FORMAL LOGIC

Abridged edition for PY1012
Spring 2017

by

STEPHEN READ

and

CRISPIN WRIGHT

University of
St Andrews

last updated 06/02/2017

Contents

Prefaces

Authors' Preface to Part I

This is the first of three booklets whose aim is to give an introduction to elementary logic — to the theory of classical first-order logic, which arose in the work of Boole and Frege in the nineteenth century, and was completed in the first few decades of this century by Russell, Post, Gentzen and others.

The focus of this part is on the semantics of classical logic, a theory of truth-functions extended by a theory of interpretations for predicates and quantifiers. We start with propositional, or zero-order logic, and show how to represent simple arguments in a formal way. We describe truth-tables and then set out the diagrammatic representation as semantic trees which they provide of the truth-functional semantics of propositional logic.

Moving onto the full power of first-order logic, with its analysis of terms, predicates and quantifiers, we explain the formal representation of arguments as sequents of predicate logic. We then show how intuitively to extend our semantics to the first-order theory of quantifiers. We add to this informal semantics an account of how the method of semantic trees can be straightforwardly extended from the simpler propositional logic theory to the case of first-order predicate logic.

We would like to record here our thanks to our colleagues Peter Clark and Bob Hale, for their helpful comments; to many generations of students, who have responded to changing versions of these notes; and to a succession of typists, not least Janet Kirk and Anne Cameron, for their indispensable contribution.

Stephen Read
Crispin Wright
December 1994

A few typographical mistakes have been rectified in this reprinting, and the earlier terminology of 'sound' for instances of valid argument-forms (deriving from Lemmon) has been replaced by the more usual 'valid'.

Stephen Read
December 1995

Authors' Preface to Part II

This is the second of three booklets whose aim is to give an introduction to elementary logic — to the theory of classical first-order logic, which arose in the work of Boole and Frege in the nineteenth century, and was completed in the first few decades of this century by Russell, Post, Gentzen and others.

The focus of this part is the proof theory of classical logic, in particular, on a natural deduction presentation of the rules of inference. These rules stem from work by Gerhard Gentzen in the 1930s, separating the proof rules into pairs of rules for each connective, one rule, the introduction-rule, giving the grounds for asserting a wff containing the connective as main connective, the other, the elimination-rule, using that information to draw further inferences. This separation property, of separate rules for each connective, is most clearly expressed in Gentzen's multiple-conclusion sequent calculus; but even in its natural deduction form for classical logic it creates a symmetry and elegance which is attractive.

We set out these paired rules of inference for the conditional, conjunction and disjunction, and then deal with the case of negation, where three rules are needed. This provides the proof theory for propositional logic. We then describe how what is in effect a form of Gentzen's sequent calculus can be used as a formalisation of strategy and tactics in the construction of proofs. This strategy is implemented in the computer program MacLogic, written in St Andrews during a project funded by the Computers in Teaching Initiative in 1987-89, and extended and improved since that time. (The program is used for formal teaching from this booklet. Further information on the availability of the program to other sites is available from Dr R. Dyckhoff, Division of Computer Science, University of St Andrews: rd@cs.st-and.ac.uk.) Comparison is also made between these tactics and the method of semantic trees introduced in Part I of this series of booklets.

The proof theory of propositional logic is later rounded out by discussion of the biconditional, of the paradoxes of material implication, and of the method of sequent-introduction. Finally, the rules are extended by pairs of rules for each of the quantifiers, universal and existential, so that by the end of the book students have a complete method of proof for first-order classical predicate logic.

We would like to record here our thanks to our colleagues Peter Clark and Bob Hale, for their helpful comments; to many generations of students, who have responded to changing versions of these notes; and to a succession of typists, not least Janet Kirk and Anne Cameron, for their indispensable contribution.

Stephen Read
Crispin Wright
December 1994

Editor's Preface

A few more typographical mistakes have been rectified in this re-typesetting, although new ones may have been introduced inadvertently. If you notice any mistakes, please contact me. Some formatting styles also have been changed. I have mostly avoided underlining, although this was frequently used in the original text. Instead, technical uses of terms and terms that are being defined are now typeset in **bold**. Emphasis is indicated with *italics*. Stipulations in translation keys are made using '$\triangleright$', e.g.: $Fx \triangleright x$ is wise. Finally, I have omitted some portions of the text. This is noted where it occurs, except where I have deleted some bracket-elimination conventions (adjusting formulae throughout the text accordingly).

This document is typeset in LaTeX using Garamond for plain text (with the garamondx package) and using Euler for most logical and mathematical symbols (with the eulervm package). The qtree and nd3 packages are used extensively for truth-trees and natural deduction, respectively.

Ephraim Glick
February 2016

PART I

FORMS AND SEMANTICS

1 What is formal logic?

What is logic? In ordinary speech, “logical” thought is rationally persuasive thought. And logic might correspondingly be defined as the science of reasoning, of rational persuasion. Consider the following examples:

(1) No virus can reproduce outside the environment of a plant or animal cell. The common cold is caused by a virus. So the common cold-causing virus cannot reproduce outside the environment of a plant or animal cell.

(2) No-one who knowingly and avoidably endangers their health is rational. All who consume large amounts of alcohol knowingly and avoidably endanger their health. So no-one who drinks large amounts of alcohol is rational.

(3) Violets are purple; some of the flowers I saw were purple. So some of the flowers I saw were violets.

(4) No animals are intelligent. All humans are intelligent. So no humans are animals.

These are all **arguments**: each of them has two premises (an incidental feature) and a conclusion, which is what they aim to persuade you of. Logic, in the ordinary sense referred to, is concerned with the question of which arguments ought to be regarded as persuasive. Which of (1)-(4) come into that category? Well, notice that in order to decide, we have to ask two different questions: are the premises **true**; and does the conclusion **follow from** the premises? Only if we may take it that the answer to both questions is ‘yes’ is an argument rationally persuasive.

What does ‘follow from’ mean here? This is a fundamental notion for the philosophy of logic whose clarification is a difficult and profound issue. We shall by the end of the course have clarified at least one explicit answer, applicable to a wide class of arguments. For the moment however, we can gesture at the notion only by example. Thus in (1), (2) and (4) of the above list, the conclusion does follow from the premises; and in (3) it does not. Why not? Because in (3) we can tell a story of how it could be that the premises were all true, but the conclusion false. (Make sure you can see how that could be so.) But in (1), (2) and (4) there is no such story to tell.

Roughly, then: a conclusion follows from a set of premises if it is impossible that the premises should be all true, yet that conclusion false. (Logicians are accustomed to using various terminological alternatives to saying that a conclusion follows from a certain set of premises: e.g. that the premises **entail** the conclusion, that the conclusion is a **logical consequence** of the premises, that the **inference** from the premises to the conclusion is **valid**, that the **argument** from the premises to the conclusion is valid.)

Notice, however, that the idea of **impossibility** to which we are appealing is a very strong one: when a conclusion follows from a set of premises, the impossibility that it should be false while they are true does not depend upon the laws of science, or anything like that — it is not at all like the impossibility of long-jumping the Tay at Dundee or building a perpetual motion machine. The idea is essentially that the falsity of the conclusion implicitly **contradicts**, or is **inconsistent with**, the idea that the premises are all true. (Spell out that implicit contradiction in the examples (1), (2) and (4) from the above list.) Our preferred terminology will be that an argument is **valid** if the conclusion follows from the premises.

We should now modify the suggestion that logic is properly concerned with the theory of **rationally persuasive** argument: for rationally persuasive arguments are valid arguments with nothing but true, or reasonably acceptable, premises, and whether the premises of a given argument are true, or reasonably acceptable, is in general not to be decided by logic but by doing natural science or just ordinary empirical investigation. (Logicians often call valid arguments with true premises, **sound** arguments.) Thus (1) and (2) above are, but (3) and (4) are not, rationally persuasive. (3) is not rationally persuasive because it is not valid, and (4) is not so because ...? But the truth of the premises of (1) and (2) is not something logic can settle. All that logic can do is to say, given certain premises, what follows from them.

So: logic is concerned with truth, but only by way of study of the kinds of arguments that *preserve* truth, that transmit truth from the premises to the conclusion i.e. with valid arguments, where the conclusion follows from the premises.

Notice that valid arguments can have false premises and true conclusions and false premises and false conclusions. Thus, for example:

> George Bush is Welsh; all Welshmen are over 6′ tall; therefore George Bush is over 6′ tall

has false premises and a true conclusion, whereas

> George Bush is Welsh; all Welshmen are good rugby players; therefore George Bush is a good rugby player

has false premises and a (presumably) false conclusion. Also, valid arguments can have mixed premises (i.e. some of the premises true and some of the premises false) and false conclusions. Thus:

> George Bush is American; all Americans are Democrats; therefore George Bush is a Democrat.

Again, valid arguments can have mixed premises and true conclusions. For example:

> George Bush is American; all Americans are Republicans; therefore George Bush is a Republican.

And of course, they can have true premises and a true conclusion:

> George Bush is American; all Americans are entitled to free speech; therefore George Bush is entitled to free speech.

WHAT VALID ARGUMENTS CANNOT HAVE IS TRUE PREMISES AND A FALSE CONCLUSION

Why? Because, according to our intuitive account, a valid argument is precisely one where it is impossible that the premises should be all true and the conclusion be false. Notice too that an argument can be invalid even though its premises and conclusion are all true. For example:

> Ronald Reagan is American; some Americans are of Irish descent; therefore Ronald Reagan is of Irish descent.

Why is this invalid? Because the premises *could be* true while the conclusion was false — the falsity of the conclusion is consistent with the truth of the premises. Consider these examples:

(i) Either I shall contract heart disease or I shall not. If I do, then all the precautions I have taken have been ineffectual. If I do not, then all the precautions I have taken have been unnecessary. If all the precautions I have taken have been either ineffectual or unnecessary, then I have been wasting my time taking them. Therefore, I have been wasting my time taking them.

(ii) Either Catherine is at home or she is not at home. If she is then she did not get my message at work. If she isn't, then she won't have seen the note I left on the front door. If she either did not get my message at work or hasn't seen the note I left on the front door, she won't be expecting me. Therefore, she won't be expecting me.

Let us note some facts about both these arguments. They are both valid (though hardly rationally persuasive). And it is striking that despite their very different subject matter, they have something in common. What they share is their **form**. We can exhibit the form of both arguments by first schematising the sentences they contain like this:

Argument (i)

P	$\triangleright$	I shall contract heart disease
Q	$\triangleright$	All the precautions I have taken have been ineffectual
R	$\triangleright$	All the precautions I have taken have been unnecessary
S	$\triangleright$	I have been wasting my time taking them

Argument (ii)

P	$\triangleright$	Catherine is at home
Q	$\triangleright$	She did not get my message at work
R	$\triangleright$	She won't have seen the notice I left on the front door
S	$\triangleright$	She won't be expecting me

The common form of the two arguments is then naturally presented as:

(iii) P or not P; if P, then Q; if not P then R; if Q or R, then S; therefore S.

Now clearly, what *makes* these two arguments valid is certainly not anything to do with their respective subject matters. Rather it is has to do with their form, as exhibited. One can intuitively see that if the premises are true, it is quite impossible for the conclusion to be false simply because of the form they take. Let us take two further examples. Consider :

(iv) If the points are damp, no spark will be generated. If no spark is generated, the engine will not start. Therefore, the engine will start only if the points are dry.

(v) If I am snowed in, I will miss classes. If I miss classes the course will get behind schedule. Therefore, the course will stay on schedule only if I am not snowed in.

Here the schematisation may go as follows:

Argument (iv)

P	▷	The points are damp
Q	▷	No spark will be generated
R	▷	The engine will not start

Argument (v)

P	▷	I am snowed in
Q	▷	I will miss class
R	▷	The course will get behind schedule

The common form is then:

(vi) If P, then Q; if Q then R; therefore not R only if not P.

Again, despite the difference in subject matter, these two arguments are both valid and their validity depends upon their common form. No matter what we write in for P, Q or R — as long as we do so consistently — the resulting argument will be valid. The argument-form exhibited by (vi) has the property that, in any instance of it, it is impossible for the premises to be true and the conclusion to be false.

This idea is crucial. To fix it, take the following very simple example:

(vii) If it is wet, then it is raining. It is wet. Therefore, it is raining.

Here, the form of the argument may be exhibited by:

If P then Q; P; therefore Q.

Clearly, (vii) is a valid argument. But it is intuitively evident that its validity depends not upon its subject matter (the connection between wetness and rain) but simply upon its having the form (viii). If the first premise is true, then whenever P holds Q does; but the second premise affirms precisely that P holds; so Q must hold also.

The property on which we are focussing — being a valid argument whose validity depends only on its form — is a feature of a large class of valid arguments. It might be expected that the validity or otherwise of an argument will *always* depend upon its form, but this is not so. There are numerous examples of arguments which are valid but which are not valid in virtue merely of form. For example:

My hat is red. Therefore, my hat is coloured.

This is valid; it is certainly a conceptual impossibility that the conclusion should be false while the premise be true. But the form of the argument is simply "P, therefore Q", and there is obviously no shortage of instances of that form which are invalid!

We can now introduce our preferred use of the notion of **validity**. We shall treat validity as a property of argument forms. We say that an argument is of a valid form just in case *every* argument of the same form is valid. Validity is therefore a function of form. And **formal logic**, unsurprisingly, is concerned essentially with the study of valid forms.

We shall begin our study of formal logic with examination of those arguments whose form may be exhibited like those above, in terms of sentential clauses and expressions — **connectives** — which may be applied to them like "if ... then ...", "... and ...", "... or ...", "it is not the case that ...". Later on in the course we shall extend our investigation of logical form into that class of arguments whose form cannot be exhibited using only sentences and these connectives. But initially we shall be concerned only with arguments which can be formalised within **sentential logic** (or **propositional logic** as we are calling it in this book).

The first ability you will require, then, is to be able to represent the form of an argument. This is very much an acquired skill. There are no completely effective mechanical rules for formalisation, because it is often necessary to appeal to one's intuitive understanding of sentences and phrases in order to recognise negation and synonyms. But the basic move is, as illustrated, to look for recurrent clauses and to replace them with an appropriate type of schematic letter. In essence, it is a matter of common sense and, in the case of valid arguments of doing justice to one's intuitions about what the validity of that particular argument depends upon.

A final remark on the relation between form and validity. We cannot say that any argument of the same form as a valid argument must be valid, because form is not an absolute notion. One and the same argument can be represented as having several different forms — some valid, some not so — depending upon the depth of our formal analysis — the extent to which we try to render the form explicit. For example, it would not be incorrect, though it would be pointlessly inexplicit to represent the common form of arguments (i) and (ii) as

(iii*) P; Q; R; S; therefore T,

in which the fact that each premise differs from the others and from the conclusion gets the crudest possible reflection via the selection of five distinct propositional (or sentence) letters and no attempt is made to capture the features of the internal structure of their premisses and conclusion on which the validity of the two examples depends.

What we can say is that, once we have displayed the validity of an argument by displaying it as being of a particular form, then any argument of *that* form is likewise valid.

Logic is a subject which people sometimes experience difficulty in getting into. Don't be disheartened if matters seem complex and unfamiliar in the early stages. The chapters which follow will steer you carefully through the issues involved and give you plenty of exercises in order to enable you to become fluent at formalisation and the other essential skills which build upon it.

2 Formalisation

We can now begin a more systematic study of propositional logic. We shall divide the vocabulary of our theory into two parts. The **logical vocabulary** will consist of a number of connectives such as "… and …", "if … then…", "… or …", etc. The remaining vocabulary of the theory will consist of a list of letters 'P', 'Q', 'R', 'S', …, which will schematise individual sentences. In fact, the letters 'P', 'Q', 'R', etc. may be thought of as standing to particular sentential clauses much as 'x', 'y', 'z', etc., stand to particular decimal numerals — '0', '3', '25', etc — in schoolroom algebra.

Now consider the following argument:

> If terrorism increases, public opinion will move to the right. If public opinion moves to the right, there will be legislation effecting a major increase in police powers. If there is legislation effecting a major increase in police powers, the rights of ordinary citizens will be further eroded. Therefore the rights of the ordinary citizen will be maintained only if terrorism does not increase.

Let us try to formalise this argument. Our rule of thumb is to locate the recurrent clauses, together with any that do not recur, and then to locate explicit or implicit negations. The recurrent clauses are

'Terrorism increases'
— which we represent by 'P'
'Public opinion will move to the right'
— which we represent by 'Q'
'There will be legislation effecting a major increase in police powers.'
— which we represent by 'R'
'The rights of ordinary citizens will be eroded'
— which we represent by 'S'

Moreover 'the rights of the ordinary citizen will be maintained' is an implicit negation of 'The rights of ordinary citizens will be further eroded', and may therefore be represented by 'not R'. So the interim result is:

If P then Q, if Q then R, if R then S $\therefore$ not S only if not P.

(Note the occurrence of ' $\therefore$ '; we shall use it as roughly an equivalent to the English 'therefore', or 'so', to mark the actual drawing of a conclusion.)

That gives the basic skeleton of our argument which, however, still retains a number of English expressions. And these also come in for symbolic re-expression. The logical vocabulary with which we shall be working in general includes the items in the table below.

Logical notion	*English expression*	*Logical Symbol*
The Conditional	if … then …	$\rightarrow$
Conjunction	… and …	&
Disjunction	… or …	$\vee$
Negation	it is not the case that	$\neg$

Table 2.1: Logical vocabulary and symbols

The symbols '$\rightarrow$', '&', '$\vee$', and '$\neg$' are sometimes known as **logical constants** — because their interpretation is held constant when we analyse sentences and arguments into their logical form. But we shall continue to refer to them as logical connectives since they mostly enable us to connect sentential clauses together to make new sentences. Using logical symbols for the relevant connectives we can now give the logical form of the above example as follows:

$$P \rightarrow Q,\ \ Q \rightarrow R,\ \ R \rightarrow S \therefore \neg S \text{ only if } \neg P.$$

Likewise, the logical forms of arguments (iii) and (vi) in Ch. 1, p. 7 are

$$P \vee \neg P,\ \ P \rightarrow Q,\ \ \neg P \rightarrow R,\ \ (Q \vee R) \rightarrow S \therefore S$$

and

$P \to Q,\ Q \to R \therefore \neg R$ only if $\neg P$.

We are now on the way to an algebra of argumentation, a pure science of forms of valid reasoning.

English, and indeed all natural languages, abound with expressions of the types which we are calling connectives, that is, devices for making new sentences out of existing sentences. For example, 'it is doubtful whether …', 'X knows that …', 'it is possible that …', 'although, …' etc. Most of these are of no interest to the logician, because they do not generate distinctive patterns of valid argument, and so do not count as **logical** connectives. But others besides the four we listed are associated with their own distinctive logical powers. Thus, for example, 'unless' can indeed be regarded as a logical connective. Let us see it in operation. Here is a simple case:

> Unless it snows, I shall go for a walk. It is not going to snow. Therefore I shall go for a walk.

We may compare this example with:

> I shall go for a walk only if it is fine. It is not fine. Therefore, I shall not go for a walk.

Do we need to introduce new special symbols for the connectives 'unless', and 'only if'? The answer is no. For we can define these and interestingly many other logically significant connectives in terms of the list of connectives which we already have. Thus we have the translations in the table below.

(For the last, 'P if and only if Q', we often write '$P \leftrightarrow Q$'. '$\leftrightarrow$' is called the biconditional. Logicians often abbreviate 'if and only if' to 'iff'.)

It is of course true that a number of these expressions, in particular 'but' and 'although' and 'unless', may carry a different *rhetorical* force from the logical paraphrase in the right-hand column. For instance, 'P although Q' generally carries a suggestion that the truth of Q might be expected to work against that of P, an implication quite missing from the logical conjunction of P and Q. Nevertheless, in respect of basic logical properties — what follows from what — the translations provided in the right-hand column are accurate enough. Notice in particular the paraphrase of 'only if': putting an 'only' in front of the 'if' has the effect of reversing, so to speak, the direction of the arrow. In terms of the

English Connectives	*Logical Symbol*
P but Q	$P \& Q$
P although Q	$P \& Q$
Unless P, Q	$\neg P \to Q$
Only if Q, P	$P \to Q$
P, provided that Q	$Q \to P$
P if and only if Q	$(P \to Q) \& (Q \to P)$

Table 2.2: Connectives

terminology of **necessary** and **sufficient** conditions, 'if P then Q' says that P is sufficient for Q; and 'only if P, Q' says that P is necessary for Q. We can now complete the formalisation of the terrorism example as:

$$P \to Q,\ Q \to R,\ R \to S\ \therefore\ \neg S \to \neg P$$

Consider one more example:

> Either the sun will shine and the match will be played, or it will be overcast and the light will be bad. If the light is bad the match will not be played. If the match is played England will not win; England will square the series only if they win. Unless the match is played England will not square the series. Therefore, England will not square the series.

The recurrent clauses may be schematised as:

P	$\rhd$	the sun will shine
Q	$\rhd$	the match will be played
R	$\rhd$	the light will be bad
S	$\rhd$	England will win
T	$\rhd$	England will square the series

And, replacing the connectives by their logical symbols, the resulting formalisation is

$$(P \& Q) \vee (\neg P \& R),\ R \to \neg Q,\ Q \to \neg S,\ T \to S,\ \neg Q \to \neg T\ \therefore\ \neg T.$$

We should note here a point about negation. Consider this example:

> If John hurries, Mary won't. So if Mary hurries, John will not.

How should we formalise the premise? Suppose we take as our key:

P	$\triangleright$	John hurries
Q	$\triangleright$	Mary won't,

and represent the premise as '$P \to Q$'? How should we then represent the conclusion? Does it have the form '$\neg Q \to \neg P$'? No. For '$\neg Q$' then represents 'It's not the case that Mary won't (hurry)'. And this is not strictly the same as 'Mary will hurry', or 'Mary hurries' True, it arguably has the same **logical force**. But it does not *mean* the same. It is a matter of *logical theory* that '$\neg\neg P$' and P are equivalent. And we should not prejudge this point of theory when we formalise. (Indeed, in intuitionistic logic, discussed in Part III, Ch. 21, 'P' and '$\neg\neg P$' are *not* equivalent.) In order correctly to formalise the above example, we should take the key:

P	$\triangleright$	John hurries
Q	$\triangleright$	Mary hurries,

and then give as the form,

$$P \to \neg Q \therefore Q \to \neg P.$$

It may seem to the alert reader that we are still being rather cavalier about tense in this example — and in others in Chapter 1 — making no distinction between 'Mary will hurry' and 'Mary hurries'. Sometimes tense distinctions matter to an argument — and a **tense logic** must be used. But tense logic, though a fascinating subject, is somewhat complex, and beyond the scope of this book. The formalisation above depends for its correctness on the presupposition that, typically used, the sentence, 'If John hurries, Mary won't', implicitly concerns the future throughout, so that it might be clumsily but correctly paraphrased as 'If John *will* hurry, Mary will not hurry'. One can imagine a use for 'If John hurries, Mary won't', in which 'John hurries' is to be understood as couched in the habitual present tense. For that use, the suggested formalisation would be inadequate.

Finally, three matters of terminology: argument patterns represented in a purely formal way as above are known as **sequents**. A sequent gives the structure of a possible argument. And an argument is an instance of a sequent if and only if it can be arrived at by **uniform** substitutions for the sentential variables, where a uniform substitution is one in which the same variable always gets the same sentence substituted for it. The constituents of sequents, such as '$\neg P$' and '$P \to Q$', are called **wffs**, short for 'well-formed formulae'.

Exercises

(1) Formalise the following sentences as explicitly as possible using sentence letters and the symbolic connectives of the logical vocabulary of propositional logic. Give a key to your formalisation.

(a) If Mr Watson is happy, Mrs Watson is unhappy, and if Mr Watson is unhappy, Mrs Watson is unhappy.

(b) A necessary and sufficient condition for the sheikh to be happy is that he has wine, women and song.

(c) Florence goes to the cinema only if a disaster movie is playing.

(d) The bribe will be paid if and only if the Excise officer comes alone and the money is provided in used notes or the Excise officer is not alone and the police have not been informed.

(e) If the paper turns red, then the solution is acid, and the paper only turns red if the solution is acid.

(f) You will get a room provided you have no pets.

(g) If the battery is flat, then the starter is dead, and you won't get the car started unless we push it.

(h) If the State schools lack adequate space, then the private schools, providing that they give a good education, will ease the burden on the State's facilities.

(i) Either taxes are inflationary, but wage increases are not or wage increases are inflationary and monetary policy is inadequate.

(2) Formalise the following arguments as explicitly as possible in sequent notation:

(a) The mother will die unless the doctor kills the child. If the doctor kills the child, the doctor will be causing death. If the mother dies, the doctor will be causing death. So, either way, the doctor will be causing death.

(b) Either the vicar or the butler shot the earl. If the butler shot the earl, then the butler wasn't drunk at nine o'clock. Unless the vicar is a liar, the butler was drunk at nine o'clock. So either the vicar is a liar, or he shot the earl.

(c) If the terrorists are tired, then they are on edge. If the terrorists are armed and on edge, then the hostages are in danger. The terrorists are armed and tired. So the hostages are not in danger

(3) Using the Key given, translate the following wffs into sentences of English:

Key

P	$\triangleright$	John is happy
Q	$\triangleright$	Mary is happy
R	$\triangleright$	All's well with the world
S	$\triangleright$	John is cross

(a) $P \& Q$

(b) $P \vee Q$

(c) $P \rightarrow (Q \rightarrow R)$

(d) $\neg Q \vee S$

(e) $\neg(P \& Q) \vee \neg(R \& S)$

(f) $\neg(P \vee Q)$

3 Truth-Tables

Two questions now arise: first, how can we *characterise* which sequents are valid, that is, have only valid instances? And, second and relatedly, how can we *tell* whether a sequent is valid?

Our intuitive characterisation amounts to this: a sequent is valid if and only if each instance of it is such that it is impossible, in the conceptual sense, for the premises to be true and the conclusion false. But can we always easily tell whether that is so? How about this:

$$(P \to Q) \to R,\ R \vee T,\ S \to P,\ T \to \neg R,\ (S \to (Q \to P)) \vee \neg T \ \therefore\ Q \vee R \text{ ??}$$

To try to determine informally whether this sequent was valid or not would be to invite muddle and confusion. So we need a better test for validity than simple good sense and intuitive reasoning — though whatever we come up with must be grounded in good sense and intuitive reasoning.

A better test will also, hopefully, give us a better characterisation of validity. Obviously we cannot possibly give such a characterisation just by listing all the valid sequents, because even for those just containing four propositional variables and four basic connectives, there are literally infinitely many of them — infinitely many valid sequents. (Can you give a simple demonstration that this is so?)

So, how are we to solve these two problems? The goals of formal logic are ambitious: our target is to construct a formal system in which *all* forms of argument whose validity depends only upon the recurrence of whole clauses and the meanings of 'not', 'and', 'or', 'if .. then', and 'if and only if' (cf. Ch. 1, p. 8) can be established.

In other words: if our system of propositional logic achieves its target, then every valid argument of the appropriate type can be shown to be so by using the methods of propositional logic; and moreover, for any invalid propositional argument, we wish to show that it is invalid — we want an effective method of **disproof**. (Reliance on intuition, or ingenuity in thinking up counterexamples, is not an effective such method — i.e. it cannot be guaranteed success wherever success is possible.)

We will develop two methods in this Part of *Formal Logic*, namely, the method of truth-tables and the method of semantic trees. In Part II, a different approach is taken, namely, the method of proofs. First: truth-tables. To begin with, we give a more exact account of the conditions under which statements built up using '&', '$\vee$', '$\neg$', '$\rightarrow$' and '$\leftrightarrow$' are true or false:

> '$\neg A$' is true if and only if A is false; otherwise false.
>
> '$A \,\&\, B$' is true if and only if both A and B are true; otherwise false.
>
> '$A \vee B$' is true if and only if either A, or B, or both are true; otherwise false.
>
> '$A \rightarrow B$' is true if and only if it is not the case that A is true and B is false; otherwise false.
>
> '$A \leftrightarrow B$' is true if and only if A and B are both true, or both false; otherwise false.

These stipulations exactly determine the truth-conditions, and hence (on one influential philosophical view) the meanings, of our five connectives. We can give equivalent stipulations in tabular form like this:

A	$\neg A$
T	F
F	T

A	B	$A \,\&\, B$
T	T	T
T	F	F
F	T	F
F	F	F

A	B	$A \vee B$
T	T	T
T	F	T
F	T	T
F	F	F

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

A	B	$A \leftrightarrow B$
T	T	T
T	F	F
F	T	F
F	F	T

These should be more or less self-explanatory. To the left in each table we list all the possible combinations of truth-values which the connected wffs can have; to the right, the truth-value, for each of those combinations, of the wff in question.

Using those stipulations we can now evaluate wffs in a very natural way. Here first are some examples of the method of evaluation:

(i) $P \vee Q \to (P \to Q)$

P	Q	(P	∨	Q)	→	(P	→	Q)
T	T	T	T	T	T	T	T	T
T	F	T	T	F	F	T	F	F
F	T	F	T	T	T	F	T	T
F	F	F	F	F	T	F	T	F
		1	2	1	3	1	2	1

(ii) $(P \leftrightarrow \neg Q) \to \neg Q$

P	Q	(P	↔	¬	Q)	→	¬	Q
T	T	T	F	F	T	T	F	T
T	F	T	T	T	F	T	T	F
F	T	F	T	F	T	F	F	T
F	F	F	F	T	F	T	T	F
		1	3	2	1	4	2	1

(iii) $(P \& Q) \vee (P \& \neg Q)$

P	Q	(P	&	Q)	∨	(P	&	¬	Q)
T	T	T	T	T	T	T	F	F	T
T	F	T	F	F	T	T	T	T	F
F	T	F	F	T	F	F	F	F	T
F	F	F	F	F	F	F	F	T	F
		1	3	1	4	1	3	2	1

(iv) $(P \to \neg P) \to P$

P	(P	→	¬	P)	→	P
T	T	F	F	T	T	T
F	F	T	T	F	F	F
	1	3	2	1	4	1

The procedure, it will be apparent, is first to list all possible assignments of the values T (True) and F (False) to the individual variables of the wff in question. If there are n variables, there will be 2^n such assignments. This we do on the left. Next (stage 1) we enter the corresponding columns of Ts and Fs under each occurrence of these variables in the wff. Finally we proceed to evaluate each sub-formula in that wff in accordance with our stipulations for the connectives, building up until we have evaluated the wff itself. (The successive stages of the work thus correspond to the successive stages we would pass through if we were constructing the wff in stages — see Ch. 4, p. 36.) The numbers under the columns indicate the order of successive evaluations — but they are only an expository device and should be omitted from formal work.) Work over these examples and make sure you follow what is happening.

OK. Now let's evaluate a couple of three-variable wffs. First:

(v) $P \to ((Q \vee R) \to (P \to Q))$

P	Q	R	P	→	((Q	∨	R)	→	(P	→	Q))
T	T	T	T	T	T	T	T	T	T	T	T
T	T	F	T	T	T	T	F	T	T	T	T
T	F	T	T	F	F	T	T	F	T	F	F
T	F	F	T	T	F	F	F	T	T	F	F
F	T	T	F	T	T	T	T	T	F	T	T
F	T	F	F	T	T	T	F	T	F	T	T
F	F	T	F	T	F	T	T	T	F	T	F
F	F	F	F	T	F	F	F	T	F	T	F
			1	4	1	2	1	3	1	2	1

Notice first that, since there are three variables, there are $2^3 = 8$, possible truth-value assignments. So the **truth-table** has 8 rows. (Note too the foolproof way of making sure we have listed all the assignments: we "double" the rate at which the Ts and Fs oscillate, so to speak; so that for the first variable they oscillate once, for the second, they oscillate twice, for the next they oscillate four times, and so on. The final evaluation, as before is given in the highest-numbered column, numbered 4.

(vi) $(P \vee (Q \vee R))\,\&\,R$

P	Q	R	(P	∨	(Q	∨	R))	&	R
T	T	T	T	T	T	T	T	T	T
T	T	F	T	T	T	T	F	F	F
T	F	T	T	T	F	T	T	T	T
T	F	F	T	T	F	F	F	F	F
F	T	T	F	T	T	T	T	T	T
F	T	F	F	T	T	T	F	F	F
F	F	T	F	T	F	T	T	T	T
F	F	F	F	F	F	F	F	F	F
			1	3	1	2	1	4	1

We now define a **tautology** of the language of propositional logic as a wff whose evaluation is T for all possible truth-value assignments to its individual propositional variables. Example:

(vii) $P \rightarrow (Q \rightarrow Q)$

P	Q	P	→	(Q	→	Q)
T	T	T	T	T	T	T
T	F	T	T	F	T	F
F	T	F	T	T	T	T
F	F	F	T	F	T	F
		1	3	1	2	1

Likewise an **inconsistency** (sometimes called a "contradiction") of the language of propositional logic is a wff whose evaluation is F for all possible assignments to its individual propositional variables. Thus:

(viii) $(P \to \neg P) \& P$

P	(P	→	¬	P)	&	P
T	T	F	F	T	F	T
F	F	T	T	F	F	F
	1	3	2	1	4	1

We now have an effective method for evaluating wffs. But our goal, recall, was a method for evaluating sequents, so as to recognise the invalid ones. We can easily extend our method to apply to sequents too: we simply evaluate all the assumptions of $A_1, \ldots, A_n \models B$ and its conclusion simultaneously; and if there is any truth-value assignment under which all of $A_1, \ldots, A_n$ come out true while B comes out false, then the sequent is invalid. (Intuitively, the method is exactly what we want; for a valid argument-form is one every instance of which is valid, and a valid argument is precisely one where it is impossible that the premises should all be true but the conclusion false.)

A couple of examples, then:

P	Q	R	P	→	Q,	Q	→	R	⊨	R	→	P	
T	T	T	T	T	T	T	T	T		T	T	T	
T	T	F	T	T	T	T	F	F		F	T	T	
T	F	T	T	F	F	F	T	T		T	T	T	
T	F	F	T	F	F	F	T	F		F	T	T	
F	T	T	F	**T**	T	T	**T**	T		T	**F**	F	⇐
F	T	F	F	T	T	T	F	F		F	T	F	
F	F	T	F	**T**	F	F	**T**	T		T	**F**	F	⇐
F	F	F	F	T	F	F	T	F		F	T	F	
			1	2	1	1	2	1		1	2	1	

We look for any rows where both assumptions are true but the conclusion is false: and there are two such rows — the fifth and the seventh. So the sequent is invalid. Next:

P	Q	R	P	$\rightarrow$	(Q	$\rightarrow$	R)	$\models$	Q	$\rightarrow$	(P	$\rightarrow$	R)
T	T	T	T	T	T	T	T		T	T	T	T	T
T	T	F	T	F	T	F	F		T	F	T	F	F
T	F	T	T	T	F	T	T		F	T	T	T	T
T	F	F	T	T	F	T	F		F	T	T	F	F
F	T	T	F	T	T	T	T		T	T	F	T	T
F	T	F	F	T	T	F	F		T	T	F	T	F
F	F	T	F	T	F	T	T		F	T	F	T	T
F	F	F	F	T	F	T	F		F	T	F	T	F
			1	3	1	2	1		1	3	1	2	1

In this case there are no rows where the assumption is true but the conclusion false. The only row where the conclusion is false is the second; but there the assumption is false also. So the sequent passes the test.

Here are a couple of further examples of our technique for evaluating the validity of propositional logic sequents:

P	$\rightarrow$	Q,	R,	P	$\models$	Q	&	R
T	T	T	T	T		T	T	T
T	T	T	F	T		T	F	F
T	F	F	T	T		F	F	T
T	F	F	F	T		F	F	F
F	T	T	T	F		T	T	T
F	T	T	F	F		T	F	F
F	T	F	T	F		F	F	T
F	T	F	F	F		F	F	F

The sequent is invalid if and only if, under some assignment of Ts and Fs, the conclusion comes out false while the assumptions are all true. But here the only line at which the assumptions are all true is the first; and there the conclusion is also true. So the sequent is valid. Next:

P	→	(Q	∨	R),	¬	Q	⊨	R	→	P	
T	T	T	T	T	F	T		T	T	T	
T	T	T	T	F	F	T		F	T	T	
T	T	F	T	T	T	F		T	T	T	
T	F	F	F	F	T	F		F	T	T	
F	T	T	T	T	F	T		T	F	F	
F	T	T	T	F	F	T		F	T	F	
F	T	F	T	T	T	F		T	F	F	⇐
F	T	F	F	F	T	F		F	T	F	
1	3	1	2	1	2	1		1	2	1	

Here the conclusion is false at lines five and seven. And at line seven both $P \to (Q \vee R)$ and $\neg Q$ are true. So the sequent is invalid.

The full truth-table method for evaluating wffs and sequents is straightforward, but cumbersome. Even eight-line tables are tedious to construct. But it is apparent that we can very often simplify matters considerably. Consider this case:

P	→	Q,	¬	Q	⊨	P	→	R
T	F	F	T	F		T	F	F

Remember: the sequent is invalid if and only if it can be falsified, i.e. we can bring it about that the conclusion is false while the assumptions are all true. Here we have tried to do that directly: the only way $P \to R$ can be false is if P gets T and R gets F; but if $\neg Q$ is to get T, Q must get F; so $P \to Q$ cannot be true under this assignment, since its antecedent gets T and its consequent F. But this is the only assignment under which the conclusion of the sequent gets F; so there is no assignment under which the conclusion gets F while all the assumptions get T. Hence the sequent is valid.

Two more examples:

P	→	Q	&	R	⊨	Q	→	(¬	R	→	¬	P)
T	F	T	F	F		T	F	T	F	F	F	T
8	12	9	11	10		2	1	4	5	3	6	7

(P	→	Q)	∨	(R	→	S)	⊨	(P	→	S)	∨	(R	→	Q)
T	F	F	F	T	F	F		T	F	F	F	T	F	F
8	10	9	14	11	13	12		3	2	4	1	6	5	7

The numerals indicate the order in which assignments of truth-values are made in our attempt to falsify the sequent. Thus, in the first example:

(1) : the conclusion has to get F. Since it is a conditional, that requires

(2) : that Q, its antecedent, gets T and

(3) : that its consequent gets F. This consequent is itself a conditional, so

(4) : its antecedent, ¬R, gets T (whence

(5) : R gets F) while

(6) : its consequent, ¬P, gets F (whence

(7) : P gets T). Hence

(8) : the occurrence of P in the assumption gets T;

(9) : the occurrence of Q in the assumption gets T (cf. stage 2); and

(10) : the occurrence of R in the assumption gets F (cf. stage 5). So

(11) : Q & R gets F, whence

(12) : the assumption gets F.

Thus the only possible way of falsifying the sequent fails. If the conclusion is to be F (false), the assumption must be F too. So it is impossible for the assumption to be T and the conclusion at the same time (i.e., for the same assignment of T and F to P, Q and R) to be false. So the sequent is valid.

Check through the sequence of assignments in the second example for yourself (noting that we have saved ourselves the labour of a table of sixteen lines!).

Compare those examples with what happens if we apply this short method to an invalid sequent:

P	→	Q,	R	→	S	⊨	P	→	S
T	T	T	F	T	F		T	F	F
4	7	6	8	9	5		2	1	3

Here the attempt to falsify the sequent goes mechanically as far as stage 5. But there is then no difficulty in finding an assignment (6) under which the first assumption comes out true (7) and an assignment (8) under which the second assumption comes out true (9). So the attempt succeeds, and the sequent is invalid. (Notice that we have again saved ourselves a sixteen-line table). The counterexample is: P: T, Q: T, R: F and S: F.

It probably won't have escaped your notice that what is making this simplified method of evaluation possible is that we are testing sequents whose *conclusions* can be *false* in only one way, with the result that the attempt to falsify the sequents is able, at least initially, to follow only one path. So in which cases, in general, is the simplified method, as illustrated, appropriate? Well, conjunctions and biconditionals can be false in more than one way; but any negation, disjunction, or conditional of individual propositional variables can be false in only one way — and any disjunction of statements that can be false in only one way can itself be false in only one way. So always check to see if the conclusion meets any of those descriptions; if so, a one-line evaluation of the type illustrated will be possible.

A one-line evaluation will be possible in another type of case: sequents whose *assumptions* can simultaneously be *true* in only one way, e.g.:

P	&	Q,	R	&	S	$\models$	(P	$\to$	R)	&	(Q	$\to$	S)
T	T	T	T	T	T		T	T	T	T	T	T	T
2	1	3	5	4	6		7	9	8	13	10	12	11

Here the conclusion is a conjunction of conditionals which can be false in lots of ways — would in fact be so in twelve of the sixteen lines in a full truth-table. But the only way the assumptions can both be true is if P, Q, R, S are all true; so the question, whether the sequent is valid, is just the question whether the conclusion is true under that assignment. Thus in this kind of case, our attempt to falsify the sequent starts with the truth of the assumptions, rather than with the falsity of the conclusion; and, as the numerals indicate, we work from left to right rather than, as before, from right to left. The sequent is valid.

When will this second short method work? Obviously, at least when all the assumptions of the sequent to be evaluated are conjunctions of individual variables, or conjunctions of other wffs that can be true in only one way; and those, of course, will include the negations of all wffs that can be false in only one way. Here are some further examples:

$\neg$	(P	$\to$	Q),	Q	$\vee$	(R	&	S)	$\models$	R	&	S
T	T	F	F	F	T	T	T	T		T	T	T
1	4	3	5	6	2	8	7	9		10	12	11

— so if the assumptions are true, so is the conclusion — so the sequent is valid.

$\neg$	$(P$	$\vee$	$Q),$	$\neg$	$(\neg$	Q	$\rightarrow$	$\neg$	$(R$	$\leftrightarrow$	$S))$	$\models$	R	$\rightarrow$	S
T	F	F	F	T	T	F	F	F		T				T	
1	5	3	6	2	7	8	4	9		10				11	

— sequent valid. (Notice that we did not need to assign values to R and S.)

$\neg$	$(P$	$\vee$	$Q),$	P	$\&$	R	$\models$	Q	$\rightarrow$	R
T	**F**	F		**T**	T					
1	4	3		5	2					

— sequent valid! We do not need to proceed further, since we have discovered that there is no way of bringing about the simultaneous truth of the assumptions — a fortiori, there is no way of bringing about their truth simultaneously with the falsity of the conclusion.

P	$\&$	$\neg$	$(Q$	$\rightarrow$	$R),$	$\neg$	$(S$	$\rightarrow$	$R)$	$\models$	$\neg$	P	$\leftrightarrow$	$\neg$	S
T	T	T	T	F	F	T	T	F	F		F	T	T	F	T
3	1	4	6	5	7	2	9	8	10		13	11	15	14	12

— sequent valid.

Sometimes it is possible to get by without the labour of a full truth-table, although not in one line. Here is an example:

$(P$	$\&$	$\neg$	$Q)$	$\vee$	$(Q$	$\&$	$\neg$	$P)$	$\models$	P	$\leftrightarrow$	Q
T	T	T	F	T	F	F	F	T		T	F	F
F	F	F	T	T	T	T	T	F		F	F	T
4	10	9	5	12	6	11	8	7		2	1	3

Here there are two ways for the conclusion to be false; but the assumption comes out true either way — sequent invalid. One counterexample: P: T, Q: F.

We noted that, given a wff with n propositional variables, P, Q, R, ..., etc., there will be 2^n ways of assigning truth-values to those variables (so 2^n lines in the full truth-table.) And, for each of these possible overall assignments, the whole wff will either be true or false — so there are 2^{2^n} such possibilities: 2^{2^n}

ways of determining the truth-conditions of the wff. This will be clearer if we consider the case where n = 2, i.e. a wff with two variables. Then, familiarly, there are $2^2 = 4$ ways of assigning truth-values to those variables:

P	Q
T	T
T	F
F	T
F	F

and $2^{2^2} = 2^4 = 16$ ways of writing in a truth table for those assignments. Here they are:

P	Q	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
T	T	T	T	T	T	T	T	T	T	F	F	F	F	F	F	F	F
T	F	T	T	T	T	F	F	F	F	T	T	T	T	F	F	F	F
F	T	T	T	F	F	T	T	F	F	T	T	F	F	T	T	F	F
F	F	T	F	T	F	T	F	T	F	T	F	T	F	T	F	T	F

Some of these columns are familiar; for example, column 2 gives the truth-table for $P \vee Q$, 5 that for $P \rightarrow Q$, 7 that for $P \leftrightarrow Q$, and 8 that for $P \,\&\, Q$. But what about the others — do we have any way, using our original connectives, of expressing these other columns, of stating the propositions whose truth-conditions are stipulated by those columns? If not, classical propositional logic will be importantly incomplete: there will be valid arguments whose validity depends only on the connectives they contain and the recurrence of certain whole propositions — i.e. valid arguments of the type propositional logic was designed to capture — which we cannot even express in our chosen vocabulary. For example, let us write the proposition determined by column 15 as $P \downarrow Q$. That is,

P	Q	$P \downarrow Q$
T	T	F
T	F	F
F	T	F
F	F	T

'$\downarrow$' is called **joint denial**. Then

$$P \downarrow Q \therefore \neg P$$

is, by inspection of the truth-table, valid. So if we can't express $P \downarrow Q$ using our original connectives, there is at least one valid pattern of argument which propositional logic, as we have formulated it, fails to capture.

Well, the answer is that our original vocabulary is not inadequate in this way; indeed it supplies the means not merely to express the sixteen listed propositions in two variables, but all possible purely truth-functional propositions, i.e. propositions whose truth values are determined purely by the truth-values of their constituent propositions. The point is complex to demonstrate in full generality — as we saw, where n such constituent propositions are involved (so n propositional variables) there will be 2^{2^n} such possible truth-functional propositions — so 256 in the case of three constituent variables, and a massive 65,536 in the case of four constituent variables! For illustration, though, here are the sixteen possibilities tabulated above for two variables all expressed using, at most, just '$\neg$', '&', 'P' and 'Q':

(1) $\neg(P \& \neg P)$

(2) $\neg(\neg P \& \neg Q) = P \vee Q$

(3) $\neg(\neg P \& Q) = Q \rightarrow P$

(4) P

(5) $\neg(P \& \neg Q) = P \rightarrow Q$

(6) Q

(7) $\neg(P \& \neg Q) \& \neg(\neg P \& Q) = P \leftrightarrow Q$

(8) $P \& Q$

(9) $\neg(P \& Q)$

(10) $\neg(\neg(P \& \neg Q) \& \neg(\neg P \& Q))$

(11) $\neg Q$

(12) $P \& \neg Q$

(13) $\neg P$

(14) $\neg P \& Q$

(15) $\neg P \,\&\, \neg Q = P \downarrow Q$

(16) $P \,\&\, \neg P$

Thus '¬' and '&' are together **adequate** in the following sense: all possible truth-functional propositions with at most two variables can be expressed in terms of them. (And, as hinted, that remains true if the phrase, "with at most two variables", is omitted. We said that the demonstration is complex — but in fact it isn't all that complex; can you see how it might go?) We call each of these abstract propositions — as, for example, each column in the table on the previous page — **truth-functions**. They are functions which map truth-values to truth-values. For example, '&' is a two-place truth-function mapping pairs of truth-values to a truth-value. Its value for the pair ⟨T,T⟩ is T, for the pair ⟨T,F⟩ is F, and so on.

Other adequate sets of connectives are '¬' with '∨', and '¬' with '→'. Exercise: Verify that these three sets are adequate at least for two-variable truth-functional propositions by expressing the sixteen possible columns in terms of them, in the way illustrated for '¬' and '&' above.

Two further striking points:

First, '↓', as defined above, is adequate all on its own! Verify, by constructing truth-tables, that the first eight columns can be expressed as follows:

(1) : $(P \downarrow (P \downarrow P)) \downarrow (P \downarrow (P \downarrow P))$

(2) : $(P \downarrow Q) \downarrow (P \downarrow Q)$

(3) : $(P \downarrow (Q \downarrow Q)) \downarrow (P \downarrow (Q \downarrow Q))$

(4) : P

(5) : $((P \downarrow P) \downarrow Q) \downarrow ((P \downarrow P) \downarrow Q)$

(6) : Q

(7) : $((P \downarrow Q) \downarrow ((P \downarrow P) \downarrow (Q \downarrow Q))) \downarrow ((P \downarrow Q) \downarrow ((P \downarrow P) \downarrow (Q \downarrow Q)))$

(8) : $(P \downarrow P) \downarrow (Q \downarrow Q)$

What $A \downarrow B$ says is: 'both A and B are false', i.e. 'neither A nor B'. Thus $\neg A$ can be expressed as: $A \downarrow A$. It is therefore a simple matter to construct the remaining columns 9-16 by reflecting that 9 negates 8, 10 negates 7, 11 negates 6, and so on; so that all we have to do is negate the wffs we already have — which we achieve either by doubling (e.g. column 13 will "double" column 4 into $P \downarrow P$) or halving (e.g. column 12 will "halve" column 5 into $(P \downarrow P) \downarrow Q$), depending on whether or not we are transforming a wff which is already "doubled".

Secondly, let column 9 be expressed as '$P \mid Q$', i.e.

P	Q	P | Q
T	T	F
T	F	T
F	T	T
F	F	T

$A \mid B$ thus says: 'either A, or B, is false', i.e. 'either not A or not B'. '$\mid$' is known as Sheffer's Stroke (after the eponymous American logician), or **alternative denial** and, like '$\downarrow$', it is singly adequate. (Again, verify that it is so for the two-variable case.) This fact greatly impressed the early Wittgenstein who, in his *Tractatus Logico-Philosophicus*, believed that all statements of definite sense were either direct unanalysable pictures of reality or (in effect) combinations of such using only the connective, '$\mid$'. If that view were correct, Sheffer's Stroke would be the cement of all analysable discourse.

Exercises

(1) Give truth-table evaluations of the following sequents, setting out the full truth-table. Specify counterexamples where needed:

(a) $P \& Q \therefore P \vee Q$

(b) $P, P \rightarrow Q \therefore P \& Q$

(c) $P, \neg P \therefore Q$

(2) Give truth-table evaluations of the following sequents, using the short method. Specify counterexamples where needed:

(a) $P \rightarrow (Q \vee R), R \rightarrow P \therefore Q \rightarrow P$

(b) $(P \vee S) \rightarrow R, R \& \neg S \therefore P$

(c) $\therefore ((P \rightarrow Q) \rightarrow P) \rightarrow Q$

(3) Formalise the following argument and determine whether or not it is valid by a truth-table evaluation (again, either by a full truth-table or the short method):

> If interest rates rise, businesses will be short of capital. If businesses are short of capital, they will not reinvest. But failure to reinvest is sufficient for diminished competitiveness only if there is economic expansion overseas. So, provided the recession is worldwide, rising interest rates will not lead to diminished competitiveness.

(4) Decide using truth-tables (full or short) which of the following are **tautologies**, which are **contradictions**, and which are **neither**.

(a) $P \rightarrow (Q \rightarrow P)$

(b) $P \leftrightarrow (Q \& \neg P)$

(c) $\neg(P \& \neg Q) \leftrightarrow \neg(P \rightarrow Q)$

(d) $(P \mid P) \mid ((P \mid P) \mid (P \mid P))$

(e) $(P \downarrow P) \downarrow ((P \downarrow P) \downarrow (P \downarrow P))$

4 Semantic Trees for Propositional Logic

It is now time to adopt a rather more rigorous approach to the *grammar* of the language of classical sentential logic. There is, for example, an intuitive difference between

a) $P \rightarrow (Q \vee R)$
b) $P \rightarrow \vee Q \neg ($

and

c) $\leftrightarrow$ P $\vee$ $\rightarrow$ $\neg$

Only the first says anything — even in the rather skeletal sense in which the sentences of sentential logic can be said to say something. Of the other two, b) is gibberish, while c) is just a jumble. How can we define this intuitive difference more precisely?

First we define the primitive vocabulary ("words") of the language. We have

brackets: '(' and ')';

connectives:
one-place (i.e. taking one argument) '$\neg$'
two-place (i.e. taking two arguments) '&', '$\vee$', '$\rightarrow$', '$\leftrightarrow$'

and an infinite stock of

propositional variables: 'P', 'Q', 'R', 'S', 'T', ...

(We need infinitely many variables because there is no limit to the complexity of the arguments which we wish, in principle, to be able to formalise.)

Next we define a **formula** of propositional logic as any **sequence** of elements from the primitive vocabulary. Don't confuse a sequence with a sequent. A sequence is any finite string, that is, a list of elements, a first, a second, and so on. Thus b) above is at least a formula, whereas c) is not, since there is no saying which is, e.g., the third, or the fifth.

But how can we give an exact definition of the formulae which make sense, the class of which a) is illustrative? A natural strategy is the following: intuitively, the ones that make sense are the ones that are built up in a certain way, so that, e.g., '&' is always written *between* two formulae, '$\neg$' always *prefixes* a formula, and so on. So we can define the formulae in question — the **well-formed formulae**, as we shall call them (**wffs** for short) — by characterising the ways in which they can be constructed. Thus:

(1) a formula consisting of an individual propositional variable is a wff;

(2) if A is a wff, then $\neg$A is a wff.

(3) if A and B are both wffs, so is $(A \& B)$.

(4) if A and B are both wffs, so is $(A \vee B)$.

(5) if A and B are both wffs, so is $(A \rightarrow B)$.

(6) if A and B are both wffs, so is $(A \leftrightarrow B)$.

(7) no formula is a wff unless its being so can be shown by 1) - 6).

The technique is thus simply to construct the wff systematically out of its constituents. Notice that a pair of brackets gets introduced with each application of any of 3) - 6), though none are introduced by an application of 2). Thus every wff has an even number of brackets. (We count zero as even.)

We define the **scope** of an occurrence of a particular connective in a wff as the *smallest* wff in which it occurs there. E.g., in a) above, the scope of '$\vee$' is '$Q \vee R$', while the scope of '$\rightarrow$' is the whole formula. We define the **main connective** of a wff to be that occurrence of a connective whose scope is the whole wff.

All that, then, renders more precise certain intuitive ideas with which we have been working. Now, however, having just given exact rules for the formation of wffs, we proceed to give ourselves a license to break them (in a restricted, harmless way) by introducing a **bracket elimination convention**:

The outermost pair of brackets in a wff may always be eliminated.

We now turn to the main business of this chapter: the notion of semantic trees. Recall that an argument is valid if and only if the truth of the premises guarantees the truth of the conclusion. In other words, it is valid if and only if it is impossible for the premises to be true and the conclusion false. An argument-form is valid iff all its instances are valid. Hence an argument-form is valid iff however the letters are interpreted, it is impossible for the premises to be true and the conclusion false. Indeed, the main purpose of exhibiting such particular interpretations (interpreting the basic symbols in our vocabulary) is to show that certain forms of argument are invalid. We put forward the interpretation as a counterexample to the argument's validity, showing that under at least one interpretation the premises are true but the conclusion is not true. It is therefore not the case that the conclusion is true in every model of the premises.

This definition does not, however, provide any systematic procedure for finding such counter-examples. In the last chapter, we developed such a procedure in the construction of a truth-table for the wff(s). We can work through all possible assignments of truth-values to the propositional letters occurring in the wff(s), to see if there is such an assignment (a row) which constitutes a counterexample. But such a procedure is long-winded; as you will remember, the indirect method (despite its name) attacks the problem directly, by supposing there is such a row, and looking to see whether that supposition is sustainable.

Our task now is to develop an alternative approach which can be applied to any wff, to any sequent, and which will also be capable of extension to full predicate logic, which we will meet in the next chapter. It not only duplicates the truth-table procedure we already know, but will be seen to be in essence the same technique.

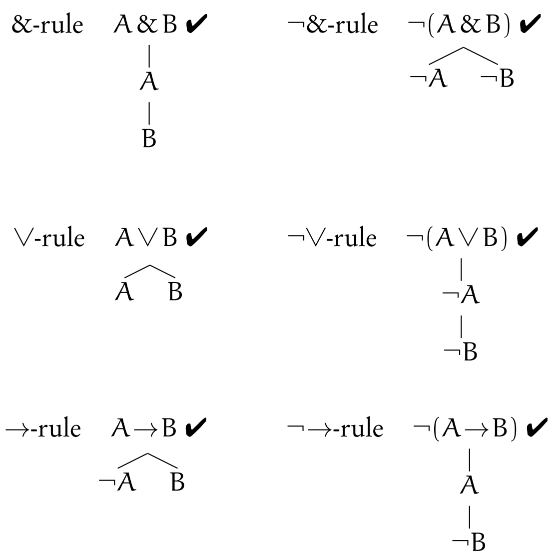
Let us start with a method for finding a model of a single propositional wff, C. That is, we wish to find an interpretation under which C is true, or to show that there is no such interpretation. Our method consists in constructing a tree of wffs with C at its root.

The metaphor of a tree used here may, when you see one, seem perverse: however, the idea is that a tree consists of a set of **nodes** connected by **branches**. The tree with C at its root branches *downwards* from C — the tree is written, one might say, upside down. What is the principle according to which the tree is constructed?

We are trying to find an interpretation under which C is true. We require, therefore, to examine the wff C, and consider how it may be made true. Suppose C has the form ‘A & B’: then C will be true provided both A and B are true. We therefore assign both A and B to the branch leading from C. However, suppose C has the form ‘A ∨ B’: then C will be true provided either A is true or B is true. We therefore assign each of A and B to its own branch leading from the node C. That is, the tree **splits** into two **branches** at a node containing ‘A ∨ B’; but it does not split at a wff ‘A & B’.

C may also have the forms A → B, A ↔ B, ¬(A & B), ¬(A ∨ B), ¬(A → B), ¬(A ↔ B), and ¬¬A. Otherwise C is a propositional letter, or its negation. Our procedure is to continue to generate branches from nodes until such time as each branch in the tree terminates in a propositional letter or its negation. We therefore need to formulate rules for the nine cases of the form of the wff at a node.

We present each rule by showing the one or two branches which lead from a node containing a wff of the respective form. Having constructed the branch or branches leading from a node, we put a tick (✔) against the wff to show that we have dealt with it:



$$\leftrightarrow\text{-rule}\quad \begin{array}{c} A \leftrightarrow B\ \checkmark \\ \begin{array}{cc} A & \neg A \\ | & | \\ B & \neg B \end{array} \end{array} \qquad\qquad \neg\leftrightarrow\text{-rule}\quad \begin{array}{c} \neg(A \leftrightarrow B)\ \checkmark \\ \begin{array}{cc} A & \neg A \\ | & | \\ \neg B & B \end{array} \end{array}$$

$$\text{DN-rule}\quad \begin{array}{c} \neg\neg A\ \checkmark \\ | \\ A \end{array}$$

Suppose we construct the tree for some wff, A: what do we do with it? The reason for constructing the tree was to find a model of A, that is an interpretation for which A is true. The rationale of the tree rules is that they set out the conditions of truth by reducing the question of A's truth to that of the truth of its parts. Suppose therefore that we have completed the tree, so that in each branch every wff which is not a propositional letter or its negation has been checked off. Then two possible cases arise:

(1) Every branch contains at least one wff and its negation; or

(2) At least one branch does not contain any wff together with its negation.

In case 1) we say that each branch **closes**, and we mark each such branch with a cross (✘) at its tip. In case 2) we say that that particular branch is **open**, and from it we read off a model for A as follows:

> Let $\mathscr{I}$ be that interpretation which assigns to each propositional letter occurring in the branch, the value T, and to each propositional letter whose negation occurs in the branch, the value F. Then A is true under $\mathscr{I}$.

Let us work some examples:

(i) $\neg P \mathbin{\&} \neg(P \rightarrow Q)$

Tree:

$$
\begin{array}{c}
\neg P \,\&\, \neg(P \to Q) \checkmark \\
| \\
\neg P \\
| \\
\neg(P \to Q) \checkmark \\
| \\
P \\
| \\
\neg Q \\
\times
\end{array}
$$

The only branch closes, and so there is no interpretation under which $\neg P \,\&\, \neg(P \to Q)$ is true.

(ii) $(\neg P \vee \neg Q) \to (P \,\&\, \neg Q)$

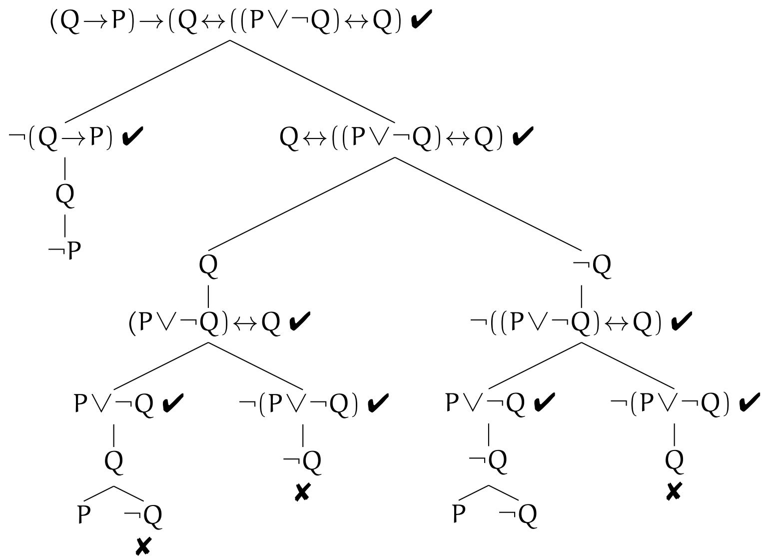
Tree:

$$
\begin{array}{ccc}
& (\neg P \vee \neg Q) \to (P \,\&\, \neg Q) \checkmark & \\
\swarrow & & \searrow \\
\neg(\neg P \vee \neg Q) \checkmark & & P \,\&\, \neg Q \checkmark \\
| & & | \\
\neg\neg P \checkmark & & P \\
| & & | \\
\neg\neg Q \checkmark & & \neg Q \\
| & & \\
P & & \\
| & & \\
Q & &
\end{array}
$$

Both branches are open, and each specifies an interpretation making '$(\neg P \vee \neg Q) \to (P \,\&\, \neg Q)$' true. (Indeed, each branch corresponds to a row of the truth-table assigning T to '$(\neg P \vee \neg Q) \to (P \,\&\, \neg Q)$'.) So taking the left-hand branch, '$(\neg P \vee \neg Q) \to (P \,\&\, \neg Q)$' will be true if we make P and Q both true. For then '$\neg P$' and '$\neg Q$' are both false, so '$\neg P \vee \neg Q$' is false, and so '$(\neg P \vee \neg Q) \to (P \,\&\, \neg Q)$' is true. The right-hand branch shows that '$(\neg P \vee \neg Q) \to (P \,\&\, \neg Q)$' will equally be true if we make P true and Q false.

(iii) $(Q \to P) \to (Q \leftrightarrow ((P \vee \neg Q) \leftrightarrow Q)$

Tree:



Here three branches close, but four remain open. If we take, e.g., the leftmost branch, which is open, we can use it to describe an interpretation under which the wff is true: we require that P be false and Q true; and when P is false and Q is true, the original wff $(Q \to P) \to (Q \leftrightarrow ((P \vee \neg Q) \leftrightarrow Q))$ is true. (Check it.) Or if we take the rightmost open branch, in which Q is false and P arbitrary (false, too, say), then the original wff is true. (Check it.)

We are now in a position to describe the techniques for finding a counterexample to purported claims to logical consequence. Take a sequent:

$$A_1, \ldots, A_n \therefore A$$

If this is invalid, there is an interpretation under which each of $A_1, \ldots, A_n$ is true and A is false, i.e. $\neg A$ is true. To find such an interpretation, we construct a semantic tree with the wffs $A_1, \ldots, A_n, \neg A$ at its root:

$$\begin{array}{c} A_1 \\ | \\ A_2 \\ | \\ \vdots \\ | \\ A_n \\ | \\ \neg A \end{array}$$

Applying the rules, we will either find that every branch closes, or that at least one branch is open. In the first case, it follows that there is no interpretation under which all of $A_1, \ldots, A_n$ and $\neg A$ are true, and so we can infer that the sequent $A_1, \ldots, A_n \therefore A$ is valid. On the other hand, if there is an open branch, then there is an interpretation under which all of $A_1, \ldots, A_n$ and $\neg A$ are true, and so the sequent is invalid — there is a counterexample to its validity. If every branch closes, we write

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

that is, A is a logical (or semantic) consequence of $A_1, \ldots, A_n$.

More examples:

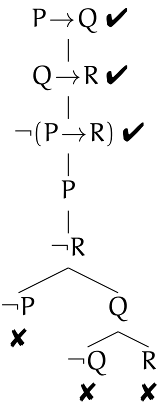
(iv) $P \vee Q, \neg P \therefore Q$

Tree:
$$\begin{array}{c} P \vee Q \\ | \\ \neg P \\ | \\ \neg Q \\ \diagup \quad \diagdown \\ P \quad\quad Q \\ \text{✘} \quad\quad \text{✘} \end{array}$$

Each branch closes, and so there is no interpretation making each of $P \vee Q$, $\neg P$ and $\neg Q$ true; that is, Q is true in every model of $P \vee Q$ and $\neg P$, i.e. $P \vee Q, \neg P \models Q$.

(v) $P \to Q, Q \to R \therefore P \to R$

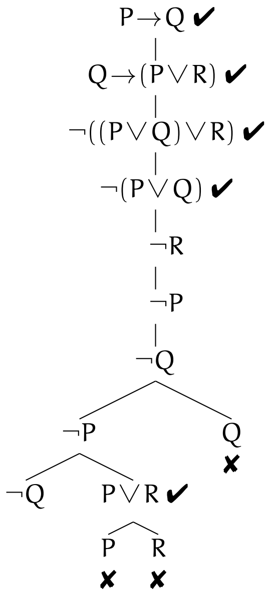
Tree:



Each branch closes, and so there is no interpretation making each of $P \to Q$ and $Q \to R$ and $\neg(P \to R)$ true; that is, $P \to R$ is true in every model of $P \to Q$ and $Q \to R$, i.e. $P \to Q,\ Q \to R \models P \to R$.

(vi) $P \to Q,\ Q \to (P \vee R)\ \therefore\ (P \vee Q) \vee R$

Tree:



Note two things: first, we close the right hand branch before working out all wffs: for we can see it must close, as soon as both Q and $\neg$Q have appeared in it;

secondly, we work out $\neg((P \vee Q) \vee R)$ first, since it does not demand branching. Both these manoeuvres economise on what has to be written.

Three of the branches close; but one does not, and so there is an interpretation I under which the premises are true and the conclusion false, i.e.

$$P \to Q,\ Q \to P \vee R \not\models (P \vee Q) \vee R.$$

A counterexample is specified by making P, Q and R all false.

We say that a set $\{A_1, \ldots, A_n\}$ is **satisfiable** iff $A_1, \ldots, A_n$ have a model. We can now perform a tree on a single propositional wff (to discover whether it is satisfiable), on a set of wffs and the negation of a further wff (to discover whether the latter wff is a logical consequence of the former set of wffs), and on a set of wffs (to discover whether the set is satisfiable).

$\models A$, i.e. A is a **tautology**, iff A is true under every interpretation — that is, iff A is false under no interpretation, in other words, iff $\neg A$ is false under every interpretation. i.e. $\neg A$ is not satisfiable. To discover whether A is a tautology, therefore, we need to perform a tree on $\neg A$. If the tree closes, that is, if no branch is open, then A is a tautology; otherwise it is not a tautology.

E.g., is $((P \to Q) \to P) \to P$ a tautology? We perform a tree on its negation:

(vii) $\neg(((P \to Q) \to P) \to P)$

Tree:

$$\begin{array}{ccc}
 & \neg(((P \to Q) \to P) \to P)\ ✔ & \\
 & | & \\
 & (P \to Q) \to P\ ✔ & \\
 & | & \\
 & \neg P & \\
 \swarrow & & \searrow \\
\neg(P \to Q)\ ✔ & & P \\
| & & ✘ \\
P & & \\
| & & \\
\neg Q & & \\
✘ & &
\end{array}$$

The tree closes, since both branches close. Hence Peirce's Law, $((P \to Q) \to P) \to P$, is a tautology: $\models ((P \to Q) \to P) \to P$.

Another example: is $((P \to Q) \to Q) \to P$ a tautology?

(viii) $((P\to Q)\to Q)\to P$

$$
\text{Tree:}\quad
\begin{array}{c}
\neg(((P\to Q)\to Q)\to P)\ \checkmark \\
| \\
(P\to Q)\to Q\ \checkmark \\
| \\
\neg P \\
\begin{array}{cc}
/ & \backslash \\
\neg(P\to Q)\ \checkmark & Q \\
| & \\
P & \\
| & \\
\neg Q & \\
\times &
\end{array}
\end{array}
$$

One branch closes, but the other does not. Hence there is a countermodel to the Abelian principle, $((P\to Q)\to Q)\to P$, and so it is not a tautology. $\not\models ((P\to Q)\to Q)\to P$. Let P be F and Q be T.

As we apply the tree rules, we check off more complex wffs, and introduce shorter less complex wffs into the tree. Hence we must in a finite time eventually check off all wffs in every branch except propositional letters or their negations. At that point we can check through the finite number of finitely long branches, and note whether any branch is open, or if they are all closed. If the former, the wff which sits at the root of the tree is not a tautology; if the latter, it is a tautology.

We have, therefore, an effective procedure for testing whether a wff is a tautology. An effective procedure is a purely mechanical method which when applied is guaranteed to yield in a finite time the answer 'yes' or 'no' to the initial question. Familiar effective procedures, or algorithms, are addition, multiplication, and prime factorization on the natural numbers. We all know — though you may never have thought of it this way! — methods of calculating in a finite time, guaranteed to yield an answer one way or the other, whether, for example, a given number is the sum or product of two others, or is a prime number, or is a factor of another.

Of course, we clearly had an effective method for being a tautology in propositional logic in Chapter 3, in the form of truthtables. The present tree method is, however, generalisable, so that we can extend it to give an effective method for testing validity (as we will there call it) in monadic predicate calculus. However, when we turn to full predicate logic, we will find that a third possibility arises for the tree. Besides

(1) every branch closes; or

(2) after a finite time every wff in every branch which is not checked off is atomic or negation of atomic, and at least one branch is not closed,

a third case will arise:

(3) at least one branch continues to generate new wffs indefinitely — in other words, the tree is infinite.

In both cases 2) and 3) the wff at the root is invalid: in both cases a countermodel will be describable. However, at any finite point we may not know if the branch will indeed continue for ever, or will eventually close. Hence our method is not guaranteed, for full predicate logic, to give an answer in a finite time. Indeed, there is no such method available. Of course, if the sequent is valid (i.e. in propositional logic, a tautology), the tree will close — full predicate logic is semi-decidable, as it is called. But whether it is valid is precisely what we do not know.

Any property for which there is an effective method of testing is called **decidable**. Thus validity (being a tautology) for propositional logic, and for monadic predicate logic, is decidable. However, validity in full predicate logic is undecidable. The tree method is called a **decision procedure** for propositional logic.

Exercises

(1) Test the following wffs for satisfiability by the tree method. If satisfiable, specify an interpretation which satisfies them:

(a) $P \& \neg P$

(b) $((P \to Q) \to Q) \to P$

(c) $(P \to Q) \vee (Q \to P)$

(d) $(P \vee \neg Q) \& \neg(\neg P \to \neg Q)$

(e) $(P \leftrightarrow Q) \leftrightarrow (P \leftrightarrow \neg Q)$

(f) $((P \leftrightarrow Q) \leftrightarrow \neg Q) \leftrightarrow Q$

(2) Test the following sequents for logical validity (or consequence) by the tree method. If invalid, describe all counterexamples.

(a) $Q, P \leftrightarrow Q \therefore P$

(b) $Q \to R \therefore (P \vee Q) \to (P \vee R)$

(c) $(P \& Q) \to R, P \to R \therefore Q \to R$

(d) $P \to Q, R \to S \therefore (P \vee R) \to (Q \vee S)$

(e) $(P \& Q) \to R \therefore ((P \& \neg Q) \to R) \vee ((Q \& \neg P) \to R)$

(f) $P \to Q, R \to S, P \vee R \therefore Q \& S$

(g) $P \leftrightarrow Q, Q \to R, \neg R \vee S, \neg P \vee S \therefore S$

(h) $(S \to P) \& (T \to R), S \vee T, S \to \neg R, T \to \neg P \therefore R \leftrightarrow \neg P$

(i) $(Q \& R) \to S, (S \& T) \to P, \neg U \to (T \& \neg P) \therefore (Q \& R) \to U$

(j) $S \vee T, Q \to \neg(P \vee U), R \to \neg(W \& V), (\neg V \& T) \& P \therefore Q$

(k) $P \to (Q \& R), S \leftrightarrow (Q \& T), Q \to (\neg P \vee S) \therefore \neg R \to T$

5 The Syntax of Predicate Logic

It is time now to widen our concerns to those arguments which seem to be valid in virtue of their form but for whose analysis the techniques of propositional logic are too crude: arguments whose validity depends upon the recurrence in them of sub-sentential expressions, i.e. words and phrases smaller than whole clauses. Examples:

(i) All men are mortal
All mortals fear dying
———
All men fear dying

(ii) All violets are purple flowers
All purple flowers attract butterflies
———
All violets attract butterflies

(iii) Some logicians are good philosophers
Every good philosopher enjoys an argument
———
Some logicians enjoy an argument

(iv) Some sweet peas are pink
All pink flowers attract bees
———
Some sweet peas attract bees

(v) Giscard is a banker
No bankers are generous
———
Some logicians enjoy an argument

(vi) 28 is a perfect number
No perfect numbers are odd
———
28 is not odd

Here (i) and (ii) intuitively share their logical form; as do (iii) and (iv), and (v) and (vi). And each of the three forms is intuitively valid. How can we characterise these forms, so as to bring out the validity of the informal arguments, (i)-(vi)? The best approach to the question is to ask on what, intuitively, the validity of those arguments depends. Take (i): clearly the meanings of 'men', 'mortal' and 'fear dying' don't matter at all — what is important is the pattern of recurrence which those predicates display in the argument. This is what (i) shares with (ii); schematically, the pattern is:

> all Fs are G
> <u>all Gs are H</u>
> all Fs are H

On the other hand, the meaning of 'all' is clearly crucial — replace it by, for example, 'no', or 'some', and the result is an invalid argument pattern.

Likewise, the form of (iii) and (iv) is naturally represented as:

> some Fs are G
> <u>all Gs are H</u>
> some Fs are H,

and again it is intuitively clear that any argument of this form is going to be valid, so long as the interpretation of 'some' and 'all' is held constant; on the other hand, if 'no' is substituted for 'some', for example, the form ceases to be a valid one.

These words, 'all', 'some', 'no' ... etc., are standardly called **quantifiers**. The rationale for the term is, roughly, this: quantifiers typically specify the extent of the range of things to which a predicate may truly be applied. Thus:

everything is F	:	F has universal application
nothing is F	:	F has zero application
something is F	:	F has non-empty application
few things are F	:	F has small application
many things are F	:	F has wide application

and so on. English abounds with quantifiers. Sometimes, as above, they play the role of noun-phrases (others are the personal 'everyone', 'no-one', 'someone'); and sometimes they play the role of adverbs, e.g., 'always', 'nowhere', 'everywhere' — which quantify over times and places, i.e. indicate the range of times and places in which a predicate is true of a subject, or, more generally, in which a sentence is true. Thus: Jones is never bored = 'is bored' is true of Jones at no time.

Propositional logic was concerned with all valid patterns of argument whose validity is on account of patterns generated by repetition of whole clauses and the meanings of the connectives. **Predicate logic**, (or **quantification theory**, or **first-order logic**, as it is variously called), is going to be concerned not only with that but also with valid patterns of argument generated by repetition of subject and predicate phrases and by the meanings of quantifiers.

We shall take as primitive two quantifiers:

The universal quantifier, expressing the universal application of the predicate to which it is attached. We write this using an upside-down 'A' followed by a small letter (x, y , or z, usually), which for the moment you should think of as a kind of pronoun, in the following style:

$(\forall x)Fx$ — every item, x, is such that it (x) is F,
i.e. F has universal application.

$(\forall y)Gy$ — every item, y, is such that it (x) is G,
i.e. G has universal application.

The existential quantifier, expressing the non-empty application of the predicate to which it is attached. We write this in a similar style, using a backwards 'E' and the same kind of small letter:

$(\exists x)Gx$ — at least one item, x, is such that it is G,
i.e. G has non-empty application.

$(\exists y)Fy$ — at least one item, y, is such that it is F,
i.e. F has non-empty application.

This is, in essentials, the apparatus that the founder of modern logic, Gottlob Frege introduced about 100 years ago. The small letters, 'x', 'y ', 'z', are **individual variables**; they stand in the same relation to particular names ('Andrew', 'Thames', 'Seven') and definite descriptions (roughly, descriptions of particular objects prefixed by the definite article, e.g. 'the Queen of England', 'the oldest man in the room', 'the clock on the shelf') as our old propositional variables, P,Q,R, etc., stand to individual statements in natural language. We shall focus on the purpose and advantages of this quantifier/variable notation shortly.

Our vocabulary will for the present consist of the following:

(1) the two quantifiers, '$(\forall \ldots)\ldots$' and '$(\exists \ldots)\ldots$'

(2) an indefinite stock of individual variables, 'x', 'y' , 'z', ...

(3) an indefinite stock of letters representing predicates, 'F', 'G', 'H', ...

(4) an indefinite stock of proper names, 'a', 'b', 'c', ... (we call these individual constants).

(5) all the vocabulary of propositional logic.

We call the individual constants and variables, **terms**. In a minute we shall see how to formalise the examples (i)-(vi) using this apparatus. But first, a response to the following question which may have occurred to you: we said that English abounds with quantifier expressions — yet our system contains only two. So how is the system going to be able to cope? Given that the validity of the kind of arguments in which we are now interested depends on the meaning of quantifiers, may there not be many valid arguments which we cannot represent because we cannot express the quantifiers featuring in them?

The situation is similar to that which arose concerning our chosen connectives for propositional logic and the wide variety of connectives in natural language (cf. Ch. 2, p. 13); and the answer is similar also — we'll be able to define most of the quantifiers in terms of our chosen two; and those we cannot define prove not to be of central logical interest. Thus:

(a) 'Everything', 'anything', 'each thing', 'all things', when they occur in initial position in a sentence (e.g. 'Everything is F', 'all things are F', etc.) are synonymous and can be directly expressed by the universal quantifier.

(b) 'Nothing' can be defined in terms of 'everything' thus: to say that nothing is F is to say that everything is not F.

(c) 'Something', 'at least one thing', 'there are things which' can all be directly expressed by the existential quantifier.

(d) The personal, positional, and temporal quantifiers (like 'everybody', 'nowhere', 'sometimes') can all be paraphrased by using a neutral quantifier and introducing an appropriate predicate. Thus 'everybody is beautiful', becomes: 'everything which is a person is beautiful'; 'it is sunny nowhere' becomes: 'nothing is a place at which it is sunny'; and so on.

(e) We can't, admittedly, cope with the family: 'many (things)', 'most (things)', 'few (things)', 'often', 'usually', 'seldom', etc. But there are very few valid forms of argument which depend just on the meanings of these quantifiers and on sub-sentential structure. (See if you can think of some.)

Let us now attempt to formalise the examples (i)-(vi). In (i) and (ii) premises and conclusions are all of the type: all things with a certain property have another specific property. We therefore immediately confront a problem: we know how to represent the structure of 'all things have a certain property' — that is just '$(\forall x)Fx$' — but it is not straightaway clear how to capture the more complex kind of 'all' statement involved in (i) and (ii) in which the quantification involves not one predicate but two.

Evidently the solution must consist in finding a predicate such that to say that all Fs are G is to say that all things are characterised by it — and remember that we still have the resources of propositional logic to help us to define such a predicate. A moment's reflection yields an answer: to say that all Fs are G is just to say that everything, x, is characterised by the following complex predicate: if it is F, then it is G; i.e. it is true of anything that if it is F, it is also G. And conversely: if we know that it is true of absolutely anything that if it is F, it is also G, then it must be true that all Fs are G.

It is worth being a bit more exact about the way we are understanding the notion of a predicate at this point, since we are extending the orthodox grammarian's usage. For us, a predicate will be anything that results from a complete declarative sentence (a "statement") by deleting one or more occurrences of some particular singular term. Thus for us

'… is bald',
'… is held in contempt by the radical left',
'… likes Andrew better than Harry',
'… likes Harry better than Harry likes … ' and
'… is rich or … is good-looking',

are all predicates; for they respectively result from deleting one or more occurrences of 'Bill' in:

'Bill is bald'
'Bill is held in contempt by the radical left',
'Bill likes Andrew better than Harry',
'Bill likes Harry better than Harry likes Bill',
'Bill is rich or Bill is good looking'.

But probably only the first, and perhaps also the second and third, would rank as grammarians' predicates.

Our solution to the formalisation of (i) and (ii), then, is to recognise the implicit occurrence of a conditional predicate: 'if … is F, then … is G', in 'all Fs are G'. So the formalisation comes out as:

$$(\forall x)(Fx \rightarrow Gx),\ (\forall x)(Gx \rightarrow Hx)\ \therefore\ (\forall x)(Fx \rightarrow Hx)$$

where $Fx \rhd x$ is a man / a violet
$Gx \rhd x$ is mortal / a purple flower
$Hx \rhd x$ fears death / attracts butterflies

N.B. the use of brackets to enclose complex predicates to which quantifiers are being applied. This will be important if ambiguity is to be avoided in more elaborate constructions.

How do we fare, now, with (iii) and (iv)? Both have an initial premise of the form, 'some Fs are G'; and this immediately presents a formalisation problem analogous to that we just coped with. We know how to represent 'some things are F', viz. '$(\exists x)Fx$', but that is no help unless we can discern a complex predicate in 'some Fs are G' such that to affirm that sentence is to affirm that something is characterised by that predicate. A natural ploy would be to interpret 'some Fs are G' as involving the same conditional predicate as 'all Fs are G' — after all, the two sentences differ only in that one has 'some' where the other has 'all', so it looks as though they differ only in the quantifiers which they involve. But this can't be right, for a simple reason: 'all Fs are G' does not entail 'all Gs are F'; and this we have reflected in our formalisation, since (because $P \rightarrow Q$ does not entail $Q \rightarrow P$!), $(\forall x)(Fx \rightarrow Gx)$ won't entail $(\forall x)(Gx \rightarrow Fx)$. But, in contrast, 'some Fs are G' does entail 'some Gs are F' — if some apes are intelligent, then there are some intelligent creatures which are apes; and this entailment will be forfeited if we formalise 'some Fs are G' as: $(\exists x)(Fx \rightarrow Gx)$ — i.e., there is at least one x such that if it is F, it is G. Indeed, $(\exists x)(Fx \rightarrow Gx)$ is a very odd expression: what object is it supposed to be that is such that if it is an ape it is intelligent? Think of the paradoxes of implication: since you are intelligent, you are such that if you are an ape you are intelligent — or at least, such that you are an ape $\rightarrow$ you are intelligent. So *you* make $(\exists x)(Fx \rightarrow Gx)$ true! But that fact about you is not enough to make true the statement which we ordinarily mean by 'Some apes are intelligent'.

So, we have to find some other complex predicate in 'some Fs are G' — one that sustains the entailment to 'some Gs are F'. If we reflect also on the evident entailment to both 'something is F' and 'something is G', the solution is clear: the predicate which we want is '... is F *and* ... is G'. To say 'some Fs are G' is to say that something is both F and G.

The formalisation of (iii) and (iv) thus proceeds:

$$(\exists x)(Fx \,\&\, Gx),\ (\forall x)(Gx \rightarrow Hx)\ \therefore\ (\exists x)(Fx \,\&\, Hx)$$

where $Fx \triangleright x$ is a logician / a sweet pea
$Gx \triangleright x$ is a good philosopher / a pink flower
$Hx \triangleright x$ enjoys an argument / attracts bees

Finally (v) and (vi). Here the initial premises contain proper names ('Giscard' and '28') so we shall need for the first time to use one of our individual constants, a, b, c, ... Also we have to cope with the quantifier, 'no', in the second premise. Here, again, a little care is necessary. Earlier we noted that to say that nothing is F is to say that everything is not F (which of course is not the same as saying that not everything is F, i.e. $(\forall x)\neg Fx \neq \neg(\forall x)Fx$). But that, once more, is no immediate help with the construction, 'no Fs are G'. So, in order to use our definition of 'no' in terms of 'all', we need to find a predicate such that to say that no Fs are G is to say that all things fail to be characterised by that predicate, i.e. all things are characterised by its negation. (We will understand the negation of a predicate in the most obvious way: viz. what you get when you first negate the declarative sentence from which the predicate is formed and then omit all occurrences of the relevant proper name.) The answer is not far to seek: 'no Fs are G' is, intuitively, the negation of 'some Fs are G'; so to affirm 'no Fs are G' is to affirm that everything fails to be characterised by that predicate which 'some Fs are G' claims to characterise at least one thing, i.e. the predicate 'F... & G...'. So 'no Fs are G is: $(\forall x)\neg(Fx \,\&\, Gx)$; and, since $\neg(P \,\&\, Q) \leftrightarrow P \rightarrow \neg Q$, we can take the equivalent form:

$$(\forall x)(Fx \rightarrow \neg Gx),$$

which will be more convenient to handle in proofs.

The formalisation of (v) and (vi) is now straightforward:

$$(\exists x)(Fx \,\&\, Gx),\ (\forall x)(Gx \rightarrow Hx)\ \therefore\ (\exists x)(Fx \,\&\, Hx)$$

where $a \rhd$ Giscard / 28
$Fx \rhd x$ is a banker / is a perfect number
$Gx \rhd x$ is generous / odd.

We are now in position to represent each of the four forms of propositions dealt with in the traditional Aristotelian Syllogistic:

All Fs are G	=	$(\forall x)(Fx \rightarrow Gx)$	—	A-proposition
No Fs are G	=	$(\forall x)(Fx \rightarrow \neg Gx)$	—	E-proposition
Some Fs are G	=	$(\exists x)(Fx \,\&\, Gx)$	—	I-proposition
Some Fs are not G	=	$(\exists x)(Fx \,\&\, \neg Gx)$	—	O-proposition

The Syllogistic was just a catalogue of valid argument patterns among propositions of those four kinds, attention being restricted to arguments with two premises which collectively featured three predicates (or 'terms' as they were traditionally and confusingly called). Thus in syllogistic notation (i) and (ii) are of the famous form:

FaG
$\underline{GaH}$ — aaa, known as 'Barbara'!
FaH

while (iii) and (iv) are of the form:

FiG
$\underline{GaH}$ — iai
FiH

(For more on syllogisms, see Part III, Ch.1. The letter 'a' is used for A-propositions, 'i' for I-propositions, and so on.) Arguments (v) and (vi), however, are not syllogisms because they feature proper names in a structurally essential way. So we can already claim the modest advantage over Aristotelian logic that we can express valid patterns of argument of this sort.

In fact, though, the advantages of Frege's notation are much more than that. Consider this example:

(vii) All fish have gills
All gills work by filtration
All fish have something which works by filtration

Clearly the argument is valid. And clearly its being so is a consequence of its being of a valid form — the form it shares with this:

(viii) All mathematicians use pencils
All pencils contain lead
All mathematicians use something which contains lead

But (vii) and (viii) are not syllogisms; and if we wish to capture their form, we cannot do it by construing their premises as being of the form FaG — the attempt to do so will result in our discerning four different predicates ('terms') in the premises (i.e., for (viii) 'x is a mathematician', 'x uses a pencil', 'x is a pencil', and 'x contains lead') and thus losing grip on the essential recurrence of 'gills' and 'pencils'.

What Frege realised is that valid forms involving quantifiers are not restricted to forms of argument containing only sentences with a single, initial quantifier. The conclusions of (vii) and (viii), for example, contain a 'something' embedded in the initial 'all'. We need to be able to formalise not only predicates but also **relations**, a distinction which the Aristotelian notion of a 'term' blots out.

Consider for example the argument:

(ix) Andrew loves Brenda
Someone loves someone

This is trivially valid, and has the same form as

(x) Peter is older than Paul
Someone is older than someone

But we can't represent this form if we confine ourselves to discerning the occurrence of predicates, in the sense defined earlier. For the only predicates occurring in (ix)'s premise are '... loves Brenda' and 'Andrew loves ...' — and neither of those recurs in the conclusion. What does recur is the word 'loves'; but this cannot rank as a predicate since it results from the premise from deleting the occurrence not of one proper name but of two. Hence the need for the idea of a relation, which simply generalises our earlier account of a predicate.

Thus:

(1) a predicate is any expression which results from a declarative sentence by deletion of one or more occurrences of *one* proper name (or other singular term, e.g. a definite description).

(2) a two-place relation is any expression resulting from a declarative sentence by deletion of one or more occurrences of *two* proper names (or other singular terms).

(3) a three-place relation is any expression resulting from a declarative sentence by deletion of one or more occurrences of *three* proper names (or other singular terms).

and so on.

We can now start to define a wff of predicate logic. A **term** is an individual variable or constant. An **atomic wff** consists of a predicate or relation letter followed by the appropriate number of terms. If A and B and wffs, so are $\neg A$, $(A \& B)$, $(A \vee B)$, $(A \rightarrow B)$, $(A \leftrightarrow B)$. Finally, if $A(\xi)$ is a wff containing one or more occurrences of the variable 'ξ' (the Greek letter 'xi') then $(\forall\xi)A(\xi)$ and $(\exists\xi)A(\xi)$ are **wffs**. 'ξ' here stands for any variable, 'x', 'y ', 'z' etc.

Thus (ix) and (x) both contain a premise involving a two-place relation; and our formalisation, if it is to capture their form, must reflect the fact. We therefore simply extend our notation to include relation letters (usually R, S, T ...) which, depending on the number n of terms which the relation in question takes, will then be followed by n occurrences of individual constants or variables. Thus the common form of the premises in (ix) and (x) will just be: Rab.

(ix) and (x) also raise another point. How should we formalise the conclusion? If we wrote

$$(\exists x)(\exists y)Rxy,$$

where Rxy means 'x loves y', that would represent "Something loves something". We want to restrict it to people. There are two ways to do this.

One way is to introduce an explicit predicate, 'Px', for 'x is a person'. The conclusion is

$$(\exists x)(Px \,\&\, (\exists y)(Py \,\&\, Rxy)),$$

i.e., there is a person and there is (another) person such that the first loves the second.

But that is very clumsy. Far smoother is to introduce the notion of a **universe of discourse**. What we note is that the argument only speaks of people — Andrew, Brenda and people in general in (ix), and Peter, Paul and people in general in (x). When there is a restricted universe of discourse in this way, we can add to our key the note that the universe of discourse is, here say, people. Then

$$(\exists x)(\exists y)Rxy$$

will indeed represent "Someone loves someone", relative to this (restricted) key, because we are understood only to be speaking of people, in our universe of discourse.

This move will not always work, however. For example, we could not use it in (viii), for the argument speaks of pencils, as well as of mathematicians. But then we don't really need it, for the predicates 'mathematician', 'pencil' etc already define what we are speaking of.

So the form of (ix) is

$$Rab \;\therefore\; (\exists x)(\exists y)Rxy$$

Key: Universe of discourse: People

$Rxy \rhd x$ loves y

$a \rhd$ Andrew

$b \rhd$ Brenda

(x) has the same form. What is the appropriate key?

Let us now return to our remarks about the use of letters for relations. We get an initial glimmering of how this innovation is going to help with (vii) and (viii) if we reflect that 'have' in (vii) and 'use' in (viii) intuitively indicate relations. But, before we address ourselves to formalising those arguments, there are several points to note:

(a) the order of the terms following a relation letter affects the interpretation. E.g., if a = Andrew and b = Brenda, then 'Rab' will convey that Andrew

R's Brenda whereas 'Rba' will convey that Brenda R's Andrew. We follow the convention which prevails in English: the earlier a singular term occurs in the English sentence which we are formalising, the earlier its formal counterpart will occur in the formalisation. Thus the four-place relational sentence, 'London is nearer to Paris than Edinburgh is to Birmingham' comes out as:

$$Rabcd,$$

where a = London, b = Paris, c = Edinburgh, and d = Birmingham.

(b) the same point applies, of course, if the terms following the relation letter include individual variables. Thus $(\forall x)Rax$ correponds to 'Andrew likes everyone' whereas $(\forall x)Rxa$ corresponds to 'Everyone likes Andrew'. Likewise $(\forall x)(\exists y)Rxy$ corresponds to 'Everyone likes someone' whereas $(\forall x)(\exists y)Ryx$ correponds to 'Everyone is liked by someone'. Just here is where the great advantages of the quantifier/variable notation are apparent: by shifting order of the variables after a relation letter, we shift the interpretation — and that gives us wide expressive power. But also, the other side of the coin, we can disambiguate unclear natural language constructions, e.g. 'Everyone likes anyone who cleans his shoes'; to whom does the 'his' refer — the liker or the liked? Using the syntax of quantifiers, variables, and relations, we can give a sharp formulation of the difference. (See if you can work out exactly how.) A tremendous variety of natural language constructions — transitive verbs, comparatives, active and passive voices, possessives — come within our compass once we have relation letters; and the quantifier/variable notation then provides a pellucid way of capturing the effect of 'all', 'some', etc., as they feature in those natural language constructions.

(c) Just as varying the order of the terms after a relation varies the interpretation, so does varying the order of initial quantifiers if they are mixed. Thus $(\forall x)(\exists y)Rxy$ and $(\exists y)(\forall x)Rxy$ do not mean the same. The easiest way to see this is just to paraphrase them back into English in the most simpleminded fashion: the second says 'There is an object of which it is true that every x Rs it'; whereas the first says 'For every object x there is an object y which it Rs'. And the difference is just that what the first says can be true even if there is no one y which is Rd by everything; whereas the second can't. Thus let Rxy = x is a parent of y. Then $(\forall x)(\exists y)Rxy$ says: everyone is a parent of someone (i.e. no-one is childless — which is false but biologically possible) whereas $(\exists y)(\forall x)Rxy$ says: there is someone of whom everyone is a parent (a remarkable and presumably biologically impossible feat of collective procreation!). Notice that to give this significance to the variation in the order of mixed initial quantifiers is not something which we are doing optionally, in order to increase the expressive power of our notation. Rather, the differences are imposed on us by our basic readings of the quantifiers. For $(\forall x)Fx$ simply means: everything is characterised by F; and $(\exists x)Fx$ simply means: at least one thing is characterised by F. So $(\forall x)(\exists y)Rxy$ has to mean: everything is characterised by the predicate '$(\exists y)R\ldots y$', while $(\exists y)(\forall x)Rxy$ has to mean: something is characterised by the predicate '$(\forall x)Rx\ldots$' — and, as you will see, the different readings just sketched are then immediately imposed.

So: putting together this point about varying the order of the quantifiers with the point (b) about the significance of varying the order of the variables, it emerges that none of the following forms is equivalent:

(1) $(\forall x)(\exists y)Rxy$

(2) $(\exists y)(\forall x)Rxy$

(3) $(\forall x)(\exists y)Ryx$

(4) $(\exists y)(\forall x)Ryx$

In particular where it is understood that the universe of discourse is the natural numbers, 0, 1, 2, etc., and $Rxy \rhd x$ is greater than y, the correct readings are respectively:

(1) Every number is greater than some number.
— [false: 0 is a counterexample]

(2) Some number is such that every number is greater than it.
— [false]

(3) Every number is such that some number is greater than it.
— [true]

(4) Some number is greater than every number. — [false]

(d) According to our original explanation of what a predicate is, the result of deleting both occurrences of 'Andrew' in 'Andrew likes Andrew', viz. '… likes …', will count as a predicate; for the result is achieved by deleting two occurrences of one proper name. It is evident, however, that '… likes …' can also result from 'Andrew likes Bill' by deleting occurrences of two proper names — and so qualifies as a two-place relation, as well; which we take it to be depends, therefore, on the character of the sentence in which it is embedded. Well, there is no harm in drawing these distinctions in that particular way. But notice that where we are concerned with a predicate which results from deleting more than one occurrence of a particular proper name from a sentence, it will generally be necessary to reflect that fact in the way we formalise that sentence if we are to capture logical form. Thus consider

(xi) Andrew likes Andrew
Someone likes himself

This inference is plainly valid in virtue of form. The form is shared with:

(xii) Venice is equalled only by Venice
Something is equalled only by itself

But we cannot represent the form of their respective premises merely as: 'Fa'. For how in that case, do we represent the conclusion? '$(\exists x)Fx$', perhaps? But then what expression, occurring in both premise and conclusion, are we representing by '$F\ldots$'? Clearly the representation we want is rather:

$$\frac{Raa}{(\exists x)Rxx}$$

In the key for (xi), we will specify a universe of discourse, people. That for (xii) does not need it. Only thereby can we both capture the affinity between, and validity of, (xi) and (xii); and only thereby can we explain why 'Andrew likes Andrew' and, 'Andrew likes Bill' both entail 'Someone likes someone'.

Let us now set out the grammar of predicate logic formally, as we did in Ch. 4 for propositional logic:

Logical constants:

$\rightarrow$ implication
$\&$ conjunction
$\vee$ disjunction
$\neg$ negation
$\leftrightarrow$ the biconditional
$\exists$ the existential quantifier
$\forall$ the universal quantifier

Individual variables:
'R', ..., 'z'

Individual constants:
'a', ..., 'e', 'l', ..., 'Q'

Predicate, relation and propositional letters:
'A', ..., 'Z'

A **term** consists of an individual variable or an individual constant.

An **atomic formula** consists of a predicate, relation or propositional letter followed by a finite, possibly zero, number of terms.

Well-formed formulae (wffs) are defined inductively as follows:

(a) any atomic formula is a wff;
(b) if A is a wff, then $\neg A$ is a wff;
(c) if A and B are wffs, then $(A \to B)$ is a wff;
(d) if A and B are wffs, then $(A \,\&B)$ is a wff;
(e) if A and B are wffs, then $(A \vee B)$ is a wff;
(f) if A and B are wffs, then $(A \leftrightarrow B)$ is a wff;
(g) let A be a wff containing the variable 'ξ', but not containing a quantifier of the form $(\forall\xi)$ or $(\exists\xi)$; then $(\forall\xi)A$ and $(\exists\xi)A$ are wffs.

Note that (g) does not allow:

repeated quantification, as in '$(\forall x)(\forall x)Fxx$', and
vacuous quantification, as in '$(\exists x)(\forall y)Fxx$'.

Now, to fix ideas, carefully run through the following series of examples together with their formalisations (where Rxy = x likes y, and a = Andrew):

(a) 'Andrew likes someone' —

$(\exists x)Rax$

Key: Universe of discourse: People
$Rxy \rhd x$ likes y
$a \rhd$ Andrew

(b) 'No-one likes Andrew' —

$(\forall x)\neg Rxa$ (same key)

(c) ‘No-one dislikes himself’ —

$(\forall x)\neg\neg Rxx$ (same key)

(d) ‘Andrew does not like all girls’ — intuitively this means that there are some girls whom Andrew does not like, so

$(\exists x)(Gx \,\&\, \neg Rax)$ add to key: $Gx \rhd x$ is a girl

(e) ‘Andrew does not like any girls’ — this means, rather, that Andrew dislikes every girl, so

$(\forall x)(Gx \rightarrow \neg Rax)$ (same key)

(Notice, incidentally, how ‘all’ and ‘any’ diverge in meaning when grammatically governed by a negative.)

(f) ‘Everyone except the wise likes themselves’ — what does ‘everyone except the wise’ mean? Presumably, ‘everyone who is not wise’, so:

$(\forall x)(\neg Wx \rightarrow Rxx)$ *Key*: Universe of discourse: People
$Wx \rhd x$ is wise
$Rxy \rhd x$ likes y

(g) ‘Everyone except the wise likes the rich’ — this is pretty ambiguous. Taken one way it is tantamount to ‘Rich people are liked only by those who are not wise’; taken a second way it comes to ‘Foolish people like all rich people’. The first reading is

$(\forall x)(Fx \rightarrow (\forall y)(Ryx \rightarrow \neg Wy))$ add to key: $Fx \rhd x$ is rich

— ‘Any rich x is such that any y who likes x is not wise.’ The second reading is

$(\forall x)(\neg Wx \rightarrow (\forall y)(Fy \rightarrow Rxy))$

— ‘Any x which is not wise is such that any y who is rich is liked by x’.

(h) ‘Andrew likes everyone except the rich’ — that is, anyone who isn’t rich is liked by Andrew; so

$(\forall x)(\neg Fx \rightarrow Rax)$ (same key)

(i) ‘Andrew likes only the wise’ — how do we handle ‘only’ here? Well, recall that ‘P only if Q’ says: the truth of Q is necessary for the truth of P. Likewise i) says: x’s being wise is necessary if Andrew is to like x. So

$(\forall x)(Rax \rightarrow Wx)$ (same key)

(j) ‘Only people whom Andrew likes are wise’ — this just means: all wise people are people whom Andrew likes. For it affirms that being liked by Andrew is necessary in order to be wise. So

$(\forall x)(Wx \rightarrow Rax)$ (same key)

(k) Contrast ‘The only people whom Andrew likes are wise’ — this just means: all people whom Andrew likes are wise. (The proof is to think about it!) So it is just another way of saying the same as i).

(l) ‘Only the unwise like the rich’ — this has an ambiguity similar to that of g). Taken one way it is, indeed, tantamount to g) under its first reading, i.e., ‘Any rich man is liked only by a fool’. But it can also be read as saying: any one who likes everyone rich is a fool. This second reading comes to

$(\forall x)((\forall y)(Fy \rightarrow Rxy) \rightarrow \neg Wx)$ (same key)

— ‘Any man x such that every rich y is liked by x is not wise’.

Do work through these formalisations patiently and try to get a sense of the principles which inform their construction. The following schematic table may prove useful:

All Fs are G	being F is sufficient for being G	$(\forall x)(Fx \rightarrow Gx)$
Only Fs are G	being F is necessary for being G	$(\forall x)(Gx \rightarrow Fx)$
No Fs are G	being F is sufficient for not being G	$(\forall x)(Fx \rightarrow \neg Gx)$
The only Fs are G	all Fs are G	$(\forall x)(Fx \rightarrow Gx)$
All Fs except Hs are G	being F and not-H is sufficient for being G	$(\forall x)((Fx \& \neg Hx) \rightarrow Gx)$
Some Fs are G	there is at least one thing which is both F and G	$(\exists x)(Fx \& Gx)$
Some Fs are not G	there is at least one thing which is F but not G	$(\exists x)(Fx \& \neg Gx)$

Formalising into the language of classical predicate logic is not, in general, a trivial or mechanical matter. It can be difficult to get a clear view of the sense of the great variety of idioms which English uses. Practice is essential.

Let us finally formalise arguments (vii) and (viii):

(1) All fish have gills.

(2) All gills work by filtration

(3) All fish have something which works by filtration.

The second premise characterises things which the first premise says all fish have. And the relevant predicate 'works by filtration' occurs in the conclusion, and is nowhere broken up. That suggests:

$Gx \rhd x$ is a gill/a pencil
$Wx \rhd x$ works by filtration/contains lead

and 2) therefore becomes : $(\forall x)(Gx \rightarrow Wx)$. 'Fish' recurs too, so let us have

$Fx \rhd x$ is a fish/a mathematician

Then 1) is

$(\forall x)(Fx \rightarrow x$ has gills); i.e.
$(\forall x)(Fx \rightarrow x$ has something which is G), whence
$(\forall x)(Fx \rightarrow (\exists y)(Gy \,\&\, Rxy))$, where $Rxy \rhd x$ has y (uses y)

Hence 3) is

$$(\forall x)(Fx \rightarrow (\exists y)(Wy \,\&\, Rxy)).$$

So the whole form is

$$\begin{array}{l} (\forall x)(Fx \rightarrow (\exists y)(Gy \,\&\, Rxy)) \\ (\forall x)(Gx \rightarrow Wx) \\ \hline (\forall x)(Fx \rightarrow (\exists y)(Wy \,\&\, Rxy)) \end{array}$$

Remember that, as we noted above, there is no restricted universe of discourse here. The key simply specifies the reading of Gx, Wx and Rxy.

Note, before closing this chapter, however, one point about fineness of analysis, which was also true in propositional logic, but is even more worthy of remark here. Correct formalisation is a skill, requiring judgment. One principle which is important is: do not represent more detail than is necessary. For example, consider the argument:

> Everyone who lives in Edinburgh and owns a car passes a garage on the way to Haddington. Anyone who passes a garage on the way to Haddington knows a tall hardworking farmer. So everyone who lives in Edinburgh and owns a car knows a tall hardworking farmer.

We could represent this with the following key:

Rxy	$\triangleright$	x lives in y
Sxy	$\triangleright$	x owns y
Fx	$\triangleright$	x is a car
$Txyz$	$\triangleright$	x passes y on the way to z
Gx	$\triangleright$	x is a garage
Uxy	$\triangleright$	x knows y
Hx	$\triangleright$	x is tall
Ix	$\triangleright$	x is hardworking
Jx	$\triangleright$	x is a farmer
a	$\triangleright$	Edinburgh
b	$\triangleright$	Haddington

and so obtain the form:

$$\begin{array}{l}(\forall x)((Rxa \,\&\, (\exists y)(Fy \,\&\, Sxy)) \rightarrow (\exists y)(Gy \,\&\, Txyb)),\\ \underline{(\forall x)((\exists y)(Gy \,\&\, Txyb) \rightarrow (\exists y)(Hy \,\&\, Iy \,\&\, Jy \,\&\, Uxy))}\\ \therefore\ (\forall x)(Rxa \,\&\, (\exists y)(Fy \,\&\, Sxy) \rightarrow (\exists y)(Hy \,\&\, Iy \,\&\, Jy \,\&\, Uxy)).\end{array}$$

But that detail is totally unnecessary! Much better would be to take this key:

Fx	$\triangleright$	x lives in Edinburgh and owns a car
Gx	$\triangleright$	x passes a garage on the way to Haddington
Hx	$\triangleright$	x knows a tall hardworking farmer.

The form is then:

$$\frac{\begin{array}{l}(\forall x)(Fx \to Gx),\\ (\forall x)(Gx \to Hx)\end{array}}{(\forall x)(Fx \to Hx),}$$

which is a valid form, and so shows the original argument to be valid without unnecessary detail.

Exercises

(1) Formalise the following arguments. Give a key to your formalisation, and specify a universe of discourse where appropriate:

(a) All whales are mammals. No fish are mammals. So no whales are fish.

(b) William loves Mary. Mary is loved by no-one who is Welsh. Hence William is not Welsh.

(c) Only the French produce good wine. No Tunisian wine is French. So no Tunisian wine is good.

(d) All Spaniards except the Basques and the Catalans are loyal to the King. Hence any Spaniard who is not loyal to the King is either a Basque or a Catalan.

(2) Formalise the following sentences. Give a key to your formalisation, and specify a universe of discourse where appropriate:

(a) Frank and Gloria are married but not living together.

(b) Some people fear everyone.

(c) There's always somebody tougher than you.

(d) No-one is liked by everyone.

(e) Each of us is lonely some of the time.

(f) The Lord accepts all who accept Him.

(g) All life comes from something living.

(h) No detergent gets out everything.

(i) All provable sentences are true sentences, but there are some sentences which are true sentences but which are not provable.

(j) Not all students are apathetic about everything.

(k) There's no chain lock that someone couldn't break.

(3) Using the Keys given, translate the following wffs into sentences of English:

(a) *Key*: universe of discourse: people

Fx	$\triangleright$	x is happy
Gx	$\triangleright$	x is wise
Rxy	$\triangleright$	x is married to y
a	$\triangleright$	John
b	$\triangleright$	Sarah

(i) $(\forall x)(Fx \rightarrow Gx)$

(ii) $(\forall x)(\forall y)(\forall z)(Rxy \rightarrow (Ryz \rightarrow Rxz))$

(iii) Fa

(iv) $(\forall x)(Fx \rightarrow Rxb)$

(v) $(\forall x)(Rax \rightarrow Fx)$

(vi) $\neg(\exists x)(Rxa \& Gx)$

(b) *Key*: Universe of discourse: Numbers

Fx	$\triangleright$	x is even
Gx	$\triangleright$	x is odd
Rxy	$\triangleright$	x is less than or equal to y
a	$\triangleright$	zero

(i) $(\forall x)(Fx \rightarrow \neg Gx)$

(ii) $(\forall x)(\forall y)(\forall z)(Rxy \& Ryz \rightarrow Rxz)$

(iii) $\neg Ga$

(iv) $(\exists x)(\forall y)(Ryx \rightarrow Fy)$

(v) $(\forall x)Rax$

(vi) $\neg(\exists x)(\forall y)Ryx$

6 Models and Counterexamples

How can we show which sequents of predicate logic are valid, and which invalid? When we considered this point with respect to propositional logic, we turned to the development of the truth-table test. The truth-table test provided a mechanical and effective way of showing, of any sequent of propositional logic, whether it was valid or not. Truth-tables not only demonstrated invalidity, but could also be used to show validity as well.

> *Decidability of propositional logic*:
> Every sequent can be shown to be valid or invalid
> by a mechanical method (truth-tables).

However, the situation is entirely different for predicate logic. Since 1936 we have known that there is no effective mechanical method of showing whether an arbitrary sequent of predicate logic is valid or not. Nonetheless, the concept should be clear. An argument is invalid if it is possible for the premises to be true and the conclusion false; and a form is invalid if it has an invalid instance.

So if we suspect a form of being invalid, our aim must be to find an instance, that is, an interpretation of the letters in it, which has true premises and a false conclusion. If we could succeed in finding this, we would have shown the form was invalid. In fact, this is essentially what the truth-table test made mechanical. We took the pain out of finding a counterexample, that is, an instance with true premises and false conclusion, by simply and directly showing that we could interpret the premises as true and the conclusion as false, without bothering about finding an actual instance with these values. What made it relatively easy was that we were dealing only with unanalysed propositions, and so did not need to consider anything more than the truth-value of the constituents and the truth-functional nature of the logical connectives which combine the atomic propositions into more complex ones.

At the predicate level, two differences emerge. First, we have now analysed propositions into constituent parts — predicates, terms and quantifiers — and so the interpretation of the elementary expressions is that much more complex. Even so, we can still abstract away from particular instances, where, for example, '$(\forall x)(Fx \to Gx)$' is interpreted as saying that every man is mortal, and concentrate only on the general way in which it would need to be interpreted in order to be true (or false), that is, that we find properties, ϕ and ψ, say, such that (for the sentence to be true), everything with the property ϕ has the property ψ. So a semantic theory for predicate logic will go beyond the theory of truth-values, truth-tables and truth-functions (which, of course, will still be there to continue to deal with the truth-functional connectives, '$\neg$', '&', '$\vee$', '$\to$' and '$\leftrightarrow$'), and will specify interpretations for predicates and terms and how those elements build up to provide truth-values for complex wffs.

The second difference concerns the undecidability of predicate logic. Monadic predicate logic (the logic of one-place predicates) is decidable. It is the theory of relations — the ability to formalise arguments about horses' heads, etc. — which removes the possibility of an effective mechanical method of deciding of an arbitrary sequent whether it is valid or not. Even then, full predicate logic is semi-decidable; that is, there is a mechanical method which, given a valid sequent, will terminate in a demonstration that it is valid.

Undecidability of predicate logic:
Predicate logic is not decidable; it is semi-decidable.

The catch is that, presented with an invalid sequent, the procedure may not terminate. It may: it may come to an end, giving an interpretation of the elementary expressions under which the premises are true and the conclusion false (just as the truth-table method always does). But the threat is that, for certain sequents, a class which we cannot in advance characterise, the method may not terminate, and so, at any stage of the procedure, we will not know whether further application of the procedure will terminate (in either a proof or a counterexample) or will continue forever.

We will not pursue those hard ideas here. What we shall do is, first, consider briefly and rather generally, what an interpretation of predicate logic requires; and secondly, how we can, in an informal way, show certain more obviously invalid sequents to be invalid. Then, in the next chapter, we will develop a semi-decision method — the technique of semantic trees.

Models of predicate logic

We've noted that the difference in syntax in predicate logic — analysing wffs into a combination of predicates, terms and quantifiers, besides the further combinations with connectives — requires a more complex system of interpretation. Wffs only need truth-values for their interpretation. But what of these subsentential components?

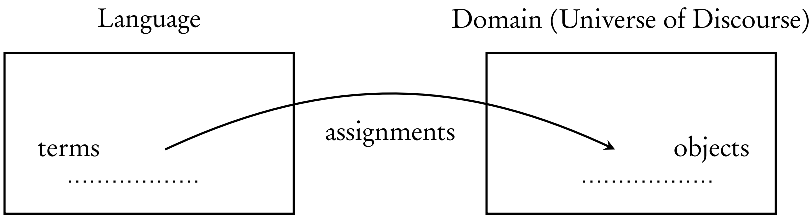
The obvious interpretation for the terms is a domain of objects (which earlier we called the "universe of discourse"). Central to classical predicate logic is the idea that the domain of interpretation must be non-empty. One consequence will be that the formula

$$(\exists x)(Fx \vee \neg Fx)$$

will come out valid. If we could interpret our logic in an empty domain, then there would be nothing which was either F or not F — there would be nothing at all. Since we restrict interpretations to those with non-empty domains, '$(\exists x)(Fx \vee \neg Fx)$' is a valid wff.

We have both constants and variables. Different interpretations will interpret the constants differently — for nothing in the *logic* of predicate logic takes note of which object a constant denotes. But then how do constants differ from variables, and if constants range in different interpretations over any of the objects in the domain, what (different from that) will the interpretation of variables be?

If you think about it, the only syntactic difference between constants and variables is that variables can be bound by quantifiers. When they are not so bound, they are no different in behaviour from constants; and when they are bound, it matters not what variable it is — e.g., '$(\forall x)Fx$' has the same force as (is equivalent to) '$(\forall y)Fy$'; '$(\exists x)Gx$' is equivalent to '$(\exists w)Gw$'. We'll see later how this fact is reflected in the semantics. For now, it is enough to have noted that terms denote members of some non-empty domain of objects.



For example, if our universe of discourse is people, 'Fx' is read as 'x is human', 'Gx' is interpreted as 'x is male', and 'y' is taken to be Ronald Reagan, then the premises of

$$(\forall x)Fx,\ Gy \therefore (\forall x)(Fx \& Gx)$$

are true and the conclusion false. It reads: 'Everyone is human. Reagan is male. So everyone is a male human.' So the sequent is invalid. The same reasoning would have shown the invalidity of

$$(\forall x)Fx,\ Gb \therefore (\forall x)(Fx \& Gx)$$

if 'b' were the name of Reagan too.

What of the predicates and relations? Obviously, they must denote properties and relations holding of and between the objects in the domain. Then an atomic formula, which consists of a predicate or relation letter followed by a number of terms, $\phi t_1 \ldots t_n$, say, will be true if the property or relation which interprets ϕ holds of the objects which $t_1, \ldots, t_n$ denote. E.g. suppose we have an interpretation $\mathscr{I}_1$ whose domain is the set of all human beings, that 'F' is a one-place predicate, interpreted as being an inhabitant of Athens, 'R' is a two-place relation interpreted as being married to, 'a' and 'b' denote Socrates and Xanthippe, his wife, and that 'x' and 'y' denote Alexander and Philip of Macedon, his father. Then

Fa, Fb, Rab	are all true
Fx, Fy, Rby, and so on	are false

Propositional letters are a special case, which are simply interpreted directly by truth-values.

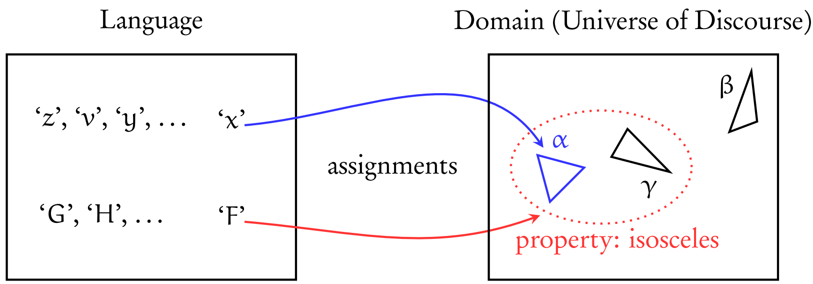
Thus our interpretation of terms over a non-empty domain of objects, and of predicates and relations over properties, relations (and truth-values, for the propositional letters) fixes the truth-values of atomic wffs. It remains to consider how the values assigned to atomic wffs extend to complex wffs. In the case of propositional logic, truth-tables did all the work, mapping (truth-functionally) the truth-values of atomic wffs (i.e. propositional letters) to the truth-values of the complex propositional wffs. The truth-tables will continue to work on the wffs of predicate logic. So, e.g., '$Fa \rightarrow Ray$' will be true if 'Fa' is false or 'Ray' is true, and false only if 'Fa' is true and 'Ray' false. '$(\forall x)Gx \vee (\exists x)(\forall y)(Gx \& Gy \rightarrow Rxy)$' will be true if either '$(\forall x)Gx$' is true or '$(\exists x)(\forall y)(Gx \& Gy \rightarrow Rxy)$' is true.

But what of the quantifiers? This is where the semantics of predicate logic becomes most difficult, and where we must be most sketchy. Note that the point of the truth-tables is to relate the truth-value of a complex wff to the truth-values of its constituents. In that way, the truth-values of simpler wffs (ultimately, of the atomic wffs) determine the truth-values of the more complex wffs. The same approach must apply here — to do the same with quantified wffs: we need some way of relating the truth-value of a wff such as '$(\forall x)Fx$' to the truth-value of 'Fx'. Yet not just that; for the value of '$(\forall x)Fx$' is the same as that of '$(\forall y)Fy$' and '$(\forall z)Fz$', so it must relate equally to 'Fy' and 'Fz'.

There are two main ways of giving the semantics of quantified wffs. One way in fact dispenses with truth-values altogether, replacing it with a concept of satisfaction. Its author, Alfred Tarski, was unhappy with saying that a wff such as 'Fx', containing a free variable, was simply true or false. Better, he thought, to say of some object, such as Socrates, that it **satisfied** the wff 'Fx' (under a certain interpretation). We will sketch his solution in Part III, Ch. 5. But another approach finds this subtlety unnecessary, and trades only in truth-values. This we will now consider — so-called "truth-value semantics".

According to truth-value semantics, '$(\forall x)Fx$' is true under a particular interpretation if 'Fx' is true under that interpretation *regardless of how* 'x' *is interpreted*. That is, although, in our earlier example, 'F' is interpreted as being an inhabitant of Athens and 'x' as denoting Alexander, the value of '$(\forall x)Fx$' under that interpretation depends only on the interpretation of 'F', not of 'x'. For it is true if 'Fx' is true, whatever 'x' denotes. But since 'x' can (and does) denote Alexander, who was not an inhabitant of Athens, '$(\forall x)Fx$' is false under that interpretation. Hence '$(\forall x)Fx$' is false.

Take another example. Let the domain of interpretation $\mathscr{I}_2$ be that of triangles. Let 'F' mean 'isosceles' (i.e. two equal sides), and let 'x' be assigned to some equilateral triangle (3 equal sides) — call it α.



Properties are not objects in the domain, but subsets of the domain — subsets containing those objects which have the property. Then 'Fx' means under interpretation $\mathscr{I}_2$ that α is isosceles. Since α is equilateral, α is also isosceles. So 'Fx' is true under this interpretation, $\mathscr{I}_2$. But '$(\forall x)Fx$' is false. For although 'Fx' is true, there is another interpretation, $\mathscr{I}_3$, under which 'Fx' is false. What is it? It differs from our interpretation $\mathscr{I}_2$ only in that 'x' is assigned to some scalene triangle (all sides and angles different), β. Now (under this different interpretation, $\mathscr{I}_3$), 'Fx' means 'β is isosceles'. And β is not isosceles. So 'Fx' is false under $\mathscr{I}_3$. Hence '$(\forall x)Fx$' is false under $\mathscr{I}_2$. Not every triangle is isosceles.

In contrast, '$(\exists x)Fx$' is sensibly interpreted as true under an interpretation if 'Fx' is true under *some interpretation differing only in what it assigns to* 'x'. Reverting to the interpretation among inhabitants of Athens, $\mathscr{I}_1$, 'Fx' is false (under that particular interpretation); but it could be true, if 'x' denoted, say, Aristotle or Socrates — indeed, it is true under that different interpretation whose domain is the same, whose interpretation of 'F' is the same, but which interprets 'x' as Aristotle. So '$(\exists x)Fx$' is true under our earlier interpretation, $\mathscr{I}_1$, as indeed it should be, for it means "Someone lives in Athens".

In propositional logic, each row of the truth-table marks a different interpretation. To decide on validity, we need look only at the 2^n different rows, where n is the number of constituent propositional variables. In predicate logic, finite domains are possible in interpretations. But so too are infinite domains. Indeed, it is the possibility of infinite domains which leads to the undecidability of predicate logic which we have mentioned at the start of this chapter.

We now have a rough idea of what determines the truth-value of each wff under different possible interpretations. We can, therefore, define the notion of a valid sequent. We say that a **model** of a set X of wffs is an interpretation under which every member of X is true. We say that the sequent $X \therefore A$, where X consists of zero or more wffs, is valid if A is true whenever every member of X is true. That is, $X \therefore A$ is valid if A is true in every model of X. We write

$$X \models A$$

to signify that the sequent $X \therefore A$ is valid. If X is empty, and A is true under every interpretation, we say that A is (universally) valid, in symbols, $\models A$.

Counterexamples

The semantics of predicate logic is very abstract. It goes straight to the interpretation of wffs of predicate logic bypassing any translation of those wffs into English. We can perhaps get a better grasp of the idea of models and counterexamples if we think of the interpretation as a translation into English — as the taking of an instance.

For example, consider our example of the Athenians, Socrates and Xanthippe, and the Macedonians, Philip and Alexander, of the last section. That interpretation, $\mathscr{I}_1$, of the wffs of predicate logic consists essentially in translating, or instancing:

Fa	as	Socrates lives in Athens	— True
Rab	as	Socrates is married to Xanthippe	— True
Fx	as	Philip lives in Athens	— False
Rby	as	Xanthippe is married to Alexander	— False

and so on.

In fact, we have come full circle. We started with arguments in natural language — English. We formalised them into predicate calculus, and asked which forms were valid, and so whether their instances were valid. We are now using instances of those forms to understand the idea of counterexamples. In other words, an argument is invalid if it is not materially valid, and every form of which it is an instance is invalid (cf. Ch. 1); and those forms are invalid if there is an interpretation under which the premises are true and the conclusion false, that is, if it is possible for the premises to be true and the conclusion false — that is, if there is an invalid instance. Hence we can show a form is invalid by instancing a counterexample.

Take, for example, the sequent:

$$Rab \therefore Rba.$$

If we interpret 'R' as 'is married to', then the conclusion is true whenever the premise is. But if we take the same domain and interpret 'R' as 'is husband of' and 'a', 'b' as Socrates and Xanthippe again, we get a true premise and a false conclusion. So the sequent is invalid. We can specify the counterexample:

Key: Universe of discourse: Human beings

$Rxy \ \triangleright \ x$ is a husband of y

$a \ \triangleright$ Socrates

$b \ \triangleright$ Xanthippe

Then 'Rab' means "Socrates is husband of Xanthippe" and is true, while 'Rba' means "Xanthippe is husband of Socrates" and is false.

Here is a more complex example:

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

Let 'F' mean 'is a husband' and 'G' mean 'is married'; then the premise means 'Every husband is married', while the conclusion means 'Everyone who is married is a husband'. This interpretation, therefore, provides a counterexample to the validity of the sequent: under this interpretation, the premise is true and the conclusion false, and so the sequent is invalid. To specify the counterexample, we take:

Key: Universe of discourse: Human beings

$Fx \ \triangleright \ x$ is a husband

$Gx \ \triangleright \ x$ is married

Let's try the famous quantifier-shift fallacy:

$$(\forall x)(\exists y)Rxy \therefore (\exists y)(\forall x)Rxy$$

Let us take as domain this time, the set of integers ..., -2, -1, 0, 1, 2, ..., and interpret 'R' as 'is greater than'. Then the premise means 'Every number is greater than some number' (which is true), while the conclusion means 'Some number is less than every number' (which is false — though of course it is true of a domain like the natural numbers, 0, 1, 2, ...). Thus the sequent is invalid. It is a classic and notorious fallacy.

Key: Universe of discourse: Integers

$Rxy \ \triangleright \ x > y$

A final example:

$$(\forall x)(Fx \rightarrow Gx),\ (\exists x)(Gx \,\&\, Hx) \therefore (\exists x)(Fx \,\&\, Hx)$$

Key: Universe of discourse: Human beings
$Fx \quad \triangleright \quad x$ is a husband
$Gx \quad \triangleright \quad x$ is married
$Hx \quad \triangleright \quad x$ is a woman

Then the premises are true and the conclusion is false (under this interpretation — i.e. with this key).

The semantics of predicate logic is admittedly more complicated than that of propositional logic. The language of predicate logic is much richer, containing predicates, terms and quantifiers, as well as propositional connectives. So the interpretations have to be correspondingly more complex. In particular, there is no mechanical way to show invalidity. Nonetheless, with a little ingenuity and practice, we can find counterexamples to invalid sequents, by setting out a key interpreting the predicates and terms over a suitable non-empty domain. In the next chapter we will develop a technique which, though it cannot guarantee to generate a counterexample to every invalid sequent (because of the undecidability of predicate logic), removes that mind-numbing need for ingenuity which the above examples required. It is a formal technique, like that of truth-tables. In the meantime, use a little ingenuity to tackle the following exercises.

Exercises

(1) Show, by giving a counterexample (a domain, and instances for the predicate letters and terms) that the following sequents are invalid:

(a) $(\forall x)(Fx \to (\exists y)(Gy \,\&\, Rxy)) \therefore (\exists y)(Gy \,\&\, (\forall x)(Fx \to Rxy))$

(b) $(\forall x)(Fx \to Gx), (\forall x)(Fx \to Hx) \therefore (\exists x)(Gx \,\&\, Hx)$

(c) $(\exists y)(Fy \to Q), (\exists y)Fy \therefore Q$

(d) $\therefore (\forall x)Fx \lor \neg(\exists x)Fx$

(e) $(\forall x)(Fx \to Gax), (\exists x)\neg Gxa \therefore \neg Fa$

(2) How would you show that the following argument is invalid?

All cats are animals.
Some animals have tails.
So some cats have no tails.

7 Semantic Trees for Predicate Logic

[This chapter is omitted in this version of these notes.]

PART II

PROOF THEORY

1 Proofs and the Conditional

In Part I, Ch. 3, we posed two questions to ourselves, *viz* which sequents are valid, and how can we tell this? We there developed two methods for testing the validity of sequents — the method of truth-tables, and that of semantic trees. Both these methods are semantic. In this Part of *Formal Logic*, we will develop a very different method, the method of proof. Semantics deals with meaning: it refers to what expressions denote, what their meaning is, which statements are true and which false. Our new method will not use any semantic notions; it will be, as we say, purely *syntactic*. It will relate solely to the *shape* or *form* of the wffs we are dealing with.

Here is an analogy: there are infinitely many true arithmetical statements of the form: $m+n=k$. How can we characterise which they are, and how can we tell which they are? One answer is: by laying down rules of addition which we can use to verify or falsify any such statement. That is exactly what calculation is — a way of sorting out the true arithmetical statements from the false. The point about these rules is that they operate on formulae simply in virtue of the *form* and *shape* of those formulae.

What we are going to do is give an essentially similar answer to our problem about logic. We shall lay down rules of argument — standardly called **rules of inference** — such that (or so we intend) by applying these rules we can tell of any valid sequent expressible in the vocabulary of sentential logic that it is valid. And we will then be able to characterise the valid sequents as being precisely those which we can vindicate by our rules of inference.

The next few chapters will be devoted to giving a full and rigorous treatment of a suitable set of rules of inference for sentential logic.

Our strategy is going to be to lay down rules which we will apply to whole sequents to give us new whole sequents; and the intention will be that the rules will have the property that, provided the sequents we apply them to are valid, the resulting sequents will be valid also. Broadly speaking, we shall need rules of two types for each of our connectives: rules telling us what we can get from sequents in whose conclusion that connective is the main connective — what logicians generally call **elimination rules** — and rules telling us when we can advance to a sequent in whose conclusion that connective is the main connective — what logicians call **introduction rules**. (Recall the notion of the 'main connective' in a formula from Part I.) We will formulate our rules by using the letters A, B, ..., to talk about formulae of propositional logic, just as we are using P, Q, R to talk about sentences of English. Remember that a sequent consists of a string of formulae — the assumptions — and another formula — the conclusion, separated by the turnstile, '$\vdash$'. In general, an arbitrary sequent will be shown as $X \vdash A$ or $Y \vdash B$ for a string of formulae X, Y, and a particular formula A or B. And always remember, we are trying to design our rules to ensure that they generate only valid sequents when applied to valid sequents.

First, then, the **elimination** rule for the conditional, '$\rightarrow$':

$$\frac{X \vdash A \rightarrow B \qquad Y \vdash A}{X, Y \vdash B} \quad \rightarrow E$$

The rule $\rightarrow$E (often referred to as Modus Ponendo Ponens, or simply, Modus Ponens), so schematised, could be equivalently characterised thus:

> Given a pair of sequents, one of which concludes in a conditional statement on certain assumptions, X, and the other of which concludes in the antecedent of that conditional on certain (not necessarily different) assumptions, Y, we may infer a sequent whose conclusion consists of the consequent of that conditional and whose assumptions consist of the **pool** of the assumptions of the original two sequents.

(The antecedent of a conditional, '$A \rightarrow B$' is the proposition A, and its consequent is the proposition B.) If you don't find this explanation immediately perspicuous, reflect on how much clearer it is to represent the rule schematically, as above.

Here is a simple example: suppose '$A \rightarrow B$' is '$(Q \vee R) \rightarrow S$', and A is '$Q \vee R$' itself. We had an example of the form, '$(Q \vee R) \rightarrow S$' in Part I, Ch. 1, p. 6: : suppose we know that if all the precautions I have taken have been either ineffectual or unnecessary, then I have been wasting my time taking them. If I now discover that indeed all the precautions I have taken have been either ineffectual or unnecessary, then I may conclude that I have been wasting my time taking them. '$(Q\vee R)\rightarrow S$' is one premise; the other premises entail '$Q\vee R$'; and so, by $\rightarrow$E (Modus Ponens) I may infer S, that is, that I have been wasting my time.

Now for the **introduction** rule for '$\rightarrow$', $\rightarrow$I. First, we need some more notation. Suppose we have a string of formulae, X. Then '$X\backslash A$' denotes the result of removing all self-standing occurrences of the formula A from the string X, if there are any. We can read it as 'X minus A'. For example, if X is:

$$(P\rightarrow Q)\rightarrow R,\ R\vee T,\ S\rightarrow P,\ T\rightarrow\neg R,\ (S\rightarrow(Q\rightarrow P))\vee\neg T,\ S$$

and A is 'S', then $X\backslash A$ is:

$$(P\rightarrow Q)\rightarrow R,\ R\vee T,\ S\rightarrow P,\ T\rightarrow\neg R,\ (S\rightarrow(Q\rightarrow P))\vee\neg T$$

Notice that there are still two occurrences of 'S' remaining; but we leave them in place because they are not self-standing but occur as parts of larger formulae.

(It may occur to the tidy-minded reader to wonder how "$X\backslash A$" should be read in case X does not contain the formula A. For the moment we make no pronouncement about this, but it is actually an issue of some significance to which we shall return below, in Ch. 7, p. 156.)

We now schematise the conditional-introduction rule as follows:

$$\frac{X\vdash B}{X\backslash A\vdash A\rightarrow B}\quad\rightarrow\mathrm{I}$$

Reflect on the intuitive import of this rule. What it allows is that, given any valid sequent in which a conclusion B is drawn from certain assumptions X, we can validly pass to a sequent in which any one of those assumptions — any one you care to choose — is lifted out and made into the antecedent of a conditional whose consequent is the original conclusion B; and this conditional then constitutes the conclusion of the new sequent whose assumptions consist in the remaining assumptions in the original sequent. And the intuitive justification is completely clear: the original sequent says that, given the truth of all the formulae in X, the truth of B follows; so evidently, given the truth of *all but one* of the formulae X, it follows that, if the remaining one is true, then so is B.

The standard term when we "lift out" an assumption in this way is that the assumption has been **discharged**. Several of our rules will involve a similar discharge of assumptions.

Think about it and you will see that the principle involved is really completely trivial. And that is as it should be: the more trivially obvious our rules, the better founded the resulting system of inference is going to be.

In due course we'll introduce rules for '$\vee$' and '&', and later for '$\neg$'. But first let's practice the workings of what we already have.

Obviously we cannot do very much with these rules as they stand, because we don't yet have sequents to apply them to! Each of them gives us a rule for passing to valid sequents *from* valid sequents; but neither of them gives us a starting point, some initial valid statements on which to go to work. We therefore adopt the following convention — often called the "rule of assumptions" (**Asmp**):

> Any sequent of the form $A \vdash A$, may be simply assumed.

This hardly stands in need of an intuitive justification! It just says that given any assumption A, A follows.

Now let's see what we can do. Suppose we are asked to prove (= show to be valid) the following sequent:

$$\mathsf{P}, \mathsf{P} \to \mathsf{Q} \vdash \mathsf{Q}.$$

We can proceed like this:

$$\begin{array}{rcll} \mathsf{P} & \vdash & \mathsf{P} & \quad \text{Asmp} \\ \mathsf{P} \to \mathsf{Q} & \vdash & \mathsf{P} \to \mathsf{Q} & \quad \text{Asmp} \\ \mathsf{P}, \mathsf{P} \to \mathsf{Q} & \vdash & \mathsf{Q} & \quad \to\text{E} \end{array}$$

Note how at the third line, after two examples of our allowable type of assumption-sequent we have straightforwardly applied $\to$E , pooling the assumptions on the left as indicated. And the third line is precisely what we were trying to prove.

Let's try another example:

$$\mathsf{P}, \mathsf{P} \to \mathsf{Q}, \mathsf{Q} \to \mathsf{R} \vdash \mathsf{R}$$

$P \to Q$	$\vdash$	$P \to Q$	Asmp
$Q \to R$	$\vdash$	$Q \to R$	Asmp
P	$\vdash$	P	Asmp
$P, P \to Q$	$\vdash$	Q	$\to$E
$P, P \to Q, Q \to R$	$\vdash$	R	$\to$E

Here we first apply $\to$E to the first and third lines to get the fourth line; then we apply it to the second and fourth lines to get the concluding line — which is what we set out to prove.

Check through these proofs carefully, and make sure you follow how we are keeping faith with our schematisation of the rules.

Note that we have slipped into the following terminological conventions:

> **formulae**: are expressions formed from propositional letters and connectives.
>
> **assumptions**: are the things that feature to the left of '$\vdash$' in a sequent. Every sequent thus has a number of assumptions and a conclusion.
>
> **premises**: are the things we apply our rules of inference to; thus they are whole sequents. E.g. the premises for an application of $\to$E will be a pair of sequents of the form: $X \vdash A \to B$; $Y \vdash A$.
>
> **assumption-sequents**: are sequents of the form $A \vdash A$ (we shall say that the formula, A, is assumed by writing such a sequent into a proof).
>
> **conclusions**: we shall use the term to denote both formulae that feature to the right of '$\vdash$' in a sequent and whole sequents resulting from an application of one or more of our rules of inference to certain premisses. Thus the conclusion of a sequent is a formula; and the conclusion of a proof is a sequent. Usually context will resolve any ambiguity. Where not, we shall speak of conclusion-sequents to denote the latter.

N.B. This is not everyone's terminology. Nonetheless, it has the advantage of enabling us to keep clear that we apply our rules of inference to sequents, not to their assumptions (= what some people call their premises!). We will continue with this terminology throughout this book.

The above proofs are in what we may call basic symbolism. We will now make various changes to simplify our presentation of proofs and, at the same time, make them easier to check. The result will be our preferred working symbolism for proofs:

(i) We leave out all occurrences of '$\vdash$' in the proof and number the lines consecutively instead.

(ii) We use the line number of an assumption-sequent to refer to its assumption — the formula, recall, to the left of the '$\vdash$' — in later lines which depend on the same assumption. In other words, we replace the left-hand column of formulae by a column of numbers and strings of numbers, the line numbers where those formulae were first assumed.

(iii) On the right, prior to citing the rule used in getting a certain line, we cite the numbers of the lines to which the rule is being applied and the line numbers of any assumptions discharged (two in the case of $\rightarrow$E , but later rules will require one, two, or even in one case, five lines to be cited).

Thus the second proof above becomes:

1	(1)	$P \rightarrow Q$	Asmp
2	(2)	$Q \rightarrow R$	Asmp
3	(3)	P	Asmp
1,3	(4)	Q	1,3 $\rightarrow$E
1,2,3	(5)	R	2,4 $\rightarrow$E

(You convert the other proof.) Note that in citing the premises of $\rightarrow$E we adopt the convention of first citing the line where the conditional occurred, and then the line where its antecedent was the conclusion. Nonetheless, the assumption numbers in the left-hand column are given in ascending order.

Now let's try to construct some proofs working in our preferred symbolism from the outset:

$$P,\ Q \rightarrow \neg P,\ P \rightarrow R,\ R \rightarrow Q \vdash \neg P$$

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

1,3	(5)	R	1,3 $\rightarrow$E
1,3,4	(6)	Q	4,5 $\rightarrow$E
1,2,3,4	(7)	$\neg$P	2,6 $\rightarrow$E

Notice that every line of the proof corresponds to a valid sequent: all we have to do to recover that sequent is specify the numbered assumptions on the left and replace the line number by '$\vdash$'. For example, the final line converts back to basic symbolism as

$$P, Q\rightarrow\neg P, P\rightarrow R, R\rightarrow Q \vdash \neg P \text{ ND}$$

which is exactly the sequent we set out to prove.

Now for a number of examples which involve $\rightarrow$I:

P$\rightarrow$Q, Q$\rightarrow$R $\vdash$ P$\rightarrow$R
(known as 'Transitivity of the Conditional')

1	(1)	P$\rightarrow$Q	Asmp
2	(2)	Q$\rightarrow$R	Asmp
3	(3)	P	Asmp
1,3	(4)	Q	1,3 $\rightarrow$E
1,2,3	(5)	R	2,4 $\rightarrow$E
1,2	(6)	P$\rightarrow$R	3,5 $\rightarrow$I

Notice that at line 3 we for the first time use an assumption-sequent other than one corresponding to the assumptions in the sequent we are trying to prove. This is OK because, using $\rightarrow$I we are going to discharge assumption 3 at line 6.

N.B. (i) how the number of assumptions drops by one at line (6); (ii) the convention for citing the lines on the right in the case of $\rightarrow$I: we cite first the line where the antecedent of the inferred conditional is first assumed, and then the line of the sequent which is providing the premise for the application of the rule; (iii) there are three sequents we could have proved at line 6 by $\rightarrow$I. The other two are:

1,3	(6)	(Q$\rightarrow$R)$\rightarrow$R	2,5 $\rightarrow$I

2,3	(6)	$(P\rightarrow Q)\rightarrow R$	1,5 $\rightarrow$I

Can you recover these sequents back into basic symbolism? Both of them are perfectly valid — it is just that they are not the sequent we set out to prove. However, note that if we had intended to derive either of these sequents, we would have begun the proof differently. This is an important part of strategy. First, put down the assumptions of the sequent to be proved. Then, if the conclusion to be proved is a conditional, assume its antecedent. The alternative moves at line 6 disrupt this tactical approach.

It is very important to realise that only assumptions may be discharged. For example, the correct use of $\rightarrow$I to prove a sequent,

$$X \vdash A\rightarrow B,$$

involves first assuming the formula, A, and then proceeding to derive the formula, B, from it together with the formulae in X. We may then apply $\rightarrow$I, discharging the assumption A, to conclude $A\rightarrow B$.

Some more examples to try:

$P\rightarrow Q,\ Q\rightarrow\neg R \vdash P\rightarrow\neg R$
$P\rightarrow(Q\rightarrow R) \vdash Q\rightarrow(P\rightarrow R)$
$P\rightarrow((Q\rightarrow P)\rightarrow R) \vdash (Q\rightarrow P)\rightarrow(P\rightarrow R)$
$P\rightarrow Q \vdash (Q\rightarrow R)\rightarrow(P\rightarrow R)$

Now work through the following proofs, noting the way the rules are applied and the strategies they exemplify:

(i) $P\rightarrow\neg\neg Q,\ P \vdash \neg\neg Q$

1	(1)	$P\rightarrow\neg\neg Q$	Asmp
2	(2)	P	Asmp
1,2	(3)	$\neg\neg Q$	1,2 $\rightarrow$E

(ii) $P \to (Q \to R) \vdash (P \to Q) \to (P \to R)$

1	(1)	$P \to (Q \to R)$	Asmp
2	(2)	$P \to Q$	Asmp
3	(3)	P	Asmp
1,3	(4)	$Q \to R$	1,3 $\to$E
2,3	(5)	Q	2,3 $\to$E
1,2,3	(6)	R	4,5 $\to$E
1,2	(7)	$P \to R$	3,6 $\to$I
1	(8)	$(P \to Q) \to (P \to R)$	2,7 $\to$I

Here again, going for a conditional conclusion, we assume its antecedent (line 2), and look for a derivation of its consequent. But that, in turn, is a conditional; so we assume its antecedent (line 3), and look for a proof of its consequent on that (among other) assumptions. Success is met with at line (6); two steps of $\to$I then do the trick.

(iii) $P \vdash (P \to Q) \to Q$

1	(1)	P	Asmp
2	(2)	$P \to Q$	Asmp
1,2	(3)	Q	2,1 $\to$E
1	(4)	$(P \to Q) \to Q$	2,3 $\to$I

(What other sequents could we have derived at line 4 by $\to$I? — again, ignoring thoughts about what we intended to do at line 4 when we started the proof as we did.)

(iv) $P \to (Q \to (R \to S)) \vdash R \to (P \to (Q \to S))$

1	(1)	$P \to (Q \to (R \to S))$	Asmp
2	(2)	R	Asmp
3	(3)	P	Asmp
4	(4)	Q	Asmp
1,3	(5)	$Q \to (R \to S)$	1,3 $\to$E
1,3,4	(6)	$R \to S$	5,4 $\to$E
1,2,3,4	(7)	S	6,2 $\to$E
1,2,3	(8)	$Q \to S$	4,7 $\to$I
1,2	(9)	$P \to (Q \to S)$	3,5 $\to$I
1	(10)	$R \to (P \to (Q \to S))$	2,9 $\to$I

Again, notice our basic $\to$I strategy in operation; and the careful attention to bracketing which is necessary.

Exercises

(1) Prove the examples from p. 92 using the rules Asmp, →I and →E :

(a) $P \to Q, Q \to \neg R \vdash P \to \neg R$

(b) $P \to (Q \to R) \vdash Q \to (P \to R)$

(c) $P \to (Q \to R) \vdash (P \to Q) \to (P \to R)$

(d) $P \to Q \vdash (Q \to R) \to (P \to R)$

(2) Prove the following sequents using the rules Asmp, →I and →E :

(a) $P \to (Q \to R), Q \vdash P \to R$

(b) $R \to S \vdash (P \to R) \to (P \to S)$

(c) $P \to (P \to Q) \vdash P \to Q$

(3) Formalise the following argument, and show that the sequent form is valid.

> It is sufficient for John to pass the logic exam that he does the logic exercises. He does the logic exercises provided that he is not working on his history essay. So unless he works on his history essay he will pass the logic.

(4) Formalise the following argument in sequent notation and show that it is of a valid form, and is therefore sound:

> Unless the timing is correctly set, the engine will misfire. If the engine misfires, it will overheat; So, if the timing was not correctly set, you didn't service the engine properly, for you can't have done if it overheats.

2 Conjunction and Disjunction

We now add four more rules of inference to our stock: introduction and elimination rules for both '&' and '$\vee$'. There are two rules for '&', schematised in the usual way.

First the introduction rule:

$$\frac{X \vdash A \qquad Y \vdash B}{X, Y \vdash A \,\&\, B} \qquad \text{Conjunction Introduction (\&I)}$$

This rule is perfectly straightforward. It says simply that, given any pair of sequents as premises, we may validly advance to a sequent which pools their assumptions and whose conclusion is the conjunction of their conclusions. (Remember that X and Y stand here for strings of formulae — one, many or even none.) And the rule is evidently well justified; for if A follows from the formulae X, and B follows from the formulae Y, then if every member of X and Y is true, A and B must be true as well — hence, since it is sufficient for the truth of a conjunction that both its conjuncts are true, the conjunction of A with B must be true.

The elimination rule has two cases:

$$\frac{X \vdash A \,\&\, B}{X \vdash A} \qquad \text{Conjunction Elimination (\&E)}$$

and

$$\frac{X \vdash A \,\&\, B}{X \vdash B} \qquad \text{Conjunction Elimination (\&E)}$$

This rule says simply that given as premise a sequent whose conclusion is a conjunction, we may validly advance to a sequent which retains the same assumptions but whose conclusion is one of the two conjuncts on its own. The intuitive justification of the rule is again completely clear: if, given certain assumptions, the truth of a conjunction follows, then — since, intuitively, a conjunction is true only if both its conjuncts are true — the truth of each of the conjuncts individually must follow from the same assumptions.

Let's exemplify the use of these two rules in a couple of proofs:

$P \to R \vdash (P \& Q) \to R$

1	(1)	$P \to R$	Asmp
2	(2)	$P \& Q$	Asmp
2	(3)	P	2 &E
1,2	(4)	R	1,3 →E
1	(5)	$(P \& Q) \to R$	2,4 →I

Since the conclusion of the sequent which we are trying to prove is a conditional, the obvious strategy is to assume its antecedent (line 2) and look for a derivation of its consequent, preparatory to a →I step. The proof is then quite straightforward. Note the application of &E at line 3 — as our schema indicates, the pool of assumptions remains the same (viz just '2'); on the right, we cite just one line, viz the premise to which the rule has been applied.

$P, Q, R \vdash (P \& Q) \& R$

1	(1)	P	Asmp
2	(2)	Q	Asmp
3	(3)	R	Asmp
1,2	(4)	$P \& Q$	1,2 &I
1,2,3	(5)	$(P \& Q) \& R$	4,3 &I

Here we build up the conclusion by successive steps of &I in the obvious way, pooling the assumptions of the premises to which the rule is applied on the left and citing the line numbers of both premises on the right. (Relate the two applications of &I to our schema and make sure they conform to the pattern which it dictates.) Note that we again adopt a convention, this time of citing the premises on the right in the order in which they are conjoined — so '4,3' conjoins 'P & Q' on line 4 with R on line 3 in that order.

Now for an example involving both conjunction rules:

$P \to Q,\ R \to S \vdash (P \& R) \to (Q \& S)$

1	(1)	$P \to Q$	Asmp
2	(2)	$R \to S$	Asmp
3	(3)	$P \& R$	Asmp
3	(4)	P	3, &E
3	(5)	R	3, &E
1,3	(6)	Q	1,4 →E
2,3	(7)	S	2,5 →E
1,2,3	(8)	$Q \& S$	6,7 &I
1,2	(9)	$(P \& R) \to (Q \& S)$	3,8 →I

Here again, trying to prove a sequent with a conditional conclusion, we go for the basic →I strategy, taking the appropriate assumption sequent at line 3. Then, using &E, we break it up at lines 4 and 5, so as to get to work with the '→'-elimination rules at lines 6 and 7. Next we build up the consequent of the aimed-at conditional by &I at line 8; finally a step of →I does the trick. (Check over the proof carefully, verifying that the correct assumptions for the &E and &I steps are given on the left and the correct lines are cited on the right.)

Next we introduce the two rules for '∨'. Again, the first rule has two cases:

$$\frac{X \vdash A}{X \vdash A \vee B} \qquad \text{Disjunction Introduction } (\vee\text{I})$$

and

$$\frac{X \vdash B}{X \vdash A \vee B} \qquad \text{Disjunction Introduction } (\vee\text{I})$$

This rule says in effect that, given any sequent, we may advance to any sequent which has the same assumptions but whose conclusion consists of the disjunction of the conclusion of the first sequent with any other proposition. Notice that the original conclusion can be either the first or the second disjunct in the new conclusion. The rule is intuitively well justified provided we understand '$\vee$' to convey the inclusive sense of 'or', where 'A or B' allows for the truth of either or both disjuncts; for, trivially, the truth of A is sufficient for the truth of the claim that at least one, and possibly both, of A, B are true — which is exactly what '$A \vee B$', interpreted in the inclusive sense, says.

The other disjunction rule is:

$$\frac{X \vdash A \vee B \quad Y \vdash C \quad Z \vdash C}{X, Y \backslash A, Z \backslash B \vdash C} \qquad \text{Disjunction Elimination } (\vee E)$$

This looks, and is, a little complicated. It is the most complicated, in fact, of all the rules of classical logic, and it is the only one which uses three premises. (Remember that $Y \backslash A$ means: the string of formulae Y with A deleted.) But the intuitive principle involved is not hard to grasp. Suppose you know that $A \vee B$ is true; then you know that at least one of its disjuncts is true, though you may not know which. Suppose you also know that C follows from each of the disjuncts, that C follows from A and that C follows from B. Then you know that C must be true — for either A is true or B is true and, whichever it is that is true, C follows from it! Here is a diagram:

$$\begin{tikzcd}
 & A \vee B \arrow[dl] \arrow[dr] & \\
A \arrow[d] & & B \arrow[d] \\
C \arrow[dr] & & C \arrow[dl] \\
 & C &
\end{tikzcd}$$

Starting from $A \vee B$, we can go either by route A to C, or by route B to C, so we can go to C either way. That, in essence, is all the $\vee$E rule comes to. But our statement of it has to be more involved because

(i) knowledge of $A \vee B$ may be contingent on certain assumptions, X;

(ii) there may be extra assumptions, Y (or strictly, $Y \backslash A$), involved in the derivation of C from A; and

(iii) there may be extra assumptions, $Z\backslash B$, involved in the derivation of C from B.

In that case your knowledge that C is true is contingent, strictly, on the truth of *all* of X, $Y\backslash A$ and $Z\backslash B$; so, in its most general form, the proper result of a $\vee$E step takes the form we have schematised. (In practice, though, X is often $A \vee B$ itself; while $Y\backslash A$ and $Z\backslash B$ are often empty.)

Some sample proofs, then. First, one using $\vee$I:

$P \& Q \vdash P \vee Q$

1	(1)	$P \& Q$	Asmp
1	(2)	P	1 &E
1	(3)	$P \vee Q$	2 $\vee$I

Note that we cite just one line on the right for the $\vee$I step, viz that of the premise to which the rule is applied. Another example:

$\neg Q \vdash P \rightarrow ((P \& \neg Q) \vee R)$

1	(1)	$\neg Q$	Asmp
2	(2)	P	Asmp
1,2	(3)	$P \& \neg Q$	2,1 &I
1,2	(4)	$(P \& \neg Q) \vee R$	3 $\vee$I
1	(5)	$P \rightarrow ((P \& \neg Q) \vee R)$	2,4 $\rightarrow$I

Note that the application of $\vee$I to the compound formula '$P \& \neg Q$' requires the introduction of brackets at line 4; and that a second pair of brackets is needed to avoid ambiguity in line 5. Note too that the basic $\rightarrow$I strategy is once again successful at proving a sequent with a conditional conclusion.

Now for a couple of examples involving Disjunction Elimination:

$P \vee Q \vdash (P \vee R) \vee (Q \vee R)$

1	(1)	$P \vee Q$	Asmp
2	(2)	P	Asmp
2	(3)	$P \vee R$	2 $\vee$I
2	(4)	$(P \vee R) \vee (Q \vee R)$	3 $\vee$I
5	(5)	Q	Asmp
5	(6)	$Q \vee R$	5 $\vee$I
5	(7)	$(P \vee R) \vee (Q \vee R)$	6 $\vee$I
1	(8)	$(P \vee R) \vee (Q \vee R)$	1,2,4,5,7 $\vee$E

This proof should be studied carefully. First the overall strategy: we are trying to show that a disjunctive conclusion follows from a particular assumption, so the most direct approach would be to derive one of the disjuncts from that assumption, and then use $\vee$I to get what we want. But intuitively, that direct strategy cannot work, since neither of the disjuncts '$P \vee R$' and '$Q \vee R$', actually does follow from '$P \vee Q$'. (Why not?) Reflecting, then, that our assumption is itself a disjunction, what we have done is show that one half of it entails one disjunct of the conclusion (at line 3) and that the other half entails the other disjunct (at line 6); so each half of the assumption entails the whole conclusion — by $\vee$I (lines 4 and 7). So, since the assumption is true only if at least one of its halves is true, the truth of the assumption guarantees the truth of the conclusion, by $\vee$E (line 8). Points to note:

(i) the use of bracketing to avoid ambiguity at lines 4, 7 and 8;

(ii) the formula '$(P \vee Q) \vee (P \vee R)$' occurs three times in the proof, but the job isn't over till line 8 since at line 4 we have proved $P \vdash (P \vee R) \vee (Q \vee R)$, which isn't what was asked, and at line 7 we have proved... (Well, what have we proved at line 7? Write line 7 in basic symbolism to make clear to yourself what it says);

(iii) a step of $\vee$E involves, alas, five citations on the right — one for each of its premises (here, lines 1, 4 and 7), plus a citation of each of the two lines (here, lines 2 and 5) where the relevant disjuncts are assumed. Note the strict order again: first, the line containing the disjunction, '$P \vee Q$'; then the line where P is assumed followed by the line where the conclusion, viz. $(P \vee R) \vee (Q \vee R)$ was derived from A; then the line where Q was assumed followed by, fifth and finally, the line where $(P \vee R) \vee (Q \vee R)$ was derived from B.

If we relate the $\vee$E-step in this proof back to our schema for $\vee$E it looks like this:

$$\dfrac{\dfrac{\text{line 1}}{P \vee Q \vdash P \vee Q} \qquad \dfrac{\text{line 4}}{P \vdash (P \vee Q) \vee (Q \vee R)} \qquad \dfrac{\text{line 7}}{Q \vdash (P \vee R) \vee (Q \vee R)}}{P \vee Q \vdash (P \vee Q) \vee (Q \vee R)}$$

Here, then, as we foretold would often happen, X is '$P \vee Q$' itself, while $Y \backslash A$ and $Z \backslash B$ are empty — so just one assumption is carried forward to make up the left hand side of the derived sequent.

Here's another example, combining $\vee$E with the basic $\rightarrow$I strategy:

$P \rightarrow S,\ Q \rightarrow S \vdash (P \vee Q) \rightarrow S$

1	(1)	$P \rightarrow S$	Asmp
2	(2)	$Q \rightarrow S$	Asmp
3	(3)	$P \vee Q$	Asmp
4	(4)	P	Asmp
1,4	(5)	S	1,4 $\rightarrow$E
6	(6)	Q	Asmp
2,6	(7)	S	2,6 $\rightarrow$E
1,2,3	(8)	S	3,4,5,6,7 $\vee$E
1,2	(9)	$(P \vee Q) \rightarrow S$	3,8 $\rightarrow$I

In order to show that the conditional conclusion follows from the given assumptions, we assume its antecedent in the usual way (line 3); but since the antecedent is a disjunction, we now have to show that the consequent follows from each of the two disjuncts in turn. Then we will be able to use $\vee$E. So at line 4 we assume the first disjunct and show the desired consequent follows from that (line 5); next we do the same for the second disjunct (lines 6 and 7).

Line 8 — the $\vee$E step — then says: "So, either way, whether it's P or Q that is true, the truth of S follows". Relating the moves back to the $\vee$E schema, what we have is:

$$\dfrac{\dfrac{\text{line 3}}{P \vee Q \vdash P \vee Q} \qquad \dfrac{\text{line 5}}{P \rightarrow S,\, P \vdash S} \qquad \dfrac{\text{line 7}}{Q \rightarrow S,\, Q \vdash S}}{P \vee Q,\, P \rightarrow S,\, Q \rightarrow S \vdash S}$$

Thus X is '$P \vee Q$'; Y\A is '$P \rightarrow S$' and Z\B is '$Q \rightarrow S$'; and each of these is carried forward into the assumption pool in the derived sequent. Again, note the five citations at line 8: the three premises, plus the line numbers of the assumption-sequents $P \vdash P$ and $Q \vdash Q$, where P and Q, the two relevant disjuncts, are assumed.

Disjunction elimination can seem a hard rule to master, but, with practice, it will soon become completely familiar.

Exercises

(1) Find proofs of the following sequents.

(a) $P \vdash Q \rightarrow (P \& Q)$

(b) $P \& (Q \& R) \vdash (P \& Q) \& R$

(c) $P \rightarrow Q,\ P \rightarrow R \vdash P \rightarrow (Q \& R)$

(d) $P \rightarrow Q \vdash P \rightarrow (Q \vee R)$

(e) $P \vee (Q \& R) \vdash (P \vee Q) \& (P \vee R)$

(f) $P \& (Q \vee R) \vdash (P \& Q) \vee (P \& R)$

(g) $(P \& Q) \vee (P \& R) \vdash P \& (Q \vee R)$

(h) $(P \vee Q) \& (P \vee R) \vdash P \vee (Q \& R)$

(i) $(P \vee Q) \rightarrow R \vdash (P \rightarrow R) \& (Q \rightarrow R)$

(j) $P \vee P \vdash P$

(Examples (e) – (h) are known as the Distributivity Laws. The proof of (h) is a little more subtle than the general run of proofs you have so far encountered, and some students may find it challenging.)

(2) Provide for the following sequent a proof in our usual style and a proof in basic symbolism.

$$P \rightarrow (Q \rightarrow R) \vdash (P \& Q) \rightarrow R$$

3 Negation

Negation also has an introduction and an elimination rule. Informally, the import of the introduction rule can be expressed as follows. Suppose that a **contradiction** may be shown to follow from a pool of assumptions. Then they cannot all be true — for, if they were, the contradiction would be true, which is absurd. But if they cannot all be true, then, if all but one are true, the last one must be false: so, on the assumption of all but one of the original pool, we can conclude that the negation of the last one holds true.

What for these purposes should should we take to be a contradiction? The simplest answer would be: a formula of the form '$B \& \neg B$'. But notice that it does not at all matter, for the purpose of the intuitve justification of the rule just sketched, *which* contradiction is shown to follow from a pool of assumptions — it does not matter what formula B is. *Any* contradiction will do; for no matter what contradiction is shown to follow from a pool of assumptions, it will be necessarily false and we shall have to conclude that, by the same token, they cannot all simultaneously be true. A device often employed in modern logic is to have a piece of notation for an *arbitrary* contradiction. We follow this practice here and introduce a special symbol '$\perp$', to denote an arbitrary contradiction. You may think of $\perp$ as, so to speak, the epitome of all contradictory propositions, the ultimate absurdity proposition. Using this notation, we may schematise our introduction rule for '$\neg$' thus:

$$\frac{X \vdash \perp}{X \backslash A \vdash \neg A} \quad \neg I$$

If a set of assumptions, X, yields absurdity, then one of them is false. We can choose whichever member A of X we wish. Then $\neg A$ follows from the remainder of X with A deleted — the assumption A has been discharged.

Of course, the question immediately arises, how do we establish the premise for a negation introduction step — how do we go about showing that $X \vdash \bot$? The answer is simple. We stipulate that a pool of asumptions entails $\bot$ just when it entails the separate ingredients of some contradiction. I.e., we add what is in effect an elimination rule for '$\neg$':

$$\frac{X \vdash \neg A \qquad Y \vdash A}{X, Y \vdash \bot} \quad \neg E$$

(Note that the elimination rule for '$\neg$' can be thought of as an introduction rule for '$\bot$'; and the introduction rule for '$\neg$' can be thought of as an elimination rule for '$\bot$'.)

The purpose of a negation introduction rule is, of course, to give us a way of proving sequents with negative conclusions. And the basic strategy associated with $\neg$I as formulated will be clear. Suppose we are set to prove $Y \vdash \neg A$: then all we have to do is derive a contradiction from Y together with A; i.e. prove a sequent of the form $Y, A \vdash \bot$, and then blame A for the contradiction in the fashion indicated.

Let's try a few examples:

(i) $P \to Q, \neg Q \vdash \neg(P \& R)$

1	(1)	$P \to Q$	Asmp
2	(2)	$\neg Q$	Asmp
3	(3)	$P \& R$	Asmp
3	(4)	P	3 &E
1,3	(5)	Q	1,4 $\to$E
1,2,3	(6)	$\bot$	2,5 $\neg$E
1,2	(7)	$\neg(P \& R)$	3,6 $\neg$I

Note: Line (3) is the formula whose negation we want to prove; and line (6) is a contradiction depending on the assumptions of the sequent we aim to prove, viz 1 and 2, and the one we aim to"reduce to absurdity", viz 3).

Notice how assumption number 3 disappears, being discharged at line (7); and note the lines cited on the right for an $\neg$ step; they are, first, the line where the negated formula first appears in an assumption-sequent; second, the line where we obtain the contradiction. Note too the convention for citing the lines in a $\neg$E step: we cite the two lines at which the premisses for the step occur, that with a negated conclusion being cited first.

Next:

(ii) $\neg(P \vee Q) \vdash \neg(P \& Q)$.

(This **contraposes** a sequent we proved a while back, viz $P \& Q \vdash P \vee Q$. The **contra-positive** of a sequent $A \vdash B$, is $\neg B \vdash \neg A$; likewise the contrapositive of a conditional $A \to B$, is $\neg B \to \neg A$.)

1	(1)	$\neg(P \vee Q)$	Asmp
2	(2)	$P \& Q$	Asmp
2	(3)	P	2 &E
2	(4)	$P \vee Q$	3 $\vee$I
1,2	(5)	$\perp$	1,4 $\neg$E
1	(6)	$\neg(P \& Q)$	2,5 $\neg$I

This is a nice, simple and typical $\neg$I proof: aiming for the conclusion $\neg(P \& Q)$, we take the assumption-sequent corresponding to the opposite at line (2). A contradiction straightforwardly and speedily follows (line (5)), with the result we want at line (6).

(iii) $P \to R, Q \to R, \neg R \vdash \neg(P \vee Q)$

1	(1)	$P \to R$	Asmp
2	(2)	$Q \to R$	Asmp
3	(3)	$\neg R$	Asmp
4	(4)	$P \vee Q$	Asmp
5	(5)	P	Asmp
1,5	(6)	R	1,5 $\to$E

1,3,5	(7)	$\bot$	3,6 $\neg$E
8	(8)	Q	Asmp
2,8	(9)	R	2,8 $\rightarrow$E
2,3,8	(10)	$\bot$	3,9 $\neg$E
1,2,3,4	(11)	$\bot$	4,5,7,8,10 $\vee$E
1,2,3	(12)	$\neg(P \vee Q)$	4,11 $\neg$I

Here, having listed the assumption-sequents corresponding to our given assumptions (lines (1)-(3)), we next take, at line 4, the opposite of the conclusion we want, aiming for a contradiction, $\bot$. But this assumption is a disjunction — so, in order to derive $\bot$ from it, we have to show that that $\bot$ follows from each of its disjuncts. Lines (5)-(7) exhibit the derivation of $\bot$ from the first disjunct, P; lines (8)-(10) do the same for the second disjunct, Q. Thus at line (11) we are in a position, using $\vee$E, to attribute the derivation of $\bot$ to $P \vee Q$. (Note again the working of our conventions for line citations in the case of $\vee$E: five lines, *viz* original disjunction, first disjunct, desired conclusion from first disjunct, second disjunct, desired conclusion from second disjunct.) At line (12) using $\neg$I, we blame the contradiction $\bot$ on $P \vee Q$ — and that gives us the sequent we set out to prove.

Here is a sequent which looks similar to ii), but whose conclusion is stronger:

(iv) $\neg(P \vee Q) \vdash \neg P \& \neg Q$

1	(1)	$\neg(P \vee Q)$	Asmp
2	(2)	P	Asmp
2	(3)	$P \vee Q$	2 $\vee$I
1,2	(4)	$\bot$	1,3 $\neg$E
1	(5)	$\neg P$	2,4 $\neg$I
6	(6)	Q	Asmp
6	(7)	$P \vee Q$	6 $\vee$I
1,6	(8)	$\bot$	1,7 $\neg$E
1	(9)	$\neg Q$	6,8 $\neg$I

1	(10)	$\neg P \& \neg Q$	5,10 &I

Study this proof carefully. We aim to derive a conjunctive conclusion; so the obvious strategy is to derive each conjunct separately and then use &I to bring them together. But each conjunct is a negated formula; so the obvious strategy is to try to prove each by ¬I. And this is just what we have done: lines (2)-(4) reduce P to absurdity, giving us ¬P at line (5); and lines (6)-(8) reduce Q to absurdity, giving us ¬Q at line (9); all that remains is to conjoin ¬P and ¬Q together at line (10). Quite a smart proof — but one whose strategy is dictated perfectly straightforwardly just by reflecting on the structure of the conclusion of the sequent we are aiming for. Try to get into the habit of letting the structure of what you are trying to prove dictate how you go about it — just as we did when proving sequents with conditional conclusions. The basic approach is: try to see how to get the conclusion from the assumptions using the introduction rule corresponding to the main connective in the conclusion; and let the same strategy govern your attempt to derive any subsidiary conclusions which you then have to go for.

Now try to prove this: $P \vee Q \vdash \neg(\neg P \& \neg Q)$. It is obviously interestingly related to what we have just proved; try to be self-conscious about the strategy, in the way just sketched. We shall give a proof later — don't look for it now! Do it yourself. You can see later if yours is the same.

(Recall that we remarked that the rule of inference →E was traditionally called **Modus Ponendo Ponens** (MPP) — (literally, the way of asserting (putting — Latin: ponens) the consequent by asserting (by putting — ponendo) the antecedent. Another traditional mode was **Modus Tollendo Tollens** (MTT) — the way of denying (taking away — tollens) the antecedent by denying (by taking away — tollendo) the consequent:

(v) $P \to Q, \neg Q \vdash \neg P$ MTT

We can now easily show the validity of this sequent:

1	(1)	$P \to Q$	Asmp
2	(2)	$\neg Q$	Asmp
3	(3)	P	Asmp
1,3	(4)	Q	1,3 →E
1,2,3	(5)	$\perp$	2,4 ¬E

1,2	(6)	$\neg P$	3,5 $\neg$I

We have now introduced no less than nine rules in all — introduction and elimination rules for each of our four connectives, plus the rule of Assumptions. However, classical propositional logic involves one further rule. We can lead up to it by raising, as an aside, the following interesting question. What should be our attitude to the sequent

(vi) $P \& \neg P \vdash \neg Q$?

Ought it to be regarded as valid? Well, according to our original intuitive account, a sequent is valid just in case all instances of it are sound; and a sound argument is one where it is impossible that the premises should be all true and the conclusion be false. Now, when the premisses of an argument embody a contradiction, it is — of course — impossible that they should all be true, so — in consequence — impossible that they should be true while the conclusion, whatever it is, is false. So it seems that any instance of (vi) should indeed be regarded as sound, and (vi) is therefore valid. If we buy this line, we will want to be able to prove (vi) using our basic rules. But can we?

Yes, using a ploy that will strike you as subtle or perverse, according to temperament. We take what we are given

1	(1)	$P \& \neg P$	Asmp

and then proceed:

2	(2)	Q	Asmp
1,2	(3)	$(P \& \neg P) \& Q$	1,2 &I
1,2	(4)	$P \& \neg P$	3 &E
1,2	(5)	P	4 &E
1,2	(6)	$\neg P$	4 &E
1,2	(7)	$\perp$	6,5 $\neg$E
1	(8)	$\neg Q$	2,7 $\neg$I

Look over this proof carefully. The key move is the combined play with &I and &E at lines (3) and (4) which has the effect that Q is insinuated into the pool of assumptions on which, at line (4), $P \& \neg P$ depends.

We shall return to the wider implications of this stratagem in chapter 7. But now reflect that the informal justification for the sequent just proved proceeded in a way which was entirely independent of the identity of its conclusion. The conclusion happened to be $\neg Q$ but it might just as well have been Q, for the informal justification showed — if it showed anything — that from a contradiction anything follows. Just as it is impossible for $P \& \neg P$ to be true and $\neg Q$ false (since it is impossible for $P \& \neg P$ to be true), so too it is impossible for $P \& \neg P$ to be true and Q false. So Q follows from $P \& \neg P$ just as $\neg Q$ does.

But what of $P \& \neg P \vdash Q$? Obviously, we cannot derive it by a proof of the same structure as the above, a $\neg$I proof, since it doesn't have a negated conclusion. And in fact we cannot yet derive this sequent in our system. To remedy matters, we need a final rule, the rule of **double negation elimination**, DN:

$$\frac{X \vdash \neg\neg A}{X \vdash A} \qquad \text{DN}$$

With DN on board, essentially the same proof strategy does succeed:

(vii) $P \& \neg P \vdash Q$

1	(1)	$P \& \neg P$	Asmp
2	(2)	$\neg Q$	Asmp
1,2	(3)	$(P \& \neg P) \& \neg Q$	1,2 &I
1,2	(4)	$P \& \neg P$	3 &E
1,2	(5)	P	4 &E
1,2	(6)	$\neg P$	4 &E
1,2	(7)	$\perp$	6,5 $\neg$E
1	(8)	$\neg\neg Q$	2,7 $\neg$I
1	(9)	Q	8 DN

Here $\neg Q$ simply replaces Q throughout the proof of (vi); we consequently arrive at a doubly negated conclusion at line (8), and an application of DN then turns the trick.

The sequent,

$$A, \neg A \vdash B$$

is a famous one. It is often known as **Ex Falso Quodlibet**, or EFQ for short. Given A and $\neg A$, anything (quodlibet) follows.

Note that we need DN in a wide range of cases. E.g.

(viii) $\neg P \rightarrow \neg Q \vdash Q \rightarrow P$

1	(1)	$\neg P \rightarrow \neg Q$	Asmp
2	(2)	Q	Asmp
3	(3)	$\neg P$	Asmp
1,3	(4)	$\neg Q$	3,1 $\rightarrow$E
1,2,3	(5)	$\perp$	4,2 $\neg$E
1,2	(6)	$\neg\neg P$	3,5 $\neg$I
1,2	(7)	P	6 DN
1	(8)	$Q \rightarrow P$	2,7 $\rightarrow$I

It is straightforward to establish the converse of DN:

(ix) $P \vdash \neg\neg P$

1	(1)	P	Asmp
2	(2)	$\neg P$	Asmp
1,2	(3)	$\perp$	2,1 $\neg$E
1	(4)	$\neg\neg P$	2,3 $\neg$I

Two further examples for practicing strategic thinking:

(x) $\neg P \vdash \neg(P \,\&\, Q) \vee R$

We take what we are given

1	(1)	$\neg P$	Asmp

and now reflect: the main connective in the sought-for conclusion is disjunction — that suggests we try to derive one of the two disjuncts separately, and then use $\vee$I to finish off. But the second disjunct, R, obviously doesn't follow from $\neg P$ so if the strategy is going to work we'll have to go for the first disjunct: $\neg(P \,\&\, Q)$. This is a negated formula — that suggests we take an assumption-sequent corresponding to its opposite and aim at a contradiction, using $\neg$I to finish off. So let's do so:

2	(2)	$P \,\&\, Q$	Asmp

We are now looking for a contradiction, $\perp$ — and a glance at what we have so far shows the way: get P from $P \,\&\, Q$ by &E, which contradicts $\neg P$. The proof now comes out quite mechanically:

2	(3)	P	2 &E
1,2	(4)	$\perp$	1,3 $\neg$E
1	(5)	$\neg(P \,\&\, Q)$	2,4 $\neg$I
1	(6)	$\neg(P \,\&\, Q) \vee R$	5 $\vee$I

Next:

(xi) $P \to \neg P \vdash \neg P$

1	(1)	$P \to \neg P$	Asmp
2	(2)	P	Asmp
1,2	(3)	$\neg P$	1,2 $\to$E
1,2	(4)	$\perp$	3,2 $\neg$E
1	(5)	$\neg P$	2,4 $\neg$I

Finally, a remark on two well-known fallacies, i.e. mistaken forms of inference, beloved by politicians and other free thinkers: **Affirming the Consequent** and **Denying the Antecedent**. Schematically, they are cock-eyed versions of $\to$E (Modus Ponendo Ponens) and Modus Tollendo Tollens respectively:

FALLACY! $$\frac{X \vdash A \to B \qquad Y \vdash B}{X, Y \vdash A}$$ Affirming the Consequent

FALLACY! $$\frac{X \vdash A \to B \qquad Y \vdash \neg A}{X, Y \vdash \neg B}$$ Denying the Antecedent

Don't *ever* slip into these patterns of inference! They are quite spurious. (Think up your own counterexamples — or scan *Hansard* and the leader pages of the tabloid press.)

Exercises

(1) Find proofs of the following sequents:

(a) $P \rightarrow S,\ Q \rightarrow \neg S \vdash \neg(P \,\&\, Q)$

(b) $P \rightarrow Q,\ Q \rightarrow \neg P \vdash \neg P$

(c) $P \rightarrow (Q \vee R),\ \neg Q,\ \neg R \vdash \neg P$

(d) $\neg P \rightarrow P \vdash P$

(2) Formalise the following argument and show that the argument is sound:

> Unless Kenyon fails to prolong the first game beyond half an hour, he has a good chance. But if Briars achieves his best form, Kenyon has no chance. Briars will win the first game only if he settles down quickly. So, provided that Kenyon prolongs the first game over the half hour or Briars is slow to settle, the latter will lose the first game or fail to achieve his best form.

4 Strategic Thinking

We have often spoken of the (basic) $\rightarrow$I tactic, and of the $\neg$I tactic. It is now time to reflect at greater length and in a formally more explicit fashion on the basic strategies needed in finding and constructing a proof.

The goal of a proof is to find a succession of sequents the last one of which is the particular sequent we wish to prove. Each sequent in the proof is either an assumption-sequent, or follows from earlier sequents as the conclusion of a rule of inference with those earlier sequents as premises.

With the exception of the rules of double negation (DN) and of Assumptions, the rules of inference so far introduced pair off into introduction rules and elimination rules. Introduction rules introduce an occurrence of a connective; the conclusion contains a formula whose main connective is introduced by the rule. Thus, e.g., the conclusion of &I is the sequent

$$X, Y \vdash A \,\&\, B,$$

where the displayed occurrence of '&' is "introduced" by the rule. All the other occurrences of connectives in the premises of the rule,

$$X \vdash A$$
$$Y \vdash B,$$

occur again in the conclusion; in addition, the displayed occurrence of '&' is created by the application of &I. A similar story can be told of $\vee$I, $\rightarrow$I and $\neg$I.

Elimination rules work to opposite effect, and eliminate or remove displayed occurrences of connectives. E.g., let us take, without loss of generality,

$$\frac{X \vdash A \,\&\, B}{X \vdash A}$$

as an instance of &E. The effect of &E is to remove the occurrence of '&' displayed in '$A \,\&\, B$'. (It also, of course, removes all occurrences of connectives in B, which are not displayed.) Again, a similar story can be told for $\vee$E, $\rightarrow$E and $\neg$E.

To each rule of inference there corresponds a particular tactic in the overall strategy of finding a proof. Let us consider an example. Suppose we wish to find a proof of the sequent

$$P \rightarrow Q,\ P \rightarrow \neg Q \vdash \neg P,$$

to derive $\neg P$ from the formulae $P \rightarrow Q$ and $P \rightarrow \neg Q$. We will achieve this goal if we assume

1	(1)	$P \rightarrow Q$	Asmp
2	(2)	$P \rightarrow \neg Q$	Asmp

and attempt to fill in the steps of a proof which terminates in the sequent

1,2	(n)	$\neg P$	?

To find the proof, we can start from either end, both working back from the goal, and working forwards from the assumption-sequents. The technique is often called the "top-down" approach, because one starts at, so to say, the top level, attempting to derive $\neg P$ from 1 and 2, and works "down", or better, into the proof, identifying subgoals more deeply embedded in the structure of the proof. For example, we here see that we might first start with a $\neg$I tactic, realising that we will be able to use the $\neg$I rule to obtain our goal sequent if we could first achieve the sub-goal of deriving '$\bot$' from the assumption of P (given $P \rightarrow Q$ and $P \rightarrow \neg Q$ as well):

1	(1)	$P \rightarrow Q$	Asmp
2	(2)	$P \rightarrow \neg Q$	Asmp
3	(3)	P	Asmp

		$\vdots$	
1,2,3	$(n-1)$	$\bot$	?
1,2	(n)	$\neg P$	3, $n-1$ $\neg$I

Our subgoal is now to find a proof of line $n-1$, that is, to derive '$\bot$' from lines 1, 2 and 3.

How is this subsidiary goal to be achieved? One way results from two applications of the $\to$E tactic. This time, instead of working from the end of the proof backwards (that is, by an introduction rule), we are working forwards from the beginning of the proof, using elimination rules:

1	(1)	$P \to Q$	Asmp
2	(2)	$P \to \neg Q$	Asmp
3	(3)	P	Asmp
1,3	(4)	Q	1,3 $\to$E
2,3	(5)	$\neg Q$	2,3 $\to$E
		$\vdots$	
1,2,3	$(n-1)$	$\bot$	?
1,2	(n)	$\neg P$	3, $n-1$ $\neg$I

Finally, we realise that the $\neg$E tactic will complete the proof, for an application of $\neg$E to lines 4 and 5 will give us line $n-1$ — which now becomes line 6: that is, we realise that line $n-1$ results immediately from lines 4 and 5 by $\neg$E:

1	(1)	$P \to Q$	Asmp
2	(2)	$P \to \neg Q$	Asmp
3	(3)	P	Asmp
1,3	(4)	Q	1,3 $\to$E
2,3	(5)	$\neg Q$	2,3 $\to$E

1,2,3	(6)	$\bot$	4,5 $\neg$E
1,2	(7)	$\neg$P	3,6 $\neg$I

The proof strategy breaks up, then, into the successive application of, in this case, a fairly obvious succession of tactics:

$\neg$I tactic
$\to$E tactic
$\to$E tactic
$\neg$E tactic

Not all proofs respond so well to the strategic approach. Nonetheless, adopting a strategy will, in general, constitute the most promising way of finding a proof.

We now schematise the tactics corresponding to each of the introduction and elimination rules. Consider our usual mode of schematising, say, $\vee$I:

$$\frac{X \vdash A}{X \vdash A \vee B} \quad \vee\mathrm{I}$$

Here the elements of the list X of formulae are all **assumptions**. But when it comes to characterising tactics, we need to take account of a difference. For our tactical goal, at a particular point in a proof, may not only involve deriving a certain formula from assumptions, but also deriving a formula from various other **derived formulae**. For example, at line 6 in the above proof, we derived $\bot$ from $\neg$Q and Q. But $\neg$Q and Q were not assumptions, but were themselves derived from assumptions by application of $\to$E . In other words, our goal at line $n-1$ had become, after use of the $\to$E tactics, that of deriving $\bot$ from $P \to Q$, $P \to \neg Q$, P, Q, and $\neg$Q. So we shall represent tactics by breaking a problem, or goal

$$F \Rightarrow A$$

into one or more subproblems, or sub-goals,

$$F_1 \Rightarrow A_1,\ F_2 \Rightarrow A_2, \ldots$$

where F, F_1, F_2 etc. are (what we shall call) **fact-lists**, that is, lists of formulae which may be assumptions, or may equally be themselves derived formulae. The full set of tactics for a particular proof will be called a **derivation**.

First, we shall give the introduction rules and the corresponding tactics, used for working backwards from the goal and the subgoals:

	Rule	Tactic
&I	$\dfrac{X \vdash A \qquad Y \vdash B}{X, Y \vdash A \& B}$	To show $F \Rightarrow A \& B$, attempt to show both $F \Rightarrow A$ and $F \Rightarrow B$.
∨I	$\dfrac{X \vdash A}{X \vdash A \vee B}$ $\dfrac{X \vdash B}{X \vdash A \vee B}$	To show $F \Rightarrow A \vee B$, attempt to show either $F \Rightarrow A$ or $F \Rightarrow B$
→I	$\dfrac{X \vdash B}{X \backslash A \vdash A \rightarrow B}$	To show $F \Rightarrow A \rightarrow B$, attempt to show $A, F \Rightarrow B$
¬I	$\dfrac{X \vdash \bot}{X \backslash A \vdash \neg A}$	To show $F \Rightarrow \neg A$, attempt to show $A, F \Rightarrow \bot$

That should seem reasonably self-explanatory. Now we need to characterise the tactics associated with the elimination rules, which we use for working *forwards* from the beginning of a proof, or from the elements of a fact-list, towards its conclusion. Care is needed here. For our goal may not result from an immediate application of the elimination rule corresponding to this tactic, which may merely take us (we hope) closer to that goal. For example, in the proof above, we applied the →E tactic when our goal was

$$P \rightarrow Q,\ P \rightarrow \neg Q,\ P \Rightarrow \bot$$

If we were to characterise the tactic involved as follows:

$$\begin{array}{rl} \text{to show} & F \Rightarrow B, \\ \text{check that} & A \rightarrow B \in F \text{ (i.e. that } A \rightarrow B \text{ is a member of } F\text{)} \\ \text{and attempt to show} & F \Rightarrow A, \end{array}$$

we would presuppose the occurrence a formula in F (which here consists of P, $P \rightarrow Q$, and $P \rightarrow \neg Q$) of the form $A \rightarrow \bot$ for some A. So then the tactic would be inapplicable, for there is no such formula. Instead, we must characterise the →E tactic in more general terms. We take the goal to be

$$F \Rightarrow C$$

to obtain which

we check that	$A \rightarrow B \in F$, for some A and B,
and attempt first to show	$F \Rightarrow A$,
and then to show	$B, F \Rightarrow C$.

In the particular example we first show that we can derive A from F (which in the case above is immediate, since $A \rightarrow B$ is $P \rightarrow Q$, and P is a member of the fact list in question), and then go onto the second subgoal — $B, F \Rightarrow C$ — which in this case is

$$Q, P, P \rightarrow Q, P \rightarrow \neg Q \Rightarrow \bot,$$

which we achieved by a further application of the $\rightarrow$E tactic, followed by a final application of the tactic for $\neg$E.

Note that the goal $F \Rightarrow A$ is achieved immediately if $A \in F$. For A is the formula to be derived (from certain assumptions), and F is a list of the things so far derived (including, and all ultimately dependent on, our assumptions). In fact, there will often be redundancy in F. Think of the &I-tactic given above: even if everything in the current fact-list F is needed to derive $A \,\&\, B$, it is very unlikely that they will all be needed for each of A and B separately. Nonetheless, we do not know which parts of F are needed to show A, and which to show B. So we set ourselves the goals $F \Rightarrow A$ and $F \Rightarrow B$, knowing full well that not all of F will in each case be needed.

This is a particular aspect of an important logical point. If the sequent $X \vdash A$ is derivable, then surely if every member of X is also a member of Y — that is, if Y either is X or differs just by containing additional formulae — then $Y \vdash A$ should be derivable too. (We write this condition as $X \subseteq Y$, that is, X is contained in Y.) We say that Y is a **thinning** of X, or that $Y \vdash A$ is obtained from $X \vdash A$ by **Thinning** (sometimes called "Weakening" or "Dilution"). In particular, if $X \vdash A$, then $X, B \vdash A$ for any formula B.

Conversely, we may treat $F \Rightarrow B$ as a subgoal for the goal $F, A \Rightarrow B$ by the *tactic* of Thinning. If we can achieve the subgoal (i.e. $F \Rightarrow B$) then certainly we can achieve the goal (i.e. $F, A \Rightarrow B$), since $F \subseteq F, A$. (In particular, if $A \in F$, the goal $F \Rightarrow A$ is immediately satisfied by Thinning. That is, clearly the goal $A \Rightarrow A$ is achievable: given A, derive A; hence, by Thinning, the goal $F \Rightarrow A$ is achievable, for $A \in F$. We say that $F \Rightarrow A$ is **immediate** if $A \in F$.)

Here is a somewhat abstract, summary statement of the tactics corresponding to each of the elimination rules, interspersed with some more concrete commentary:

	<u>Rule</u>	<u>Tactic</u>	
&E	$\dfrac{X \vdash A \& B}{X \vdash A}$ $\quad$ $\dfrac{X \vdash A \& B}{X \vdash B}$	To show	$F \Rightarrow C$,
		check that	$A \& B \in F$ for some A and B,
		and attempt to show	$A, B, F \Rightarrow C$.

Although the rule &E has two cases, we cover them both in a single &E tactic. What it says is that, if in pursuit of some goal formula C we already know $A \& B$ (i.e. $A \& B$ is one of our "facts"), then we may add A and B separately to what we know. With the rule, we must choose which of A and B to extract — or if we want both, to take them out successively. With the tactic, we do it all at once.

Often, C will already be one of the conjuncts in $A \& B$, and so the resulting goal will be immediate. Consider, for example, the proof of $\neg(P \vee Q) \vdash \neg(P \& Q)$ on p. 109 above. After applying the $\neg$I and $\neg$E tactics, our goal is

$$P \& Q, \neg(P \vee Q) \Rightarrow P \vee Q$$

(where the "fact" $\neg(P \vee Q)$ is redundant); and so applying the $\vee$I-left tactic, we want to show

$$P \& Q, \neg(P \vee Q) \Rightarrow P.$$

Applying the &E tactic, we split $P \& Q$ up into its components:

$$P, Q, P \& Q, \neg(P \vee Q) \Rightarrow P.$$

This is immediate. We wanted P, and the &E tactic gave it to us.

It will not always be so easy. For example, in the proofs on p. 98, &E is followed in the proof by $\rightarrow$E . This means that if we apply the &E tactic to the subgoal

$$P \& Q, P \rightarrow R \Rightarrow R$$

we obtain

$$P,\ Q,\ P \,\&\, Q,\ P \rightarrow R \Rightarrow R$$

and an application of the $\rightarrow$E tactic is now needed to complete the derivation.

Now for the $\vee$E tactic:

	Rule	Tactic
$\vee$E	$\dfrac{\begin{array}{c} X \vdash A \vee B \qquad Y \vdash C \\ Z \vdash C \end{array}}{X,\ Y\backslash A,\ Z\backslash B \vdash C}$	To show $F \Rightarrow C$, check that $A \vee B \in F$ for some A and B, attempt to show $A,\ F \Rightarrow C$ *and* $B,\ F \Rightarrow C$.

Once again, the general form of the tactic may look intimidating, just as the full form of the rule does on first acquaintance. However, if we look at a particular case, it should be easier to follow. For example, take the proof of $P \rightarrow S,\ Q \rightarrow S \vdash (P \vee Q) \rightarrow S$ on p. 103. After the initial $\rightarrow$I tactic, we obtain as our subgoal:

$$P \vee Q,\ P \rightarrow S,\ Q \rightarrow S \Rightarrow S.$$

Applying the $\vee$E tactic to this yields two subgoals, corresponding to the two subproofs

$$P,\ P \rightarrow S,\ Q \rightarrow S \Rightarrow S$$

(the goal corresponding to the step from lines 1, 2 and 4 to line 5 of the proof) and

$$Q,\ P \rightarrow S,\ Q \rightarrow S \Rightarrow S$$

(corresponding to the step from lines 1, 2 and 6 to line 7). In other words, if one has a goal where some disjunction $A \vee B$ is in the fact-list, one must achieve two subgoals, one for A and the other for B.

The $\rightarrow$E tactic looks like this:

Rule		Tactic	
$\to$E	$\dfrac{X \vdash A \to B \quad Y \vdash A}{X, Y \vdash B}$	To show check that attempt to show *and*	$F \Rightarrow C$, $A \to B \in F$ for some A and B, $F \Rightarrow A$ $B, F \Rightarrow C$.

It may at first seem puzzling that applying the $\to$E tactic should generate two subgoals. However, recall what was said at p. 124. To use a "fact" $A \to B$, we first need to establish its antecedent A; that is the point of the first subgoal, $F \Rightarrow A$. $A \to B$ and A together yield B (by the *rule* $\to$E), so we now have B in our fact-list. But then we must go back to our original goal $F \Rightarrow C$, and (with B added) try to achieve C; i.e. the second subgoal is $B, F \Rightarrow C$. Consider, for example, the proof of $P, P \to Q, Q \to R \vdash R$ on p. 88. The goals read as follows:

Original goal: $\quad P, P \to Q, Q \to R \Rightarrow R$

We first apply the $\to$E tactic to $P \to Q$, to obtain

$$P, P \to Q, Q \to R \Rightarrow P$$

which is immediate, for we already have P in our fact-list, and so can add Q to the list:

$$Q, P, Q \to R \Rightarrow R$$

(we can drop $P \to Q$, since we have used it). We now apply the $\to$E tactic again, this time to $Q \to R$, and obtain first

$$Q, P, Q \to R \Rightarrow Q$$

(which again is immediate) and secondly

$$R, Q, P \Rightarrow R,$$

which completes the derivation.

Lastly, we have the $\neg$E tactic:

	Rule	Tactic	
$\neg$E	$\dfrac{X \vdash \neg A \qquad Y \vdash A}{X, Y \vdash \bot}$	To show	$F \Rightarrow C$,
		check that	$\neg A \in F$ for some A
		attempt to show	$F \Rightarrow A$,
		and	$\bot, F \Rightarrow C$.

Recall that ¬E can be seen equally well as an introduction rule for ⊥ — it's just a way of recording that we have derived the separate components of some contradiction ⊥. In fact, in almost every case of application, the goal C will actually be ⊥ itself. Consider for example, the proof of $P \to Q, \neg Q \vdash \neg P$ on p. 111. At line 5, our goal (following initial application of the ¬I tactic) is

$$P,\ P \to Q, \neg Q \Rightarrow \bot.$$

So there is a formula $\neg Q$ in our fact-list. The ¬E tactic then suggests first that we should try to prove Q:

$$P,\ P \to Q, \neg Q \Rightarrow Q$$

(which we will get by the →E tactic), so that, Q and $\neg Q$ being established, we can add ⊥ to our fact-list by the ¬E *rule* and revert to our original goal — which will now be immediate:

$$\bot, P,\ P \to Q \Rightarrow \bot.$$

The above rules and tactics define a logic often referred to as **minimal logic**. It is strictly weaker than classical logic, **NK**. (One logic is weaker than another if every sequent provable in the first is also provable in the second, yet in the second additional sequents are provable.) Every sequent provable in minimal logic is provable in **NK**. But additional sequents — for example, $A, \neg A \vdash B$ — are provable in **NK** which, as we know, contains one further rule, DN. Here is a schematisation of the corresponding tactic:

	Rule	Tactic	
DN	$\dfrac{X \vdash \neg\neg A}{X \vdash A}$	To show	$F \Rightarrow C$,
		either attempt to show	$F \Rightarrow \neg\neg C$
		or check that	$\neg\neg A \in F$, for some A,
		and attempt to show	$A, F \Rightarrow C$

In other words, the DN tactic is either to work backwards directly on the goal, C, setting a new goal, $\neg\neg C$; or to work forwards from some already derived formula $\neg\neg A$, by applying DN as an elimination rule.

There are two cases here. For the moment, we can set the second aside. It will only be needed later. The first is straightforward, and makes explicit the tactic we followed in Ch. 3, p. 114. Consider the proof of viii) $\neg P \rightarrow \neg Q \vdash Q \rightarrow P$ there. After the initial $\rightarrow$I tactic, we need to solve the subgoal

$$Q, \neg P \rightarrow \neg Q \Rightarrow P.$$

Applying the DN tactic, we convert this goal into

$$Q, \neg P \rightarrow \neg Q \Rightarrow \neg\neg P,$$

which we then solve by the $\neg$I, $\rightarrow$E and $\neg$E tactics.

Finally, before we return to our first example to apply what we have learned: note that we can also work out from the tactical derivation the maximal length of our proof. Each tactic generates a predictable number of lines of proof. The only exception is the &E tactic, which will generate either one or two lines of proof, depending on whether or not both conjuncts are actually used in the proof. Moreover, sometimes a proof will be shorter than predicted, as we will see in due course.

The &I, $\vee$I, $\rightarrow$E , $\neg$E and DN tactics will each generate 1 line of proof, where they are applied to their premises, once they belong to the fact-list. The &E tactic generates either 1 or 2 lines. It is more subtle with the remaining tactics, for they each require the generation of an assumption. The $\vee$E tactic will generate 3 lines, 1 for each of the disjuncts assumed and another for the application of the rule; the $\rightarrow$I and $\neg$I tactics will generate 2 lines each, 1 for the assumption and 1 for the application. Finally, there will be 1 line for each assumption in the sequent to be proved. Here is a table:

Tactic	Number of lines generated
&I	1
&E	1 or 2
∨I	1
∨E	3
→I	2
→E	1
¬I	2
¬E	1
DN	1

Let us now return to our proof that $P\to Q, P\to\neg Q \vdash \neg P$, and follow through the tactics as set out above. To show the trivial — the **immediate** — goals (where the goal formula already belongs to the fact-list), we write: ■.

? $P\to Q, P\to\neg Q \Rightarrow \neg P$
> Using tactic for ¬I
> ? $P, P\to Q, P\to\neg Q \Rightarrow \bot$
>> Using tactic for →E
>> ? $P, P\to Q, P\to\neg Q \Rightarrow P$■
>> ? $Q, P, P\to\neg Q \Rightarrow \bot$
>>> Using tactic for →E
>>> ? $Q, P, P\to\neg Q \Rightarrow P$ ■
>>> ? $\neg Q, Q, P \Rightarrow \bot$
>>>> Using tactic for ¬E
>>>> ? $\neg Q, Q, P \Rightarrow Q$ ■
>>>> ? $\bot, Q, P \Rightarrow \bot$ ■

We quickly calculate, therefore, that the corresponding proof will have 7 lines — 2 assumptions, plus 2 for ¬I and 1 each for two applications of →E and one of ¬E. We saw it before, on p. 121:

1	(1)	$P \to Q$	Asmp
2	(2)	$P \to \neg Q$	Asmp
3	(3)	P	Asmp
2,3	(4)	$\neg Q$	2,3 $\to$E
1,3	(5)	Q	1,3 $\to$E
1,2,3	(6)	$\bot$	4,5 $\neg$E
1,2	(7)	$\neg P$	3,6 $\neg$I

Here are two more examples, first giving the tactical derivations, followed by the proofs interlaced with remarks about strategy:

$$P \to (Q \& R) \vdash (P \to Q) \& (P \to R)$$

Here is the tactical derivation:

? $P \to (Q \& R) \Rightarrow (P \to Q) \& (P \to R)$
- Using tactic for &I
- ? $P \to (Q \& R) \Rightarrow P \to Q$
  - Using tactic for $\to$I
  - ? $P, P \to (Q \& R) \Rightarrow Q$
    - Using tactic for $\to$E
    - ? $P, P \to (Q \& R) \Rightarrow P$ ■
    - ? $Q \& R, P \Rightarrow Q$
      - Using tactic for &E
      - ? $Q, R, P \Rightarrow Q$ ■
- ? $P \to (Q \& R) \Rightarrow P \to R$
  - Using tactic for $\to$I
  - ? $P, P \to (Q \& R) \Rightarrow R$
    - Using tactic for $\to$E
    - ? $P, P \to (Q \& R) \Rightarrow P$ ■
    - ? $Q \& R, P \Rightarrow R$
      - Using tactic for &E
      - ? $Q, R, P \Rightarrow R$ ■

A quick survey will lead us to expect a proof of 10 lines — 1 assumption, 1 for &I, 2 for $\to$I, 1 for $\to$E and 1 for the first &E (only the left conjunct is used), then 2 more for the next →I, and 1 each for →E and &E (this time the right conjunct only is used). But actually it will be shorter. For each →I generates an assumption of P — and we need only assume it once; moreover, each →E generated is the same, so we need only apply it once. Hence we obtain a proof of 8 lines. First we take what we are given

1	(1)	$P \rightarrow (Q \,\&\, R)$	Asmp

and now reflect: the main connective in the sought-for conclusion is conjunction — that suggests we try to derive the two conjuncts separately and then use &I to finish off, i.e. &I tactic. Each conjunct is a conditional — that suggests we assume their antecedent(s), then aim to derive their consequents, and use →I to finish off — →I tactic. So we assume their (common) antecedent:

2	(2)	P	Asmp

Now we are looking for derivations of Qand of R — the two consequents. A glance at the assumption of the sequent we are ultimately aiming for shows the way: get $Q \,\&\, R$ by →E, then get Q and R separately by &E. The proof is now quite mechanical:

1,2	(3)	$Q \,\&\, R$	1,2 →E
1,2	(4)	Q	3 &E
1	(5)	$P \rightarrow Q$	2,4 →I
1,2	(6)	R	3 &E
1	(7)	$P \rightarrow R$	2,6 →I
1	(8)	$(P \rightarrow Q) \,\&\, (P \rightarrow R)$	5,7 &I

Lastly:

$$P \rightarrow S,\ Q \rightarrow R \vdash (P \vee Q) \rightarrow (R \vee S)$$

Here is the tactical derivation:

? $P \rightarrow S, Q \rightarrow R \Rightarrow (P \vee Q) \rightarrow (R \vee S)$
- Using tactic for $\rightarrow$I
- ? $P \vee Q, P \rightarrow S, Q \rightarrow R \Rightarrow R \vee S$
  - Using tactic for $\vee$E
  - ? $P, P \rightarrow S, Q \rightarrow R \Rightarrow R \vee S$
    - Using tactic for $\vee$I
    - ? $P, P \rightarrow S, Q \rightarrow R \Rightarrow S$
      - Using tactic for $\rightarrow$E
      - ? $P, P \rightarrow S, Q \rightarrow R \Rightarrow P$ ■
      - ? $S, P, Q \rightarrow R \Rightarrow S$ ■
  - ? $Q, P \rightarrow S, Q \rightarrow R \Rightarrow R \vee S$
    - Using tactic for $\vee$I
    - ? $Q, P \rightarrow S, Q \rightarrow R \Rightarrow R$
      - Using tactic for $\rightarrow$E
      - ? $Q, P \rightarrow S, Q \rightarrow R \Rightarrow Q$ ■
      - ? $R, Q, P \rightarrow S \Rightarrow R$ ■

So we can expect at most 11 lines of proof: 2 assumptions, 2 for $\rightarrow$I, 3 for $\vee$E, and 1 for each of $\vee$I, $\rightarrow$E , $\vee$I and $\rightarrow$E again. We take what we are given

1	(1)	$P \rightarrow S$	Asmp
2	(2)	$Q \rightarrow R$	Asmp

and now reflect: the main connective in the sought-for conclusion is the conditional — that suggests we should take the assumption-sequent corresponding to its antecedent, aim to derive its consequent, and then use $\rightarrow$I to finish off ($\rightarrow$I tactic). But the antecedent is a disjunction — that suggests use of $\vee$E: assume each disjunct in turn and get the desired consequent from them separately ($\vee$E tactic). So let's take the appropriate assumptions and then think again:

3	(3)	$P \vee Q$	Asmp
4	(4)	P	Asmp

OK; so far we have done nothing but list assumption-sequents. Better do some work soon! Now we are aiming to get the original consequent, $R \vee S$, from both P and Q separately. But the consequent is also a disjunction; that suggests we aim simply to get either disjunct — R, or S — from P and from Q, and use $\vee$I to finish off ($\vee$I tactic). (And, of course, it needn't be the same disjunct in both cases.) The proof now falls into place:

1,4	(5)	S	1,4 $\rightarrow$E
1,4	(6)	R$\vee$S	5 $\vee$I
7	(7)	Q	Asmp
2,7	(8)	R	2,7 $\rightarrow$E
2,7	(9)	R$\vee$S	8 $\vee$I
1,2,3	(10)	R$\vee$S	3,4,6,7,9 $\vee$E
1,2	(11)	(P$\vee$Q)$\rightarrow$(R$\vee$S)	3,10 $\rightarrow$I

Understanding formal logic requires two stages of learning — two stages which can be usefully likened to two stages in learning chess. In chess, one must first learn which pieces are allowed to move in which particular way. But having learned those rules, or moves, one will still not know how to play a good game of chess. That requires a second stage of learning, learning how to assemble those moves into a strategy for winning a game. Similarly in logic: first, one must learn the rules of inference. But having learned those rules, one will still not know how to find a proof. That requires a second stage of learning, learning how to assemble applications of those rules into a strategy for finding the desired proof. To become a chess-player, one must understand the strategies needed to win a game. To become a logician, one must understand the strategies needed to find a proof. Try to get into the habit of letting your search for a proof be informed by a fully self-conscious strategy, and practice setting out explicitly the corresponding tactical derivation.

Exercises

(1) Set out the tactical derivation for examples x) and xi) from chapter 3, namely:

(a) $\neg P \vdash \neg(P \& Q) \vee R$

(b) $P \rightarrow \neg P \vdash \neg P$

(2) Find proofs of the following sequents, setting out the tactical derivation explicitly as above:

(a) $P \rightarrow Q \vdash (Q \rightarrow R) \rightarrow (P \rightarrow R)$

(b) $P \rightarrow Q \vdash (R \rightarrow P) \rightarrow (R \rightarrow Q)$

(c) $\neg(P \& Q), P \vdash \neg Q$

(d) $P \rightarrow Q, Q \rightarrow R \vdash \neg R \rightarrow \neg P$

(e) $P \vee Q, P \rightarrow Q \vdash Q$

(f) $(P \& Q) \rightarrow R \vdash (P \rightarrow Q) \rightarrow (P \rightarrow R)$

5 Trees and Tactics Compared

[This chapter is omitted in this version of these notes.]

6 The Biconditional

Now for a further but modest technical innovation. 'P only if Q' says that the truth of Q is necessary for the truth of P, i.e. that if Q is false, so will P be. So 'P only if Q' is tantamount to $\neg Q \rightarrow \neg P$, which in turn is tantamount to $P \rightarrow Q$ (since $\neg Q \rightarrow \neg P \vdash P \rightarrow Q$ and $P \rightarrow Q \vdash \neg Q \rightarrow \neg P$ are both valid sequents). So 'P only if Q' reverses the arrow in 'P if Q': the latter is $Q \rightarrow P$ and the former is $P \rightarrow Q$. 'P **if and only if** Q', that is 'P if Q and P only if Q', is thus equivalent to 'if P then Q and if Q then P': $(P \rightarrow Q) \,\&\, (Q \rightarrow P)$.

The **biconditional**, 'if and only if', is quite an important connective for logical purposes. Following usual practice, we will write it, suggestively, as '$\leftrightarrow$' in proofs.

We will treat '$\leftrightarrow$' simply as a **definitional abbreviation**, i.e., we don't lay down rules of inference for it but simply treat '$A \leftrightarrow B$' as a shorthand for '$(A \rightarrow B) \,\&\, (B \rightarrow A)$'. That means that whenever a formula of the form '$A \leftrightarrow B$' occurs in the course of a proof, we can simply replace it by the appropriate formula of the form '$(A \rightarrow B) \,\&\, (B \rightarrow A)$'; conversely, wherever a formula of the form '$(A \rightarrow B) \,\&\, (B \rightarrow A)$' occurs in a proof, we can abbreviate it to '$A \leftrightarrow B$'; in either case, our authority for doing so is the definition of '$\leftrightarrow$':

$$A \leftrightarrow B =_{\mathrm{df}} (A \rightarrow B) \,\&\, (B \rightarrow A).$$

(This is rather similar to what happens in arithmetic when, for example, we give the definition:

$$m \times n = m + m + \ldots + m, \text{ n times}$$

which licenses replacing any expression of the form "$m \times n$" by one of the form "$m + m + \ldots + m$" in which "m" occurs n times.)

Here are some examples of the way '$\leftrightarrow$' can now feature in proofs:

$P, P \leftrightarrow Q \vdash Q$

1	(1)	P	Asmp
2	(2)	$P \leftrightarrow Q$	Asmp
2	(3)	$(P \rightarrow Q) \& (Q \rightarrow P)$	2 df.$\leftrightarrow$
2	(4)	$P \rightarrow Q$	3 &E
1,2	(5)	Q	1,4 $\rightarrow$E

$P \leftrightarrow Q \vdash (P \rightarrow Q) \vee (P \rightarrow \neg Q)$

1	(1)	$P \leftrightarrow Q$	Asmp
1	(2)	$(P \rightarrow Q) \& (Q \rightarrow P)$	1 df.$\leftrightarrow$
1	(3)	$P \rightarrow Q$	2 &E
1	(4)	$(P \rightarrow Q) \vee (P \rightarrow \neg Q)$	3 $\vee$I

$P \rightarrow Q \vdash (Q \rightarrow P) \rightarrow (P \leftrightarrow Q)$

1	(1)	$P \rightarrow Q$	Asmp
2	(2)	$Q \rightarrow P$	Asmp
1,2	(3)	$(P \rightarrow Q) \& (Q \rightarrow P)$	1,2 &I
1,2	(4)	$P \leftrightarrow Q$	3 df.$\leftrightarrow$
1	(5)	$(Q \rightarrow P) \rightarrow (P \leftrightarrow Q)$	2,4 $\rightarrow$I

The tactical derivation is:

$?\ P \rightarrow Q \Rightarrow (Q \rightarrow P) \rightarrow (P \leftrightarrow Q)$
> Using tactic for $\rightarrow$I
> $?\ Q \rightarrow P,\ P \rightarrow Q \Rightarrow P \leftrightarrow Q$
> > Using tactic for Definition
> > $?\ Q \rightarrow P,\ P \rightarrow Q \Rightarrow (P \rightarrow Q) \,\&\, (Q \rightarrow P)$
> > > Using tactic for &I
> > > $?\ Q \rightarrow P,\ P \rightarrow Q \Rightarrow P \rightarrow Q$ ■
> > > $?\ Q \rightarrow P,\ P \rightarrow Q \Rightarrow Q \rightarrow P$ ■

Lastly,

$P \rightarrow (Q \leftrightarrow R) \vdash (P \,\&\, Q) \rightarrow R$

1	(1)	$P \rightarrow (Q \leftrightarrow R)$	Asmp
2	(2)	$P \,\&\, Q$	Asmp
2	(3)	P	2 &E
1,2	(4)	$Q \leftrightarrow R$	1,3 $\rightarrow$E
1,2	(5)	$(Q \rightarrow R) \,\&\, (R \rightarrow Q)$	4 df.$\leftrightarrow$
1,2	(6)	$Q \rightarrow R$	5 &E
2	(7)	) Q	2 &E
1,2	(8)	R	6,7 $\rightarrow$E
1	(9)	$(P \,\&\, Q) \rightarrow R$	2,8 $\rightarrow$I

Here are the tactics used:

? $P\rightarrow(Q\leftrightarrow R) \Rightarrow (P\,\&\,Q)\rightarrow R$
> Using tactic for $\rightarrow$I
> ? $P\,\&\,Q,\ P\rightarrow(Q\leftrightarrow R) \Rightarrow R$
>> Using tactic for &E
>> ? $P,\ Q,\ P\rightarrow(Q\leftrightarrow R) \Rightarrow R$
>>> Using tactic for $\rightarrow$E
>>> ? $P,\ Q,\ P\rightarrow(Q\leftrightarrow R) \Rightarrow P$ ■
>>> ? $Q\leftrightarrow R,\ P,\ Q \Rightarrow R$
>>>> Using tactic for Definition
>>>> ? $(Q\rightarrow R)\,\&\,(R\rightarrow Q),\ P,\ Q \Rightarrow R$
>>>>> Using tactic for &E
>>>>> ? $Q\rightarrow R,\ R\rightarrow Q,\ P,\ Q \Rightarrow R$
>>>>>> Using tactic for $\rightarrow$E
>>>>>> ? $Q\rightarrow R,\ R\rightarrow Q,\ P,\ Q \Rightarrow Q$ ■
>>>>>> ? $R,\ R\rightarrow Q,\ P,\ Q \Rightarrow R$ ■

These are all quite straightforward. Notice that the pool of assumptions on the left is always unchanged by an application of df.$\leftrightarrow$ (naturally, since we have simply changed the form of expression of the conclusion of the sequent operated on). Note also the continuing success of the basic $\rightarrow$I-strategy in the third and fourth examples. Let's try another — the **transitivity of the biconditional**:

$P\leftrightarrow Q,\ Q\leftrightarrow R \vdash P\leftrightarrow R$

1	(1)	$P\leftrightarrow Q$	Asmp
2	(2)	$Q\leftrightarrow R$	Asmp
3	(3)	P	Asmp for $\rightarrow$I at line 10
1	(4)	$(P\rightarrow Q)\,\&\,(Q\rightarrow P)$	1 df.$\leftrightarrow$
2	(5)	$(Q\rightarrow R)\,\&\,(R\rightarrow Q)$	2 df.$\leftrightarrow$
1	(6)	$P\rightarrow Q$	4, &E
2	(7)	$Q\rightarrow R$	5 &E
1,3	(8)	Q	3,6 $\rightarrow$E
1,2,3	(9)	R	7,8 $\rightarrow$E

1,2	(10)	$P \to R$	3,9 $\to$I
11	(11)	R	Asmp for $\to$I at line 16
2	(12)	$R \to Q$	5 &E
1	(13)	$Q \to P$	4 &E
2,11	(14)	Q	11,12 $\to$E
1,2,11	(15)	P	13,14 $\to$E
1,2	(16)	$R \to P$	11,15 $\to$I
1,2	(17)	$(P \to R) \& (R \to P)$	10,16 &I
1,2	(18)	$P \leftrightarrow R$	17 df.$\leftrightarrow$

The strategy here is simple enough: we are aiming to prove a sequent with a biconditional conclusion, so we have to derive both its constituent conditionals (lines 10 and 16) prior to conjoining them (line 17) and applying df.$\leftrightarrow$ (line 18). Each constituent conditional is proved by the basic $\to$I strategy (lines 3 to 9, and lines 11 to 15). The result is a straightforward, but admittedly cumbersome proof. In particular, we have laboriously to break down the assumptions of the proved sequent using df.$\leftrightarrow$ and &E (lines 4, 5, 6, 7, 12 and 13).

We could make our handling of '$\leftrightarrow$' less long-winded. We could introduce explicit rules of inference for '$\leftrightarrow$', so treating it just like the other connectives. Here are a couple of suitable rules:

$$\leftrightarrow\text{I} \qquad \frac{X \vdash A \qquad Y \vdash B}{X\backslash B,\ Y\backslash A \vdash A \leftrightarrow B}$$

$$\leftrightarrow\text{E} \qquad \frac{X \vdash A \leftrightarrow B \qquad Y \vdash A}{X, Y \vdash B} \quad \text{and} \quad \frac{X \vdash A \leftrightarrow B \qquad Y \vdash B}{X, Y \vdash A}$$

You can use these rules if you wish. They give some saving in the length of proofs (how many lines in the case of the proof of transitivity of the biconditional?). However, we will continue to treat '$\leftrightarrow$' as introduced and eliminated by definition.

Here are four more examples to illustrate the use of the $\leftrightarrow$-defn.

$P \leftrightarrow Q \vdash \neg P \leftrightarrow \neg Q$

We take what we are given

1	(1)	$P \leftrightarrow Q$	Asmp

and then reflect: aiming for a biconditional conclusion, the obvious ploy is to prove each of its constituents and then use df.$\leftrightarrow$; and since each constituent is, naturally, a conditional, we go for the basic $\rightarrow$I strategy twice over. First, to show $\neg P \rightarrow \neg Q$. So we take the appropriate assumption-sequent, $\neg P$:

2	(2)	$\neg P$	Asmp

and then aim to derive $\neg Q$ from it. Thus we break up (1):

1	(3)	$(P \rightarrow Q) \& (Q \rightarrow P)$	1 df.$\leftrightarrow$
1	(4)	$Q \rightarrow P$	3 &E

and proceed:

5	(5)	Q	Asmp
1, 5	(6)	P	4,5 $\rightarrow$E
1, 2, 5	(7)	$\perp$	2,6 $\neg$E
1, 2	(8)	$\neg Q$	5,7 $\neg$I
1	(9)	$\neg P \rightarrow \neg Q$	2,8 $\rightarrow$I

Now we use the same strategy to obtain the other conjunct, $\neg Q \rightarrow \neg P$:

10	(10)	$\neg Q$	Asmp
11	(11)	P	Asmp
1	(12)	$P \rightarrow Q$	3 &E
1, 11	(13)	Q	11,12 $\rightarrow$E
1, 10, 11	(14)	$\perp$	10,13 $\neg$E
1, 10	(15)	$\neg P$	11,14 $\neg$I

1	(16)	$\neg Q \rightarrow \neg P$	10,15 $\rightarrow$I
1	(17)	$(\neg P \rightarrow \neg Q) \& (\neg Q \rightarrow \neg P)$	9,16 &I
1	(18)	$\neg P \leftrightarrow \neg Q$	17 df.$\leftrightarrow$

Next: $(P \vee Q) \leftrightarrow P \vdash Q \rightarrow P$

1	(1)	$(P \vee Q) \leftrightarrow P$	Asmp
2	(2)	Q	Asmp
1	(3)	$((P \vee Q) \rightarrow P) \& (P \rightarrow (P \vee Q))$	1 df.$\leftrightarrow$
1	(4)	$(P \vee Q) \rightarrow P$	3 &E
2	(5)	$P \vee Q$	2 $\vee$I
1,2	(6)	P	4,5 $\rightarrow$E
1	(7)	$Q \rightarrow P$	2,6 $\rightarrow$I

This needs little comment: it is again, just a matter of letting the main connective in the conclusion dictate strategy (here the basic $\rightarrow$I strategy) and breaking up the given assumption in accordance with the appropriate elimination rule (here df.$\leftrightarrow$ and $\rightarrow$E). After that, the proof more or less constructs itself.

Next: $P \leftrightarrow Q \vdash (P \vee R) \leftrightarrow (Q \vee R)$

This time, let's go through the strategy first and then just let the proof fall into place. To begin with, since we are going for a biconditional conclusion, we shall have to derive each constituent conditional — each, hopefully, by the basic $\rightarrow$I strategy. Each of those constituent conditionals has a disjunctive consequent, so it will do to derive either of that consequent's disjuncts and then use $\vee$I. First though, we break up the assumption of the sequent we are trying to prove, using df.$\leftrightarrow$:

1	(1)	$P \leftrightarrow Q$	Asmp
1	(2)	$(P \rightarrow Q) \& (Q \rightarrow P)$	1 df.$\leftrightarrow$
1	(3)	$P \rightarrow Q$	2 &E
1	(4)	$Q \rightarrow P$	2 &E

Now we implement our strategy:

5	(5)	$P \vee R$	Asmp
6	(6)	P	Asmp
1,6	(7)	Q	3, 6 $\rightarrow$E
9	(8)	R	Asmp
9	(9)	$Q \vee R$	9 $\vee$I
1,5	(10)	$Q \vee R$	5,6,8,9,10 $\vee$E
1	(11)	$(P \vee R) \rightarrow (Q \vee R)$	5,11 $\rightarrow$I
13	(12)	$Q \vee R$	Asmp
14	(13)	Q	Asmp
1,14	(14)	P	4, 14 $\rightarrow$E
1,14	(15)	$P \vee R$	15 $\vee$I
17	(16)	R	Asmp
17	(17)	$P \vee R$	17 $\vee$I
1,13	(18)	$P \vee R$	13,14,16,17,18 $\vee$E
1	(19)	$(Q \vee R) \rightarrow (P \vee R)$	13, 19 $\rightarrow$I
1	(20)	$((P \vee R) \rightarrow (Q \vee R))\ \&$ $((Q \vee R) \rightarrow (P \vee R))$	12, 20 &I
1	(21)	$(P \vee R) \leftrightarrow (Q \vee R)$	21 df.$\leftrightarrow$

(Strictly, there is no need to assume R again at line 17, having already done so at line 9. We do so merely to preserve the strategic symmetry of the proof.) The result is a proof which looks lengthy and complicated but which, viewed in terms of the strategy worked out in advance, is straightforward enough.

Finally, a simple proof of the **symmetry of the biconditional**:

$P \leftrightarrow Q \vdash Q \leftrightarrow P$

1	(1)	$P \leftrightarrow Q$	Asmp
1	(2)	$(P \rightarrow Q)\ \&\ (Q \rightarrow P)$	1 df.$\leftrightarrow$
1	(3)	$P \rightarrow Q$	1 &E

1	(4)	$Q \rightarrow P$	1 &E
1	(5)	$(Q \rightarrow P) \& (P \rightarrow Q)$	3,4 &I
1	(6)	$Q \leftrightarrow P$	2,3 df.$\leftrightarrow$

Exercise

Prove the following sequents:

(a) $P \leftrightarrow Q \vdash \neg(P \& \neg Q)$
(b) $P \leftrightarrow Q,\ Q \leftrightarrow \neg R \vdash R \leftrightarrow \neg P$
(c) $P \& Q \vdash \neg(P \leftrightarrow \neg Q)$

7 The Paradoxes of Implication

Let us start with a couple of more advanced proofs which will teach us something about the character of the system of propositional logic which we have adopted. First we show:

$$P \rightarrow Q \vdash \neg(P \& \neg Q)$$

Strategy: we are aiming for a negative conclusion, so the appropriate introduction rule, determining our overall strategy, will be ¬I. We will therefore take the assumption sequent corresponding to the opposite of what we are trying to prove, and look for a contradiction by applying the relevant elimination rules (in this case, &E and those for the conditional) to that sequent and to our first assumption:

1	(1)	$P \rightarrow Q$	Asmp

The procedure is straightforward:

1	(1)	$P \rightarrow Q$	Asmp
2	(2)	$P \& \neg Q$	Asmp
2	(3)	P	2 &E
1,2	(4)	Q	1,3 →E
2	(5)	$\neg Q$	2 &E
1,2	(6)	$\perp$	4,5 ¬E
1	(7)	$\neg(P \& \neg Q)$	2,6 ¬I

Next we prove the converse:

$$\neg(P \& \neg Q) \vdash P \rightarrow Q$$

Here we use a refinement of the standard $\neg$I tactic. Standardly, when we want to prove:

$$X \vdash A \rightarrow B$$

we assume the premises of the sequent, X, the antecedent, A, and then try directly to derive B on those assumptions. With $\neg$I and DN in our armoury, however, we can go at this indirectly, via $\neg\neg$B as a sub-goal. That is, we assume $\neg$B and on the basis of X, A and $\neg$B aim to obtain a contradiction, i.e.,

$$X, A, \neg B \vdash \bot.$$

We can then use $\neg$I to obtain $X, A \vdash \neg\neg B$, and so use DN to obtain what we want, viz $X, A \vdash B$.

Putting it in terms of our tactics, it runs like this:

? $X \Rightarrow A \rightarrow B$
> Using tactic for $\rightarrow$I:
> ? $A, X \Rightarrow B$
>> Using tactic for DN:
>> ? $A, X \Rightarrow \neg\neg B$
>>> Using tactic for $\neg$I:
>>> ? $\neg B, A, X \Rightarrow \bot$

So to show $X, A \Rightarrow B$ we try to show

$$X, A, \neg B \Rightarrow \bot.$$

Now to apply this strategy explicitly to the example: $\neg(P \& \neg Q) \vdash P \rightarrow Q$. We are aiming for a conditional conclusion so the appropriate introduction rule, determining our overall strategy, will be $\rightarrow$I. We therefore proceed:

1	(1)	$\neg(P \& \neg Q)$	Asmp
2	(2)	P	Asmp

and now look for a proof of Q on 1 and 2 as assumptions. Evidently, though, Q does not follow directly by any of our elimination rules — we therefore need further assumption-sequents. The strategy above for this kind of situation tells us to look to DN and ¬I to bale us out; we should therefore take an assumption sequent corresponding to the negation of Q:

3	(3)	$\neg Q$	Asmp

and look for a contradiction. Clearly we don't have to look far:

2,3	(4)	$P \& \neg Q$	2,3 &I
1,2,3	(5)	$\bot$	1,4 ¬E

The rest is straightforward:

1,2	(6)	$\neg\neg Q$	3,5 ¬I
1,2	(7)	Q	6 DN
1	(8)	$P \rightarrow Q$	2,7 →I

The result of the last two proofs is that $P \rightarrow Q$ and $\neg(P \& \neg Q)$ are **interderivable** — we can derive either from the other. We express the situation like this:

$$P \rightarrow Q \dashv\vdash \neg(P \& \neg Q)$$

Interderivability is an important relation; for when formulae are interderivable they are *interchangeable in any valid sequent without affecting its validity*. (Can you, informally, explain why?)

Here is another interderivability result:

$$\neg(P \& \neg Q) \dashv\vdash \neg P \vee Q$$

First, we prove it from left to right:

1	(1)	$\neg(P \& \neg Q)$	Asmp

Strategy: we are aiming for a disjunctive conclusion, so the appropriate direct strategy would be to derive one of its disjuncts and then use $\vee$I. But that isn't going to work here because neither $\neg P$ nor Q individually follows from $\neg(P \& \neg Q)$; all that follows is that they won't both be false (which is precisely what the sequent says). So, since direct methods fail us, we resort, as always in such circumstances, to $\neg$I:

2	(2)	$\neg(\neg P \vee Q)$	Asmp

However, no contradictions seem to be immediately in the offing. Clearly, we need more assumption-sequents. A good rule of thumb is to take the one that will most speedily give us what we want. A moment's reflection suggests this:

3	(3)	$\neg P$	Asmp

whence

3	(4)	$\neg P \vee Q$	3 $\vee$I

so that

2,3	(5)	$\perp$	2,4 $\neg$E

What should we blame this contradiction on? Clearly it is no use blaming it on 2; all we shall get by that route — with DN — is a proof of our desired conclusion on 3 as assumption, which would be quite irrelevant. So let's see how things look if we blame it on 3 instead:

2	(6)	$\neg\neg P$	3,5 $\neg$I
2	(7)	P	6 DN

Now we have to step back and think again. We are trying to get a contradiction from 1 and 2 together, and we now have P from 2; what more do we need in order to get something which contradicts 1? Clearly, we need $\neg Q$; then we can get $P \& \neg Q$, of which 1 is the negation. So the obvious tactic is another $\neg$I:

8	(8)	Q	Asmp
8	(9)	$\neg P \vee Q$	8 $\vee$I
2,8	(10)	$\bot$	2,9 $\neg$E
2	(11)	$\neg Q$	8,10 $\neg$I

The rest now falls into place:

2	(12)	$P \& \neg Q$	7,11 &I
1,2	(13)	$\bot$	1,12 $\neg$E
1	(14)	$\neg\neg(\neg P \vee Q)$	2,13 $\neg$I
1	(15)	$\neg P \vee Q$	14 DN

That is as devious a proof as any; fortunately most of them are much simpler.

Will that be true of the converse sequent — the right-to-left ingredient — which we are about to prove?

1	(1)	$\neg P \vee Q$	Asmp

Strategy: we are aiming at a negative conclusion $\neg(P \& \neg Q)$, so the obvious strategy is $\neg$I. But the disjunctive character of 1 suggests that the overall strategy will be $\vee$E, with $\neg$I playing a role on both branches. Let's see how it goes:

2	(2)	$P \& \neg Q$	Asmp
3	(3)	$\neg P$	Asmp
2	(4)	P	2 &E
2,3	(5)	$\bot$	3,4 $\neg$E
3	(6)	$\neg(P \& \neg Q)$	2,5 $\neg$I

There, then, is the desired conclusion from one disjunct of $\neg P \vee Q$; now for the other half:

7	(7)	Q	Asmp
2	(8)	$\neg Q$	2 &E
2,7	(9)	$\perp$	8,7 $\neg$E
7	(10)	$\neg(P \& \neg Q)$	2,9 $\neg$I

All that remains to do is:

1	(11)	$\neg(P \& \neg Q)$	1,3,5,7,10 $\vee$E

We now have, then, the two interderivability sequents:

(i) $P \to Q \dashv\vdash \neg(P \& \neg Q)$
(ii) $\neg(P \& \neg Q) \dashv\vdash \neg P \vee Q$.

Bearing in mind the significance of interderivability: viz interderivable formulae can always be exchanged in a sequent in a validity-preserving way, we therefore know that the following sequent must be valid:

(iii) $P \to Q \dashv\vdash \neg P \vee Q$

for it results from exchanging the two halves of (ii) in the right hand side of (i). In particular, therefore, reading (iii) from right to left, we know that $\neg P \vee Q$ entails $P \to Q$.

It will therefore be straightforward to prove both these sequents in our system:

(iv) $\neg P \vdash P \to Q$
(v) $Q \vdash P \to Q$,

since both $\neg P \vdash \neg P \vee Q$ and $Q \vdash \neg P \vee Q$ are immediate by $\vee$I. But are not these results unwelcome? Sequents (iv) and (v) seem bizarre: (iv) says, intuitively, that the falsity of its antecedent is sufficient for the truth of any conditional; and (v) says the same about the truth of the consequent. And this seems incorrect. For example, Margaret Thatcher is, one assumes, English; but that is hardly sufficient, it seems, for the acceptability of:

> If Margaret Thatcher and Ronald Reagan are the same nationality, then Margaret Thatcher is English.

Likewise we are not, as it happens, going to boil up any water in a carafe today; but that hardly seems sufficient for the acceptability of all conditionals of the form:

> If we boil up some water in a carafe today, it will boil at n°C,

for every n. On the contrary, we feel, only one of those conditionals is acceptable: the one that correctly describes the temperature at which the water would actually boil.

On reflection, we ought not to be surprised at the situation. For recall the "subtle" (or perverse) manoeuvre we used to establish $P \& \neg P \vdash \neg Q$ (chapter 3).That involved a ploy with the two &-rules which can be applied to provide quite short proofs of the two worrying sequents (iv) and (v). First,

$\neg P \vdash P \to Q$

1	(1)	$\neg P$	Asmp
2	(2)	P	Asmp
3	(3)	$\neg Q$	Asmp
2,3	(4)	$P \& \neg Q$	2,3 &I
2,3	(5)	P	4 &E
1,2,3	(6)	$\bot$	1,5 ¬E
1,2	(7)	$\neg\neg Q$	3,6 ¬I
1,2	(8)	Q	7 DN
1	(9)	$P \to Q$	2,8 →I

Study this proof carefully. Note the ploy at lines (4) and (5) with the two &-rules, contriving a "dependence" of P on $\neg Q$, and the use once again of ¬I and DN in combination.

Next, still more simply:

$Q \vdash P \rightarrow Q$

1	(1)	Q	Asmp
2	(2)	P	Asmp
1,2	(3)	P & Q	1,2 &I
1,2	(4)	Q	3 &E
1	(5)	$P \rightarrow Q$	2,4 $\rightarrow$I

We will say some more about the significance of these two sequents — known as the Paradoxes of Material Implication — in a moment. But first let's see why their derivability effectively settles a question we left open in Chapter 3, p. 87, viz. how to understand 'X\A' when the formula A does not occur in the string X. For what follows is that it cannot make any difference — cannot alter the class of sequents we can prove — if we stipulate that when the formula A does not occur in the string X, X\A is simply to be X itself.

How can we be sure? Well, we have to demonstrate that anything we can prove from our rules once we so stipulate could have been proved without the stipulation. Three rules — $\rightarrow$I, ¬I and ∨E — are potentially affected, since they are the only rules whose schematisation uses the notation 'X\A'. And what we need to show is that if in their respective schematisations we replace all occurrences of expressions of the type 'X\A' by ones of the corresponding type 'X', the resulting patterns still represent inferences which our rules — without the stipulation — enable us to make. (Make sure you see why this suffices.) We have therefore to consider three patterns

$$\text{(i)} \qquad \frac{X \vdash A \vee B \quad Y \vdash C \quad Z \vdash C}{X, Y, Z \vdash C}$$

which results from the ∨E schematisation by replacing the occurrences of 'Y\A' and 'Z\A' in its lower line by 'Y' and 'Z' respectively;

$$\text{(ii)} \qquad \frac{X \vdash B}{X \vdash A \rightarrow B}$$

which results from the $\rightarrow$I schematisation by replacing the occurrences of 'X\A' in its lower line by 'X'; and finally

(iii) $$\frac{X \vdash \bot}{X \vdash \neg A}$$

which in turn results from the ¬I schematisation by replacing the occurrences of ‘X\A’ in its lower line by ‘X’. It will be clear that each of (i)–(iii) represents a transition which we can already make in the system: (i) is simply a case of Thinning on either the second or third premise, and the characteristic Thinning move is something whose effect we can get in any case, using the &I/&E ploy. (Can you see how?) Likewise the move licensed by (ii) is in effect exactly what we establish by the proof of the second Paradox, the sequent $Q \vdash P \rightarrow Q$.

Only (iii) is a bit more complicated. Here is the essential point. Reflect that if we have established $X \vdash \bot$, we must have proved both $X \vdash B$ and $X \vdash \neg B$ for some formula B, and then used ¬E. But the type of proof given of the first Paradox — the sequent $\neg P \vdash P \rightarrow Q$ — could obviously be used to get from the second of those, viz. $X \vdash \neg B$, to a proof of $X \vdash B \rightarrow \neg A$. And that, coupled with $X \vdash B$, would give us $X \vdash \neg A$ by an →E step. Think about it.

Not only is no change effected by the proposed stipulation in which sequents we are able to prove, but proof construction often becomes appreciably simpler. (There are also certain advantages at a more theoretical level, when we come to reason in a formal way not within but *about* the system we have developed; but we leave those aside for the present.) Consider the second Paradox again. With ‘X\A’ read as proposed, we may simply proceed like this:

$P \vdash Q \rightarrow P$

1	(1)	P	Asmp
2	(2)	Q	Asmp
1	(3)	$Q \rightarrow P$	1,2 →I

There is no need for the ploy with the two & -rules. We merely faithfully apply the schematisation of →I in the light of the stipulation, taking ‘Q’ as *A* and ‘P’ as both X and — although it contains no occurrence of ‘Q’ — as X\A. But notice that, as a matter of protocol, we still insist that the antecedent of the conditional to be proved be taken as an assumption — line (2) — even though it is not “used” in the proof.

Another example:

$P \to Q, Q \to R \vdash P \to (P \to R)$

1	(1)	$P \to Q$	Asmp
2	(2)	$Q \to R$	Asmp
3	(3)	P	Asmp
1,3	(4)	Q	1,3 $\to$E
1,2,3	(5)	R	2,4 $\to$E
1,2	(6)	$P \to R$	3,5 $\to$I
1,2	(7)	$P \to (P \to R)$	3,6 $\to$I

This proof is perfectly orthodox as far as line (6) where P, having been used to achieve R at line (5), is discharged from the assumption pool by the $\to$I step. But the move to line (7) depends on our new stipulation, with 'P' taken as A and the string consisting of '$P \to Q$' and '$Q \to R$' taken as both X and $X \backslash A$. Notice that the point of protocol we mentioned is satisfied by the assumption of P at line (3) — it makes no difference that that assumption was also used earlier in the proof, and there is no need to assume P twice.

Next an example of how the stipulation can simplify $\neg$I proofs. Recall the nine line proof of Ex Falso Quodlibet (Ch. 3). Now we can proceed as follows

$P \& \neg P \vdash Q$

1	(1)	$P \& \neg P$	Asmp
2	(2)	$\neg Q$	Asmp
1	(3)	P	1 &E
1	(4)	$\neg P$	1 &E
1	(5)	$\perp$	4,3 $\neg$E
1	(6)	$\neg\neg Q$	2,5 $\neg$I
1	(7)	Q	6 DN

Here we take '$\neg Q$' as A and '$P \& \neg P$' as both X and $X \backslash A$, and have simply and faithfully applied the $\neg$I schematisation.

In sum: the stipulation makes no difference to what we can prove, but simplifies proof construction and has certain theoretical advantages. It is also in a way more honest, since it makes the provability of the Paradoxes transparent — makes it clearer what kind of system we have. So we now adopt it.

In any case, we can derive the so-called Paradoxes of Material Implication. Is that a disaster, or can we live with the fact? What is uncomfortable about it is that we originally set ourselves the task of codifying a set of rules of inference which would enable us to prove all the valid patterns of argument essentially involving 'not', 'and', 'or' and 'if. . . , then. . . '. Our intention was therefore to respect the meanings of those English connectives. And that we seem somehow to have failed to do, despite taking the greatest care to ground our rules in the ordinary understanding we have of those words. If we take the view that

> If Margaret Thatcher and Ronald Reagan are the same nationality, then Margaret Thatcher is English,

does *not* follow logically from from

> Margaret Thatcher is English,

and that

> If we boil up some water in a carafe today, it will boil at 44°C,

does *not* follow logically from

> We are not going to boil any water in a carafe today,

then there is no alternative, so long as we stick to our intention that our logic should respect the properties of the ordinary language connectives, to regarding the system we have characterised as unsound and our project as, so far, a failure.

If we take this line, there are two things that have to be done. The first is to say where we went wrong — what it is about the rules we laid down which surreptitiously betrayed the meaning of 'if. . . , then. . . '. But don't imagine that will be easy. It is clear that the trouble will afflict any system which has both $\rightarrow$I and Thinning (either as rules laid down from the start or as principles whose effect can be captured indirectly — the situation of Thinning in our system as characterised by the original ten rules.) Thinning allows that we may augment the assumptions in any derivable sequent without affecting derivability: that where $A_1, \ldots, A_n \vdash A$ is derivable, so is $A_1, \ldots, A_n, B \vdash A$ for any additional assumption B. But if that is so, then obviously — in the presence of $\rightarrow$I — we can always prove $A_1, \ldots, A_n \vdash B \rightarrow A$ whenever we can prove $A_1, \ldots, A_n \vdash A$. Consequently we cannot avoid the paradoxes if we accept that both $\rightarrow$I and the augmentation of assumptions allowed by Thinning are faithful to the ordinary meaning of 'if…, then…'. But the case for both is, intuitively, strong. That for $\rightarrow$I was made in Ch. 1, p. 87. That for augmentation is simply this: if B follows from A — i.e. if the truth of A logically guarantees the truth of B — then surely B is guaranteed by any stronger set of assumptions which includes A!

The second thing we have to do if we regard the situation as intolerable is to devise a system of propositional logic which fares better — which correctly codifies the patterns of inference sustained by the intuitive meaning of 'if…, then…'. There are two main alternative conceptions, (so-called) conditional logic on the one hand, and relevance logic on the other. Both have their advocates, and both have well-developed theories. Nonetheless, their claims are certainly heterodox, and their formal development requires a logical sophistication beyond what we are presupposing here. The interested reader should make a point of considering them once the theories in this book have been mastered.

Different reactions to the Paradoxes are possible. One, especially interesting, argues that the two troublesome sequents are not really unfaithful to the meaning of 'if…, then…'; that the impression to the contrary comes from confusing features of the use of an expression which strictly and literally belong to its meaning with features which are conditioned by pragmatics and social convention. Thus, the claim will be, although

> If we boil up some water in a carafe today, it will boil at 44°C,

is *true* when we are not going to boil any water, it is not *properly assertible* on that basis — asserting it simply on the strength of our intention not to boil water would flout various conventions which govern our talk but do not strictly have to do with meaning. This line has been argued for with some ingenuity, and it would be nice if it could succeed. But most philosophers of logic do not believe it.

We shall take a different line. First, reflect that the exceptions are limited: they cover sequents with conditional conclusions, and conditionals embedded in conclusions; and contraposed versions of these, that is, sequents with negated conditionals among their premises. Apart from these cases, there is no suggestion that any invalid sequents — sequents instances of which may pass from true assumptions to a false conclusion — can be proved in our system. It might be feared that if we can derive conditionals from assumptions they do not really follow from, then — using those conditionals in turn as imput for e.g. modus ponens and modus tollens inferences — we will be able to derive *non-conditional* conclusions from assumptions they do not really follow from, and Chaos will break out. But this is not the situation: for instance, if 'If P, then Q' is taken to follow from 'Not P' and the latter is taken to be true, then we do indeed have to regard the former as true. But there is no consequent risk of a bogus argument via modus ponens that 'Q' must also be true. That would be a catastrophe, no doubt. However by taking 'Not P' to be true, we have already assumed that 'P' — the other assumption the modus ponens will need — is false. So no case for the truth of 'Q' can be made. No: the only doubtful sequents we have to reckon with are ones which explicitly involve '$\rightarrow$'; and they are doubtful only on the assumption that the meanings of 'if…, then…' and '$\rightarrow$' perfectly coincide.

Second, even if '$\rightarrow$' fails to capture the meaning of the ordinary English conditional — whatever that is — it is clearly *very similar* in meaning. After all, its use, as reflected in the rules of inference which govern it, was motivated by thinking about the ordinary conditional, and it can be agreed on all sides that it has many points of affinity with the latter. In particular, and crucially, it has the modus ponens property: as with the ordinary conditional, the truth of '$A \rightarrow B$', together with that of its antecedent, entails the truth of its consequent. This means that, as far as their role as input in potential arguments is concerned, statements of the form '$A \rightarrow B$' serve all the same purposes as ordinary conditionals.

These two points make it possible to view '$\rightarrow$' as at least an idealisation, or simplification of the ordinary conditional, a connective which coincides in important logical respects with ordinary 'if…, then…' but whose meaning is fixed by the rules governing it in our formal system and diverges in certain respects from that of 'if…, then…'. Even so the analogies between '$\rightarrow$' and 'if…, then…' are enough to ensure the interest and utility of our system as it stands.

Exercises

(1) Establish the following equivalence: $P \to (P \to Q) \dashv\vdash P \to Q$

(2) Find a seven line proof of the first Paradox of Material Implication, $\neg P \vdash P \to Q$, by exploiting our stipulation for '$X \backslash A$'.

(3) Formalise the following argument and show that it is sound:

> The Christian God, as we conceive Him, is omnipotent, omniscient, and morally perfect. If he has the first two attributes, nothing happens without his consent; in which case, if he is also morally perfect, everything that happens is necessarily for the best. But prayer has point only if we can thereby secure a better train of events than would otherwise occur; and that we cannot do if everything that happens is necessarily for the best. So prayer is pointless — unless God somehow falls short of our conception of Him.

(4) Provide for the following sequents a proof in our usual style, a specification of the tactical derivation *and* a proof in basic symbolism.

(a) $(P \to Q) \& (Q \to R), \neg R \vdash \neg P$

(b) $P \to Q, \neg Q \vdash \neg\neg\neg P \vee Q$

(c) $P \& (P \to Q) \vdash Q$

(d) $(P \to Q) \to P \vdash (P \to Q) \to Q$

(e) $P \vee Q, R \vdash (P \& R) \vee (Q \& R)$

8 Further Proofs and the de Morgan Laws

[This chapter is omitted in this version of these notes]

9 Sequent-Introduction

[This chapter is abridged in this version of these notes]

Normally, when we make an application of $\rightarrow$I or $\neg$I, the pool of assumptions on the left falls by one. What happens if we push this process right to the limit? E.g.:

1	(1)	P	Asmp
2	(2)	Q	Asmp
1,2	(3)	$P \,\&\, Q$	1,2 &I
1	(4)	$Q \rightarrow (P \,\&\, Q)$	2,3 $\rightarrow$I
—	(5)	$P \rightarrow (Q \rightarrow (P \,\&\, Q))$	1,4 $\rightarrow$I

Here at line (5) we seem to have proved $P \rightarrow (Q \rightarrow (P \,\&\, Q))$ on no assumptions at all — and that is just what we have done! The truth of that wff has been shown to depend on no assumptions; it is unconditionally true. We write the result of the proof as a sequent with nothing to the left '$\vdash$', thus:

$$\vdash P \rightarrow (Q \rightarrow (P \,\&\, Q))$$

Such sequents are known as **theorems**. That is, theorems are *provable* sequents *whose left-hand side is empty*. Here are some further examples:

$$\vdash ((P \rightarrow Q) \,\&\, (P \rightarrow \neg Q)) \rightarrow \neg P$$

1	(1)	$(P \rightarrow Q) \,\&\, (P \rightarrow \neg Q)$	Asmp

1	(2)	$P\to\neg Q$	1 &E
1	(3)	$P\to Q$	1 &E
4	(4)	P	Asmp
1,4	(5)	Q	3,4 $\to$E
1,4	(6)	$\neg Q$	2,4 $\to$E
1,4	(7)	$\bot$	5,6 $\neg$E
1	(8)	$\neg P$	4,7 $\neg$I
–	(9)	$(P\to Q)\,\&\,(P\to\neg Q)\to\neg P$	1,8 $\to$I

$\vdash P\to P$ (Identity)

1	(1)	P	Asmp
–	(2)	$P\to P$	1,1 $\to$I !

$\vdash (\neg P\to P)\to P$ (Consequentia mirabilis)

1	(1)	$\neg P\to P$	Asmp
2	(2)	$\neg P$	Asmp
1,2	(3)	P	1,2 $\to$E
1,2	(4)	$\bot$	2,3 $\neg$E
1	(5)	$\neg\neg P$	2,4 $\neg$I
1	(6)	P	5 DN
–	(7)	$(\neg P\to P)\to P$	1,6 $\to$I

$\vdash (P\to Q)\to(\neg Q\to\neg P)$ (Contraposition)

1	(1)	$P\to Q$	Asmp
2	(2)	$\neg Q$	Asmp
3	(3)	P	Asmp
1,3	(4)	Q	1,3 $\to$E

1,2,3	(5)	$\bot$	2,4 $\neg$E
1,2	(6)	$\neg P$	3,5 $\neg$I
1	(7)	$\neg Q \rightarrow \neg P$	2,3 $\rightarrow$I
—	(8)	$(P \rightarrow Q) \rightarrow (\neg Q \rightarrow \neg P)$	1,4 $\rightarrow$I

$\vdash P \vee \neg P$ (The Law of Excluded Middle)

1	(1)	$\neg(P \vee \neg P)$	Asmp
2	(2)	P	Asmp
2	(3)	$P \vee \neg P$	2 $\vee$I
1,2	(4)	$\bot$	1,3 $\neg$E
1	(5)	$\neg P$	2,4 $\neg$I
1	(6)	$P \vee \neg P$	5 $\vee$I
1	(7)	$\bot$	1,6 $\neg$E
—	(8)	$\neg\neg(P \vee \neg P)$	1,7 $\neg$I
—	(9)	$P \vee \neg P$	8 DN

Now for one more new concept — then we'll do some work with them all. We define a **substitution-instance** of a wff as follows:

> A wff A is a **substitution-instance** of a wff B iff A results from B by uniformly replacing one or more of the variables in B by wffs.

Here are some examples:

initial formula: $P \to (Q \to (P \& Q))$
substitution key: $(R \vee S)/P;\ T/Q$
(i.e. '$R \vee S$' for 'P' and 'T' for 'Q')
resulting formula: $(R \vee S) \to (T \to ((R \vee S) \& T))$

initial formula: $P \to (Q \to (P \& Q))$
substitution key: $R/P;\ (\neg S \vee P)/Q$
resulting formula: $R \to ((\neg S \vee P) \to (R \& (\neg S \vee P)))$

initial formula: $\neg P \to (P \to Q)$
substitution key: $Q \& R/P;\ S \leftrightarrow T/Q$
resulting formula: $\neg(Q \& R) \to ((Q \& R) \to (S \leftrightarrow T))$

Intuitively, the idea of a uniform substitution should be perfectly clear. We can summarise it by giving four rules, two of them imperative and two permissive:

(1) Don't substitute for anything but individual variables — you are NOT allowed e.g. to do this:

initial formula: $\neg P \to (P \to Q)$
substitution key: ! $R/\neg P$!; S/P; $P \& Q/Q$ NO!
resulting formula: $R \to (S \to (P \& Q))$

(2) if any occurrences of a particular variable are substituted for, *all* must be — and all must be replaced by the *same* wff. So this is not on:

initial formula: $\neg P \to (P \to Q)$
substitution key: $R \to P/P;\ T/Q$ NO!
resulting formula: $\neg(R \to P) \to (!\ P\ ! \to T)$

and neither is this:

initial formula: $\neg P \rightarrow (P \rightarrow Q)$
substitution key: ! $R \rightarrow P/P$!; ! $Q \vee S/P$!; T/Q NO!
resulting formula: $\neg(R \rightarrow P) \rightarrow ((Q \vee S) \rightarrow T)$

(3) You may substitute with absolutely any wffs, however simple or complicated, including individual variables.

(4) Although, by (2), you may not substitute with distinct wffs for the same variable, you *may* substitute with the same wff for distinct variables. So this is OK:

initial formula: $\neg P \rightarrow (P \rightarrow Q)$
substitution key: R/P; R/Q
resulting formula: $\neg R \rightarrow (R \rightarrow R)$

Now here is a very striking thought: any proof of a theorem is, in effect, a proof of *any substitution instance* of that theorem; for all we have to do to construct the proof of the substitution instance is carry out the appropriate substitutions uniformly throughout the original proof, e.g.:

Theorem: $\vdash P \rightarrow (P \vee Q)$

Proof:

1	(1)	P	Asmp
1	(2)	$P \vee Q$	1 $\vee$I
–	(3)	$P \rightarrow (P \vee Q)$	1,2 $\rightarrow$I

Substitution instance: $\vdash (S \rightarrow Q) \rightarrow ((S \rightarrow Q) \vee (R \& P))$

Proof:

1	(1)	$S \rightarrow Q$	Asmp
1	(2)	$(S \rightarrow Q) \vee (R \& P)$	1 $\vee$I
–	(3)	$(S \rightarrow Q) \rightarrow ((S \rightarrow Q) \vee (R \& P))$	1,2 $\rightarrow$I

Uniform substitution throughout a proof always preserves the correctness of the proof. The reason is simple: the correctness of a proof depends only on the overall structure of the wffs involved and not on the details of the way they exemplify that structure.

A second bright thought about theorems is this: since they are proved on zero assumptions, it can often be a good tactic to prove a convenient theorem in the course of a proof; for it can then serve as a premise for further moves without having any repercussions on the assumptions pool. For example, seeking to prove

$$P \to Q, \neg P \to Q \vdash Q,$$

we could first give the proof of $\vdash P \vee \neg P$ illustrated above, and then proceed by $\vee$E as follows:

10	(10)	P	Asmp
11	(11)	$P \to Q$	Asmp
10, 11	(12)	Q	11,10 $\to$E
13	(13)	$\neg P$	Asmp
14	(14)	$\neg P \to Q$	Asmp
13, 14	(15)	Q	14,13 $\to$E
11, 14	(16)	Q	9,10,12,13,15 $\vee$E

On the other hand, why should it be necessary to have to prove $\vdash P \vee \neg P$ all over again? Is it not enough that it has already been proved? Obviously so. Accordingly, we now introduce the following further pair of rules of proof:

> TI: **Theorem Introduction**: it is permissible, at any stage in a proof, to introduce a line consisting of a previously proved theorem. No assumptions are cited on the left; and the citation on the right will be: TI, plus the name, or number, of the theorem.

TI(S): **Theorem Introduction (Substitution)**: it is permissible, at any stage in a proof, to introduce a line consisting of a substitution instance of a previously proven theorem. No assumptions are cited on the left; and the citation on the right will be: TI(S), plus the name, or number, of the theorem, and the appropriate substitution key.

N.B. These two rules are not increasing our proving power; they don't enable us to prove anything we couldn't prove before — indeed, that is why we are justified in introducing them. What they do is simplify the presentation of proofs. We could read an application of TI, or TI(S), as in effect a shorthand saying: "insert the proof of the appropriate theorem at this stage." Here are four useful theorems by name, three of which we proved above.

$\vdash P \lor \neg P$	:	Law of Excluded Middle (LEM)
$\vdash P \to P$	:	(Law of) Identity
$\vdash (P \to Q) \to (\neg Q \to \neg P)$	:	Contraposition
$\vdash \neg(P \,\&\, \neg P)$	:	(Law of) Non-Contradiction

Now for a further strengthening of our derivational muscles. Corresponding to every provable sequent there is — we trust — a valid form of argument: the form of argument that infers its conclusion from its assumptions. But for the reasons rehearsed when we introduced TI and TI(S), the proof of that sequent will remain correct under all uniform substitutions for variables in the wffs which it contains (uniform substitutions, that is, throughout the proof). So, if we generalise the notion of a substitution-instance to apply not just to individual wffs but to sequents like this:

> $B_1, \ldots, B_n \vdash B$ is a substitution-instance of $A_1, \ldots, A_n \vdash A$ just in case B_1 results from $A_1, \ldots, B_n$ results from A_n, and B results from A by some one overall uniform substitution for variables in $A_1, \ldots, A_n \vdash A$,

it is evidently correct to assert that a proof can be given for any substitution-instance of a provable sequent.

So two things follow: first, since our provable sequents correspond — we hope — to valid forms of argument, there is no reason why we should not use them, in addition to our primitive rules, to derive further sequents. Secondly, since we can prove any substitution instance of a provable sequent, there is no reason why we should not use any substitution instance of a proved sequent to prove further sequents. We therefore introduce two further conventions:

> SI: **Sequent Introduction**: Suppose we have previously proved some sequent, $A_1, \ldots, A_n \vdash B$; and suppose that, in the course of another

proof, each of its assumptions, A_1, …, A_n, figures as the conclusion of some line. Then SI says we may immediately derive a line whose conclusion is B and whose assumptions consist of the pool of all the assumptions of the lines in which A_1, …, A_n figure as conclusions.

Example: suppose we have previously proved:

$P, Q \to \neg P \vdash \neg Q$ $\quad (*)$

and now consider the following proof of $R, R \to P, Q \to \neg P \vdash \neg Q$:

1	(1)	R	Asmp
2	(2)	$R \to P$	Asmp
1,2	(3)	P	2,1 $\to$E
4	(4)	$Q \to \neg P$	Asmp
1,2,4	(5)	$\neg Q$	3,4 SI on (*) (sequent above)

The assumptions of our previously proved sequent figure as conclusions at lines (3) and (4); what SI permits us to do is go straight to the conclusion of that sequent at line 5, pooling the assumptions of lines (3) and (4) on the left.

Another example. Suppose we have previously proved:

$P \to Q, R \to \neg Q, P \vdash \neg R$ $\quad (**)$

and consider the following proof of $P, P \to (P \to Q), Q \to \neg R \vdash \neg R$:

1	(1)	P	Asmp
2	(2)	$P \to (P \to Q)$	Asmp
3	(3)	$Q \to \neg R$	Asmp
1,2	(4)	$P \to Q$	1,2 $\to$E
5	(5)	R	Asmp

6	(6)	Q	Asmp
3,6	(7)	$\neg R$	3,6 $\rightarrow$E
3,5,6	(8)	$\bot$	7,5 $\neg$E
3,5	(9)	$\neg Q$	6,8 $\neg$I
3	(10)	$R\rightarrow\neg Q$	5,9 $\rightarrow$I
1,2,3	(11)	$\neg R$	4,10,1 SI on (**), our previously proved sequent

The assumptions of our previously proved sequent figure as conclusions at lines (4), (8) and (1); so we go straight to the conclusion of that sequent at line (9), pooling the assumptions of (4), (8) and (1) on the left.

Now for SI(S): **Sequent Introduction (Substitution)**: This corresponds to SI exactly as TI(S) corresponds to TI. Thus whereas TI licenses us to introduce any previously proved theorem into a proof, and TI(S) licenses us to introduce any substitution-instance of a previously proved theorem, SI licenses us to use any previously proved sequent as a rule of inference and SI(S) licenses us so to use any substitution instance of a previously proved sequent. More formally, suppose we have previously proved some sequent, $A_1, \ldots, A_n \vdash B$, and that, in the course of a proof, $B_1, \ldots, B_n$ figure as the conclusions of n of the lines, where some one substitution key yields each of $B_1, \ldots, B_n$ from the corresponding $A_1, \ldots, A_n$. Then SI(S) says we may immediately derive a line whose conclusion is the corresponding substitution-instance of B and whose assumptions consist of the pool of all the assumptions of the lines in which $B_1, \ldots, B_n$ figure as conclusions.

Example: Suppose again, that we have previously proved

$$P,\ Q\rightarrow\neg P \vdash \neg Q \qquad (***)$$

and consider the following proof of $P\vee R,\ P\vee R\rightarrow S\vee T,\ R\rightarrow\neg(S\vee T) \vdash \neg R$:

1	(1)	$P\vee R$	Asmp
2	(2)	$P\vee R\rightarrow S\vee T$	Asmp
3	(3)	$R\rightarrow\neg(S\vee T)$	Asmp

1,2	(4)	$S \vee T$	2,1 $\rightarrow$E
1,2,3	(5)	$\neg R$	4,3 SI(S) on our previously proved sequent $(* * *)$, $S \vee T/P$, R/Q

Here lines (3) and (4) are substitution-instances of the assumptions of our previously proved sequent. Accordingly at line (5) we pass straight to the corresponding substitution-instance of the conclusion of that sequent, pooling the assumptions of (3) and (4) on the left. On the right we specify our authority for the step, mentioning the sequent we are appealing to, the lines where the substitution-instances of its assumptions figure, and the key to the relevant substitution.

Between them, TI, TI(S), SI, and SI(S), hugely facilitate proof construction. Note that TI is a special case of SI, for theorems are simply provable sequents with empty left-hand side; and SI itself is a special case of SI(S), with a vacuous substitution. So in future, instead of distinguishing TI, SI and so on, we will simply write 'SI', and call them all, equally, **sequent-introduction**.

In particular, we shall operate with a short-list of named sequents parallel to the short-list of theorems given above. So, to begin with, we take:

De Morgan's Laws (DEM):

$$
\begin{aligned}
\neg(P \vee Q) &\vdash \neg P \,\&\, \neg Q \\
\neg P \,\&\, \neg Q &\vdash \neg(P \vee Q) \\
\neg P \vee \neg Q &\vdash \neg(P \,\&\, Q) \\
\neg(P \,\&\, Q) &\vdash \neg P \vee \neg Q \\
\neg(\neg P \,\&\, \neg Q) &\vdash P \vee Q \\
\neg(\neg P \vee \neg Q) &\vdash P \,\&\, Q \\
P \vee Q &\vdash \neg(\neg P \,\&\, \neg Q) \\
P \,\&\, Q &\vdash \neg(\neg P \vee \neg Q)
\end{aligned}
$$

the Paradoxes of Material Implication (PMI):

$$\neg P \vdash P \rightarrow Q$$
$$Q \vdash P \rightarrow Q$$

Modus Tollendo Tollens (MTT):

$$P \rightarrow Q, \neg Q \vdash \neg P$$

and Ex Falso Quodlibet (EFQ):

$$P, \neg P \vdash Q.$$

Now let's see what we can do:

$\vdash P \vee (P \rightarrow Q)$

—	(1)	$P \vee \neg P$	SI, LEM
2	(2)	P	Asmp
2	(3)	$P \vee (P \rightarrow Q)$	2 $\vee$I
4	(4)	$\neg P$	Asmp
4	(5)	$P \rightarrow Q$	4 SI, PMI
4	(6)	$P \vee (P \rightarrow Q)$	5 $\vee$I
	(7)	$P \vee (P \rightarrow Q)$	1,2,3,4,6 $\vee$E

$\vdash (P \rightarrow Q) \vee (Q \rightarrow R)$

	(1)	$Q \vee \neg Q$	SI, LEM, Q/P
2	(2)	Q	Asmp
2	(3)	$P \rightarrow Q$	2 SI, PMI
2	(4)	$(P \rightarrow Q) \vee (Q \rightarrow R)$	3 $\vee$I
5	(5)	$\neg Q$	Asmp
5	(6)	$Q \rightarrow R$	5 SI, PMI, Q/P, R/Q.
5	(7)	$(P \rightarrow Q) \vee (Q \rightarrow R)$	6 $\vee$I
	(8)	$(P \rightarrow Q) \vee (Q \rightarrow R)$	1,2,4,5,7 $\vee$E

Both those proofs make nice use of Excluded Middle and the Paradoxes. Study them carefully, noting in particular the citation conventions for our new proof techniques.

Next: $P \vee Q \vdash \neg P \rightarrow Q$

Here is yet another way of proving this basic fact:

1	(1)	$P \vee Q$	Asmp
2	(2)	$\neg P$	Asmp
1	(3)	$\neg(\neg P \& \neg Q)$	1 SI, DEM
4	(4)	$\neg Q$	Asmp
2,4	(5)	$\neg P \& \neg Q$	2,4 &I
1,2,4	(6)	$\perp$	3,5 ¬E
1,2	(7)	$\neg\neg Q$	4,6 ¬I
1,2	(8)	Q	7 DN
1	(9)	$\neg P \rightarrow Q$	2,8 →I

Now: $P \vee Q, \neg P \vdash Q$: known as Modus Tollendo Ponens (MTP)

1	(1)	$P \vee Q$	Asmp
2	(2)	$\neg P$	Asmp
1	(3)	$\neg P \rightarrow Q$	1 SI, on the sequent we have just proved!
1,2	(4)	Q	2,3,→E

Finally: $\neg(P \& Q), P \vdash \neg Q$: known as Modus Ponendo Tollens (MPT)

1	(1)	$\neg(P \& Q)$	Asmp
2	(2)	P	Asmp

1	(3)	$\neg P \vee \neg Q$	1 SI, DEM
4	(4)	$\neg P$	Asmp
4	(5)	$P \rightarrow \neg Q$	4, SI, PMI, P/P, $\neg Q/Q$
2,4	(6)	$\neg Q$	2,4 $\rightarrow$E
7	(7)	$\neg Q$	Asmp
1,2	(8)	$\neg Q$	3,4,6,7,7 $\vee$E

This proof is, of course, needlessly involved. Its purpose is merely further to illustrate the use of the new techniques. (Give a simple primitive proof of the same sequent.) We may now add MTP and MPT to our stock of named results.

Exercises

(1) (a) Write out the formulae which result from the formulae given according to the substitution specified:

(i) formula: $P \to Q$
substitution: R/P; S/Q

(ii) formula: $P \to Q$
substitution: R/P; R/Q

(iii) formula: $P \to Q$
substitution: $P \to Q/P$

(iv) formula: $P \to Q \vee R$
substitution: $Q \,\&\, S/P$; $Q \,\&\, \neg S/Q$; S/R

(b) Write out the sequents which result from the sequents given according to the substitution specified:

(i) sequent: $P \to Q, P \vdash Q$
substitution: R/P; S/Q

(ii) sequent: $P \to (P \to Q) \vdash P \to Q$
substitution: $P \to Q/P$

(iii) sequent: $\vdash P \vee \neg P$
substitution: $\neg(P \,\&\, (Q \vee R))/P$

(iv) sequent: $\neg(A \,\&\, B) \vdash \neg A \vee \neg B$
substitution: P/A; $\neg Q/B$

(2) Prove the following theorems of propositional logic using only primitive rules:

(a) $\vdash \neg(P \,\&\, \neg P)$

(b) $\vdash (P \,\&\, \neg P) \to Q$

(3) Using SI, on either of the paradoxes of material implication or the law of excluded middle, prove the following sequents:

(a) $P \to \neg(Q \to R) \vdash (P \to Q) \,\&\, (P \to \neg R)$

(b) $\vdash (P \to Q) \to ((\neg P \to Q) \to Q)$

(c) $\neg(P \to (Q \vee R)) \vdash (Q \vee R) \to P$

10 Rules for the Universal Quantifier

We now turn to the inference rules for quantified logic. As with the connectives of propositional logic, we shall need two types of rule for each quantifier, in particular, a rule to take us *from* a sequent with a universally quantified conclusion, i.e. an **elimination** rule; and a rule to take us to a sequent with a universally quantified conclusion, i.e. an **introduction** rule. Both these rules for the universal quantifier are reasonably straightforward.

We start with the introduction rule; when can we validly infer a sequent of the following form?

$$X \vdash (\forall\xi)(\ldots\xi\ldots)$$

The answer, evidently, is: when, given X, we have shown that the predicate to which the quantifier is attached — symbolised by the dots around the variable 'ξ' (the Greek letter 'xi') — is true of any particular object we care to choose. That is, if υ (upsilon) is any object you like to choose, and we can guarantee to be able to show that $X \vdash \ldots\xi\ldots$, irrespective of which object υ you choose, then the assumptions X must entail $(\forall\xi)(\ldots\xi\ldots)$ — since they entail all the particular cases.

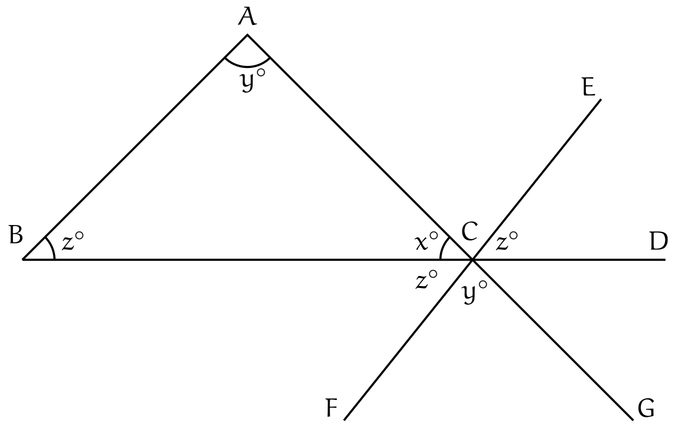
This mode of inference is familiar from Euclidean geometry. Suppose we want to prove that all triangles have interior angles whose sum is 180°. And suppose we have, as data,

(1) that the angle on a straight line is 180°;

(2) that intersecting straight lines enclose equal opposite angles; and

(3) that a straight line intersecting two parallel lines encloses equal complementary angles.

Euclid reasoned like this. Let υ be any triangle with points A, B, C: and let BC be extended to D, and a line EF, parallel to AB, be run through C. Finally let AC be extended to G.



Then by (3), the angle ECD $= z°$. Hence by (2) the angle BCF $= z°$. And by (3) the angle FCG $= y°$. So by (1), the angle on the straight line ACG, $x° + y° + z° = 180°$. Hence, υ's interior angles add up to 180°. But α could have been *any* triangle — we made no special assumptions about it except that it was a triangle; so our reasoning would apply to any triangle. So all triangles have interior angles whose sum is 180°.

That is a classic intuitive application of the rule which we are going to have: $\forall$-**Introduction** — $\forall$I. The rule says that given that we have shown that a particular object, υ, is characterised by a certain predicate, and *given that the way we have shown this would work for any other object* — i.e. the proof makes no special assumptions about υ — we can infer that everything is characterised by that predicate. Schematically:

$$\frac{X \vdash \ldots\upsilon\ldots}{X \vdash (\forall\xi)(\ldots\xi\ldots)} \quad \forall\mathrm{I}$$

where every occurrence of 'υ' in $(\ldots\upsilon\ldots)$ is replaced by one of 'ξ', and provided none of the formulae in X make any special assumptions about υ.

But what does that proviso mean?

Let us try to say precisely what it means. Recall that if $A(\xi)$ is a wff containing one or more occurrences of the variable 'ξ', and not already containing a quantifier $(\forall\xi)$ or $(\exists\xi)$, $(\forall\xi)A(\xi)$ and $(\exists\xi)A(\xi)$ are wffs. We say that $(\forall\xi)A(\xi)$ and $(\exists\xi)A(\xi)$ is in each case the **scope** of the quantifier $(\forall\xi)$ or $(\exists\xi)$ shown. All occurrences of 'ξ' lying within the scope of a quantifier $(\forall\xi)$ or $(\exists\xi)$ with the same variable 'ξ' are said to be **bound**. Any occurrence of 'ξ' in a wff which does not lie in the scope of a quantifier of the form $(\forall\xi)$ or $(\exists\xi)$ is called **free**.

For example, in

$$(\forall x)Rxy$$

both occurrences of 'x' are bound, that of 'y' is free. In

$$(\forall y)Ryy \,\&\, Gxy$$

the first three occurrences of 'y' are bound, the fourth free (it is not within the scope of the quantifier), and the only occurrence there of 'x' is free.

Note the contrast in what is said by $(\forall x)Rxx$ and $(\forall y)Ryx$. $(\forall x)Rxx$ says that every thing has R to itself. It speaks of any and everything, not of any particular thing. In contrast, $(\forall y)Ryx$ says that everything has R to x, to a *particular* object x. Free occurrences of a variable relate to a particular thing — whatever x or y is. Bound occurrences, on the other hand, do not. To say that for every x, x has R to itself, is the same as to say that for every y, y has R to itself. There's nothing particular to x or y in the meanings of $(\forall x)Rxx$ and $(\forall y)Ryy$.

This distinction solves our problem. It is exactly this notion of there being nothing particular in using one bound variable rather than another that we need in order to express the proviso that in X we make no particular or special assumptions about 'υ'. In brief, we can express the proviso like this: the variable 'υ' may occur *bound* in the assumptions, X, without harm. But if the $\forall$I step is to be valid, it must not appear there *free*.

Let 'ξ', 'υ' be any variables, and $A(\xi)$ a formula with at least one free occurrence of 'ξ'. Let $A(\upsilon)$ result from replacing *all* free occurrences of 'ξ' by 'υ'. Then the $\forall$I rule has the form

$$\frac{X \vdash A(\upsilon)}{X \vdash (\forall\xi)A(\xi)} \quad \forall\mathrm{I}$$

provided 'υ' does not occur *free* in any formula in X nor in $(\forall\xi)A(\xi)$.

We call 'υ' the **eigenvariable**, or the **parametric variable**, of the inference.

So much for motivational remarks. They are designed to give you some idea why we frame the ∀I rule in the way we have. But you may not fully understand them yet. Indeed, they are only motivational. That this is the *correct* form of the rule is a much harder question, requiring a proof of soundness and completeness which, as in the case of propositional logic, goes well beyond the scope of this course.

Next, ∀-**Elimination — ∀E** . This rule says simply that, given that it follows from certain assumptions that *everything* is characterised by a certain predicate, those same assumptions will entail that *any particular thing* is characterised by that predicate; i.e.,

$$\frac{X \vdash (\forall\xi)(\ldots\xi\ldots)}{X \vdash \ldots\tau\ldots} \qquad \forall\mathrm{E}$$

for any variable 'ξ' and for any term (i.e., variable or constant) 'τ' (tau) you care to choose. Here τ must replace every occurrence of ξ in $(\ldots\xi\ldots)$. There may, however, already be occurrences of τ in $(\ldots\xi\ldots)$. The rule is intuitively absolutely self-evident: it simply relies on the principle that if everything has some property, so does each particular thing.

There is a complication, however. Consider the wff $(\forall x)(\exists y)Rxy$. If we were to substitute 'y' for 'x' in applying ∀E to this wff, we would obtain

$$(\exists y)Ryy.$$

But the premise does not warrant the conclusion. That everything is related by R to something, does not entail that something is related by R to *itself*. The trouble is that the term substituted for 'x' has become inadvertently *bound*, by the quantifier $(\exists y)$. But the rule is only valid if the term substituted for the variable 'x' is either a constant, or, if a variable, occurs as a *free* variable.

We therefore introduce the notion **free for** 'υ': the variable 'υ' is free for 'ξ' in a wff $A(\xi)$ if 'ξ' does not occur in $A(\xi)$ within the scope of a quantifier of the form $(\forall\upsilon)$ or $(\exists\upsilon)$. Then, in ∀E , we require that 'τ' be such that 'τ' is free for 'υ' in $(\ldots\xi\ldots)$. Thus the full form of the rule is:

$$\frac{X \vdash (\forall \xi)A(\xi)}{X \vdash A(\tau)} \quad \forall E$$

provided τ is free for ξ in $A(\xi)$. Here $A(\tau)$ results from $A(\xi)$ by replacing all occurrences of ξ by τ.

Make sure you are clear when an occurrence of a variable 'ξ' in a wff is **free**, when it is **bound**, and when a variable is **free for** another variable 'τ'.

Here are some examples of ∀E in use:

(a) $(\forall x)Fx \vdash Fa$

1	(1)	$(\forall x)Fx$	Asmp
1	(2)	Fa	1 ∀E

(b) $\neg Ga, (\forall x)(Fx \rightarrow Gx) \vdash \neg Fa$

1	(1)	$\neg Ga$	Asmp
2	(2)	$(\forall x)(Fx \rightarrow Gx)$	Asmp
2	(3)	$Fa \rightarrow Ga$	2 ∀E
1,2	(4)	$\neg Fa$	1,3 SI, MTT, Fa/P, Ga/Q

(c) $(\forall x)(Fx \rightarrow Gx), (\forall x)(Gx \rightarrow Hx) \vdash Fa \rightarrow Ha$

1	(1)	$(\forall x)(Fx \rightarrow Gx)$	Asmp
2	(2)	$(\forall x)(Gx \rightarrow Hx)$	Asmp
3	(3)	Fa	Asmp
1	(4)	$Fa \rightarrow Ga$	1 ∀E
2	(5)	$Ga \rightarrow Ha$	2 ∀E
1,3	(6)	Ga	3,4 →E
1,2,3	(7)	Ha	5,6 →E
1,2	(8)	$Fa \rightarrow Ha$	3,7 →I

(d) $(\forall x)(Fx \vee P), \neg P \vdash Fa$

1	(1)	$(\forall x)(Fx \vee P)$	Asmp
2	(2)	$\neg P$	Asmp
1	(3)	$Fa \vee P$	1 ∀E
1,2	(4)	Fa	2,3 SI, MTP, Fa/P, P/Q

Points to note:

(1) The ∀E step just consists in deleting the quantifier and replacing all occurrences of the relevant variable by the chosen term.

(2) In a ∀E step just one line is cited on the right — that of the premise-sequent — and the assumption pool remains the same.

(3) The rules, including derived rules, and vocabulary, including propositional variables, of propositional logic are all in play; and all the old proof strategies remain useful. Thus c) is basically a →I proof; the only novelty lies in its having quantified assumptions.

Let us now try some examples involving both ∀E and ∀I, to see how the rules work in practice.

(e) $(\forall x)(Fx \rightarrow Gx), (\forall x)\neg Gx \vdash (\forall x)\neg Fx$

1	(1)	$(\forall x)(Fx \rightarrow Gx)$	Asmp
2	(2)	$(\forall x)\neg Gx$	Asmp

Strategy: we want to prove a universal conclusion; so we'll try to prove the special case, $\neg Fx$, say, from assumptions 1 and 2 — for, since there are no free occurrences of 'x' in 1 or 2, ∀I will then give us $(\forall x)\neg Fx$ on the same assumptions. So, a couple of steps of ∀E first:

1	(3)	$Fx \rightarrow Gx$	1 ∀E
2	(4)	$\neg Gx$	2 ∀E
1,2	(5)	$\neg Fx$	3,4 SI, MTT, Fx/P, Gx/Q

(Trivially, 'x' is free for 'x' in $Fx \rightarrow Gx$ and in $\neg Gx$. Indeed, 'ξ' is always free for 'ξ' in $A(\xi)$). Think about it.) So finally,

1,2	(6)	$(\forall x)\neg Fx$	5 $\forall$I

The rule $\forall$I is thus simplicity itself to operate. All we have to do, whenever we have a line containing a free variable, is check whether the assumptions of that line contain any free occurrences of that variable. If, and only if, they don't, we can universally quantify the line; i.e. replace all occurrences of the variable by any other variable not already in the wff, or leave it unchanged, and then prefix the universal quantifier for that variable. $\forall$I gives the basic strategy for proving sequents with universally quantified conclusions. Another example:

(f) $(\forall x)(Fx \rightarrow \neg Gx) \vdash (\forall x)(Gx \rightarrow \neg Fx)$

1	(1)	$(\forall x)(Fx \rightarrow \neg Gx)$	Asmp

Strategy: we'll aim to prove a special case of the conclusion prior to a $\forall$I step — say, $Gy \rightarrow \neg Fy$ (just to ring the changes on 'x'!). This is a conditional, so the basic $\rightarrow$I strategy is the obvious ploy. Hence we assume the antecedent:

2	(2)	Gy	Asmp
1	(3)	$Fy \rightarrow \neg Gy$	1 $\forall$E
4	(4)	Fy	Asmp
1,4	(5)	$\neg Gy$	3,4 $\rightarrow$E
1,2,4	(6)	$\perp$	2,5 $\neg$E
1,2	(7)	$\neg Fy$	4,6 $\neg$I
1	(8)	$Gy \rightarrow \neg Fy$	2,7 $\rightarrow$I
1	(9)	$(\forall x)(Gx \rightarrow \neg Fx)$	6 $\forall$I

After the $\forall$E at line 3, the propositional logic rules take us smoothly to line 8. The assumption on which line 8 depends does not contain 'y' free; so all is in order for the $\forall$I step at line 9. Notice the emergent pattern with these proofs: first we get rid of the quantifiers; second, we use propositional logic rules on the resulting quantifier-free wffs; third, we introduce the quantifiers again. Proofs in predicate logic do not always go like this; but very often they do so.

Let us think of example (f) tactically. The ∀I tactic looks very like the tactics in propositional logic. As an introduction tactic, it works on the right-hand side of the goal:

to show	$\Gamma \Rightarrow (\forall\xi)(\ldots\xi\ldots)$
attempt to show	$\Gamma \Rightarrow (\ldots\upsilon\ldots)$
	where υ is not free in any wff in Γ,
	and υ is free for ξ in $(\ldots\xi\ldots)$.

Note the way one must choose some variable υ *not free* in any assumptions.

The elimination tactic, ∀E -tactic, also has a familiar form:

to show	$\Gamma \Rightarrow C$
check that	$(\forall\xi)(\ldots\xi\ldots) \in \Gamma$, for some wff $(\ldots\xi\ldots)$
and attempt to show	$(\ldots\tau\ldots), \Gamma \Rightarrow C$
	where τ is free for ξ in $(\ldots\xi\ldots)$.

Again, the restriction in the rule is repeated in the tactic. This time, we work on the left-hand side of the goal, replacing a universally quantified wff by an instance.

What, then, does the tactical derivation of example (f) look like? It runs as follows:

? $(\forall x)(Fx\rightarrow\neg Gx) \Rightarrow (\forall x)(Gx\rightarrow\neg Fx)$
> Using tactic for $\forall$I (choosing 'y', not free in the LHS)
>
> ? $(\forall x)(Fx\rightarrow\neg Gx) \Rightarrow Gy\rightarrow\neg Fy$
> > Using tactic for $\rightarrow$I
> >
> > ? $Gy, (\forall x)(Fx\rightarrow\neg Gx) \Rightarrow \neg Fy$
> > > Using tactic for $\neg$I
> > >
> > > ? $Fy, Gy, (\forall x)(Fx\rightarrow\neg Gx) \Rightarrow \bot$
> > > > Using tactic for $\forall$E (choosing 'y', free for 'x' in '$Fx\rightarrow\neg Gx$')
> > > >
> > > > ? $Fy\rightarrow\neg Gy, Fy, Gy, (\forall x)(Fx\rightarrow\neg Gx) \Rightarrow \bot$
> > > > > Using tactic for $\rightarrow$E
> > > > >
> > > > > ? $Fy\rightarrow\neg Gy, Fy, Gy, (\forall x)(Fx\rightarrow\neg Gx) \Rightarrow Fy$ ■
> > > > >
> > > > > ? $\neg Gy, Fy, Gy, (\forall x)(Fx\rightarrow\neg Gx) \Rightarrow \bot$
> > > > > > Using tactic for $\neg$E
> > > > > >
> > > > > > ? $\neg Gy, Fy, Gy, (\forall x)(Fx\rightarrow\neg Gx) \Rightarrow Gy$ ■
> > > > > >
> > > > > > ? $\bot, Fy, Gy, (\forall x)(Fx\rightarrow\neg Gx) \Rightarrow \bot$ ■

We could have used 'x' for the $\forall$I -tactic, since it is not free in $(\forall x)(Fx\rightarrow\neg Gx)$ either. But in the above proof we chose to use 'y'. So that comes out in the tactics, too.

Next:

(g) $(\forall x)Fx, (\forall y)(Fy\rightarrow Gy) \vdash (\forall y)Gy$

1	(1)	$(\forall x)Fx$	Asmp
2	(2)	$(\forall y)(Fy\rightarrow Gy)$	Asmp
1	(3)	Fx	1 $\forall$E
2	(4)	$Fx\rightarrow Gx$	2 $\forall$E
1,2	(5)	Gx	3,4 $\rightarrow$E
1,2	(6)	$(\forall y)Gy$	5 $\forall$I

The only salient point in this proof is its illustration of the handling of different variables: we can replace different variables by the same variable in $\forall$E steps, provided the replacement variable is **free for** the variable to be replaced; and we can introduce whatever variable we like in a $\forall$I step. Consider this example:

(h) (∀x)(∀y)Rxy ⊢ (∀x)Rxx

1	(1)	(∀x)(∀y)Rxy	Asmp
1	(2)	(∀y)Rxy	1 ∀E ('x' unchanged)
1	(3)	Rxx	2 ∀E ('x' for 'y', since it is free for 'y')
1	(4)	(∀x)Rxx	3 ∀I ('x' unchanged)

Think what the sequent means. Read 'Rxy' as, for example, 'x loves y'.

(i) (∀x)(∀y)Rxy ⊢ (∀y)(∀x)Ryx

1	(1)	(∀x)(∀y)Rxy	Asmp
1	(2)	(∀y)Rzy	1 ∀E ('z' for 'x', since it is free for 'x')
1	(3)	Rzy	2 ∀E ('y' unchanged)
1	(4)	(∀x)Rzx	3 ∀I ('x' unchanged)
1	(5)	(∀y)(∀x)Ryx	4 ∀I ('y' for 'z')

Notice that, although 'z' is free for 'x' in (∀y)Rxy, 'y' is *not* free for 'x' there, so we could not move to (∀y)Ryy at line 2. Notice also, as illustrated in h) and i), that it is necessary to eliminate, and to introduce, quantifiers one at a time; and that we correspondingly substitute for variables one at a time. This kind of thing:

	…	(17)	(∀x)(∀y)Rxy	…
!!	…	(18)	Ryx	17 ∀E *NO!!*

is definitely forbidden!!

(j) $(\forall x)Fx,\ (\forall z)(Fz \to Hz) \vdash \neg(\forall x)\neg Hx$

1	(1)	$(\forall x)Fx$	Asmp
2	(2)	$(\forall z)(Fz \to Hz)$	Asmp

Strategy: We are going for a negative conclusion; so we'll assume the unnegated formula itself, and look to ¬I to get us home:

3	(3)	$(\forall x)\neg Hx$	Asmp
4	(4)	Fx	1 ∀E
2	(5)	$Fx \to Hx$	2 ∀E
3	(6)	$\neg Hx$	3 ∀E
1,2	(7)	Hx	4,5 →E
1,2,3	(8)	$\perp$	6,7 ¬E
1,2	(9)	$\neg(\forall x)\neg Hx$	3,8 ¬I

Here is the tactical derivation for j):

? $(\forall x)Fx,\ (\forall z)(Fz \to Hz) \Rightarrow \neg(\forall x)\neg Hx$
- Using tactic for ¬I
- ? $(\forall x)\neg Hx,\ (\forall x)Fx,\ (\forall z)(Fz \to Hz) \Rightarrow \perp$
  - Using tactic for ∀E
  - ? $Fx,\ (\forall x)\neg Hx,\ (\forall x)Fx,\ (\forall z)(Fz \to Hz) \Rightarrow \perp$
    - Using tactic for ∀E
    - ? $Fx \to Hx,\ Fx,\ (\forall x)\neg Hx,\ (\forall x)Fx,\ (\forall z)(Fz \to Hz) \Rightarrow \perp$
      - Using tactic for ∀E
      - ? $\neg Hx,\ Fx \to Hx,\ Fx,\ (\forall x)\neg Hx,\ (\forall x)Fx,\ (\forall z)(Fz \to Hz) \Rightarrow \perp$
        - Using tactic for →E
        - ? $\neg Hx,\ Fx \to Hx,\ Fx,\ (\forall x)\neg Hx,\ (\forall x)Fx,\ (\forall z)(Fz \to Hz) \Rightarrow Fx$ ■
        - ? $Hx,\ \neg Hx,\ Fx,\ (\forall x)\neg Hx,\ (\forall x)Fx,\ (\forall z)(Fz \to Hz) \Rightarrow \perp$
          - Using tactic for ¬E
          - ? $Hx,\ \neg Hx,\ Fx,\ (\forall x)\neg Hx,\ (\forall x)Fx,\ (\forall z)(Fz \to Hz) \Rightarrow Hx$ ■
          - ? $\perp,\ Hx,\ Fx,\ (\forall x)\neg Hx,\ (\forall x)Fx,\ (\forall z)(Fz \to Hz) \Rightarrow \perp$ ■

Finally, a slightly more sophisticated example:

(k) $(\forall x)Fx \vee P \vdash (\forall x)(Fx \vee P)$

First, left-to-right. Strategy: universally quantified conclusion, so we'll aim to prove a special case and then use $\forall$I ; disjunctive premise, so we'll aim to get the special case from each disjunct and then use $\vee$E . Thus:

1	(1)	$(\forall x)Fx \vee P$	Asmp
2	(2)	$(\forall x)Fx$	Asmp
2	(3)	Fx	2 $\forall$E
2	(4)	$Fx \vee P$	3 $\vee$I
5	(5)	P	Asmp
5	(6)	$Fx \vee P$	5 $\vee$I
1	(7)	$Fx \vee P$	1,2,4,5,6 $\vee$E
1	(8)	$(\forall x)(Fx \vee P)$	7 $\forall$I

Now, right-to-left. It would be natural to proceed like this:

	1	(1)	$(\forall x)(Fx \vee P)$	Asmp
	1	(2)	$Fx \vee P$	1 $\forall$E
	3	(3)	Fx	Asmp
!!	3	(4)	$(\forall x)Fx$	3 $\forall$I *NO!!*
	3	(5)	$(\forall x)Fx \vee P$	4 $\vee$I
	6	(6)	P	Asmp
	6	(7)	$(\forall x)Fx \vee P$	6 $\vee$I
	1	(8)	$(\forall x)Fx \vee P$	2,3,5,6,7 $\vee$E

But the $\forall$I step at line 4 is out of order since the assumption on which Fx depends at line 3 is Fx itself, in which 'x' occurs free. So how is the proof to go? The DN plus $\neg$I tactic is the obvious strategy to try next — and, provided we remember our propositional logic derived rules, things turn out fairly sweetly:

1	(1)	$(\forall x)(Fx \vee P)$	Asmp
2	(2)	$\neg((\forall x)Fx \vee P)$	Asmp
2	(3)	$\neg(\forall x)Fx \,\&\, \neg P$	2 SI, DEM, $(\forall x)Fx/P$, P/Q
1	(4)	$Fx \vee P$	1 $\forall$E
2	(5)	$\neg P$	3 &E
1,2	(6)	Fx	4,5 SI, MTP, $Fx/P, P/Q$
1,2	(7)	$(\forall x)Fx$	6 $\forall$I (note: neither 1 nor 2 contains 'x' free)
2	(8)	$\neg(\forall x)Fx$	3 &E
1,2	(9)	$\perp$	7,8 $\neg$E
1	(10)	$\neg\neg((\forall x)Fx \vee P)$	2,9 $\neg$I
1	(11)	$(\forall x)Fx \vee P$	10 DN

(See if you can find an alternative proof of this last result, using SI on the Law of Excluded Middle.)

Finally, a quick example of what can happen if the restriction on $\forall$I is not observed:

	1	(1)	$(\forall x)Fx$	Asmp
	2	(2)	Gy	Asmp
	1	(3)	Fy	1 $\forall$E
	1,2	(4)	$Fy \,\&\, Gy$	2,3 &I
!!	1,2	(5)	$(\forall x)(Fx \,\&\, Gx)$	4 $\forall$I *NO!!*

The $\forall$I step at line 5 is incorrect since $Fy \,\&\, Gy$ at line 4 depends on an assumption, viz. 2, which contains 'y' free. The result is the patently invalid sequent at line 5:

$$(\forall x)Fx,\ Gy \vdash (\forall x)(Fx \,\&\, Gx).$$

Think what it means. Read 'Fx' as, for example, 'x is human', 'Gx' as 'x is male' and think of y as, say, Ronald Reagan. Take as universe of discourse, people. Then it does not follow that because everyone is human and Reagan is male, that *everyone* is a male human.

Exercises

(1) (a) What are the free occurrences of variables in

(i) $Fx \rightarrow (\exists y)(Gy \,\&\, Hxy)$

(ii) $(\neg Fz \rightarrow (\forall y)\neg Lyaz) \rightarrow (\exists y)(\exists z)(Fz \,\&\, Lyaz)$

(iii) $(\forall x)(\exists y)Fxy \rightarrow Ga$

(b) Is 'y' free for 'x' in the following:

(i) $(\exists y)Fy \rightarrow Fx$

(ii) $(\forall y)(Fy \rightarrow Fx)$

(2) Formalise the following argument and show that it is sound. You may use SI on (a substitution-instance of) any tautologous sequent of propositional logic:

> Only the free are happy. No one who works in a capitalist society is free. Anyone who works in Britain works in a capitalist society. Therefore all British workers are unhappy.

(3) Prove the following sequents of classical predicate logic. You may use SI on (a substitution-instance of) any tautologous sequent of propositional logic:

(a) $(\forall x)\neg Fx,\ P \rightarrow (\forall x)Fx \vdash \neg P$

(b) $(\forall x)(Fx \rightarrow Gx),\ (\forall x)(\neg Fx \rightarrow Gx) \vdash (\forall x)Gx$

(c) $(\forall x)(Fx \rightarrow Gx),\ (\forall x)(Gx \rightarrow Hx) \vdash (\forall x)(Fx \rightarrow Hx)$

(d) $(\forall x)(Wx \rightarrow Mx),\ (\forall x)(Fx \rightarrow \neg Mx) \vdash (\forall x)(Wx \rightarrow \neg Fx)$

(e) $Lmn, (\forall x)(Wx \rightarrow \neg Lxn) \vdash \neg Wm$

(f) $(\forall x)(Gx \rightarrow Fx),\ (\forall x)(Tx \rightarrow \neg Fx) \vdash (\forall x)(Tx \rightarrow \neg Gx)$

11 Rules for the Existential Quantifier

The introduction rule for '$\exists$' is straightforward. If we ask in general terms: under what circumstances will a set of assumptions, X, entail an existential statement, $(\exists\xi)(\ldots\xi\ldots)$, the answer, evidently, is: when they entail that some particular object is characterised by the predicate in question. That suggests the rule Existential Introduction, $\exists$I:

$$\frac{X \vdash A(\tau)}{X \vdash (\exists\xi)A(\xi)} \qquad \exists\mathrm{I}$$

provided τ is free for ξ in $A(\xi)$, where 'τ' is any term — constant or variable.

The rule dictates a very simple strategy for proving sequents with existential conclusions: show that an appropriate special case follows from the given assumptions, and then use $\exists$I. Note that τ may appear in $(\exists\xi)A(\xi)$ — even if τ is a variable, it may appear free there (this was illicit in $\forall$I — see Chapter 10). For example, if we know that Russell shaves himself — Raa — we may infer not only that someone shaves himself — $(\exists x)Rxx$, but also that someone shaves Russell — $(\exists x)Rxa$ — and that Russell shaves someone — $(\exists x)Rax$. Note also, however, as with $\forall$E , that $A(\tau)$ must result from $A(\xi)$ by replacing all occurrences of ξ in $A(\xi)$ by τ.

Here are some examples of proofs using $\exists$I:

(a) $(\forall x)(Fx \to Gx), Fa \vdash (\exists x)Gx$

1	(1)	$(\forall x)(Fx \to Gx)$	Asmp
2	(2)	Fa	Asmp
1	(3)	$Fa \to Ga$	1 $\forall$E

1,2	(4)	Ga	2,3 $\rightarrow$E
1,2	(5)	$(\exists x)Gx$	4 $\exists$I

Notice that with ∃I we just cite one line — that where the special case occurs — and the assumption pool remains constant.

(b) $(\forall x)Fx \vdash (\exists x)(Fx \vee Gx)$

1	(1)	$(\forall x)Fx$	Asmp
1	(2)	Fx	1 $\forall$E
1	(3)	$Fx \vee Gx$	2 $\vee$I
1	(4)	$(\exists x)(Fx \vee Gx)$	3 $\exists$I

(c) $(\forall x)Rxx \vdash (\forall x)(\exists y)Rxy$

1	(1)	$(\forall x)Rxx$	Asmp
1	(2)	Rxx	1 $\forall$E
1	(3)	$(\exists y)Rxy$	2 $\exists$I
1	(4)	$(\forall x)(\exists y)Rxy$	3 $\forall$I

Note that step (3) is licit even though $(\exists y)Rxy$ contains 'x' free, and that step (4) is licit since 'x' is not free in (1), nor in (4).

(d) $\vdash (\exists x)(Fx \vee \neg Fx)$ — our first predicate logic theorem:

—	(1)	$Fx \vee \neg Fx$	SI, LEM, Fx/P
—	(2)	$(\exists x)(Fx \vee \neg Fx)$	1 $\exists$I

Thus we can prove that the world is non-empty — that at least one thing exists! This is sometimes felt to be a compromise of logical purity: it is no logical truth, it is felt, that the universe is not void. But that is all a muddle; the significance of the theorem is merely that validity in predicate logic is going to be relative to the (true!) assumption that at least one individual exists; nothing is implied about the status of that assumption.

Now for the rule of Existential Elimination, $\exists$E. This is the most subtle of the four quantifier rules, and the most involved to formulate exactly, as you are about to see. But anyone who has mastered $\vee$E should have no trouble with $\exists$E.

First, note that $(\exists\xi)A(\xi)$ does *NOT* entail $A(\tau)$, for any particular term 'τ'. For $A(\tau)$ can be false even though *something* (else) is such that $A(\ldots)$. So we must ask: when — intuitively — can we say that $(\exists\xi)A(\xi)$ entails some statement, B? What we need to know in order to know that $(\exists\xi)A(\xi)$ is true, is that some object, whose identity we may not know, is such that $A(\ldots)$. Call this object 'υ'. Then $(\exists\xi)A(\xi)$ entails B provided

(i) $A(\upsilon)$ entails B and

(ii) that it does so depends on no features of 'υ' other than that $A(\upsilon)$ is true — i.e. the method whereby it is shown that $A(\upsilon)$ entails B would be equally effective at demonstrating that $A(\chi)$, for *any other* variable 'χ', entailed B.

In other words, it is not to matter which object we suppose to be $A(\ldots)$; there has to be a demonstration that B follows from any assumption of the form, $A(\upsilon)$. When, and only when, that has been demonstrated can we regard the truth of $(\exists\xi)A(\xi)$ as a guarantee of the truth of B. Again, we call 'υ' the **eigenvariable**, or *parametric variable*, of the application of $\exists$E.

If you think about it, you'll see that this generality — a demonstration that B follows from any assumption of the form $A(\upsilon)$ — cannot be attained unless we make the following provisos:

First, the conclusion, B, must not contain 'υ' free, that is, B must contain no free occurrences of the eigenvariable 'υ'. For if B contains 'υ' free, it may very well follow from $A(\upsilon)$ but from no other assumption of that form. Example: $Fx \vdash Fx \vee Gx$, but $Fy \nvdash Fx \vee Gx$.

Second, any *auxiliary* assumptions used in deriving B from $A(\upsilon)$ must not contain 'υ' free. Otherwise the derivation may, again, fail if a different eigenvariable is chosen. Example: $Fx, \neg(Fx \,\&\, Q) \vdash \neg Q$, but $Fy, \neg(Fx \,\&\, Q) \nvdash \neg Q$.

These two provisos are clearly necessary. But they do not yet suffice for the generality we want — the derivability of B from *any* instance of $(\exists\xi)A(\xi)$. Remember that $A(\ldots)$ may be a predicate of any degree of complexity, which may therefore contain a variety of variables, etc. Suppose, for instance, we take $A(\ldots)$ as the predicate: '$H\ldots \& Gy$' and 'y' as 'υ'. And let B be $(\exists x)(Hx \& Gx)$. Then (i) $A(y) \vdash B$, i.e. $Hy \& Gy \vdash (\exists x)(Hx \& Gx))$; (ii) B does not contain 'y' free; and (iii) no auxiliary assumptions are involved. So if the first two provisos were enough for the validity of an ∃E step, we could infer

$$(\exists x)(Hx \& Gy) \vdash (\exists x)(Hx \& Gx),$$

which is an obviously invalid sequent, since the existence of something of which it is true that it is H while some (possibly distinct) object y is G does not entail that something is both H and G. What is the remedy? Plainly a different choice of eigenvariable — 'z' for instance — would not sustain the derivation, since $Hz \& Gy \nvdash (\exists x)(Hx \& Gx)$, and we want to ensure that that the derivation of the relevant B goes through no matter what we choose as the eigenvariable. But why does the choice make a difference in this case? Because we picked in our example the existential wff $(\exists\xi)A(\xi)$ which contained a free variable, namely 'y'. There is nothing wrong with so doing; but if we pick such an example, the logical powers of $A(\upsilon)$ are obviously going to vary with the choice of 'υ' and will be sensitive, in particular, to whether or not the free variable in question is taken as 'υ'. The way to avoid any such sensitivity is to restrict the permissible selections for 'υ' to variables which do not occur free in $(\exists\xi)A(\xi)$, i.e. we stipulate

> *Third*, $(\exists\xi)A(\xi)$ must not contain 'υ' free.

Only if these three provisos are met can we guarantee to be able to substitute any other variable for 'υ' in $A(\upsilon)$ yet still be able validly to derive B from the result. But do they suffice for the generality we want?

Not quite. Recall, once again, that $A(\ldots)$ may be a predicate of any degree of complexity — and so may contain *quantifiers*. So we have to ensure that 'υ' is so chosen that it can occur free in $A(\upsilon)$ and does not inadvertently, as it were, come within the scope of a quantifier in $A(\ldots)$. Suppose, for instance, that $(\exists\xi)A(\xi)$ is $(\exists x)(\exists y)(Hx \& Gy)$, so that $A(\ldots)$ is $(\exists y)(H\ldots \& Gy)$. And now suppose we selected 'y' as the eigenvariable, so that $A(\upsilon)$ would be $(\exists y)(Hy \& Gy)$. There is nothing to stop us taking this formula to be B as well, and hence — since it entails itself! — we have that $A(\upsilon) \vdash B$. Moreover B has no *free* occurrences of 'y', there are no auxiliary assumptions, and $(\exists x)(\exists y)(Hx \,\&\, Gy)$ has no free occurrences of 'y' either. So all three provisos so far imposed are met. Hence, if they sufficed for the validity of an ∃E step, we could infer

$$(\exists x)(\exists y)(Hx \,\&\, Gy) \vdash (\exists y)(Hy \,\&\, Gy),$$

a patently invalid sequent since its premise, unlike its conclusion, carries no implication that *some one thing* is both F and G. Clearly, then, we need to stipulate that,

> *Fourth*, the selection for 'υ' must be able to occur free in $A(\upsilon)$; that is, that 'υ' must be free for 'ξ' in $A(\xi)$.

These four stipulations *do* collectively ensure that the derivation of B from $A(\upsilon)$ has the generality sufficient to ensure that B is indeed entailed by $(\exists\xi)A(\xi)$. Look over them carefully.

So long as all four provisos are met, then, ∃E lets us pass from a sequent of the form

$$A(\upsilon) \vdash B$$

to one of the form

$$(\exists\xi)A(\xi) \vdash B.$$

But that pattern doesn't capture the rule in full generality. For one thing, $(\exists\xi)A(\xi)$ may feature in the proof in which we happen to be working not as an assumption, but as the conclusion of a sequent, $X \vdash (\exists\xi)A(\xi)$, so that what the relevant ∃E step shows, if valid, is that B is a consequence of the assumptions pooled in X. For another, as we have already noted, auxiliary assumptions may be involved in the derivation of B from $A(\upsilon)$ and we may not be able to establish $A(\upsilon) \vdash B$ simpliciter. Writing 'Y' for a pool of assumptions, including $A(\upsilon)$, for which we can show $Y \vdash B$, and recalling that '$Y \backslash A(\upsilon)$' denotes the result of removing all self-standing occurrences of the formula $A(\upsilon)$ from the string Y, a fully general schematisation of the rule ∃E is thus:

$$\frac{X \vdash (\exists\xi)A(\xi) \qquad Y \vdash B}{X, Y \backslash A(\upsilon) \vdash B} \quad \exists E$$

provided no wff among $(\exists\xi)A(\xi)$, B and $Y\backslash A(\upsilon)$ contains 'υ' free, and 'υ' is free for 'ξ' in $A(\xi)$.

N.B. any variable, 'υ', which is free for 'ξ', may be used as eigenvariable, but it must replace all occurrences of 'ξ' in '$A(\xi)$'. The formulae $Y\backslash A(\upsilon)$ are collectively called the **parametric formulae**.

Let's review how the rule works in practice.

(e) $(\exists x)Fx \vdash (\exists x)(Fx \vee Gx)$

1	(1)	$(\exists x)Fx$	Asmp
2	(2)	Fx	Asmp
2	(3)	$Fx \vee Gx$	2 $\vee$I
2	(4)	$(\exists x)(Fx \vee Gx)$	3 $\exists$I
1	(5)	$(\exists x)(Fx \vee Gx)$	1,2,4 $\exists$E

Points to note:

First: this simple proof illustrates the basic $\exists$E tactic — if you're trying to prove a sequent from existential premises, assume the appropriate special case (what we will call a 'typical disjunct' — note the similarity to $\vee$E) and derive what you want from that (line 4). Then $\exists$E will entitle you, so long as the restrictions have been observed, to refer your conclusion back to the original existential premise.

Second: that in this proof X is $(\exists x)Fx$ itself; Y is empty.

Third: in $\exists$E steps we cite three lines: the line where the original wff appears; the line where the typical disjunct is assumed; and the line where the desired conclusion is derived from the typical disjunct.

Finally (something which is necessary if the $\exists$E step is to be legitimate): the conclusion of the sequent proved at line 4 does not contain 'x' free.

(f) $(\exists x)\neg Gx, (\forall x)(Fx \rightarrow Gx) \vdash (\exists x)\neg Fx$

1	(1)	$(\exists x)\neg Gx$	Asmp
2	(2)	$(\forall x)(Fx \rightarrow Gx)$	Asmp

Strategy: we're going for an existential conclusion, so our basic strategy is to prove a special case of it from our given assumptions, and then use ∃I. One of our assumptions is existential, however, so we'll need to get the conclusion from a typical disjunct of it, and then use ∃E. Let's use 'y' instead of 'x' as eigenvariable, just for variety (we could use 'x' if we wished — or even 'z'):

3	(3)	$\neg Gy$	Asmp
2	(4)	$Fy \rightarrow Gy$	2 ∀E
2,3	(5)	$\neg Fy$	3,4 SI, MTT, Fy/P, Gy/Q
2,3	(6)	$(\exists x)\neg Fx$	5 ∃I

The conclusion of line 6 does not contain 'y' free, and neither does the auxiliary assumption 2. So we can finish off:

1,2	(7)	$(\exists x)\neg Fx$	1,3,6 ∃E

Don't forget: the assumption of the typical disjunct disappears in an ∃E step, and is replaced by the assumptions on which the original existential wff depends. But any auxiliary assumptions used in getting the desired conclusion from the typical disjunct remain. Also note that the proof was not finished at line 6. In basic symbolism, line 6 reads

$$(\forall x)(Fx \rightarrow Gx), \neg Gy \vdash (\exists x)\neg Fx$$

but what we set out to prove was

$$(\exists x)\neg Gx, (\forall x)(Fx \rightarrow Gx) \vdash (\exists x)\neg Fx$$

i.e. line 7. Another example:

(g) $(\forall x)((Fx \vee Gx) \rightarrow Hx), (\exists x)\neg Hx \vdash (\exists x)\neg Fx$

1	(1)	$(\forall x)((Fx \vee Gx) \rightarrow Hx)$	Asmp
2	(2)	$(\exists x)\neg Hx$	Asmp

3	(3)	$\neg Hx$	Asmp
1	(4)	$(Fx \vee Gx) \rightarrow Hx$	1 $\forall$E
1,3	(5)	$\neg(Fx \vee Gx)$	3,4 SI, MTT, $Fx \vee Gx/P$, Hx/Q
1,3	(6)	$\neg Fx \,\&\, \neg Gx$	5 SI, DEM, Fx/P, Gx/Q
1,3	(7)	$\neg Fx$	6 &E
1,3	(8)	$(\exists x)\neg Fx$	7 $\exists$I
1,2	(9)	$(\exists x)\neg Fx$	2,3,8 $\exists$E

Study this proof carefully. The strategy is obvious enough, but the reason for looking closely is that it is typical of so many proofs with mixed quantified premises and quantified conclusions. First, we 'get rid of' the quantifiers in the premises, by $\forall$E and assuming typical disjuncts. Then we use propositional logic to get us to a position where we can use the quantifier introduction rules, and $\exists$E if appropriate, to derive what we want.

How does this proof look when recast as a tactical derivation? The tactic for $\exists$I is straightforward, operating on the right-hand side of the goal:

to show	$F \Rightarrow (\exists\xi)(\ldots\xi\ldots)$,
attempt to show	$F \Rightarrow (\tau)$,
	where 'τ' is free for ξ in $(\ldots\xi\ldots)$.

The $\exists$E-tactic is more subtle, as was the rule. We have to eliminate an existential wff from the left-hand side, from our fact list. We do so by adding a typical disjunct as an assumption:

to show	$F \Rightarrow C$,
check that	$(\exists\xi)(\ldots\xi\ldots) \in F$, for some wff $(\ldots\xi\ldots)$,
and attempt to show	$(\ldots\upsilon\ldots), F \Rightarrow C$,
	where 'υ' is not free in any wff in F, nor in C,
	and 'υ' is free for 'ξ' in $(\ldots\xi\ldots)$.

Again, we see the restriction on the eigenvariable 'υ', that it not be free in the parametric assumptions or in the existential wff (i.e. that it nowhere occur free in the fact list, F) nor in the goal wff C.

We can then set out the tactical derivation for (g) as follows:

? $(\forall x)((Fx \vee Gx) \to Hx), (\exists x)\neg Hx \Rightarrow (\exists x)\neg Fx$
| Using tactic for ∃E (choosing 'x', not free in above wffs, as the eigenvariable)
| ? $\neg Hx, (\forall x)((Fx \vee Gx) \to Hx) \Rightarrow (\exists x)\neg Fx$
| | Using tactic for ∃I (choosing 'x', since we chose it as the eigenvariable for ∃E)
| | ? $\neg Hx, (\forall x)((Fx \vee Gx) \to Hx) \Rightarrow \neg Fx$
| | | Using tactic for ∀E (on 'x')
| | | ? $(Fx \vee Gx) \to Hx, \neg Hx, (\forall x)((Fx \vee Gx) \to Hx) \Rightarrow \neg Fx$
| | | | Let's now thin out $(\forall x)((Fx \vee Gx) \to Hx)$, which is no longer needed
| | | | ? $(Fx \vee Gx) \to Hx, \neg Hx \Rightarrow \neg Fx$
| | | | | Using tactic for SI on MTT $(A \to B, \neg B \vdash \neg A)$
| | | | | ? $(Fx \vee Gx) \to Hx, \neg Hx \Rightarrow (Fx \vee Gx) \to Hx$ ■
| | | | | ? $(Fx \vee Gx) \to Hx, \neg Hx \Rightarrow \neg Hx$ ■
| | | | | ? $\neg(Fx \vee Gx), (Fx \vee Gx) \to Hx, \neg Hx \Rightarrow \neg Fx$
| | | | | | Using tactic for SI on DEM $(\neg(P \vee Q) \vdash \neg P \& \neg Q)$
| | | | | | ? $\neg(Fx \vee Gx), (Fx \vee Gx) \to Hx, \neg Hx \Rightarrow \neg(Fx \vee Gx)$ ■
| | | | | | ? $\neg Fx \& \neg Gx, \neg(Fx \vee Gx), (Fx \vee Gx) \to Hx, \neg Hx \Rightarrow \neg Fx$
| | | | | | | Using tactic for &E
| | | | | | | ? $\neg Fx, \neg Gx, \neg(Fx \vee Gx), (Fx \vee Gx) \to Hx, \neg Hx \Rightarrow \neg Fx$ ■

We could, of course, have used SI on the tautologous sequent, $P \vee Q \to R, \neg R \vdash \neg P$, and reduced these three propositional logic steps to one. Can you see how the proof would shorten to 7 lines?

Next let's use our rules to validate a couple of Aristotelian syllogisms. First: all Fs are G; some Fs are H; therefore some Gs are H:

(h) $(\forall x)(Fx \to Gx), (\exists x)(Fx \& Hx) \vdash (\exists x)(Gx \& Hx)$

Stage 1 — quantifier elimination

1	(1)	$(\forall x)(Fx \to Gx)$	Asmp
2	(2)	$(\exists x)(Fx \& Hx)$	Asmp
1	(3)	$Fx \to Gx$	1 ∀E

4	(4)	$Fx \& Hx$	Asmp (typical disjunct)

Stage 2 — propositional logic

4	(5)	Fx	4 &E
4	(6)	Hx	4 &E
1,4	(7)	Gx	5,3 $\rightarrow$E
1,4	(8)	$Gx \& Hx$	6,7 &I

Stage 3 — quantifier reintroduction

1,4	(9)	$(\exists x)(Gx \& Hx)$	8 $\exists$I
1,2	(10)	$(\exists x)(Gx \& Hx)$	2,4,9 $\exists$E

Next: no G's are H; some F's are G; therefore some F's are not H.

(i) $(\forall x)(Gx \rightarrow \neg Hx), (\exists x)(Fx \& Gx) \vdash (\exists x)(Fx \& \neg Hx)$

Stage 1 — quantifier elimination

1	(1)	$(\forall x)(Gx \rightarrow \neg Hx)$	Asmp
2	(2)	$(\exists x)(Fx \& Gx)$	Asmp
1	(3)	$Gx \rightarrow \neg Hx$	1 $\forall$E
4	(4)	$Fx \& Gx$	Asmp (typical disjunct)

Stage 2 — propositional logic

4	(5)	Fx	4 &E
4	(6)	Gx	4 &E

1,4	(7)	$\neg Hx$	6,3 $\rightarrow$E
1,4	(8)	$Fx \& \neg Hx$	5,7 &I

Stage 3 — quantifier reintroduction

1,4	(9)	$(\exists x)(Fx \& \neg Hx)$	8 $\exists$I
1,2	(10)	$(\exists x)(Fx \& \neg Hx)$	2,4,9 $\exists$E

Now let's see if we can go one better than Aristotle and validate the form of the earlier examples about gills and pencils, etc. The sequent we have to prove is:

(j) $(\forall x)(Fx \rightarrow (\exists y)(Gy \& Hxy)), (\forall x)(Gx \rightarrow Wx) \vdash (\forall x)(Fx \rightarrow (\exists y)(Wy \& Hxy))$

1	(1)	$(\forall x)(Fx \rightarrow (\exists y)(Gy \& Hxy))$	Asmp
2	(2)	$(\forall x)(Gx \rightarrow Wx)$	Asmp

Right. Strategy: we are trying to prove a universal conclusion, so the basic ploy is to prove a special case of it and then use $\forall$I . An appropriate special case is: $Fx \rightarrow (\exists y)(Wy \& Hxy)$. This is a conditional, so we'll assume its antecedent:

3	(2)	Fx	Asmp

and now aim for its consequent. This is an existential wff, so our tactic will be to go for a special case, $Wy \& Hxy$, and then use $\exists$I. (Notice that we had to use a new variable: $Wx \& Hxx$ would have an unwanted, and in the context of the gills and pencil examples, comical reflexivity.) First, then, we carry through the quantifier elimination stage (mixed with an $\rightarrow$E):

1	(3)	$Fx \rightarrow (\exists y)(Gy \& Hxy)$	1 $\forall$E
1,3	(4)	$(\exists y)(Gy \& Hxy)$	3,4 $\rightarrow$E
6	(5)	$Gy \& Hxy$	Asmp
2	(6)	$Gy \rightarrow Wy$	2 $\forall$E

Line 6 is a typical disjunct for line 5. We are rid of all the quantifiers and can proceed to the propositional logic stage (we could of course use SI on a tautology, here):

6	(7)	Gy	6 &E
2,6	(8)	Wy	7,8 →E
2,6	(9)	Hxy	6 &E
2,6	(10)	$Wy \& Hxy$	9,10 &I

There, then, is the special case; now for the quantifier re-introduction stage. First:

2,6	(11)	$(\exists y)(Wy \& Hxy)$	11 ∃I

Next we come to the ∃E step for which we prepared at lines 6 and 7:

1,2,3	(12)	$(\exists y)(Wy \& Hxy)$	5,6,12 ∃E

This last step is well worth a long, hard look. Is it OK? — after all, the conclusion contains the variable, 'x', free. But that doesn't matter since it does not contain 'y' free; and it was 'y' that was the eigenvariable in 6. If we refer the step back to the ∃E schema, it looks like this:

$$(\forall x)(Fx \to (\exists y)(Gy \& Hxy)),\ Fx \vdash (\exists y)(Gy \& Hxy)$$

i.e. $X \vdash (\exists y)A(y)$

$$(\forall x)(Gx \to Wx),\ Gy \& Hxy \vdash (\exists y)(Wy \& Hxy)$$

i.e. $Y \vdash B$

$$(\forall x)(Fx \to (\exists y)(Gy \& Hxy)),\ (\forall x)(Gx \to Wx),\ Fx \vdash (\exists y)(Wy \& Hxy)$$

i.e. $X,\ Y \backslash A(y) \vdash B$

and none of X, $Y \backslash A(y)$ or B contains 'y' free.

Now we are ready for the →I step:

1,2	(13)	$Fx \rightarrow (\exists y)(Wy \,\&\, Hxy)$	3,13 →I

and finally ∀I :

1,2	(14)	$(\forall x)(Fx \rightarrow (\exists y)(Wy \,\&\, Hxy))$	14 ∀I

— which is all right, because neither 1 nor 2 contains ‘x’ free. Let’s repeat the proof in one go:

1	(1)	$(\forall x)(Fx \rightarrow (\exists y)(Gy \,\&\, Hxy))$	Asmp
2	(2)	$(\forall x)(Gx \rightarrow Wx)$	Asmp
3	(3)	Fx	Asmp
1	(4)	$Fx \rightarrow (\exists y)(Gy \,\&\, Hxy)$	1 ∀E
1,3	(5)	$(\exists y)(Gy \,\&\, Hxy)$	3,4 →E
6	(6)	$Gy \,\&\, Hxy$	Asmp
2	(7)	$Gy \rightarrow Wy$	2 ∀E
6	(8)	Gy	6 &E
2,6	(9)	Wy	7,8 →E
2,6	(10)	Hxy	6 &E
2,6	(11)	$Wy \,\&\, Hxy$	9,10 &I
2,6	(12)	$(\exists y)(Wy \,\&\, Hxy)$	11 ∃I
1,2,3	(13)	$(\exists y)(Wy \,\&\, Hxy)$	5,6,12 ∃E
1,2	(14)	$Fx \rightarrow (\exists y)(Wy \,\&\, Hxy)$	3,13 →I
1,2	(15)	$(\forall x)(Fx \rightarrow (\exists y)(Wy \,\&\, Hxy))$	14 ∀I

Exercises

(1) Prove the following sequents of classical predicate logic (you may use SI on any tautologous sequent of propositional logic you think may be useful):

(a) $(\exists x)(Fx \,\&\, Gx) \vdash (\exists x)Fx$

(b) $(\exists x)(Fx \,\&\, Gx) \vdash (\exists x)(Gx \,\&\, Fx)$

(c) $(\exists x)(Gx \,\&\, \neg Hx), (\forall x)(Gx \rightarrow Fx) \vdash (\exists x)(Fx \,\&\, \neg Hx)$

(d) $(\exists x)(Fx \,\&\, Gx), (\forall x)(Gx \rightarrow Hx) \vdash (\exists x)(Fx \,\&\, Hx)$

(2) Formalise the following arguments and show that they are sound.

(a) Only the brave deserve freedom. Some who deserve freedom are Christians. All Christians are blessed in the sight of God. Therefore some of the brave are blessed in the sight of God.

(b) Violets are purple. Some of the flowers I saw were violets. Therefore some of the flowers I saw were purple.

(c) All sports cars are dangerous. Jones drives a sports car. Therefore Jones drives something dangerous.

12 Further Proofs with Quantifiers

It will be salutary to remind ourselves what happens if we break the restrictions on the parametric variable for $\exists$E. First consider this case:

	1	(1)	$(\exists x)Fx$	Asmp
	2	(2)	$(\exists x)Gx$	Asmp
	3	(3)	Fx	Asmp
	4	(4)	Gx	Asmp
	3,4	(5)	$Fx \,\&\, Gx$	3,4 &I
	3,4	(6)	$(\exists x)(Fx \,\&\, Gx)$	5 $\exists$I
!!	1,4	(7)	$(\exists x)(Fx \,\&\, Gx)$	1,3,6 $\exists$E NO!!
	1,2	(8)	$(\exists x)(Fx \,\&\, Gx)$	2,4,7 $\exists$E

At line 8 we have wound up a 'proof' of the sequent $(\exists x)Fx$, $(\exists x)Gx \vdash (\exists x)(Fx \,\&\, Gx)$. The sequent is patently invalid, since it certainly doesn't follow that something is both F and G from the premises that something is F and that something is G. E.g. let F be 'greater than 20' and G be 'smaller than 10'. 26 is greater than 20 and 6 is smaller than 10, but nothing is both greater than 20 and smaller than 10. But the proof is OK up to line 6. In line 7, however, the $\exists$E step violates the restriction that the parametric formulae $Y \backslash A(\upsilon)$ should not contain the eigenvariable 'x'. Schematically we have:

$$\frac{\text{line 1: } (\exists x)Fx \vdash (\exists x)Fx; \text{ i.e. } X \vdash (\exists x)Fx \qquad \text{line 6: } Gx, Fx \vdash (\exists x)(Fx \,\&\, Gx) \text{ i.e. } Y \vdash B}{\text{line 7: } (\exists x)Fx, Gx \vdash (\exists x)(Fx \,\&\, Gx) \text{ i.e. } X, Y \backslash Fx \vdash B}$$

Notice that once we let this spurious move through, the next $\exists$E step, at line 8, is perfectly in order, since 1 does not contain 'x' free.

Now consider a case where we violate the restriction that the conclusion, B, should not contain the eigenvariable in question:

	1	(1)	$(\exists x)Fx$	Asmp
	2	(2)	Fx	Asmp
	2	(3)	$Fx \vee Gx$	2 $\vee$I
!!	1	(4)	$Fx \vee Gx$	1,2,3 $\exists$E NO!!
	1	(5)	$(\forall x)(Fx \vee Gx)$	4 $\forall$I

Here all is well up to line 3; but the $\exists$E step at 4 involves a conclusion containing the forbidden variable. If we let it go, we have no way of blocking the $\forall$I step at 5 (since 1, the assumption of line 4, does not contain 'x' free). The result is the obviously invalid sequent, $(\exists x)Fx \vdash (\forall x)(Fx \vee Gx)$. E.g., let '$Fx$' mean '$x$ is human' and 'Gx' mean 'x is a cat'.

Next, consider what can happen if the existential statement, $(\exists \xi)A(\xi)$, contains the parametric variable:

	1	(1)	$Fx \,\&\, Gy$	Asmp
	1	(2)	$(\exists x)(Fx \,\&\, Gy)$	1 $\exists$I
	3	(3)	$Fy \,\&\, Gy$	Asmp
	3	(4)	$(\exists x)(Fx \,\&\, Gx)$	3 $\exists$I
!!	1	(5)	$(\exists x)(Fx \,\&\, Gx)$	2,3,4 $\exists$E NO !!

This time we wind up with the invalid: $Fx \,\&\, Gy \vdash (\exists x)(Fx \,\&\, Gx)$. Let '$Fx$' mean '$x>20$', '$Gx$' mean '$x<10$' and let $x=26$ and $y=6$. The point is that $A(\upsilon)$ must result from $A(\xi)$ by replacing all occurrences of 'ξ' by 'υ'.

There was nothing formally, wrong, as such, with assuming ‘Fy & Gy’ at line 3 — one may assume what one likes when one likes. But heuristically, tactically, it was wrong, since it is not suitable as a typical disjunct for line 2, which contains ‘y’ free. Hence the actual formal breaking of the restriction on the rule occurs at line 5, where 3 is used as a typical disjunct. But the error, essentially, came at line 3.

Finally, consider what can go wrong if we break the restriction that in the typical disjunct, the eigenvariable ‘υ’ must be free for ‘ξ’ in $A(\xi)$. Consider the following “proof”:

	1	(1)	$(\exists x)(\forall y)Rxy$	Asmp
	2	(2)	$(\forall y)Ryy$	Asmp
!!	1	(3)	$(\forall y)Ryy$	1,2,2 $\exists$E NO!!

$(\forall y)Ryy$ does not validly follow from $(\exists x)(\forall y)Ryy$. E.g., let Rxy mean ‘x times $y=0$’. Then there is a number (namely, 0) whose product with every number is 0. But it is false that the square of every number is 0! Nonetheless, the eigenvariable ‘y’ is not free in the existential wff, (1), there are no parametric wffs, and ‘y’ is not free in the conclusion $(\forall y)Ryy$. The problem is, of course, that ‘y’ is not free in the typical disjunct either, and it should be. ‘y’ should be free for ‘x’ in $(\forall y)Rxy$, and it is not. Hence the “proof” is incorrect.

OK. Now work through the following interderivability results, noting the points of strategy which they illustrate:

(i) $(\exists x)Fx \rightarrow P \dashv\vdash (\forall x)(Fx \rightarrow P)$

First, left-to-right. Strategy: universal conclusion, so use $\forall$I; hence, subsidiary target is $Fx \rightarrow P$, a conditional; so assume its antecedent and look to $\rightarrow$I.

1	(1)	$(\exists x)Fx \rightarrow P$	Asmp
2	(2)	Fx	Asmp
2	(3)	$(\exists x)Fx$	2 $\exists$I
1,2	(4)	P	1,3 $\rightarrow$E
1	(5)	$Fx \rightarrow P$	2,4 $\rightarrow$I

1	(6)	$(\forall x)(Fx \rightarrow P)$	5 $\forall$I
		(NB: 1 does not contain 'x' free)	

Now, right-to-left. Strategy: conditional conclusion; hence assume its antecedent and look to $\rightarrow$I. The antecedent is existential; hence use $\exists$E:

1	(1)	$(\forall x)(Fx \rightarrow P)$	Asmp
2	(2)	$(\exists x)Fx$	Asmp
1	(3)	$Fx \rightarrow P$	1 $\forall$E
4	(4)	Fx	Asmp
1,4	(5)	P	3,4 $\rightarrow$E
1,2	(6)	P	2,4,5 $\exists$E
		(NB: 1, 2, and P do not contain 'x' free)	
1	(7)	$(\exists x)Fx \rightarrow P$	2,6 $\rightarrow$I

Next:

(ii) $(\forall x)Fx \vdash \neg(\exists x)\neg Fx$

Left-to-right. Strategy: negative conclusion, so assume the opposite and look to $\neg$I. The opposite is existential, so use $\exists$E.

1	(1)	$(\forall x)Fx$	Asmp
2	(2)	$(\exists x)\neg Fx$	Asmp
3	(3)	$\neg Fx$	Asmp
1	(4)	Fx	1 $\forall$E
1,3	(5)	$\bot$	3,4 $\neg$E
1,2	(6)	$\bot$	2,3 5 $\exists$E
1	(7)	$\neg(\exists x)\neg Fx$	2,6 $\neg$I

Now, right-to-left. Strategy: universal conclusion, so aim to prove Fx, then use $\forall$I.

1	(1)	$\neg(\exists x)\neg Fx$	Asmp
2	(2)	$\neg Fx$	Asmp (aiming at $\neg$I)
2	(3)	$(\exists x)\neg Fx$	2 $\exists$I
1,2	(4)	$\bot$	1,3 $\neg$E
1	(5)	$\neg\neg Fx$	2,4 $\neg$I
1	(6)	Fx	5 DN
1	(7)	$(\forall x)Fx$	6 $\forall$I

(NB: 1 does not contain 'x' free)

This result gives us in effect a definition of the universal quantifier in terms of the existential quantifier and negation. Can we also do the reverse: define the existential in terms of the universal? The following result shows that we can:

(iii) $(\exists x)Fx \vdash \neg(\forall x)\neg Fx$

Left-to-right. Strategy: negative conclusion, so assume the opposite and look to $\neg$I. Existential premise, so overall method will be $\exists$E.

1	(1)	$(\exists x)Fx$	Asmp
2	(2)	$(\forall x)\neg Fx$	Asmp
3	(3)	Fx	Asmp
2	(4)	$\neg Fx$	2 $\forall$E
2,3	(5)	$\bot$	3,4 $\neg$E
3	(6)	$\neg(\forall x)\neg Fx$	2,5 $\neg$I
1	(7)	$\neg(\forall x)\neg Fx$	1,3,6 $\exists$E

(The $\neg$I and $\exists$E steps could be reversed here.) Now right-to-left. Strategy: this is a bit more devious. The obvious ploy would be to try to prove Fx and then use $\exists$I — but this isn't going to work because Fx doesn't follow from $\neg(\forall x)\neg Fx$: it is quite consistent with its being untrue that everything is not F that x, anyway, should not be F. (Think about it!) So let's try $\neg$I.

1	(1)	$\neg(\forall x)\neg Fx$	Asmp
2	(2)	$\neg(\exists x)Fx$	Asmp

What next? Well, we want to show that something inconsistent with 1 follows from 2; so let's try to derive $(\forall x)\neg Fx$ from 2. That suggests that we derive $\neg Fx$ from 2 and then use $\forall$I ; so, again, let's assume the opposite and look to $\neg$I:

3	(3)	Fx	Asmp

Now it's fairly obvious how to proceed:

3	(4)	$(\exists x)Fx$	3 $\exists$I
2,3	(5)	$\perp$	2,4 $\neg$E
2	(6)	$\neg Fx$	3,5 $\neg$I
2	(7)	$(\forall x)\neg Fx$	1 $\forall$I
		(NB: 2 does not contain 'x' free)	
1,2	(8)	$\perp$ 1,7	$\neg$E
1	(9)	$\neg\neg(\exists x)Fx$	2,8 $\neg$I
1	(10)	$(\exists x)Fx$	9 DN

One more example:

(iv) $(\exists x)(P \,\&\, Fx) \vdash P \,\&\, (\exists x)Fx$

First, left-to-right. Strategy: conjunctive conclusion, so we'll try to prove each conjunct separately; existential premise, so each conjunct will be arrived at by $\exists$E:

1	(1)	$(\exists x)(P \,\&\, Fx)$	Asmp
2	(2)	$P \,\&\, Fx$	Asmp
2	(3)	P	2 &E
1	(4)	P	1,2,3 $\exists$E

2	(5)	Fx	2 &E
2	(6)	(∃x)Fx	5 ∃I
1	(7)	(∃x)Fx	1,2,6 ∃E
1	(8)	P & (∃x)Fx	4,7 &I

(We could economise on ∃Es here, and put the &I inside a single ∃E — can you see how?) Now right-to-left. Strategy: existential conclusion suggests we prove P & Fx and then use ∃I — but it won't be quite that simple because of the ∃E restrictions. Let's see how it develops:

1	(1)	P & (∃x)Fx	Asmp
1	(2)	P	1 &E
1	(3)	(∃x)Fx	1 &E
4	(4)	Fx	Asmp
1,4	(5)	P & Fx	2,4 &I

Now: we can't just advance to

!! 1	(6)	P & Fx	3,4,5 ∃E NO!!

for the conclusion of line 5 contains 'x' free. Hence we must first carry out the ∃I step, in order to comply with the ∃E restrictions:

1,4	(6)	(∃x)(P & Fx)	5 ∃I

Now we have something without 'X' free, so there is no objection to:

1	(7)	(∃x)(P & Fx)	3,4,6 ∃E

Learn the lesson of this last proof: often, when the ∃E restrictions seem to get in the way of a straightforward proof, we can find a way of meeting them by juggling with the order of the steps. Of course, this is only possible when the sequent is valid (cf. the examples at the beginning of this chapter).

To close, let's have a shot at formalising and validating the following pair of intuitively sound arguments:

(1) Alan owns a bicycle. All bicycles are cheaper to run than cars. Bert owns a car. Therefore Alan owns something which is cheaper to run than something which Bert owns.

(2) Everyone is loved by someone at sometime. Therefore, if Alan has not yet been loved by anyone, someone will love him in the future.

First (a). Since 'bicycles' occurs independently in the second premise, 'owns a bicycle' in the first will have to be analysed relationally. So let's say:

$a \rhd$ Alan

$Oxy \rhd x$ owns y

$Bx \rhd x$ is a bicycle.

Clearly we also need:

$b \rhd$ Bert

$Cx \rhd x$ is a car

$Rxy \rhd x$ is cheaper to run than .

What results is then:

$(\exists x)(Bx \,\&\, Oax)$: Alan owns a bicycle

$(\forall x)(\forall y)(Bx \,\&\, Cy \rightarrow Rxy)$: All bicycles are cheaper to run than cars

$(\exists x)(Cx \,\&\, Obx)$: Bert owns a car

Conclusion? Well, what it says, in effect, is that Alan owns something and Bert owns something and the first is cheaper than the second. So this will do:

$$(\exists x)(\exists y)((Oax \,\&\, Oby) \,\&\, Rxy).$$

So, can we validate it?

1	(1)	$(\exists x)(Bx \,\&\, Oax)$	Asmp
2	(2)	$(\forall x)(\forall y)(Bx \,\&\, Cy \rightarrow Rxy)$	Asmp
3	(3)	$(\exists x)(Cx \,\&\, Obx)$	Asmp

Strategy: a doubly existential conclusion suggests we go for $(Oax \& Oby) \& Rxy$ as subsidiary target, and then perform two steps of ∃I. But we're working with existential premises, so the overall strategy should be ∃E. Let's get through the quantifier elimination stage, and then see how things look:

4	(4)	$Bx \& O\alpha x$	Asmp typical disjunct of 1
5	(5)	$Cy \& Oby$	Asmp typical disjunct of 3

Remember, the parametric variable for the typical disjunct of 3 must not occur free in the parametric formulae, in particular, $Bx \& O\alpha x$, so 'x' would not work; so we choose 'y'. In other words, we use different parametric variables, 'x' for Alan's bicycle and 'y' for Bert's car — this is not just commonsense, but is dictated by the form of our subsidiary target.

2	(6)	$(\forall y)(Bx \& Cy \rightarrow Rxy)$	2 ∀E
2	(7)	$Bx \& Cy \rightarrow Rxy$	6 ∀E

Note that we must eliminate the quantifiers in line 2 one at a time. OK; now for the propositional logic stage:

4	(8)	Bx	4 &E
5	(9)	Cy	5 &E
4,5	(10)	$Bx \& Cy$	8,9 &I
2,4,5	(11)	Rxy	7,10 →E
4	(12)	$O\alpha x$	4 &E
5	(13)	Oby	5 &E
4,5	(14)	$O\alpha x \& Oby$	12,13 &I
2,4,5	(15)	$(O\alpha x \& Oby) \& Rxy$	14,11 &I

(We could use SI on an appropriate tautologous sequent, viz.

$$P \& Q,\ R \& S,\ P \& R \rightarrow T \vdash (Q \& S) \& T).$$

Right. There's the subsidiary target, so all that remains is to introduce the quantifiers and then use ∃E to refer the result back to our original assumptions:

2,4,5	(16)	$(\exists y)(Oax \,\&\, Oby) \,\&\, Rxy)$	15 ∃I
2,4,5	(17)	$(\exists x)(\exists y)((Oax \,\&\, Oby) \,\&\, Rxy)$	16 ∃I

Note that, we introduce the quantifiers one at a time.

2,4,3	(18)	$(\exists x)(\exists y)((Oax \,\&\, Oby) \,\&\, Rxy)$	3,5,17 ∃E
1,2,3	(19)	$(\exists x)(\exists y)((Oax \,\&\, Oby) \,\&\, Rxy$	1,4,18 ∃E

(*Two* existential premises require *two* uses of ∃E.) Q.E.D.!

Now for the second example. Here the crucial insight is that we need to quantify over times and to have predicates, 'past', 'future', and such like, of times, in order to capture the 'not yet' and 'in the future' in the conclusion. Let's have:

$Lxyz \rhd$ x loves y at z
$a \rhd$ Alan
$Fz \rhd$ z is future.

The premise is now straightforward: it says that for every x there is a y who will love x at some (time) z, i.e.

$$(\forall x)(\exists y)(\exists z)Lyxz.$$

What about the antecedent of the conclusion, 'Alan has not yet been loved by anyone'? Well, to say, that Alan has not yet done so-and-so is to say that for all times, z, if z is not future, then Alan is (and was) not doing so-and-so at z. So the antecedent is:

$$(\forall z)(\neg Fz \rightarrow (\forall y)\neg Lyaz)$$

And the consequent is: someone will love Alan in the future, i.e.

$$(\exists y)(\exists z)(Fz \,\&\, Lyaz).$$

The sequent to be proved is thus:

$$(\forall x)(\exists y)(\exists z)Lyxz \vdash (\forall z)(\neg Fz \rightarrow (\forall y)\neg Lyaz) \rightarrow (\exists y)(\exists z)(Fz \,\&\, Lyaz)$$

OK. Let's validate the argument:

1	(1)	$(\forall x)(\exists y)(\exists z)Lyxz$	Asmp

Strategy: we are going for a conditional conclusion, so we'll assume its antecedent:

2	(2)	$(\forall z)(\neg Fz \rightarrow (\forall y)\neg Lyaz)$	Asmp

and seek to prove the consequent. This is a double existential, so we'll go for the subsidiary target, $Fz \,\&\, Lyaz$, and then use a couple of steps of ∃I. First, however, we must eliminate the quantifiers from 1 and 2:

1	(3)	$(\exists y)(\exists z)Lyaz$	1 ∀E

Note that, for what ought to be intuitively obvious reasons, we've eliminated 'X' for Alan, so to speak, rather than for a variable.

4	(4)	$(\exists z)Lyaz$	Asmp typical disjunct of 3

Note that 'y' is (trivially) free for 'y' in $(\exists z)Lyaz$, though 'z' would not be free for it there. So to assume $(\exists z)Lzaz$ would be wrong, and would prevent use of ∃E at line 17 below.

5	(5)	$Lyaz$	Asmp typical disjunct of 4
6	(6)	$\neg Fz$	Asmp
2	(7)	$\neg Fz \rightarrow (\forall y)\neg Lyaz$	2 ∀E
2,6	(8)	$(\forall y)\neg Lyaz$	7,6 →E
2,6	(9)	$\neg Lyaz$	8 ∀E

That's all the quantifiers gone — now we exploit the contradiction at lines 5 and 9:

2,5,7	(10)	$\bot$	5,9 $\neg$E
2,5	(11)	$\neg\neg Fz$	7,10 $\neg$I
2,5	(12)	Fz	11 DN
2,5	(13)	$Fz \,\&\, Lyaz$	12,5 &I

There is the subsidiary target; so now we put the quantifiers back in:

2,5	(14)	$(\exists z)(Fz \,\&\, Laaz)$	13 $\exists$I
2,5	(15)	$(\exists y)(\exists z)(Fz \,\&\, Lyaz)$	14 $\exists$I

So there's the conclusion; it remains to use $\exists$E to refer it back to the given premise. Once again, this will have to go in two stages:

2,4	(16)	$(\exists y)(\exists z)(Fz \,\&\, Lyaz)$	4,5,15 $\exists$E
1,2	(17)	$(\exists y)(\exists z)(Fz \,\&\, Lyaz)$	3,4,16 $\exists$E

Lastly for the $\to$I step:

1	(18)	$(\forall z)(\neg Fz \to (\forall y)\neg Lyaz)$ $\to (\exists y)(\exists z)(Fz \,\&\, Lyaz)$	2,17 $\to$I

Q.E.D.

Here is the complete proof:

1	(1)	$(\forall x)(\exists y)(\exists z)Lyxz$	Asmp
2	(2)	$(\forall z)(\neg Fz \rightarrow (\forall y)\neg Lyaz)$	Asmp
1	(3)	$(\exists y)(\exists z)Lyaz$	1 ∀E
4	(4)	$(\exists z)Lyaz$	Asmp
5	(5)	$Lyaz$	Asmp
6	(6)	$\neg Fz$	Asmp
2	(7)	$\neg Fz \rightarrow (\forall y)\neg Lyaz$	2 ∀E
2,6	(8)	$(\forall y)\neg Lyaz$	7,6 →E
2,6	(9)	$\neg Lyaz$	8 ∀E
2,5,6	(10)	$\bot$	9,5 ¬E
2,5	(11)	$\neg\neg Fz$	6,10 ¬I
2,5	(12)	Fz	11 DN
2,5	(13)	$Fz \& Lyaz$	12,5 &I
2,5	(14)	$(\exists z)(Fz \& Lyaz)$	13 ∃I
2,5	(15)	$(\exists y)(\exists z)(Fz \& Lyaz)$	14 ∃I
2,4	(16)	$(\exists y)(\exists z)(Fz \& Lyaz)$	4,5,15 ∃E
1,2	(17)	$(\exists y)(\exists z)(Fz \& Lyaz)$	3,4,16 ∃E
1	(18)	$(\forall z)(\neg Fz \rightarrow (\forall y)\neg Lyaz) \rightarrow (\exists y)(\exists z)(Fz \& Lyaz)$	2,17 →I

Here is how the proof looks tactically:

? $(\forall x)(\exists y)(\exists z)Lyxz \Rightarrow (\forall z)(\neg Fz \to (\forall y)\neg Lyaz) \to (\exists y)(\exists z)(Fz \,\&\, Lyaz)$
> Using tactic for $\to$I
> ? $(\forall z)(\neg Fz \to (\forall y)\neg Lyaz),\ (\forall x)(\exists y)(\exists z)Lyxz \Rightarrow (\exists y)(\exists z)(Fz \,\&\, Lyaz)$
>> Using tactic for $\forall$E
>> ? $(\exists y)(\exists z)Lyaz,\ (\forall z)(\neg Fz \to (\forall y)\neg Lyaz),\ (\forall x)(\exists y)(\exists z)Lyxz \Rightarrow (\exists y)(\exists z)(Fz \,\&\, Lyaz)$
>>> Using tactic for $\exists$E
>>> ? $(\exists z)Lyaz,\ (\forall z)(\neg Fz \to (\forall y)\neg Lyaz),\ (\forall x)(\exists y)(\exists z)Lyxz \Rightarrow (\exists y)(\exists z)(Fz \,\&\, Lyaz)$
>>>> Using tactic for $\exists$E
>>>> ? $Lyaz,\ (\forall z)(\neg Fz \to (\forall y)\neg Lyaz),\ (\forall x)(\exists y)(\exists z)Lyxz \Rightarrow (\exists y)(\exists z)(Fz \,\&\, Lyaz)$
>>>>> Using tactic for Thin
>>>>> ? $Lyaz,\ (\forall z)(\neg Fz \to (\forall y)\neg Lyaz) \Rightarrow (\exists y)(\exists z)(Fz \,\&\, Lyaz)$
>>>>>> Using tactic for $\exists$I
>>>>>> ? $Lyaz,\ (\forall z)(\neg Fz \to (\forall y)\neg Lyaz) \Rightarrow (\exists z)(Fz \,\&\, Lyaz)$
>>>>>>> Using tactic for $\exists$I
>>>>>>> ? $Lyaz,\ (\forall z)(\neg Fz \to (\forall y)\neg Lyaz) \Rightarrow Fz \,\&\, Lyaz$
>>>>>>>> Using tactic for &I
>>>>>>>> ? $Lyaz,\ (\forall z)(\neg Fz \to (\forall y)\neg Lyaz) \Rightarrow Fz$
>>>>>>>>> Using tactic for DN
>>>>>>>>> ? $Lyaz,\ (\forall z)(\neg Fz \to (\forall y)\neg Lyaz) \Rightarrow \neg\neg Fz$
>>>>>>>>>> Using tactic for $\neg$I
>>>>>>>>>> ? $\neg Fz,\ Lyaz,\ (\forall z)(\neg Fz \to (\forall y)\neg Lyaz) \Rightarrow \bot$
>>>>>>>>>>> Using tactic for $\forall$E
>>>>>>>>>>> ? $\neg Fz \to (\forall y)\neg Lyaz,\ \neg Fz,\ Lyaz,\ (\forall z)(\neg Fz \to (\forall y)\neg Lyaz) \Rightarrow \bot$
>>>>>>>>>>>> Using tactic for Thin
>>>>>>>>>>>> ? $\neg Fz \to (\forall y)\neg Lyaz,\ \neg Fz,\ Lyaz \Rightarrow \bot$
>>>>>>>>>>>>> Using tactic for $\to$E
>>>>>>>>>>>>> ? $\neg Fz \to (\forall y)\neg Lyaz,\ \neg Fz,\ Lyaz \Rightarrow \neg Fz$ $\blacksquare$
>>>>>>>>>>>>> ? $(\forall y)\neg Lyaz,\ \neg Fz,\ Lyaz \Rightarrow \bot$
>>>>>>>>>>>>>> Using tactic for $\forall$E
>>>>>>>>>>>>>> ? $\neg Lyaz,\ (\forall y)\neg Lyaz,\ \neg Fz,\ Lyaz \Rightarrow \bot$
>>>>>>>>>>>>>>> Using tactic for $\neg$E
>>>>>>>>>>>>>>> ? $\neg Lyaz,\ (\forall y)\neg Lyaz,\ \neg Fz,\ Lyaz \Rightarrow Lyaz$ $\blacksquare$
>>>>>>>>>>>>>>> ? $\bot,\ (\forall y)\neg Lyaz,\ \neg Fz,\ Lyaz \Rightarrow \bot$ $\blacksquare$
>>>>>>>> ? $Lyaz,\ (\forall z)(\neg Fz \to (\forall y)\neg Lyaz) \Rightarrow Lyaz$ $\blacksquare$

Exercises

(1) Prove the following sequents (you may use SI on any tautologous sequent of propositional logic you think may be useful):

(a) $(\forall x)(Fx \rightarrow (\forall y)\neg Fy) \vdash \neg(\exists z)Fz$

(b) $(\exists x)(Fx \rightarrow (\exists y)\neg Fy) \vdash \neg(\forall z)Fz$

(c) $(\exists x)\neg Fx \vdash \neg(\forall x)Fx$

(d) $(\forall x)Fx \rightarrow P \vdash (\exists x)(Fx \rightarrow P)$

(e) $(\forall x)(Sx \rightarrow ((\neg Bx \& \neg Cx) \rightarrow Kx)) \vdash (\forall x)((Sx \& \neg Kx) \rightarrow (Bx \vee Cx))$

(2) Fill in all the missing steps in the proof below:

(1)		
(2)	$(\exists x)Fx$	Asmp
(3)		Asmp
(4)		1 $\forall$E
(5)		$\neg$E
(6)		$\exists$E
(7)		$\neg$I

(3) Prove the following sequents (you may use SI on any tautologous sequent of propositional logic you think may be useful):

(a) $(\exists x)(Fx \& (\forall y)(Gy \rightarrow Rxy)) \vdash (\forall x)(Gx \rightarrow (\exists y)(Fy \& Ryx))$

(b) $(\forall x)((\exists y)Rxy \rightarrow Rxx),\ Rab \vdash Raa$

(c) $(\forall x)(\forall y)(Rxy \rightarrow \neg Ryx) \vdash \neg(\exists x)Rxx$

(d) $\vdash \neg(\exists x)(\forall y)(Rxy \leftrightarrow \neg Ryy)$

(e) $\vdash (\exists x)(Fx \rightarrow (\forall y)Fy)$

(f) $\vdash (\exists x)((Fa \rightarrow Fx) \& (Fb \rightarrow Fx))$

(g) $\vdash (\exists x)((\exists y)Fy \rightarrow Fx)$