedge Gx)$, or
$\forall x(Fx \rightarrow \neg Gx)$.

This can also be embedded in more complex sentences

8. No cat owner likes every dog.



offering:

$$\neg\exists x(\exists y(Cy \wedge Oxy) \wedge \forall y(Dy \rightarrow Lxy))$$

Alternatively:

No cat owner likes every dog.
For all x: if . . . then not . . .
$\forall x(\ldots \rightarrow \neg \ldots)$

x is a cat owner	x likes every dog
$\exists y(Cy \wedge Oxy)$	$\forall y(Dy \rightarrow Lxy)$

offering:

$$\forall x(\exists y(Cy \wedge Oxy) \rightarrow \neg\forall y(Dy \rightarrow Lxy))$$

31.3 Not all

There is one final structure to complete our selection: we have $\forall$, $\exists$, $\neg\exists = \forall\neg$; the final structure to see is $\neg\forall = \exists\neg$.

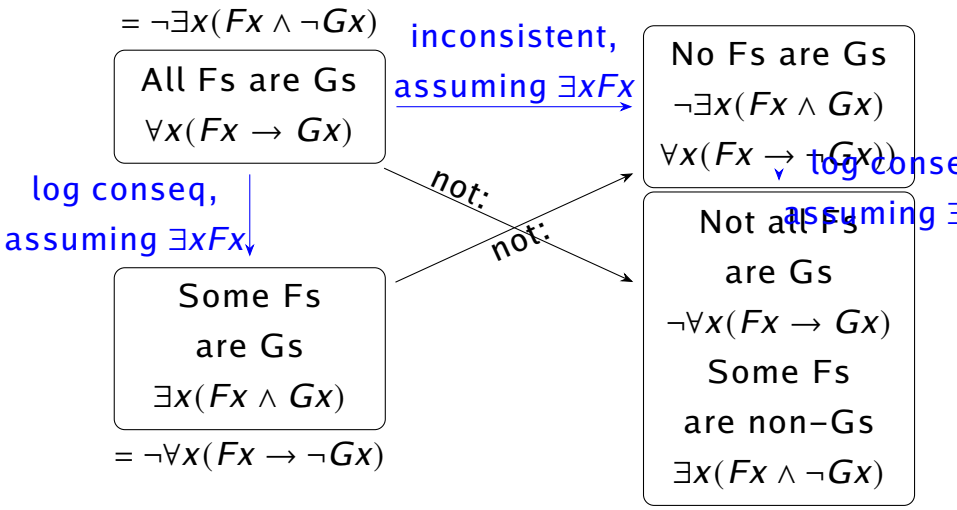
9. Not all dogs are enthusiastic.

$$\neg\forall x(Dx \to Ex)$$

In fact, this is logically equivalent to $\exists x(Dx \wedge \neg Ex)$:

10. Some dog is not enthusiastic.

We have resulted in four things we might say about whether dogs are enthusiastic, or, more generally, whether Fs are Gs.



31.4 Only

11. Only dogs are enthusiastic

We will paraphrase this as:

12. Everything which is enthusiastic is a dog.
13. Nothing which is enthusiastic is non-dog.

We can offer symbolisation options:

$$\forall x(Ex \rightarrow Dx)$$
$$\neg\exists x(Ex \wedge \neg Dx)$$

Compare this to:

14. All dogs are enthusiastic

is symbolised as

$$\forall x(Dx \rightarrow Ex)$$
$$\neg\exists x(Dx \wedge \neg Ex)$$

It can only be made false due to a dog who is not enthusiastic

15. Only dogs are enthusiastic

can only be made false due to a non-dog who is enthusiastic

If a sentence can be paraphrased in English as

'only Fs are Gs',

it can be symbolised as

$\neg\exists x(Gx \wedge \neg Fx)$, or
$\forall x(Gx \rightarrow Fx)$

Practice Exercises (Chapter 31)

I. Using the following symbolization key:

Dx: _____$_x$ is a dog
Ax: _____$_x$ is an animal
Sx: _____$_x$ likes samurai movies
Lxy: _____$_x$ is larger than _____$_y$
r: Rave
h: Shane
d: Daisy

symbolize the following sentences in FOL:

a) Rave is a dog who likes samurai movies.

(*)

b) Rave, Shane, and Daisy are all dogs.

(*)

c) Shane is larger than Rave, and Daisy is larger than Shane.

(*)

d) All dogs like samurai movies.

(*)

e) Only dogs like samurai movies.

(*)

f) There is a dog that is larger than Shane.

(*)

g) If there is a dog larger than Daisy, then there is a dog larger than Shane.

(*)

h) No animal that likes samurai movies is larger than Shane.

(*)

i) No dog is larger than Daisy.

(*)

j) Any animal that dislikes samurai movies is larger than Rave.

(∗)

k) There is an animal that is between Rave and Shane in size.

(∗)

l) There is no dog that is between Rave and Shane in size.

(∗)

m) No dog is larger than itself.

(∗)

n) Every dog is larger than some dog.

(∗)

o) There is an animal that is smaller than every dog.

(∗)

p) If there is an animal that is larger than any dog, then that animal does not like samurai movies.

(∗)

II. Using the symbolization key given, translate each English-language sentence into FOL.

Cx: $\_\_\_\_\__x$ is a candy
Hx: $\_\_\_\_\__x$ has chocolate in it.
Mx: $\_\_\_\_\__x$ has marzipan in it.
Sx: $\_\_\_\_\__x$ has sugar in it.
Tx: Boris has tried $\_\_\_\_\__x$.
Bxy: $\_\_\_\_\__x$ is better than $\_\_\_\_\__y$.

a) Boris has never tried any candy.
b) Marzipan is always made with sugar.
c) Some candy is sugar-free.
d) The very best candy is chocolate.
e) No candy is better than itself.
f) Boris has never tried sugar-free chocolate.
g) Boris has tried marzipan and chocolate, but never together.
h) Any candy with chocolate is better than any candy without it.
i) Any candy with chocolate and marzipan is better than any candy that lacks both.

Chapter 32

Ambiguity

32.1 Scope ambiguity

In chapter 17 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 '_____$_x$ glitters' and Gx is '_____$_x$ is gold', we would symbolize it as:

▹ $\forall x(Fx \rightarrow \neg Gx)$,

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:

▹ $\neg\forall x(Fx \rightarrow Gx)$

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 '¬' is in the scope of the '∀' (in the first symbolization) or '∀' is in the scope of '¬' (in the second).

Of course we can alternatively symbolize the two readings using existential quantifiers as well:

- $\neg\exists x(Fx \wedge Gx)$
- $\exists x(Fx \wedge \neg Gx)$

32.2 Ambiguous quantifier order

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:

2. 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

- $\exists x(Mx \wedge \forall y(Py \rightarrow Syx))$
- $\forall y(Py \rightarrow \exists x(Mx \wedge Syx))$

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
Syx: $\_\_\_\_\__y$ 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.

To see that these are not logically equivalent consider this model:

Dom: $\{1, 2, 3, 4\}$
$[\![P]\!]$: $\{1, 2\}$
$[\![M]\!]$: $\{3, 4\}$
$[\![W]\!]$: $\{\langle 1, 3\rangle, \langle 2, 4\rangle\}$

The premise is true (person 1 went to see movie 3, person 2 went to see movie 4), but the conclusion is false (no single movie was seen by both people).

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:

– For every person, there is some truth they cannot know. $(\forall\exists)$
$\therefore$ 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]

– Every dog has its day. (∀∃)
∴ Therefore: There is a day for all the dogs. (∃∀)

The moral is: take great care with the order of quantification.

32.3 Indefinite article used as universal quantifier

Consider

3. A lion is dangerous.

Despite starting with 'a', where we might think we need to symbolise with 'exists', instead the most natural paraphrase for this sentence is:

4. Every lion is dangerous.

[1] Thanks to Rob Trueman for the example.

which we would symbolise as $\forall x(Lx \rightarrow Dx)$ Sometimes the English indefinite article, 'a', is used generically, which we best capture with 'for all'.

This can create ambiguity, where context isn't clear enough to determine how to symbolise it.

5. A philosopher is happy.

Should it be understood to say that every philosopher is happy, $\forall x(Px \rightarrow Hx)$ or that there exists a philosopher who is happy, $\exists x(Px \wedge Hx)$? It might be ambiguous.

32.4 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

[2]The Salique law was the common law of France which prohibited the crown be passed on to female heirs.

> 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 is superior to every woman, i.e.,

$$\forall x(Mx \to \forall y(Wy \to Sxy))$$

or that some men are superior to some women,

$$\exists x(Mx \wedge \exists y(Wy \wedge Sxy)).$$

The latter is true, but so is

$$\exists y(Wy \wedge \exists x(Mx \wedge Syx)).$$

[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.

(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.,

$$\exists y(Wy \wedge \forall x(Mx \rightarrow Syx)).$$

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!

32.5 Ambiguity with 'only'

In section 31.4 we discussed how to symbolize sentences involving 'only'. Consider the sentence:

6. Only young cats are playful.

According to our schema, we would symbolize it this way:

- $\forall x(Px \rightarrow (Yx \wedge Cx))$

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 6, 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:

- $\forall x((Cx \wedge Px) \rightarrow Yx)$

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:

- $\forall x((Yx \wedge Px) \rightarrow Cx)$

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).

Practice Exercises (Chapter 32)

I. Each of the following sentences is ambiguous. Provide a symbolization key for each, and symbolize all readings.

a) Noone likes a quitter.
b) CSI found only red hair at the scene.

Part VIII

Natural deduction for FOL

Chapter 33

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 '⊢') introduced there. However, we will also need some new basic rules to govern the quantifiers.

33.1 Universal elimination

Consider:

- 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	$\forall x\, F x$	
2	$F a$	∀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:

– 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	$Cf \to Sf$	$\forall$E, 1

Note here that we have to replace two instances of 'x' with our name: 'f' or Fluffy.

The more complex thing we're saying may itself already have quantifiers.

- Fluffy is a cat.
- Every cat has a toy.

$\therefore$ Therefore: Fluffy has a toy.

This is a valid argument. And it will be allowed by our rule $\forall$E:

1	Cf	
2	$\forall x\,(Cx \to \exists y(Ty \wedge Hxy)$	
3	$Cf \to \exists y(Ty \wedge Hfy)$	$\forall$E, 2
4	$\exists y(Ty \wedge Hfy)$	$\to$E, 1, 3

We replace all instances of the variable, x with the name, f.

We can apply the rule twice:

- Every vet likes all animals
- Charlie is a vet
- Snowball is an animal.

$\therefore$ Therefore: Charlie likes Snowball.

1	$\forall x\,(V\,x \rightarrow \forall y(Ay \rightarrow L\,x\,y)$	
2	Vc	
3	As	
4	$V\,c \rightarrow \forall y(Ay \rightarrow L\,c\,y)$	$\forall$E, 1
5	$\forall y\,(A\,y \rightarrow Lc\,y)$	$\rightarrow$E, 4, 2
6	$A\,s \rightarrow Lc\,s$	$\forall$E, 5
7	Lcs	$\rightarrow$E, 6, 3

Consider the following:

- Pavel owes everyone money.

∴ Therefore: Pavel owes himself money.

We we symbolise this as:

– $\forall x Opx$
∴ Therefore: Opp

This is valid: the premise says **everything** in the domain owes money to Pavel; and names refer to things in the domain. So it implies that Pavel owes money to himself.

This argument is also directly allowed by our rule ∀E:

1	$\forall x\, Op x$	
2	$Op p$	∀E, 1

If you think that this isn't right, this might be the wrong symbolisation of 'Pavel owes everyone money'. You might think that it means: 'Pavel owes everyone else money', that is, for everyone who is not identical to Pavel, he owes them money. To symbolise this, we would need the identity

symbol. It is something which one can easily add to FOL, but we won't cover it in this unit.

To give the general rule, we introduce some particular notation:

Definition

For a formula of FOL, X, a variable v and a constant c, we define a formula:

$$X[c/v]$$

which is the formula which results from replacing every free occurrence of v with c.

All the examples of the ∀E rule we have used so far have this form:

m	$\forall v X$	
	$X[c/v]$	∀E, m

▷	1	$\forall x\, F\, x$	
		$F\, a$	∀E, 1

▹	1	$\forall x\,(C\,x \to S\,x)$	
		$C\,f \to S\,f$	∀E, 1
▹	2	$\forall x\,(C\,x \to \exists y(T y \wedge H\,x\,y)$	
		$C\,f \to \exists y(T y \wedge H\,f\,y)$	∀E, 2
▹	1	$\forall x\,(V\,x \to \forall y(A y \to L\,x\,y)$	
		$V\,c \to \forall y(A y \to L\,c\,y)$	∀E, 1
▹	5	$\forall y\,(A\,y \to L c\,y)$	
		$A\,s \to L c\,s$	∀E, 5
▹	1	$\forall x\,Op\,x$	
		$Op\,p$	∀E, 1

This is our first basic rule of natural deduction for FOL:

m	$\forall v X$	
	$X[c/v]$	∀E, m

We can now give our general rule: whenever you have a sentence $\forall v X(\ldots v \ldots v \ldots)$, for example $\forall x Fx$, $\forall x(Cx \rightarrow Sx)$, $\forall x Opx$; 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, $Cf \rightarrow Sf$ or Opp.

As with every elimination rule, you can only apply the ∀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 ∀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 **??**.)

33.2 Existential introduction

The following argument is valid:

– 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	$\exists x\, Fx$	∃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 ∃I.

This isn't restricted to simple properties.

– 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

The more complex thing we're saying may itself already have quantifiers.

- Fluffy is a cat.
- Fluffy has a toy.

$\therefore$ Therefore: Some cat has a toy.

This is a valid argument. And it will be allowed by our rule $\forall$E:

1	Cf	
2	$\exists y(Ty \wedge Hfy)$	
3	$Cf \wedge \exists y(Ty \wedge Hfy)$	$\wedge$I, 1, 2
4	$\exists x\,(Cx \wedge \exists y(Ty \wedge Hxy)$	$\exists$I, 3

We replace all instances of the variable, x with the name, f.

This rule will now allow us to carefully work through our brain teaser from section 20.4

> Three people are standing in a row looking at eachother.
>
>
> 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?

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 section 20.4 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 w
5 | | Suppose Bob is unhappy
6 | | Then happy Alice is looking at unhappy Bob
7 | | So someone who is happy is looking at someone w
8 | Therefore, someone who is happy is looking at someo

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	$Hb \vee \neg Hb$	LEM
6	$\quad Hb$	
7	$\quad Lbc \wedge \neg Hc$	$\wedge$I, 2, 6
8	$\quad Hb \wedge (Lbc \wedge \neg Hc)$	$\wedge$I, 7, 6
9	$\quad \exists y(Hb \wedge (Lby \wedge \neg Hy))$	$\exists$I, 8
10	$\quad \exists x \exists y(Hx \wedge (Lxy \wedge \neg Hx))$	$\exists$I, 9
11	$\quad \neg Hb$	
12	$\quad Lab \wedge \neg Hb$	$\wedge$I, 11, 1
13	$\quad Ha \wedge (Lab \wedge \neg Hb)$	$\wedge$I, 3, 12
14	$\quad \exists y(Ha \wedge (Lay \wedge \neg Hy))$	$\exists$I, 13
15	$\quad \exists x \exists y(Hx \wedge (Lxy \wedge \neg Hx))$	$\exists$I, 14
16	$\exists x \exists y(Hx \wedge Lxy \wedge \neg Hx)$	$\vee$E, 5, 6-10, 11-15

Here is our general rule, using the notation as we used for ∀E:

m	$X[c/v]$	
	$\exists v X$	∀E, m

Consider the following example:

– Narcissus loves himself. Lnn
∴ Therefore: Narcissus loves somebody. $\exists x Lnx$

Lnn tells us something about Narcissus. But it tells us various different things:

- That he loves himself
  - $x \mapsto$Narcissus satisfies Lxx
- That he loves Narcissus:
  - $x \mapsto$Narcissus satisfies Lxn
- That he is loved by Narcissus:
  - $x \mapsto$Narcissus satisfies Lnx

Three formulas: Lxx, Lnx and Lxn each result in the same sentence: Lnn when we replace the occurrences of x with n. The existential introduction rule will allow us to obtain any of $\exists x Lxx$, $\exists x Lxn$, $\exists x Lnx$ from Lnn.

Formula X	$X[n/x]$	$\exists x X$
Lxx	Lnn	$\exists x Lxx$
Lnx	Lnn	$\exists x Lnx$
Lxn	Lnn	$\exists x Lxn$

For example:

1	Lnn	
2	$\exists x\, Lnx\, n$	$\exists$I, 1

Usually one can simply replace any (or all) occurrence of a name with your selected variable v and add $\exists v$ at the front. The only time that care needs to be taken is if X already contains the variable v. For example

$$\exists x \forall x (Bx \to Lxx)$$

This is a sentence which you should never use when symbolising in FOL. As long as you never consider sentences which have quantifier whose scope already contains a quantifier using that variable, you don't need to worry and can simply replace your names with the new variable.

Strictly speaking, however, it is a formula of FOL, and to ensure our proof system is **complete**, that is, proves every valid argument, we need to ensure that it does appropriately apply to such sentences. For this unit, we won't ever ask about such cases and you can ignore such care.

Care in applying the rule when we have clashing quantifiers

Consider the following argument:

1	$\forall x(Bx \to Lax)$	
2	$\exists x \forall x(Bx \to Lxx)$	naughtily attempting to invoke ∃I

In fact, this final formula is logically equivalent to $\forall x(Bx \to Lxx)$. Here's the semantic reasoning:

> To evaluate $\exists x \forall x(Bx \to Lxx)$, one must consider variable assignments $x \mapsto o$ and whether $\forall x(Bx \to Lxx)$ is true under that variable assignment. But $\forall x(Bx \to Lxx)$ is a sentence with no free variables. Whether it is true or false is independent of any variable assignment, it simply depends on the model. So the quantifier $\exists x$ at the front is unimportant.

The argument is then equivalent to

– Alice likes every biologist.
∴ Therefore: Every biologist likes themselves.

which is clearly invalid.

This is blocked by taking care in our notation $X[c/v]$. The rule permits:

m	$X[c/v]$	
	$\exists v X$	$\forall$E, m

Since the conclusion of our application of the rule is $\exists x \forall x(Bx \to Lxx)$, here X is $\forall x(Bx \to Lxx)$. But then $X[a/x]$ is to replace any **free** occurences of the variable x with the name a. But there are no free occurrences, so $X[a/x]$ is simply $\forall x(Bx \to Lxx)$

33.3 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, '$\exists xFx$' follows from '$\forall xFx$', 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 chapter 26, 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:

- $\forall xFx \vdash Fa$: after all, that was $\forall$E.
- $Fa \vdash \exists xFx$: 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 $\forall xFx \vdash \exists xFx$. 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.)

33.4 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 '$\forall xFx$', then you can infer '$\forall xFx$'.

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:

$$\forall xFx \therefore \forall yFy$$

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	$\forall xFx$	
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}, ..., 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 '$\forall yFy$', thus:

1	$\forall xFx$	
2	Fa	$\forall$E, 1
3	$\forall yFy$	$\forall$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):

m	$X[c/v]$	
	$\forall vX$	$\forall$I, m

c must not occur in any undischarged assumption or premise.

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:

$$\forall xLxk \therefore \forall xLxx$$

Now, suppose we tried to offer a proof that vindicates this argument:

1	$\forall xLxk$	
2	Lkk	∀E, 1
3	$\forall xLxx$	naughtily attempting to invoke ∀I, 2

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:

1	$\forall x(Gx \to Hx)$	
2	$\forall x(Hx \to Ex)$	
3	$\quad Gd$	
4	$\quad Gd \to Hd$	$\forall$E, 1
5	$\quad Hd$	$\to$E, 4, 3
6	$\quad Hd \to Ed$	$\forall$E, 2
7	$\quad Ed$	$\to$E, 5, 6
8	$Gd \to Ed$	$\to$I, 3–5
9	$\forall z(Gz \to Ez)$	$\forall$I, 8

33.5 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 '$\exists xFx$' 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:

1	$\exists x\, F\, x$	
2	$\forall x(Fx \rightarrow Gx)$	
3	$F\, o$	
4	$Fo \rightarrow Go$	$\forall$E, 2
5	Go	$\rightarrow$E, 4, 3
6	$\exists x Gx$	$\exists$I, 5
7	$\exists x Gx$	$\exists$E, 1, 3–6

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 '$\exists x Fx$'. On this assumption, we established '$\exists x Gx$'. 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 '$\exists x Gx$' on its own.

Putting this together, we obtain the existential elimination rule (∃E):

m	$\exists v X$	
i	$X[c/v]$	
	⋮	
j	Y	
	Y	∃E, m, i–j

c must be new to the proof: it does not occur in any line $< i$;
c must not occur in Y.

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:

1	Lb	
2	$\exists x \neg L x$	
3	$\quad \neg L b$	
4	$\quad Lb \wedge \neg Lb$	$\wedge$I, 1, 3
5	$Lb \wedge \neg Lb$	naughtily attempting to invoke $\exists$E , 2

The last line of the proof is not allowed. The name that we used in our substitution instance for '$\exists x \neg Lx$' is still present on line 3, namely 'b', occurs in line 4. The following proof would be no better:

1	Lb	
2	$\exists x\, \neg L x$	
3	$\quad \neg L b$	
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 $\exists$I

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 (Chapter 33)

I. 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	Raa	
4	$\exists x Rxx$	$\exists$I, 3
5	$\exists x Rxx$	$\exists$E, 2, 3–4

II. The following three proofs are missing their citations (rule and line numbers). Add them, to turn them into bona fide proofs.

1		$\forall x \exists y (Rxy \vee Ryx)$
2		$\forall x \neg Rmx$
3		$\exists y (Rmy \vee Rym)$
4		$Rma \vee Ram$
5		$\neg Rma$
6		Ram
7		$\exists x Rxm$
8		$\exists x Rxm$

1		$\forall x (\exists y Lxy \rightarrow \forall z Lzx)$
2		Lab
3		$\exists y Lay \rightarrow \forall z Lza$
4		$\exists y Lay$
5		$\forall z Lza$
6		Lca
7		$\exists y Lcy \rightarrow \forall z Lzc$
8		$\exists y Lcy$
9		$\forall z Lzc$
10		Lcc

1		$\forall x(Jx \rightarrow Kx)$
2		$\exists x \forall y Lxy$
3		$\forall x Jx$
4		$\forall y Lay$
5		Laa
6		Ja
7		$Ja \rightarrow Ka$
8		Ka
9		$Ka \wedge Laa$
10		$\exists x(Kx \wedge Lxx)$
11		$\exists x(Kx \wedge Lxx)$

III. 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 '$\forall x(Fx \rightarrow \neg Gx)$'.

IV. 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.

V. Provide a proof of each claim.

a) $\vdash \forall x Fx \vee \neg \forall x Fx$
b) $\vdash \forall z(Pz \vee \neg Pz)$
c) $\forall x(Ax \rightarrow Bx), \exists x Ax \vdash \exists x Bx$
d) $\forall x(Mx \leftrightarrow Nx), Ma \wedge \exists x Rxa \vdash \exists x Nx$
e) $\forall x \forall y Gxy \vdash \exists x Gxx$
f) $\vdash \forall x Rxx \rightarrow \exists x \exists y Rxy$
g) $\vdash \forall y \exists x(Qy \rightarrow Qx)$
h) $Na \rightarrow \forall x(Mx \leftrightarrow Ma), Ma, \neg Mb \vdash \neg Na$
i) $\forall x \forall y(Gxy \rightarrow Gyx) \vdash \forall x \forall y(Gxy \leftrightarrow Gyx)$
j) $\forall x(\neg Mx \vee Ljx), \forall x(Bx \rightarrow Ljx), \forall x(Mx \vee Bx) \vdash \forall x Ljx$

VI. 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.

VII. Show that each pair of sentences is provably equivalent.

a) $\forall x(Ax \to \neg Bx), \neg\exists x(Ax \wedge Bx)$
b) $\forall x(\neg Ax \to Bd), \forall x Ax \vee Bd$
c) $\exists x Px \to Qc, \forall x(Px \to Qc)$

VIII. 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.

a) $\forall x Px \to Qc, \forall x(Px \to Qc)$
b) $\forall x\forall y\forall z Bxyz, \forall x Bxxx$
c) $\forall x\forall y Dxy, \forall y\forall x Dxy$
d) $\exists x\forall y Dxy, \forall y\exists x Dxy$
e) $\forall x(Rca \leftrightarrow Rxa), Rca \leftrightarrow \forall x Rxa$

IX. 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.

a) $\exists y\forall x Rxy \therefore \forall x\exists y Rxy$
b) $\forall x\exists y Rxy \therefore \exists y\forall x Rxy$
c) $\exists x(Px \wedge \neg Qx) \therefore \forall x(Px \to \neg Qx)$
d) $\forall x(Sx \to Ta), Sd \therefore Ta$
e) $\forall x(Ax \to Bx), \forall x(Bx \to Cx) \therefore \forall x(Ax \to Cx)$

f) $\exists x(Dx \vee Ex), \forall x(Dx \rightarrow Fx) \therefore \exists x(Dx \wedge Fx)$

g) $\forall x \forall y(Rxy \vee Ryx) \therefore Rjj$

h) $\exists x \exists y(Rxy \vee Ryx) \therefore Rjj$

i) $\forall x Px \rightarrow \forall x Qx, \exists x \neg Px \therefore \exists x \neg Qx$

j) $\exists x Mx \rightarrow \exists x Nx, \neg \exists x Nx \therefore \forall x \neg Mx$

Chapter 34

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 chapter 26, 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	$\forall v \neg X$	
	$\neg\exists v X$	ConvQuant, m

and

m	$\neg\exists v X$	
	$\forall v \neg X$	ConvQuant, m

Equally, we add:

m	$\exists v \neg X$	
	$\neg\forall v X$	ConvQuant, m

and

m	$\neg\forall v X$	
	$\exists v \neg X$	ConvQuant, m

Practice Exercises (Chapter 34)

I. Show in each case that the sentences are provably inconsistent:

a) $Sa \rightarrow Tm, Tm \rightarrow Sa, Tm \wedge \neg Sa$

b) $\neg\exists x Rxa, \forall x \forall y Ryx$

c) $\neg\exists x\exists y Lxy, Laa$
d) $\forall x(Px \rightarrow Qx), \forall z(Pz \rightarrow Rz), \forall y Py, \neg Qa \wedge \neg Rb$

II. Show that each pair of sentences is provably equivalent:

a) $\forall x(Ax \rightarrow \neg Bx), \neg\exists x(Ax \wedge Bx)$
b) $\forall x(\neg Ax \rightarrow Bd), \forall x Ax \vee Bd$

III. In **??**, we considered what happens when we move quantifiers 'across' various logical operators. Show that each pair of sentences is provably equivalent:

a) $\forall x(Fx \wedge Ga), \forall x Fx \wedge Ga$
b) $\exists x(Fx \vee Ga), \exists x Fx \vee Ga$
c) $\forall x(Ga \rightarrow Fx), Ga \rightarrow \forall x Fx$
d) $\forall x(Fx \rightarrow Ga), \exists x Fx \rightarrow Ga$
e) $\exists x(Ga \rightarrow Fx), Ga \rightarrow \exists x Fx$
f) $\exists x(Fx \rightarrow Ga), \forall x Fx \rightarrow Ga$

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 35

Derived rules

As in the case of TFL, we first introduced some rules for FOL as basic (in chapter 33), and then added some further rules for conversion of quantifiers (in chapter 34). In fact, the CQ rules should be regarded as **derived** rules, for they can be derived from the **basic** rules of chapter 33. (The point here is as in chapter 23.) 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	$\quad\quad\quad Da \vee \neg Da$	
5	$\quad\quad\quad Pb$	$\forall$E, 2
6	$\quad\quad\quad \neg Pb$	R, 3
7	$\quad\quad \neg(Da \vee \neg Da)$	$\neg$I, 4–6
8	$\quad \neg(Da \vee \neg Da)$	$\exists$E, 1, 3–7
9	$\quad Da \vee \neg Da$	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 (Chapter 35)

I. Offer proofs which justify the addition of the second and fourth CQ rules as derived rules.

Chapter 36

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., $\models X$. 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:

$$X_1, X_2, \ldots, X_n \vdash Y \textbf{ iff } X_1, X_2, \ldots, X_n \models Y$$

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 '⊨' and '⊢'. 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

Answers to practice exercises

Answers to Practice Exercises from Chapter 1

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.

Some of the answers to the practice questions are in chapter A, but it is important to first try yourself. The answers are provided on a best-endeavours basis. There will be gaps and mistakes. Questions with answers are indicated with stars. Mistakes can be reported to catrin.campbell-moore@bristol.ac.uk.

So here's the first exercise.

I.

a) Are arguments always presented in our standard form?
b) Do conclusions always come after the premises in arguments in texts?
c) Might premises and conclusions be clauses within sentences?
d) Can questions be premises?

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

a) It is sunny. So I should take my sunglasses.
   I should take my sunglasses.

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

c) 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!

d) 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.

e) 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.

f) 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.

g) 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**

h) 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.

i) 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

III. 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.

a) Answer:

   – 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.

b) It is sunny. So I should take my sunglasses.
   I should take my sunglasses

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

d) 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

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

f) 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.

g) 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

Answers to Practice Exercises from Chapter 2

I.

a) What kind of things are valid or invalid?
b) When is an argument said to be valid?
c) When is an argument said to be sound?

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

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

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

a)
- Everyone in group 1 handed in their homework.
- Jenny is in group 1.
- ∴ Therefore: Jenny handed in her homework.

valid

b)
- If she won the lottery then she is rich.
- She is rich.
- ∴ Therefore: She won the lottery.

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

c)
- Most people are scared of spiders.
- ∴ Therefore: Oscar is scared of spiders.

Invalid

d)
- She is a donkey.
- ∴ Therefore: She does not talk.

e)
- Socrates is a man.
- All men are carrots.
- ∴ Therefore: Socrates is a carrot. Valid

f) – Abe Lincoln was either born in Illinois or he was president.
– Abe Lincoln was not president.
∴ Therefore: Abe Lincoln was born in Illinois. Valid

g) – If I pull the trigger, Abe Lincoln will die.
– I do not pull the trigger.
∴ Therefore: 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.

h) – Abe Lincoln was either from France or from Luxembourg.
– Abe Lincoln was not from Luxembourg.
∴ Therefore: Abe Lincoln was from France.
Valid

i) – If the world were to end today, then I would not need to get up tomorrow morning.
– I will need to get up tomorrow morning.
∴ Therefore: The world will not end today.
Valid

j) – Joe is now 19 years old.

– Joe is now 87 years old.
∴ Therefore: 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.)

III. Could there be:

a) A valid argument that has one false premise and one true premise? Yes. Example: the first argument, above.
b) 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.
c) A valid argument with only false premises and a false conclusion? Yes. The same example will suffice.
d) 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).

e) 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.

Answers to Practice Exercises from Chapter 4

I. Write down (from memory) the truth rules for each of the connectives.

II.

a) Suppose X is true, what can you say about $X \vee Y$?
b) Suppose X is false, what can you say about $X \rightarrow Y$?

Answers to Practice Exercises from Chapter 6

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

a) '$\neg$', '$\wedge$', '$\vee$' and '$\rightarrow$' are the basic logical connectives of the language of TFL.
b) In a well–formed sentence for every opening bracket '(' there must be a closing bracket ')'.

c) The sentence 'snow is white' is true if and only if snow is white.
d) Kevin is a lecturer, whereas 'Kevin' names Kevin and the latter is an expression of the language (and not a lecturer).

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

a) $X \wedge Y$
the result of taking the sentence Y followed by '$\wedge$' followed by Z
b) $Y \rightarrow Z$
the result of taking the sentence Y followed by '$\rightarrow$' followed by Z
c) $\neg(Y \vee Z)$
the result of placing '$\neg($' before the sentence Y, then adding '$\vee$', then Z and finally '$)$'.

III. Consider the sentence

▹ Sharky is so-called because of his teeth.

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

Answers to Practice Exercises from Chapter 7

I. Add brackets to form a sentence of TFL.

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

a) Where the '$\rightarrow$' is the main connective
b) Where the '$\vee$' is the main connective.
c) Where the left '$\neg$' is the main connective.
d) Can it be done so that the right '$\neg$' is the main connective?

II. 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, include braces indicating the scope of each connective.

a) (A) (a) no (b) no

b) $J_{374} \vee \neg J_{374}$ (a) no (b) yes
c) $\neg\neg\neg\neg F$ (a) yes (b) yes
d) $\neg \wedge S$ (a) no (b) no
e) $(G \wedge \neg G)$ (a) yes (b) yes
f) $(A \rightarrow (A \wedge \neg F)) \vee (D \rightarrow E)$ (a) no (b) yes
g) $((A \rightarrow S) \rightarrow W) \wedge (J \vee A)$ (a) no (b) yes
h) $(F \rightarrow \neg D \rightarrow J) \vee (C \wedge D)$ (a) no (b) no

III.

a) Does $\neg A \vee B$ have the form $X \vee Y$?
b) Does $\neg(A \vee B)$ have the form $X \vee Y$?
c) Does $(A \wedge B) \vee C$ have the form $X \vee Y$?
d) Does $(A \wedge B) \vee C$ have the form $X \wedge Y$?

IV.

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

b) 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.

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

Answers to Practice Exercises from Chapter 8

I. Complete truth tables for each of the following:

a) $A \rightarrow A$

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

b) $C \rightarrow \neg C$

C	$C \to \neg C$
T	T F F T
F	F T T F

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

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

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

e) $(A \wedge B) \to (B \vee A)$

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

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

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

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

h) $\big((A \wedge B) \wedge \neg(A \wedge B)\big) \wedge C$

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

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

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

j) $\neg\big((C \vee A) \vee B\big)$

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

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

(A	→	B)	→	(¬	B	→	¬	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

l) $(C \to (D \vee E)) \wedge \neg C$

m) $\neg(G \wedge (B \wedge H)) \to (G \vee (B \vee H))$

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

a) $(D \wedge \neg D) \to G$

(D	$\wedge$	$\neg$	D)	$\rightarrow$	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

b) $(\neg P \vee \neg M) \rightarrow M$

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

$\neg$	$\neg$	($\neg$	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

d) $((D \wedge R) \rightarrow I) \rightarrow \neg(D \vee R)$

[(D	∧	R)	→	I]	→	¬	(D	∨	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

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

Answers to Practice Exercises from Chapter 10

I. Revisit your answers to section 8.3. Determine which sentences were tautologies, which were contradictions, and which were neither tautologies nor contradictions.

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

a) $A \rightarrow A$, $\neg A \rightarrow \neg A$, $A \wedge A$, $A \vee A$
b) $A \vee B$, $A \rightarrow C$, $B \rightarrow C$
c) $B \wedge (C \vee A)$, $A \rightarrow B$, $\neg(B \vee C)$
d) $A \leftrightarrow (B \vee C)$, $C \rightarrow \neg A$, $A \rightarrow \neg B$

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

a) $\neg B \wedge B$
b) $\neg D \vee D$
c) $(A \wedge B) \vee (B \wedge A)$
d) $\neg(A \rightarrow (B \rightarrow A))$
e) $A \leftrightarrow (A \rightarrow (B \wedge \neg B))$
f) $((A \wedge B) \leftrightarrow B) \rightarrow (A \rightarrow B)$

IV. 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."

a) A and $\neg A$
b) $A \wedge \neg A$ and $\neg B \leftrightarrow B$
c) $((A \vee B) \vee C)$ and $(A \vee (B \vee C))$
d) $A \vee (B \wedge C)$ and $(A \vee B) \wedge (A \vee C)$
e) $(A \wedge (A \vee B)) \rightarrow B$ and $A \rightarrow B$

V. 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."

a) $A \rightarrow A$ and $A \leftrightarrow A$
b) $\neg(A \rightarrow B)$ and $\neg A \rightarrow \neg B$
c) $A \vee B$ and $\neg A \rightarrow B$
d) $(A \rightarrow B) \rightarrow C$ and $A \rightarrow (B \rightarrow C)$
e) $A \leftrightarrow (B \leftrightarrow C)$ and $A \wedge (B \wedge C)$

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

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

b) $A \vee B$, $A \rightarrow \neg A$, $B \rightarrow \neg B$

c) $\neg(\neg A \vee B)$, $A \rightarrow \neg C$, $A \rightarrow (B \rightarrow C)$

d) $A \rightarrow B$, $A \wedge \neg B$

e) $A \rightarrow (B \rightarrow C)$, $(A \rightarrow B) \rightarrow C$, $A \rightarrow C$

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

a) $\neg B$, $A \rightarrow B$, A

b) $\neg(A \vee B)$, $A \leftrightarrow B$, $B \rightarrow A$

c) $A \vee B$, $\neg B$, $\neg B \rightarrow \neg A$

d) $A \leftrightarrow B$, $\neg B \vee \neg A$, $A \rightarrow B$

e) $(A \vee B) \vee C$, $\neg A \vee \neg B$, $\neg C \vee \neg B$

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

a) Suppose X is true and Y is true
Is $X \vee Y$ true.
true

b) Suppose X is false
Is $X \vee Y$ true?
false. If Y is also false, then $X \vee Y$ is false.

c) Suppose X is false
Is $X \rightarrow Y$ true.
true

d) 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

e) Suppose X is a tautology and $X \therefore Y$ is a valid argument.
What can you say about Y?

f) 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.

g) Suppose that X is a contradiction.
What can you say about whether $X, Y \models Z$?
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.

h) 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.

i) 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.

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

X. 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.

Answers to Practice Exercises from Chapter 11

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

a) $\neg B \wedge B$ — Contradiction

B	$\neg B \wedge B$	
T	F	F
F		F

b) $\neg D \vee D$ — Tautology

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

c) $(A \wedge B) \vee (B \wedge A)$ — 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

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

e) $A \leftrightarrow (A \rightarrow (B \wedge \neg B))$ Contradiction

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

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

g) $A \rightarrow (B \vee C)$ Neither

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

h) $(A \wedge \neg A) \rightarrow (B \vee C)$ Tautology

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

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

Answers to Practice Exercises from Chapter 15

I.

a) 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.)

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

a)

A	B	?	
T	T	T	
T	F	T	$A \vee \neg A$
F	T	T	
F	F	T	

b)

A	B	?	
T	T	F	
T	F	F	$\neg A \wedge B$
F	T	T	
F	F	F	

c)

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:

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

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

a) 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.

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

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

Answers to Practice Exercises from Chapter 16

I. 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.

a) Those creatures are not men in suits.
$\neg M$

b) Those creatures are men in suits, or they are not.
$(M \vee \neg M)$

c) Those creatures are either gorillas or chimpanzees.
$(G \vee C)$

d) Those creatures are neither gorillas nor chimpanzees.

$\neg(C \vee G)$

e) If those creatures are chimpanzees, then they are neither gorillas nor men in suits.
$(C \rightarrow \neg(G \vee M))$

f) Unless those creatures are men in suits, they are either chimpanzees or they are gorillas.
$(M \vee (C \vee G))$

II. 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.

a) Either Mister Ace or Mister Edge was murdered.
$(A \vee E)$

b) If Mister Ace was murdered, then the cook did it.
$(A \rightarrow C)$

c) If Mister Edge was murdered, then the cook did not do it.
$(E \rightarrow \neg C)$

d) Either the butler did it, or the Duchess is lying.
$(B \vee D)$

e) The cook did it only if the Duchess is lying.
$(C \rightarrow D)$

f) If the murder weapon was a frying pan, then the culprit must have been the cook.
$(F \rightarrow C)$

g) If the murder weapon was not a frying pan, then the culprit was either the cook or the butler.
$(\neg F \rightarrow (C \vee B))$

h) Mister Ace was murdered if and only if Mister Edge was not murdered.
$(A \leftrightarrow \neg E)$

i) The Duchess is lying, unless it was Mister Edge who was murdered.
$(D \vee E)$

j) If Mister Ace was murdered, he was done in with a frying pan.
$(A \rightarrow F)$

k) Since the cook did it, the butler did not.
   $(C \wedge \neg B)$
l) Of course the Duchess is lying!
   D

III. 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.

a) Ava and Harrison are both electricians.
   $(E_1 \wedge E_2)$
b) If Ava is a firefighter, then she is satisfied with her career.
   $(F_1 \rightarrow S_1)$
c) Ava is a firefighter, unless she is an electrician.
   $(F_1 \vee E_1)$
d) Harrison is an unsatisfied electrician.

$(E_2 \wedge \neg S_2)$

e) Neither Ava nor Harrison is an electrician.
$\neg(E_1 \vee E_2)$

f) Both Ava and Harrison are electricians, but neither of them find it satisfying.
$((E_1 \wedge E_2) \wedge \neg(S_1 \vee S_2))$

g) Harrison is satisfied only if he is a firefighter.

$(S_2 \rightarrow F_2)$

h) If Ava is not an electrician, then neither is Harrison, but if she is, then he is too.
$((\neg E_1 \rightarrow \neg E_2) \wedge (E_1 \rightarrow E_2))$

i) Ava is satisfied with her career if and only if Harrison is not satisfied with his.
$(S_1 \leftrightarrow \neg S_2)$

j) If Harrison is both an electrician and a firefighter, then he must be satisfied with his work.
$((E_2 \wedge F_2) \rightarrow S_2)$

k) It cannot be that Harrison is both an electrician and a firefighter.
$\neg(E_2 \wedge F_2)$

l) Harrison and Ava are both firefighters if and only if neither of them is an electrician.

$((F_2 \wedge F_1) \leftrightarrow \neg(E_2 \vee E_1))$

IV. 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.

a) Alice and Bob are both spies.
$(A \wedge B)$

b) If either Alice or Bob is a spy, then the code has been broken.
$((A \vee B) \rightarrow C)$

c) If neither Alice nor Bob is a spy, then the code remains unbroken.
$(\neg(A \vee B) \rightarrow \neg C)$

d) The German embassy will be in an uproar, unless someone has broken the code.
$(G \vee C)$

e) Either the code has been broken or it has not, but the German embassy will be in an uproar regardless.
$((C \vee \neg C) \wedge G)$

f) Either Alice or Bob is a spy, but not both.
$((A \vee B) \wedge \neg(A \wedge B))$

V. 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.

a) If there is food to be found in the pridelands, then Rafiki will talk about squashed bananas.

$(F \rightarrow R)$

b) Rafiki will talk about squashed bananas unless Simba is alive.
$(R \vee A)$

c) Rafiki will either talk about squashed bananas or he won't, but there is food to be found in the pridelands regardless.
$((R \vee \neg R) \wedge F)$

d) Scar will remain as king if and only if there is food to be found in the pridelands.
$(K \leftrightarrow F)$

e) If Simba is alive, then Scar will not remain as king.
$(A \rightarrow \neg K)$

VI. 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?

a) 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.

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

b) 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

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

c) 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

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

VII. 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 15)

a) It is going to rain soon. I know because my leg is hurting, and my leg hurts if it's going to rain.
b) **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.

c) **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.

Answers to Practice Exercises from Chapter 17

I. The following sentences are ambiguous. Give symbolization keys for each and symbolize the different readings. Show that they are not logically equivalent.

a) Haskell is a birder and enjoys watching cranes.
b) The zoo has lions or tigers and bears.
c) The flower is not red or fragrant.

Answers to Practice Exercises from Chapter 19

I. 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

II. 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. $A \rightarrow B,\ A \rightarrow C,\ A \wedge D \ \therefore\ B \wedge C$

Answers to Practice Exercises from Chapter 20

I. 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	$S \rightarrow R$	
3	S	$\wedge$E, 1
4	R	$\rightarrow$E, 2, 3
5	$R \vee E$	$\vee$I, 4

1	$A \to D$	
2	$\quad A \wedge B$	
3	$\quad A$	$\wedge$E, 2
4	$\quad D$	$\to$E, 1, 3
5	$\quad D \vee E$	$\vee$I, 4
6	$(A \wedge B) \to (D \vee E)$	$\to$I, 2–5

II. For each of the following arguments, give a proof using natural deduction.

a) $J \to \neg J \therefore \neg J$

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

b) $Q \to (Q \wedge \neg Q) \therefore \neg Q$

1	$Q \rightarrow (Q \wedge \neg Q)$	
2	$\quad Q$	
3	$\quad Q \wedge \neg Q$	$\rightarrow$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

c) $A \rightarrow (B \rightarrow C) \therefore (A \wedge B) \rightarrow C$

1	$A \rightarrow (B \rightarrow C)$	
2	$\quad A \wedge B$	
3	$\quad A$	$\wedge$E, 2
4	$\quad B$	$\wedge$E, 2
5	$\quad B \rightarrow C$	$\rightarrow$E, 1, 3
6	$\quad C$	$\rightarrow$E, 5, 4
7	$(A \wedge B) \rightarrow C$	$\rightarrow$I, 2–6

d) $(C \wedge D) \vee E \therefore E \vee D$

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

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

1	$(Z \wedge K) \vee (K \wedge M)$	
2	$K \rightarrow D$	
3	$\quad Z \wedge K$	
4	$\quad K$	$\wedge$E, 3
5	$\quad D$	$\rightarrow$E, 2, 4
6	$\quad K \wedge M$	
7	$\quad K$	$\wedge$E, 6
8	$\quad D$	$\rightarrow$E, 2, 7
9	D	$\vee$E, 1, 3–5, 6–8

f) $P \wedge (Q \vee R), P \rightarrow \neg R \therefore Q \vee E$

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

g) $\neg(P \rightarrow Q) \therefore \neg Q$

1	$\neg(P \to Q)$	
2	$\quad Q$	
3	$\quad\quad P$	
4	$\quad\quad Q$	R, 2
5	$\quad P \to Q$	$\to$I, 3–4
6	$\quad (P \to Q) \wedge \neg(P \to Q)$	$\wedge$I, 5, 1
7	$\neg Q$	$\neg$I, 2–6

h) $\neg(P \to Q) \therefore P$

1	$\neg(P \to Q)$	
2	$\quad \neg P$	
3	$\quad\quad P$	
4	$\quad\quad\quad \neg Q$	
5	$\quad\quad\quad P \wedge \neg P$	$\wedge$I, 3, 2
6	$\quad\quad Q$	$\neg$E, 4-5
7	$\quad P \to Q$	$\to$I, 3-6
8	$\quad (P \to Q) \wedge \neg(P \to Q)$	$\wedge$I, 7, 1
9	P	$\neg$E, 2-8

Answers to Practice Exercises from Chapter 22

I. 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

1		$\neg O \rightarrow L$
2		$L \vee \neg O$
		—
3		$\neg L$
		—
4		$\neg O$
5		L
6		$\neg\neg L$
7		L

1		$Z \rightarrow (C \wedge \neg N)$
2		$\neg Z \rightarrow (N \wedge \neg C)$
		—
3		$Z \vee \neg Z$
4		Z
		—
5		$C \wedge \neg N$
6		C
7		$N \vee C$
8		$Z \rightarrow (N \vee C)$
9		$\neg Z$
		—
10		$N \wedge \neg C$
11		N
12		$N \vee C$
13		$\neg Z \rightarrow (N \vee C)$
14		$N \vee C$

II. Give a proof for each of these arguments:

1. $\neg F \to G, F \to H \therefore G \vee H$

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

2. $E \vee F, F \vee G, \neg F \therefore E \wedge G$

3. $M \vee (N \to M) \therefore \neg M \to \neg N$

4. $(M \vee N) \wedge (O \vee P), N \to P, \neg P \therefore M \wedge O$

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

Answers to Practice Exercises from Chapter 23

I. Provide proof schemes that justify the addition of the third and fourth De Morgan rules as derived rules.

II. The proofs you offered in response to the practice exercises of chapter 22–25 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.)

Answers to Practice Exercises from Chapter 24

I. Use either a derivation or a truth table for each of the following.

a) Show that $A \rightarrow (((B \wedge C) \vee D) \rightarrow A)$ is a tautology.
b) Show that $A \rightarrow (A \rightarrow B)$ is not a tautology
c) Show that the sentence $A \rightarrow \neg A$ is not a contradiction.
d) Show that the sentence $A \leftrightarrow \neg A$ is a contradiction.
e) Show that the sentence $\neg(W \rightarrow (J \vee J))$ is contingent
f) Show that the sentence $\neg(X \vee (Y \vee Z)) \vee (X \vee (Y \vee Z))$ is not contingent
g) Show that the sentence $B \rightarrow \neg S$ is equivalent to the sentence $\neg\neg B \rightarrow \neg S$
h) Show that the sentence $\neg(X \vee O)$ is not equivalent to the sentence $X \wedge O$
i) Show that the sentences $\neg(A \vee B)$, C, $C \rightarrow A$ are jointly inconsistent.
j) Show that the sentences $\neg(A \vee B)$, $\neg B$, $B \rightarrow A$ are jointly consistent
k) Show that $\neg(A \vee (B \vee C)) \therefore \neg C$ is valid.
l) Show that $\neg(A \wedge (B \vee C)) \therefore \neg C$ is invalid.

II. Use either a derivation or a truth table for each of the following.

a) Show that $A \rightarrow (B \rightarrow A)$ is a tautology
b) Show that $\neg(((N \leftrightarrow Q) \vee Q) \vee N)$ is not a tautology
c) Show that $Z \vee (\neg Z \leftrightarrow Z)$ is contingent
d) show that $(L \leftrightarrow ((N \rightarrow N) \rightarrow L)) \vee H$ is not contingent
e) Show that $(A \leftrightarrow A) \wedge (B \wedge \neg B)$ is a contradiction
f) Show that $(B \leftrightarrow (C \vee B))$ is not a contradiction.
g) Show that $((\neg X \leftrightarrow X) \vee X)$ is equivalent to X
h) Show that $F \wedge (K \wedge R)$ is not equivalent to $(F \leftrightarrow (K \leftrightarrow R))$
i) Show that the sentences $\neg(W \rightarrow W)$, $(W \leftrightarrow W) \wedge W$, $E \vee (W \rightarrow \neg(E \wedge W))$ are inconsistent.
j) Show that the sentences $\neg R \vee C$, $(C \wedge R) \rightarrow \neg R$, $(\neg(R \vee R) \rightarrow R)$ are consistent.
k) Show that $\neg\neg(C \leftrightarrow \neg C), ((G \vee C) \vee G) \therefore ((G \rightarrow C) \wedge G)$ is valid.
l) Show that $\neg\neg L, (C \rightarrow \neg L) \rightarrow C) \therefore \neg C$ is invalid.

Answers to Practice Exercises from Chapter 25

I. Show that each of the following sentences is a theorem:

a) $O \rightarrow O$
b) $N \vee \neg N$
c) $J \leftrightarrow (J \vee (L \wedge \neg L))$
d) $((A \rightarrow B) \rightarrow A) \rightarrow A$

II. Provide proofs to show each of the following:

a) $C \rightarrow (E \wedge G), \neg C \rightarrow G \vdash G$
b) $M \wedge (\neg N \rightarrow \neg M) \vdash (N \wedge M) \vee \neg M$
c) $(Z \wedge K) \leftrightarrow (Y \wedge M), D \wedge (D \rightarrow M) \vdash Y \rightarrow Z$
d) $(W \vee X) \vee (Y \vee Z), X \rightarrow Y, \neg Z \vdash W \vee Y$

III. Show that each of the following pairs of sentences are provably equivalent:

a) $R \leftrightarrow E$, $E \leftrightarrow R$
b) G, $\neg\neg\neg\neg G$
c) $T \rightarrow S$, $\neg S \rightarrow \neg T$

d) $U \rightarrow I$, $\neg(U \wedge \neg I)$
e) $\neg(C \rightarrow D)$, $C \wedge \neg D$
f) $\neg G \leftrightarrow H$, $\neg(G \leftrightarrow H)$

IV. If you know that $X \vdash Y$, what can you say about $(X \wedge Z) \vdash Y$? What about $(X \vee Z) \vdash Y$? Explain your answers.

V. 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.)

Answers to Practice Exercises from Chapter 26

I. Using this symbolization key:

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:

a) If Cleo loves Bouncer, then Bouncer is a monkey.
$Lc, b \rightarrow Mb$
b) If both Bouncer and Cleo are alligators, then Amos loves them both.
$(Ab \wedge Ac) \rightarrow (La, b \wedge La, c)$
c) Cleo loves a reptile.
$\exists x(Rx \wedge Lc, x)$
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'.)
d) Bouncer loves all the monkeys that live at the zoo.
$\forall x((Mx \wedge Zx) \rightarrow Lb, x)$

e) All the monkeys that Amos loves love him back.
$\forall x((Mx \wedge La, x) \rightarrow Lx, a)$

f) Every monkey that Cleo loves is also loved by Amos.
$\forall x((Mx \wedge Lc, x) \rightarrow La, x)$

g) There is a monkey that loves Bouncer, but sadly Bouncer does not reciprocate this love.
$\exists x(Mx \wedge Lx, b \wedge \neg Lb, x)$

II. Using the following symbolization key:

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:

a) All the food is on the table.

$\forall x(Fx \rightarrow Tx)$

b) If the guacamole has not run out, then it is on the table.
$\neg Rg \rightarrow Tg$

c) Everyone likes the guacamole.
$\forall x(Px \rightarrow Lx, g)$

d) If anyone likes the guacamole, then Eli does.
$\exists x(Px \wedge Lx, g) \rightarrow Le, g$

e) Francesca only likes the dishes that have run out.
$\forall x\big[(Lf, x \wedge Fx) \rightarrow Rx\big]$

f) Eli likes anyone who likes the guacamole.
$\forall x((Px \wedge Lx, g) \rightarrow Le, x)$

g) Eli likes anyone who likes the people that he likes.
$\forall x\big[(Px \wedge \forall y[(Py \wedge Le, y) \rightarrow Lx, y]) \rightarrow Le, x\big]$

h) If there is a person on the table already, then all of the food must have run out.
$\exists x(Px \wedge Tx) \rightarrow \forall x(Fx \rightarrow Rx)$

III. Using the following symbolization key:

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:

a) All of Patrick's children are ballet dancers.
$\forall x(Cx, p \rightarrow Dx)$
b) Jane is Patrick's daughter.
$Cj, p \wedge Fj$
c) Patrick has a daughter.
$\exists x(Cx, p \wedge Fx)$
d) All of Patrick's sons dance ballet.
$\forall x\big[(Cx, p \wedge Mx) \rightarrow Dx\big]$
e) Jane is Elmer's niece.
$\exists x(Sx, e \wedge Cj, x \wedge Fj)$
f) Patrick is Elmer's brother.
$Sp, e \wedge Mp$
g) Jane is an aunt.
$Fj \wedge \exists x(Sx, j \wedge \exists y Cy, x)$

h) Everyone who dances ballet has a brother who also dances ballet.
$\forall x\big[Dx \rightarrow \exists y(My \wedge Sy, x \wedge Dy)\big]$

i) Every woman who dances ballet is the child of someone who dances ballet.
$\forall x\big[(Fx \wedge Dx) \rightarrow \exists y(Cx, y \wedge Dy)\big]$

Answers to Practice Exercises from Chapter 27

I. Identify which variables are bound and which are free.

a) $\exists xLxy \wedge \forall yLyx$
b) $\forall xAx \wedge Bx$
c) $\forall x(Ax \wedge Bx) \wedge \forall y(Cx \wedge Dy)$
d) $\forall x\exists y(Rxy \rightarrow (Jz \wedge Kx)) \vee Ryx$
e) $\forall x_1(Mx_2 \rightarrow Lx_2x_1) \wedge \exists x_2 Lx_3x_2$

Answers to Practice Exercises from Chapter 29

I. Consider the following model:

Dom: $\{1, 2\}$
$[\![A]\!]$: $\{1, 2\}$
$[\![B]\!]$: $\{2\}$
$[\![N]\!]$: $\varnothing$
$[\![c]\!]$: 1

Determine whether each of the following sentences is true or false in that model:

a) Bc — False
b) $Ac \leftrightarrow \neg Nc$ — True
c) $Nc \rightarrow (Ac \vee Bc)$ — True
d) $\forall x\, Ax$ — True
e) $\forall x \neg Bx$ — False
f) $\exists x(Ax \wedge Bx)$ — True
g) $\exists x(Ax \rightarrow Nx)$ — False
h) $\forall x(Nx \vee \neg Nx)$ — True
i) $\exists x\, Bx \rightarrow \forall x\, Ax$ — True

II. Consider the following model:

Dom: $\{$**Lemmy**, **Courtney**, **Eddy**$\}$
$[\![G]\!]$: $\{$**Lemmy**, **Courtney**, **Eddy**$\}$
$[\![H]\!]$: $\{$**Courtney**$\}$

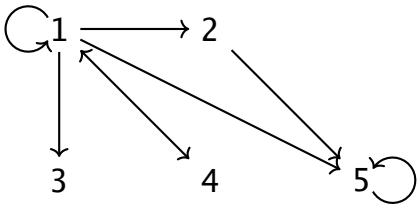
$[\![M]\!]$: {**Lemmy**, **Eddy**}
$[\![c]\!]$: **Courtney**
$[\![e]\!]$: **Eddy**

Determine whether each of the following sentences is true or false in that model:

a) Hc — True
b) He — False
c) $Mc \vee Me$ — True
d) $Gc \vee \neg Gc$ — True
e) $Mc \rightarrow Gc$ — True
f) $\exists x\, Hx$ — True
g) $\forall x\, Hx$ — False
h) $\exists x \neg Mx$ — True
i) $\exists x(Hx \wedge Gx)$ — True
j) $\exists x(Mx \wedge Gx)$ — True
k) $\forall x(Hx \vee Mx)$ — True
l) $\exists x\, Hx \wedge \exists x\, Mx$ — True
m) $\forall x(Hx \leftrightarrow \neg Mx)$ — True
n) $\exists x\, Gx \wedge \exists x \neg Gx$ — False
o) $\forall x \exists y(Gx \wedge Hy)$ — True

III. 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:

a) $\exists x\, Rxx$ — True
b) $\forall x\, Rxx$ — False
c) $\exists x \forall y\, Rxy$ — True
d) $\exists x \forall y\, Ryx$ — False
e) $\forall x \forall y \forall z((Rxy \wedge Ryz) \rightarrow Rxz)$ — False
f) $\forall x \forall y \forall z((Rxy \wedge Rxz) \rightarrow Ryz)$ — False
g) $\exists x \forall y \neg Rxy$ — True
h) $\forall x(\exists y\, Rxy \rightarrow \exists y\, Ryx)$ — True

i) $\exists x \exists y(\neg x = y \wedge Rxy \wedge Ryx)$ True

j) $\exists x \forall y(Rxy \leftrightarrow x = y)$ True

k) $\exists x \forall y(Ryx \leftrightarrow x = y)$ False

l) $\exists x \exists y(\neg x = y \wedge Rxy \wedge \forall z(Rz, x \leftrightarrow y = z))$ True

Answers to Practice Exercises from Chapter 30

I. Show that each of the following is neither a logical truth nor a contradiction:

a) $Da \wedge Db$

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

b) $\exists x\, Tx, h$

The sentence is true in this model:

Dom: Stan
$T(x, y)$: ⟨Stan, Stan⟩
$[\![h]\!]$: Stan

And it is false in this model:

Dom: Stan
$T(x, y)$:
$[\![h]\!]$: Stan

c) $Pm \wedge \neg \forall x\, Px$

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

d) $\forall z\, Jz \leftrightarrow \exists y\, Jy$

e) $\forall x(Wx, m, n \vee \exists y Lx, y)$

f) $\exists x(Gx \rightarrow \forall y\, My)$

g) $\exists x(x = h \wedge x = i)$

II. Show that the following pairs of sentences are not logically equivalent.

a) Ja, Ka

Making the first sentence true and the second false:

Dom: 0
$J(x)$: 0
$K(x)$:
$[\![a]\!]$: 0

b) $\exists x\, Jx$, Jm

Making the first sentence true and the second false:

Dom: 0, 1
$J(x)$: 0
$[\![m]\!]$: 1

c) $\forall x\, Rx, x$, $\exists x\, Rx, x$

Making the first sentence false and the second true:

Dom: 0, 1
$R(x, y)$: $\langle 0,0 \rangle$

d) $\exists x\, Px \rightarrow Qc$, $\exists x(Px \rightarrow Qc)$

Making the first sentence false and the second true:

Dom: 0, 1
$P(x)$: 0
$Q(x)$:
$[\![c]\!]$: 0

e) $\forall x(Px \to \neg Qx)$, $\exists x(Px \wedge \neg Qx)$
Making the first sentence true and the second false:

Dom: 0
$P(x)$:
$Q(x)$:

f) $\exists x(Px \wedge Qx)$, $\exists x(Px \to Qx)$
Making the first sentence false and the second true:

Dom: 0
$P(x)$:
$Q(x)$: 0

g) $\forall x(Px \to Qx)$, $\forall x(Px \wedge Qx)$
Making the first sentence true and the second false:

Dom: 0
$P(x)$:

$Q(x)$: 0

h) $\forall x \exists y\, Rx, y, \exists x \forall y\, Rx, y$
Making the first sentence true and the second false:

Dom: 0, 1
$R(x, y)$: $\langle 0, 1\rangle$, $\langle 1, 0\rangle$

i) $\forall x \exists y\, Rx, y, \forall x \exists y\, Ry, x$
Making the first sentence false and the second true:

Dom: 0, 1
$R(x, y)$: $\langle 0, 0\rangle$, $\langle 0, 1\rangle$

III. Show that the following sentences are (jointly) consistent:

a) $Ma, \neg Na, Pa, \neg Qa$
b) $Le, e, Le, g, \neg Lg, e, \neg Lg, g$
c) $\neg(Ma \wedge \exists x\, Ax), Ma \vee Fa, \forall x(Fx \rightarrow Ax)$
d) $Ma \vee Mb, Ma \rightarrow \forall x \neg Mx$
e) $\forall y\, Gy, \forall x(Gx \rightarrow Hx), \exists y \neg Iy$
f) $\exists x(Bx \vee Ax), \forall x \neg Cx, \forall x\big((Ax \wedge Bx) \rightarrow Cx\big)$
g) $\exists x\, Xx, \exists x\, Yx, \forall x(Xx \leftrightarrow \neg Yx)$
h) $\forall x(Px \vee Qx), \exists x \neg(Qx \wedge Px)$

i) $\exists z(Nz \wedge Oz, z), \forall x \forall y(Ox, y \rightarrow Oy, x)$

j) $\neg \exists x \forall y\, Rx, y, \forall x \exists y\, Rx, y$

k) $\neg Ra, a, \forall x(x = a \vee Rx, a)$

The sentences are both true in this model:

Dom: Harry, Sally

$R(x, y)$: ⟨Sally, Harry⟩

$[\![a]\!]$: Harry

l) $\forall x \forall y \forall z((x = y \vee y = z) \vee x = z), \exists x \exists y\, \neg x = y$

There are no predicates or constants, so we only need to give a domain. Any domain with 2 elements will do.

m) $\exists x \exists y((Zx \wedge Zy) \wedge x = y), \neg Zd, d = e$

IV. Show that the following arguments are invalid:

a) $\forall x(Ax \rightarrow Bx) \therefore \exists x\, Bx$

b) $\forall x(Rx \rightarrow Dx), \forall x(Rx \rightarrow Fx) \therefore \exists x(Dx \wedge Fx)$

c) $\exists x(Px \rightarrow Qx) \therefore \exists x\, Px$

d) $Na \wedge Nb \wedge Nc \therefore \forall x\, Nx$

e) $Rd, e, \exists x\, Rxd \therefore Re, d$

f) $\exists x(Ex \wedge Fx), \exists x\, Fx \rightarrow \exists x\, Gx \therefore \exists x(Ex \wedge Gx)$

g) $\forall x\, Ox, c, \forall x\, Oc, x \therefore \forall x\, Ox, x$

h) $\exists x(Jx \wedge Kx), \exists x \neg Kx, \exists x \neg Jx \therefore \exists x(\neg Jx \wedge \neg Kx)$

i) $La, b \rightarrow \forall x\, Lx, b, \exists x\, Lx, b \therefore Lb, b$

j) $\forall x(Dx \rightarrow \exists y\, Ty, x) \therefore \exists y \exists z\, \neg y = z$

Answers to Practice Exercises from Chapter 31

I. Using the following symbolization key:

Dx: $\_\_\_\_\_x$ is a dog
Ax: $\_\_\_\_\_x$ is an animal
Sx: $\_\_\_\_\_x$ likes samurai movies
Lxy: $\_\_\_\_\_x$ is larger than $\_\_\_\_\_y$
r: Rave
h: Shane
d: Daisy

symbolize the following sentences in FOL:

a) Rave is a dog who likes samurai movies.
   $Dr \wedge Sr$
b) Rave, Shane, and Daisy are all dogs.
   $Dr \wedge Dh \wedge Dd$
c) Shane is larger than Rave, and Daisy is larger than Shane.
   $Lh, r \wedge Ld, h$

d) All dogs like samurai movies.
$\forall x(Dx \rightarrow Sx)$

e) 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?

f) There is a dog that is larger than Shane.
$\exists x(Dx \wedge Lx, h)$

g) If there is a dog larger than Daisy, then there is a dog larger than Shane.
$\exists x(Dx \wedge Lx\mathbf{d}) \rightarrow \exists x(Dx \wedge Lx, h)$

h) No animal that likes samurai movies is larger than Shane.
$\forall x(Sx \rightarrow \neg Lx, h)$

i) No dog is larger than Daisy.
$\forall x(Dx \rightarrow \neg Lx, d)$

j) 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'.

k) There is an animal that is between Rave and Shane in size.
$\exists x((Lb, x \wedge Lx, h) \vee (Lh, x \wedge Lx, r))$

l) There is no dog that is between Rave and Shane in size.
$\forall x(Dx \rightarrow \neg\big[(Lb, x \wedge Lx, h) \vee (Lh, x \wedge Lx, r)\big])$

m) No dog is larger than itself.
$\forall x(Dx \rightarrow \neg Lx, x)$

n) Every dog is larger than some dog.
$\forall x(Dx \rightarrow \exists y(Dy \wedge Lxy))$
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'.

o) There is an animal that is smaller than every dog.
$\exists x\forall y(Dy \rightarrow Ly, x)$

p) If there is an animal that is larger than any dog, then that animal does not like samurai movies.
$\forall x(\forall y(Dy \rightarrow Lxy) \rightarrow \neg Sx)$
Comment: I have assumed that 'larger than any dog' here means 'larger than every dog'.

II. Using the symbolization key given, translate each English-language sentence into FOL.

Cx: _____$_x$ is a candy
Hx: _____$_x$ has chocolate in it.
Mx: _____$_x$ has marzipan in it.
Sx: _____$_x$ has sugar in it.
Tx: Boris has tried _____$_x$.
Bxy: _____$_x$ is better than _____$_y$.

a) Boris has never tried any candy.
b) Marzipan is always made with sugar.
c) Some candy is sugar-free.
d) The very best candy is chocolate.
e) No candy is better than itself.
f) Boris has never tried sugar-free chocolate.
g) Boris has tried marzipan and chocolate, but never together.
h) Any candy with chocolate is better than any candy without it.
i) Any candy with chocolate and marzipan is better than any candy that lacks both.

Answers to Practice Exercises from Chapter 32

I. Each of the following sentences is ambiguous. Provide a symbolization key for each, and symbolize all readings.

a) Noone likes a quitter.
b) CSI found only red hair at the scene.

Answers to Practice Exercises from Chapter 33

I. 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

II. The following three proofs are missing their citations (rule and line numbers). Add them, to turn them into bona fide proofs.

1	$\forall x \exists y (Rxy \vee Ryx)$	
2	$\forall x \neg Rmx$	
3	$\exists y (Rmy \vee Rym)$	
4	$Rma \vee Ram$	
5	$\neg Rma$	
6	Ram	
7	$\exists x Rxm$	
8	$\exists x Rxm$	

1	$\forall x (\exists y Lxy \rightarrow \forall z Lzx)$	
2	Lab	
3	$\exists y Lay \rightarrow \forall z Lza$	
4	$\exists y Lay$	
5	$\forall z Lza$	
6	Lca	
7	$\exists y Lcy \rightarrow \forall z Lzc$	
8	$\exists y Lcy$	
9	$\forall z Lzc$	
10	Lcc	

1		$\forall x(Jx \rightarrow Kx)$
2		$\exists x \forall y Lxy$
3		$\forall x Jx$
4		$\forall y Lay$
5		Laa
6		Ja
7		$Ja \rightarrow Ka$
8		Ka
9		$Ka \wedge Laa$
10		$\exists x(Kx \wedge Lxx)$
11		$\exists x(Kx \wedge Lxx)$

III. 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 '$\forall x(Fx \rightarrow \neg Gx)$'.

IV. 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.

V. Provide a proof of each claim.

a) $\vdash \forall x Fx \vee \neg \forall x Fx$
b) $\vdash \forall z(Pz \vee \neg Pz)$
c) $\forall x(Ax \rightarrow Bx), \exists x Ax \vdash \exists x Bx$
d) $\forall x(Mx \leftrightarrow Nx), Ma \wedge \exists x Rxa \vdash \exists x Nx$
e) $\forall x \forall y Gxy \vdash \exists x Gxx$
f) $\vdash \forall x Rxx \rightarrow \exists x \exists y Rxy$
g) $\vdash \forall y \exists x(Qy \rightarrow Qx)$
h) $Na \rightarrow \forall x(Mx \leftrightarrow Ma), Ma, \neg Mb \vdash \neg Na$
i) $\forall x \forall y(Gxy \rightarrow Gyx) \vdash \forall x \forall y(Gxy \leftrightarrow Gyx)$
j) $\forall x(\neg Mx \vee Ljx), \forall x(Bx \rightarrow Ljx), \forall x(Mx \vee Bx) \vdash \forall x Ljx$

VI. 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.

VII. Show that each pair of sentences is provably equivalent.

a) $\forall x(Ax \rightarrow \neg Bx)$, $\neg\exists x(Ax \wedge Bx)$
b) $\forall x(\neg Ax \rightarrow Bd)$, $\forall x Ax \vee Bd$
c) $\exists x Px \rightarrow Qc$, $\forall x(Px \rightarrow Qc)$

VIII. 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.

a) $\forall x Px \rightarrow Qc, \forall x(Px \rightarrow Qc)$
b) $\forall x\forall y\forall z Bxyz, \forall x Bxxx$
c) $\forall x\forall y Dxy, \forall y\forall x Dxy$
d) $\exists x\forall y Dxy, \forall y\exists x Dxy$
e) $\forall x(Rca \leftrightarrow Rxa), Rca \leftrightarrow \forall x Rxa$

IX. 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.

a) $\exists y\forall x Rxy \therefore \forall x\exists y Rxy$
b) $\forall x\exists y Rxy \therefore \exists y\forall x Rxy$
c) $\exists x(Px \wedge \neg Qx) \therefore \forall x(Px \rightarrow \neg Qx)$
d) $\forall x(Sx \rightarrow Ta), Sd \therefore Ta$
e) $\forall x(Ax \rightarrow Bx), \forall x(Bx \rightarrow Cx) \therefore \forall x(Ax \rightarrow Cx)$

f) $\exists x(Dx \vee Ex), \forall x(Dx \rightarrow Fx) \therefore \exists x(Dx \wedge Fx)$
g) $\forall x \forall y(Rxy \vee Ryx) \therefore Rjj$
h) $\exists x \exists y(Rxy \vee Ryx) \therefore Rjj$
i) $\forall x Px \rightarrow \forall x Qx, \exists x \neg Px \therefore \exists x \neg Qx$
j) $\exists x Mx \rightarrow \exists x Nx, \neg \exists x Nx \therefore \forall x \neg Mx$

Answers to Practice Exercises from Chapter 34

I. Show in each case that the sentences are provably inconsistent:

a) $Sa \rightarrow Tm, Tm \rightarrow Sa, Tm \wedge \neg Sa$
b) $\neg \exists x Rxa, \forall x \forall y Ryx$
c) $\neg \exists x \exists y Lxy, Laa$
d) $\forall x(Px \rightarrow Qx), \forall z(Pz \rightarrow Rz), \forall y Py, \neg Qa \wedge \neg Rb$

II. Show that each pair of sentences is provably equivalent:

a) $\forall x(Ax \rightarrow \neg Bx), \neg \exists x(Ax \wedge Bx)$
b) $\forall x(\neg Ax \rightarrow Bd), \forall x Ax \vee Bd$

III. In **??**, we considered what happens when we move quantifiers 'across' various logical operators. Show that each pair of sentences is provably equivalent:

a) $\forall x(Fx \wedge Ga), \forall xFx \wedge Ga$
b) $\exists x(Fx \vee Ga), \exists xFx \vee Ga$
c) $\forall x(Ga \rightarrow Fx), Ga \rightarrow \forall xFx$
d) $\forall x(Fx \rightarrow Ga), \exists xFx \rightarrow Ga$
e) $\exists x(Ga \rightarrow Fx), Ga \rightarrow \exists xFx$
f) $\exists x(Fx \rightarrow Ga), \forall xFx \rightarrow Ga$

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.

Answers to Practice Exercises from Chapter 35

I. Offer proofs which justify the addition of the second and fourth CQ rules as derived rules.

Appendix B

Symbolic notation

2.1 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 chapter 7, 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'.

2.2 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**, '¬', and the **swung dash** or **tilda**, '∼.' There are some issues typing '¬' on a keyboard, and '∼' 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 '¬' and '∼'. 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 '∨' 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**, '∧', 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 ‘$\wedge$’, 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**, '≡', 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	¬, ~
conjunction	∧, &, •
disjunction	∨
conditional	→, ⊃
biconditional	↔, ≡
universal quantifier	$\forall x$, (x)

Appendix C

Quick reference

3.1 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.

3.2 Truth Rules for Connectives in TFL

$\neg$:

T	F
F	T

$\wedge$:

T	T	T
T	F	F
F	T	F
F	F	F

$\vee$:

T	T	T
T	F	T
F	T	T
F	F	F

$\rightarrow$:

T	T	T
T	F	F
F	T	T
F	F	T

3.3 Symbolization

Rough Meaning of the TFL Connectives

symbol	name	rough meaning
¬	negation	'It is not the case that...'
∧	conjunction	'... and ...'
∨	disjunction	'... or ...'
→	conditional	'If ... then ...'
↔	biconditional	'... if and only if ...'

3.4 Symbolisation strategy

1. Can it be paraphrased in accordance with out symbolisation guide? Make sure that the '...' are complete (declarative) sentences.

2. If so, do that and then move to consider how to symbolise the constituant sentences, the '...'.

3. If not, we need to use an atomic sentence.

- If you already have an appropriate atomic sentence in your symbolisation key, use it.
- If not, choose a new atomic TFL sentence.

3.5 Symbolisation guide

Paraphrased as	Symbolised as
It is not the case that ...	$\neg X$
... and ...	$X \wedge Y$
Both ... and ...	$X \wedge Y$
Although ..., ...	$X \wedge Y$
... but ...	$X \wedge Y$
... or ...	$X \vee Y$
Either ... or ... (and not both)	$(X \vee Y) \wedge \neg(X \wedge Y)$
Neither ... nor ...	$\neg(X \vee Y)$ or $\neg X \wedge \neg Y$
If $\ldots_X$, then $\ldots_Y$	$X \rightarrow Y$
$\ldots_Y$ if $\ldots_X$	$X \rightarrow Y$
$\ldots_X$ only if $\ldots_Y$	$X \rightarrow Y$
$\ldots_X$ if and only if $\ldots_Y$	$(X \rightarrow Y) \wedge (Y \rightarrow X)$ **(can be shortened to** $X \leftrightarrow$
Unless $\ldots_X$, $\ldots_Y$	$X \vee Y$ or $\neg X \rightarrow Y$ or $\neg Y \rightarrow X$

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
Everything besides c is G	$\forall x(\neg x=c \rightarrow Gx)$
The F is G	$\exists x(Fx \wedge \forall y(Fy \rightarrow x=y) \wedge$
It is not the case that the F is G	$\neg\exists x(Fx \wedge \forall y(Fy \rightarrow x=y) \wedge$
The F is non-G	$\exists x(Fx \wedge \forall y(Fy \rightarrow x=y) \wedge \neg$

3.6 Using identity to symbolize quantities

There are at least _____ Fs.

one: $\exists xFx$

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: $\forall x_1 \forall x_2 \left[(Fx_1 \wedge Fx_2) \rightarrow x_1 = x_2 \right]$

two: $\forall x_1 \forall x_2 \forall x_3 \left[(Fx_1 \wedge Fx_2 \wedge Fx_3) \rightarrow (x_1 = x_2 \vee x_1 = x_3 \vee x_2 = x_3) \right]$

three: $\forall x_1 \forall x_2 \forall x_3 \forall x_4 \left[(Fx_1 \wedge Fx_2 \wedge Fx_3 \wedge Fx_4) \rightarrow (x_1 = x_2 \vee x_1 = x_3 \vee x_1 = x_4 \vee x_2 = x_3 \vee x_2 = x_4 \vee x_3 = x_4) \right)$

n: $\forall x_1 \ldots \forall x_{n+1} \left[(Fx_1 \wedge \ldots \wedge Fx_{n+1}) \rightarrow (x_1 = x_2 \vee \ldots \vee x_n = x_{n+1}) \right]$

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: $\exists x\left[Fx \wedge \forall y(Fy \rightarrow x = y)\right]$

two: $\exists x_1 \exists x_2\left[Fx_1 \wedge Fx_2 \wedge \neg x_1 = x_2 \wedge \forall y\big(Fy \rightarrow (y = x_1 \vee y = x_2)\big)\right]$

three: $\exists x_1 \exists x_2 \exists x_3\left[Fx_1 \wedge Fx_2 \wedge Fx_3 \wedge \neg x_1 = x_2 \wedge \neg x_1 = x_3 \wedge \neg x_2 = x_3 \wedge \forall y\big(Fy \rightarrow (y = x_1 \vee y = x_2 \vee y = x_3)\big)\right)$

n: $\exists x_1 \ldots \exists x_n\left[Fx_1 \wedge \ldots \wedge Fx_n \wedge \neg x_1 = x_2 \wedge \ldots \wedge \neg x_{n-1} = x_n \wedge \forall y\big(Fy \rightarrow (y = x_1 \vee \ldots \vee y = x_n)\big)\right)$

3.7 Basic deduction rules for TFL

Reiteration

m	*X*	
	X	R, *m*

Rules for connectives – introduction and elimination

	Introduction	Elimination
$\wedge$	$\begin{array}{l\|l} m & X \\ n & Y \\ & X \wedge Y \qquad \wedge\text{I}, m, n \end{array}$	$\begin{array}{l\|l} m & X \wedge Y \\ & X \qquad \wedge\text{E}, m \\ m & X \wedge Y \\ & Y \qquad \wedge\text{E}, m \end{array}$
$\rightarrow$	$\begin{array}{l\|l\|l} m & & X \\ & & \vdots \\ n & & Y \\ & X \rightarrow Y & \rightarrow\text{I}, m\text{–}n \end{array}$	$\begin{array}{l\|l} m & X \rightarrow Y \\ n & X \\ & Y \qquad \rightarrow\text{E}, m \end{array}$
$\neg$	$\begin{array}{l\|l\|l} m & & X \\ & & \vdots \\ k & & Y \wedge \neg Y \\ & \neg X & \neg\text{I}, m\text{–}k \end{array}$	$\begin{array}{l\|l\|l} m & & \neg X \\ & & \vdots \\ k & & Y \wedge \neg Y \\ & X & \neg\text{E}, \end{array}$
$\vee$	$\begin{array}{l\|l} m & X \\ & X \vee Y \qquad \vee\text{I}, m \\ m & Y \\ & X \vee Y \qquad \vee\text{I}, m \end{array}$	$\begin{array}{l\|l\|l} m & X \vee Y & \\ i & & X \\ & & \vdots \\ j & & Z \\ k & & Y \\ & & \vdots \\ l & & Z \\ & Z & \vee\text{E}, m, i\text{–}j, k\text{–}l \end{array}$

3.8 Derived rules for TFL

Disjunctive syllogism

m	$X \vee Y$	
n	$\neg X$	
	Y	DS, m, n

m	$X \vee Y$	
n	$\neg Y$	
	X	DS, m, n

Double-negation

m	$\neg\neg X$	
	X	DNE, m

m	X	
	$\neg\neg X$	DNI, m

Law of Excluded Middle

	$X \vee \neg X$	LEM

Modus Tollens

m	$X \to Y$	
n	$\neg Y$	
	$\neg X$	MT, m, n

Explosion

m	X	
n	$\neg X$	
	Y	Explosion, m, n

De Morgan Rules

m	$\neg(X \vee Y)$	
	$\neg X \wedge \neg Y$	DeM, m

m	$\neg X \wedge \neg Y$	
	$\neg(X \vee Y)$	DeM, m

m	$\neg(X \wedge Y)$	
	$\neg X \vee \neg Y$	DeM, m

m	$\neg X \vee \neg Y$	
	$\neg(X \wedge Y)$	DeM, m

3.9 Basic deduction rules for FOL

	Introduction	Elimination
$\forall$	$m \mid X[c/v]$ $\quad \forall v X \quad \forall\text{I}, m$ c must not occur in any premise or undischarged assumption	$m \mid \forall v X$ $\quad X[c/v] \quad \forall\text{E}, m$
$\exists$	$m \mid X[c/v]$ $\quad \exists v X \quad \exists\text{I}, m$	$m \mid \exists v X$ $i \mid\quad X[c/v]$ $\quad\quad \vdots$ $j \mid\quad Y$ $\quad Y \quad \exists\text{E}, m, i\text{–}j$ c must be new to the proof: it must not occur in any line before i; c must not occur in Y.

3.10 Derived rules for FOL

m	$\forall x \neg X$		m	$\exists x \neg X$	
	$\neg \exists x X$	ConvQuant, m		$\neg \forall x X$	ConvQuant, m

m	$\neg \exists x X$		m	$\neg \forall x X$	
	$\forall x \neg X$	ConvQuant, m		$\exists x \neg X$	ConvQuant, m

3.11 Identity rules

Identity introduction

	$c = c$	= I

Identity elimination

m	$a = b$		m	$a = b$	
n	$X(\ldots a \ldots a \ldots)$		n	$X(\ldots b \ldots b \ldots)$	
	$X(\ldots b \ldots a \ldots)$	= E, m, n		$X(\ldots a \ldots b \ldots)$	= E, m, n

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.