# Notes

## PL proofs

- The second edition of my *Introduction to Formal Logic* will have chapters on natural deduction in the main text (and supplementary chapters on truth trees, pretty much as in the first edition, are available online). So here are the draft chapters on propositional natural deduction.
- While a lot of people asked for natural deduction rather than trees, this majority consensus of course hides further dissent about the sort of natural deduction system to present. I'm offering a Fitch-style system. In a number of ways I'd prefer to do things Gentzen-style (looking forward to more formal proof-theoretic work). But I'm swayed, in part, by experience of what students find easier. You might disagree about that: but I hope we can agree that once you have met Fitch-style proofs with their elementary way of handling temporary assumptions and their discharge, it is easy to move to using Gentzen proofs. I'll also be offering a 'conversion course' as an online supplement – or at least, that's the plan!
- There will exercises added to the end of these chapters, and worked answers online. I'm not one of those who fetishizes doing a lot of exercises, teaching proof strategies etc. Basic understanding is what we should be aiming for, say I! Exercises will also add some points of detail, like the equivalence of classical reductio and (DN).
- In keeping with what happens in earlier chapters, primitive biconditionals only feature in exercises.
- In an earlier version, the Fitch-style system I offered had a liberalized ($\vee$E) rule of the kind recommended by Neil Tennant, and also treated the absurdity constant following a minority line that thinks of it as more like an exclamation mark rather than a wff in its own right. But I've been persuaded to revert to a more conservative version.
- I have included a long-winded version of the Table of Contents at the outset, so you can – if you want – see how these chapters follow on from earlier chapters in the overall structure of the book.
- Since these *are* heavily revised chapters, all comments/corrections still most gratefully received!

## General notes to readers

- What follows is an excerpt from a planned second edition of my *Introduction to Formal Logic* (CUP, 2003). This is a textbook aimed at first-year philosophy students: if all goes well, the second edition will appear in 2019 in the *Cambridge Introductions to Philosophy* series. The initial Chapters 1 to 7 introduce some key concepts in a relaxed way, and should be useful background whatever formal logic course you then go on to take. Then, from Chapter 8, we make a start on propositional logic. The Table of Contents will tell you what these chapters cover.
- Work still needs to be done on the exercise sets at the end of the chapters. There will eventually be answers to the exercises on the web.
- At this stage, all kinds of comments (other than ones that mean, in effect, 'You are writing the wrong book!') are most welcome. Obviously I want to hear about any typos you spot. But in addition, it is very useful to hear e.g. about passages that you found obscure or more difficult than others, and passages you thought were possibly misleading. And if you are not a native English-speaker, do note any words or turns of phrase that you found puzzling.
- Spelling follows British English rather than North American English conventions. So I write e.g. 'fulfil' rather than 'fulfill', 'skilful' rather than 'skillful', etc. I aim to systematically use 'ize' endings where appropriate. (But still, if you think what I've written is a typo, do say so – I'd rather you over-corrected than under-corrected!)
- One issue about punctuation. Suppose we have some introductory words followed by some indented displayed material. Should the introductory words end with a colon or not? My general line is this: if the introduction is a complete clause, I use a colon; if the thought runs on straight into the indented material, I don't.
- I use 'they'/'them' as gender-neutral singular pronouns, except when the resulting English strikes me as rebarbative, in which case I paraphrase away. (Let me know of cases that really grate.)
- Unresolved references of the kind '**??**' are forward-looking to chapters or sections that aren't included in the current chapters and will be resolved in due course.
- Comments and corrections please to ps218 at cam dot ac dot uk – if you do send any, it could be *very* helpful if you mention the date of this draft.

An Introduction to

# Formal Logic

*Second edition*

Peter Smith

May 28, 2019

# Contents

# Interlude: Why natural deduction?

(a) Let's review our explorations since the last Interlude:

We explained the 'divide and rule' approach to logic. We sidestep having to deal with the quirks of ordinary language by rendering vernacular arguments into well-behaved artificial languages. We can then investigate the validity or otherwise of the resulting formalized arguments.

We then set off to explore formalized arguments involving (tidied versions of) 'and', 'or' and 'not' as their essential logical materials.

We defined the *syntax* of our PL languages (which initially just have formal versions of these three connectives). We learnt how wffs of such languages are built up (in ways that can be displayed using parse trees). We also defined the notion of scope, the notion of a main connective, and related ideas.

Next, *semantics*. We discussed the interpretation of PL languages: we assign truth-relevant contents (senses) to atoms, and read the connectives as expressing bare conjunction, inclusive disjunction, and negation – understanding these in a way that makes the connectives truth-functional.

Given the truth-functional interpretation of the connectives, it follows that the truth-value of a wff is determined by a valuation of its atoms.

We showed that – using just the three built-in connectives – we can define *any* truth-functional connective (we can construct wffs with any desired truth-table).

This got us to the position where we could take a pivotal step, and define the notions of a tautology and of tautological validity:

A tautology is a wff true on every combinatorially possible valuation of its atoms, while a contradiction is false on every such valuation.

A PL argument is tautologically valid if every combinatorially possible valuation of the relevant atoms which makes the premisses true makes the conclusion true.

We discussed the relation between this notion of validity and the informal notion of logical validity. We argued that the notion of tautological validity gives us a cleanly defined rational reconstruction for the notion of logical validity as applied to inferences relying on truth-functional connectives.

We introduced *one* way of establishing tautological validity, namely by a brute-force truth-table test. We can in this way decide, quite mechanically, which PL inferences are tautologically valid.

Finally – after adding the absurdity constant which will be useful later – we considered the prospects for dealing with arguments involving 'if' in the same way:

> We met the so-called material conditional, which is the only possible conditional-like truth-function, and we added a special symbol for this to the formal apparatus of PL languages.
>
> However, the material conditional is at best a surrogate for the 'if' in just some indicative conditionals; and, even in these cases, there remains a serious issue about how far it captures all the meaning of 'if'.

(b) In Chapters 4 and 5, we talked about *two* informal ways of assessing inferences – namely, showing validity by giving a step-by-step *proof*, and showing invalidity by finding a *counterexample*. We have seen that a truth-table test is in effect a systematic search through all the possibilities for a counterexample to the validity of a PL inference (see §16.2). Our next task is to pick up the idea of step-by-step proofs, and to consider how to construct such proofs inside our formalized PL languages.

Take, for example, the argument

$(\mathsf{P} \land (\mathsf{Q} \land \mathsf{R})),\ (\mathsf{S} \land \mathsf{P}') \therefore\ (\mathsf{P}' \land \mathsf{Q})$.

We can run up a thirty-two line truth table to validate this little argument. But it is surely *much* more natural simply to reason as follows. We note:

(i) Given a conjunction, we can derive each conjunct; and
(ii) given a pair of propositions, we can derive their conjunction.

Then, deploying these two rules of inference for conjunctions, we can use (i) to infer '$(\mathsf{Q} \land \mathsf{R})$' and hence '$\mathsf{Q}$' from the first premiss of our argument above and then also use (i) to infer '$\mathsf{P}'$' from the second premiss. And then we can use (ii) to put these interim consequences together to get the desired conclusion.

We can next introduce rules of inference similarly governing negations, disjunctions and conditionals, and then explain how to put together natural-looking step-by-step PL deductions using these rules. And we can do all this while aiming to follow reasonably closely the modes of argument that we already find in everyday informal reasoning. In a phrase, we can aim to develop a *natural deduction proof system* for propositional logic. As we will see, this is fairly easy to do (the devil is in the details).

So this gives us a second approach to showing the validity of a correct PL inference – namely, derive its conclusion from its premisses by a step-by-step deduction.

Is it overkill to explore a proof system as well as the truth-table approach to propositional logic? Natural deduction proofs for PL inferences *are* well worth knowing about. And a key point is that, unlike the truth-table approach, *natural deduction methods can be extended from propositional logic in a smooth way to validate inferences with quantifiers*. And that's our ultimate target, to be able to deal with formal QL – quantificational logic – inferences turning on 'all', 'some', 'no', etc.

As with many other logical ideas, however, the general conception of a formal proof system is *much* more easily grasped if we meet it first in the tame context of propositional logic. So that is why we will be spending more time on PL arguments over the next few chapters before finally moving on.

# 20 PL proofs: conjunction and negation

Outside the logic classroom, when we want to convince ourselves that a supposedly valid inference really *is* valid, how do we proceed? As we just noted in the Interlude, we often try to find an informal *proof*. In other words, we aim to derive the desired conclusion by a multi-step argument, chaining together obviously truth-preserving inferences starting from the given premisses.

When working in a formal language, we can similarly aim to construct proofs to warrant inferences. So we now start developing a framework for constructing proofs in PL languages. We begin with derivations involving conjunctions and negations.

One important point before we begin. We will be laying out proofs in the vertical column-shifting style which we met informally in Chapter 4. We won't complicate matters now by considering other options (e.g. tree-shaped proofs like $\mathbf{A}^T$ in §9.2, discussed in an Appendix). But we should emphasize that there is no single 'right' way of laying out natural deduction proofs.

## 20.1 Rules for conjunction

(a) Let's start by officially adopting the following rules for arguing to and from wffs which have a bare conjunction '$\wedge$' as their main connective.

$\wedge$-*Introduction*: Given $\alpha$ and $\beta$, we can infer $(\alpha \wedge \beta)$.

$\wedge$-*Elimination*: Given $(\alpha \wedge \beta)$, we can infer $\alpha$. Equally, we can infer $\beta$.

Here, $\alpha$ and $\beta$ are of course arbitrary PL wffs.

Intuitively, we can infer a bare conjunction given both its conjuncts; and we can infer each conjunct from a conjunction. So our two inference rules, as just suggested in the Interlude, are surely faithful to the intended meaning of the connective '$\wedge$'.

Our names for the rules are standard and reasonably self-explanatory. Two particular comments on the first rule:

(i) The inputs $\alpha$ and $\beta$ used when inferring $(\alpha \wedge \beta)$ don't have to be distinct – so we are allowed e.g. to appeal to a premiss 'P' twice over to derive '(P $\wedge$ P)'.

(ii) The order in which $\alpha$ and $\beta$ appear earlier in a proof doesn't matter: we can infer $(\alpha \wedge \beta)$ either way.

With these points understood, we can display our rules diagramatically and give them brief labels like this:

*Rules for conjunction*

$$(\wedge I)\quad \begin{array}{c}\alpha\\ \vdots\\ \beta\\ \vdots\\ (\alpha\wedge\beta)\end{array} \qquad\qquad (\wedge E)\quad \begin{array}{c}(\alpha\wedge\beta)\\ \vdots\\ \alpha\end{array} \qquad \begin{array}{c}(\alpha\wedge\beta)\\ \vdots\\ \beta\end{array}$$

Note how the two rules for '$\wedge$' fit together beautifully. In an obvious sense, the elimination rule can be used to reverse an application of the introduction rule – applying ($\wedge$E) twice after invoking ($\wedge$I) takes us back to where we started.

(b) Let's have a first simple example of these official rules at work. So consider the argument

**A** Popper is a philosopher, and so are Quine and Russell. Sellars is a philosopher and Putnam is one too. Hence both Putnam and Quine are philosophers.

Regimenting this into an appropriate PL language (where 'P' means *Popper is a philosopher*, 'Q' means *Quine is a philosopher*, and so on), **A** can be rendered as

**A′** $(\mathsf{P} \wedge (\mathsf{Q} \wedge \mathsf{R})), (\mathsf{S} \wedge \mathsf{P}') \therefore (\mathsf{P}' \wedge \mathsf{Q})$.

We met this argument a few pages ago in the Interlude. So here again is the truth-preserving derivation of the conclusion from the premisses that we informally sketched there, now presented in what will be our official style:

| | | |
|---|---|---|
| (1) | $(\mathsf{P} \wedge (\mathsf{Q} \wedge \mathsf{R}))$ | (Prem) |
| (2) | $(\mathsf{S} \wedge \mathsf{P}')$ | (Prem) |
| (3) | $(\mathsf{Q} \wedge \mathsf{R})$ | ($\wedge$E 1) |
| (4) | $\mathsf{Q}$ | ($\wedge$E 3) |
| (5) | $\mathsf{P}'$ | ($\wedge$E 2) |
| (6) | $(\mathsf{P}' \wedge \mathsf{Q})$ | ($\wedge$I 5,4) |

This illustrates two basic stylistic choices:

(i) We have used a *vertical line* against which to align our column of reasoning.

(ii) And we have used a *horizontal bar* to draw a line under the premisses at the beginning of our argument.

Both these visual aids are conventional.

As an optional extra, we have also introduced on the right a rather laconic style of annotation, again pretty standard in form. 'Prem' doesn't need explanation. At later lines which result from the application of a rule of inference, we give the label for the relevant rule, and then the line number(s) of its input(s). In the case of an application of ($\wedge$I) resulting in a wff of the form $(\alpha \wedge \beta)$, we give the line numbers of the relevant preceding wffs $\alpha$ and $\beta$ in that order.

Note that we could have derived the wffs at lines (4) and (5) in the opposite order, and still gone on to derive (6). So even the simplest and shortest proof from premisses to conclusion need not be unique.

(c) That was easy! This next example is even easier. The inference

**B** $(\mathsf{P} \land \mathsf{Q}) \therefore (\mathsf{Q} \land \mathsf{P})$

is trivially valid, and has a correspondingly trivial proof:

| | | |
|---|---|---|
| (1) | $(\mathsf{P} \land \mathsf{Q})$ | (Prem) |
| (2) | $\mathsf{P}$ | ($\land$E 1) |
| (3) | $\mathsf{Q}$ | ($\land$E 1) |
| (4) | $(\mathsf{Q} \land \mathsf{P})$ | ($\land$I 3, 2) |

And the pattern of proof here can be generalized. Just replace each '$\mathsf{P}$' by $\alpha$, each '$\mathsf{Q}$' by $\beta$, and we get a derivation from the wff $(\alpha \land \beta)$ – whether as an initial premiss or midway through a longer argument – to the corresponding wff $(\beta \land \alpha)$.

(d) We have so far adopted just two basic rules for arguing with conjunctions, namely ($\land$I) and ($\land$E). But we could, with equal justification, have added other basic rules, for example the rule ($\land$C) reflecting that conjunction 'commutes':

$\land$-*Commutes*: From $(\alpha \land \beta)$, we can infer $(\beta \land \alpha)$.

However, we already know that *this* third rule would be redundant. Any application of ($\land$C) could be replaced by a little three-step dance involving ($\land$I) and ($\land$E) instead.

Other intuitively reliable additional rules for '$\land$' turn out to be similarly redundant. Of course, there would be nothing *wrong* about adding ($\land$C) or other unnecessary but truth-preserving rules to our deduction system. Economy of basic rules is a major virtue for a proof system, but it isn't the only virtue; as we will see later, there can be countervailing reasons for adding a strictly-speaking redundant rule. However, in this case, there aren't any strong reasons for adding any other basic $\land$-rules, perfectly natural and reliable though they might be. So we won't.

## 20.2 Rules for negation

(a) We next introduce rules governing negation – and things now do get a little more interesting! So consider the following informal argument:

**C** Popper didn't write a logic book. Hence it isn't the case that Popper and Quine each wrote a logic book.

This is valid. Why? Here is a natural line of reasoning, spelt out in painful detail:

> Our premiss is that Popper didn't write a logic book. Now temporarily suppose that Popper and Quine did each write a logic book. Then, trivially, we'll be committed to saying Popper in particular *did* write a logic book. But that directly contradicts our premiss. Hence our temporary supposition leads to absurdity and has to be rejected. So it *isn't* the case that Popper and Quine each wrote a logic book.

Here we are, of course, deploying a *reductio ad absurdum* (RAA) inference of the kind we first met back in §4.5. In other words, we are making a temporary assumption or supposition for the sake of argument. We find that the temporary supposition leads to absurdity, given our background premiss. So we drop or discharge that temporary supposition, and conclude that it can't be true.

(b) Let's now transpose all this into a formal mode. Argument **C** can be rendered into a PL language with the obvious glossary, giving us

**C′** $\neg P \therefore \neg(P \land Q)$.

And now we simply replicate the natural line of reasoning we used a moment ago:

| | | | |
|---|---|---|---|
| (1) | $\neg P$ | | (Prem) |
| (2) | | $(P \land Q)$ | (supposed for the sake of argument) |
| (3) | | $P$ | ($\land$E 2) |
| (4) | | $\bot$ | (absurd! – 3 and 1 are contradictory) |
| (5) | $\neg(P \land Q)$ | | (by RAA, using the subproof 2–4) |

It should be pretty clear what is going on here. But let's spell things out carefully, beginning with a general point:

(i) In a proof, we can at any step make a new temporary supposition for the sake of argument, so long as we clearly signal what we are doing. And if this isn't to become a permanent new premiss, we will later need to drop or discharge that supposition. When we drop a supposition, we can of course no longer use it or its implications as inputs to further inferences.

Next, two comments about our stylistic choices in displaying our formal proof:

(ii) The overall layout of the proof follows the convention already informally introduced in §4.5. So when we make a new supposition or additional temporary assumption for the sake of argument, we mark this by *indenting the line of argument* one column to the right. We thereby start a derivation-within-a-derivation, i.e. a *subproof*.

When we decide to finish or close off a subproof, *we shift the line of argument back* one column to the left. At this point, we say the supposition at the head of the subproof has been *discharged*.

This elegant column-shifting layout for handling temporary suppositions is primarily due to Frederic Fitch in his *Symbolic Logic* (1952).

(iii) We have decorated the proof by using another *vertical line* to mark the new column for our indented subproof. This line starts against our new supposition, and continues for as long as the supposition remains in play. (Compare: the initial premiss is in play throughout the argument, and so the vertical line starting against that premiss continues for the length of the whole proof.) We have also used another *horizontal bar* to draw a line under the temporary supposition at the beginning of the indented subproof.

Again, in using these standard visual aids, we are following Fitch.

Now a comment on the role of '$\bot$':

(iv) In our proof, we reach a blatant contradiction, with a pair of wffs of the form $\alpha$ and $\neg\alpha$ both in play. So we have reached an absurdity. We highlight this by inferring '$\bot$', using the absurdity constant which we added to the resources of PL languages in §§17.4 and 18.6, and which we have already noted is tautologically entailed by any contradictory pair.

The (RAA) rule can now be characterized like this:

(v) Suppose (i) we have a subproof which starts with a wff $\alpha$, temporarily supposed true for the sake of argument; (ii) in the we derive '$\bot$'; and then (iii) we terminate the subproof by moving back one column to the left, dropping the temporary supposition $\alpha$. We can then use (RAA) to appeal to this finished subproof to infer $\neg\alpha$.

(c) For another example, one requiring a slightly longer proof, consider the argument

**D** Popper is from Vienna. It isn't the case that both Popper and Quine are Viennese. The same goes for Popper and Russell: they aren't both from Vienna. Hence both Quine and Russell are not Viennese.

This is valid. For suppose Quine were from Vienna – then both Popper and Quine would be Viennese, contradicting the second premiss. So Quine isn't Viennese. Similarly suppose Russell were from Vienna – then both Popper and Russell would be Viennese, contradicting the third premiss. So Russell isn't Viennese either. Hence, as we want, both Quine and Russell are not Viennese.

With the obvious glossary for '$\mathsf{P}$', '$\mathsf{Q}$' and '$\mathsf{R}$', we can transcribe our informal argument into a PL language as follows:

**D′** $\mathsf{P},\ \neg(\mathsf{P} \land \mathsf{Q}),\ \neg(\mathsf{P} \land \mathsf{R})\ \therefore\ (\neg\mathsf{Q} \land \neg\mathsf{R})$.

And we can then mirror our easy informal proof as a formal PL proof:

| | | |
|---|---|---|
| (1) | $\mathsf{P}$ | (Prem) |
| (2) | $\neg(\mathsf{P} \land \mathsf{Q})$ | (Prem) |
| (3) | $\neg(\mathsf{P} \land \mathsf{R})$ | (Prem) |
| (4) | $\quad \mathsf{Q}$ | (Supp) |
| (5) | $\quad (\mathsf{P} \land \mathsf{Q})$ | ($\land$I 1, 4) |
| (6) | $\quad \bot$ | (Abs 5, 2) |
| (7) | $\neg\mathsf{Q}$ | (RAA 4–6) |
| (8) | $\quad \mathsf{R}$ | (Supp) |
| (9) | $\quad (\mathsf{P} \land \mathsf{R})$ | ($\land$I 1, 9) |
| (10) | $\quad \bot$ | (Abs 9, 3) |
| (11) | $\neg\mathsf{R}$ | (RAA 8–10) |
| (12) | $(\neg\mathsf{Q} \land \neg\mathsf{R})$ | ($\land$ 7, 11) |

We have added to our repertoire of laconic annotations in a straightforward way. 'Supp' marks, unsurprisingly, a temporary supposition starting a subproof. A reductio inference is marked by 'RAA' together with the line numbers of the beginning and end of the relevant subproof. And 'Abs' marks that we are declaring an absurdity on the basis of a blatant contradiction, and we give the line numbers of the relevant contradictory wffs $\alpha$ and $\neg\alpha$, in that order.

Two important comments about the significance of the layout here.

(i) Note that we *can* infer '$(\mathsf{P} \land \mathsf{Q})$' at line (5) even though we have shifted columns after the initial premiss '$\mathsf{P}$'. When we moved a column to the right, we *added* the new temporary supposition '$\mathsf{Q}$'; but we remain committed to the three original premisses, so we can still invoke one of them.

(ii) It is when moving columns in the other direction that we have to be really careful; for then we are *dropping* the commitment to an additional temporary assumption and its implications. For example, our supposition '$\mathsf{Q}$' made at line (4) is discharged after line (6) – and then it is no longer in play, no longer available to be used as an input to further inferences. But later at line (8) the new supposition '$\mathsf{R}$' is made and becomes available to be used, e.g. in the inference which gets us to line (9).

For more about how wffs can go in and out of play in a proof as temporary suppositions are made and discharged, see §20.6.

(d) Let's summarize the negation rules (RAA) and (Abs) which we have met so far.

> *Reductio ad absurdum*: Given a finished subproof starting with the temporary supposition $\alpha$ and concluding '$\bot$', we can infer $\neg\alpha$.
>
> *The absurdity rule*: Given wffs $\alpha$ and $\neg\alpha$, we can infer '$\bot$'.

(RAA) is our basic rule for arguing *to* a conclusion with '$\neg$' as its main connective. So we could equally well have called it our $\neg$-introduction rule. Similarly, (Abs) is our rule for arguing *from* an input with '$\neg$' as its main connective. In this sense it can be said to be our $\neg$-elimination rule.

And note that these two rules, like the introduction and elimination rules for '$\land$', also fit together beautifully. (Abs) reverses (RAA) in the sense that (Abs) recovers *from* $\neg\alpha$ the essential input used in arguing *to* $\neg\alpha$ by (RAA) – because (Abs) in effect tells us that if we have $\neg\alpha$, then from $\alpha$ we can derive absurdity.

## 20.3 A double negation rule

Our first two $\neg$-rules are intuitively faithful to our understanding of negation. Do we need any other rules?

Given our classical understanding of negation as truth-value-flipping, a wff and its double negation are equivalent, and so we will want to be able to establish this equivalence in our PL proof system. Two questions arise: (a) is there a proof of '$\neg\neg\mathsf{P}$' from '$\mathsf{P}$' using our current rules? (b) Is there a proof of '$\mathsf{P}$' from '$\neg\neg\mathsf{P}$'?

(a) The first of these questions is quickly answered.

As a generally useful rule-of-thumb, if we want to prove $\neg\alpha$, the thing to do is to suppose $\alpha$ and then aim to expose a contradiction so we can use (RAA). So let's do that here, with '$\neg\mathsf{P}$' for $\alpha$:

| | | |
|---|---|---|
| (1) | P | (Prem) |
| (2) | ¬P | (Supp) |
| (3) | ⊥ | (Abs 1,2) |
| (4) | ¬¬P | (RAA 2–3) |

Which couldn't have been easier!

And again the proof generalizes: we can always infer a wff of the form $\neg\neg\alpha$ from the corresponding $\alpha$ (wherever $\alpha$ appears in a proof). So it would be redundant to add a new basic rule to allow us to *introduce* double negations.

(b) However, a little experimentation should quickly convince you that we *do* need a new rule if we are to go in the opposite direction and infer 'P' from '¬¬P'. In §24.7 we will return to consider why this must be so. But for now we will just state a clearly reliable third ¬-rule, which will give us what we need:

*Double negation*: Given a wff $\neg\neg\alpha$, we can infer $\alpha$.

Note: this rule only allows us to eliminate double negations at the very start of a wff. For example, the inference '(P ∧ ¬¬Q) ∴ (P ∧ Q)' is valid; but to prove this, we have to apply (∧E) *before* we are in a position to apply (DN).

To introduce another simple example where we need to invoke (DN), consider the following informal argument:

**E** Popper is a philosopher. It isn't the case that Popper is a philosopher while Quine isn't. Therefore Quine is a philosopher.

This is valid. Why? Here's a simple argument:

The first premiss tells us that Popper is a philosopher. Now temporarily suppose that Quine *isn't* a philosopher. Then we'll be committed to saying Popper is a philosopher while Quine isn't. But that directly contradicts the second premiss. Hence our temporary supposition leads to absurdity and has to be rejected: so it can't be that Quine *isn't* a philosopher. Hence he *is* one.

Two comments about this:

(i) This example illustrates that sometimes, even if your target conclusion *isn't* a negated proposition, it may still be natural to establish it by a reductio argument.

(ii) But note, our reductio arguments do always end with the *negation* of some previous temporary supposition. So if that supposition is already a negated proposition, the reductio argument will yield a *doubly* negated proposition. Formally, we will need an application of our new Double Negation rule if we want to remove the double negation.

We can now render **E** into a suitable PL language like this:

**E′** P, ¬(P ∧ ¬Q) ∴ Q.

And we can regiment the informal line of reasoning just sketched into a formal proof as follows (with the obvious short-form annotation):

| | | |
|---|---|---|
| (1) | P | (Prem) |
| (2) | ¬(P ∧ ¬Q) | (Prem) |
| (3) |     ¬Q | (Supp) |
| (4) |     (P ∧ ¬Q) | (∧I 1, 3) |
| (5) |     ⊥ | (Abs 4, 2) |
| (6) | ¬¬Q | (RAA 3–5) |
| (7) | Q | (DN 6) |

(c) We will soon see more examples of our three negation rules at work. But, for the moment, let's just finish this section with a diagrammatic summary of these natural rules of inference with their short-form labels.

*Rules for negation*

(Abs)
$$\begin{array}{c}\gamma\\ \vdots\\ \neg\gamma\\ \vdots\\ \bot\end{array}$$

(RAA)
$$\begin{array}{l} \quad\big|\ \alpha\\ \quad\big|\ \vdots\\ \quad\big|\ \bot\\ \neg\alpha\end{array}$$

(DN)
$$\begin{array}{c}\neg\neg\alpha\\ \vdots\\ \alpha\end{array}$$

Note, though, that we have already learnt that the inputs for the application of a rule like (Abs) can appear in either order, and needn't be in the same column.

## 20.4 Thinking strategically

(a) Our five rules of inference tell us what we *can* infer, not what we *must* infer – the rules give us permissions, not instructions. And as we all know, permissions can allow us to do pointlessly silly things. We therefore need to keep our wits about us if we are not to waste time and effort when proof-building.

Consider for example this slightly messy inference:

**F** It isn't true that Putnam is a logician while Quine isn't. Also, it isn't the case that both Ryle and Quine are logicians. Hence it isn't true that both Putnam and Ryle are logicians.

This is valid (can you see why?). Translated into an appropriate PL language, the argument becomes

**F′** ¬(P ∧ ¬Q), ¬(R ∧ Q) ∴ ¬(P ∧ R)

So how do we get from the premisses to the conclusion using our rules of inference?

The only rule that applies to our two premisses (separately or jointly) is (∧I). And we could use the rule to infer e.g. '(¬(P ∧ ¬Q) ∧ ¬(R ∧ Q))'. This is a perfectly correct inference to make. But given our target conclusion, it would be no help at all. The permissive nature of our inference rules allows us to ramble off on pointless detours like this. So, as we said, some strategic thinking is needed.

(b) In the present case, since we want to establish '¬(P ∧ R)' from the premisses, the inviting strategy is to assume the opposite, i.e. to suppose '(P ∧ R)', and then try to show that this leads to absurdity. So our proof should take the following shape:

```
| ¬(P ∧ ¬Q)
| ¬(R ∧ Q)
|---
|  | (P ∧ R)
|  |---
|  | P
|  | R
|  | ⋮
|  | ⊥
| ¬(P ∧ R)
```

We therefore now have an easier problem. How are we going to argue towards absurdity inside that subproof?

Well, only two of our rules of inference can be applied to the five wffs we have at the beginning of our proof. We could use (∧E) again on the third wff – but what would be the point of *that*? Or we could conjoin any two of the wffs using (∧I) – but that would also be quite pointless again. So it seems that the only way forward is to make a judiciously chosen *further* supposition. But this looks promising: a simple option is to chose 'Q', as this will combine with 'R' to give us a contradiction.

So, let's make this further temporary supposition, indenting the line of proof again accordingly, and we get

```
(1)  | ¬(P ∧ ¬Q)                  (Prem)
(2)  | ¬(R ∧ Q)                   (Prem)
     |---
(3)  |  | (P ∧ R)                 (Supp)
     |  |---
(4)  |  | P                       (∧E 3)
(5)  |  | R                       (∧E 3)
(6)  |  |  | Q                    (Supp)
     |  |  |---
(7)  |  |  | (R ∧ Q)              (∧I 5 6)
(8)  |  |  | ⊥                    (Abs 7, 2)
(9)  |  | ¬Q                      (RAA 6–8)
```

And *now* we can see our way home! Another contradiction has come into view, and we can easily finish our proof:

```
(10) |  | (P ∧ ¬Q)                (∧I 4, 9)
(11) |  | ⊥                       (Abs 10, 1)
(12) | ¬(P ∧ R)                   (RAA 3-11)
```

Note the column-shifting discipline imposed in our Fitch-style proof system. In this proof, we make two temporary suppositions, with the second one introduced while the first is still active. That is why, at line (6), the column of reasoning has to be indented a second time. Then we eventually discharge our suppositions one at a time, as we jump back leftwards just one column at a time. (Do compare this proof with the proof for **D′**, which also involves two reductio arguments. Make sure you understand why, in the proof for **D′**, those arguments appear one *after* the other, while here in the proof for **F′**, they appear one *inside* the other.)

## 20.5 Understanding proofs, discovering proofs

An important general remark, worth pausing over before continuing any further.

This is a logic book for philosophers, not a mathematics book. You should aim to *understand* what is going on in Fitch-style proofs. Being able to *discover* proofs like the last one is much less important. We will work through more examples, talking strategies for finding proofs as we go. But if you still find proof-discovery a bit challenging, don't panic! Philosophically speaking, it is *not* the crucial thing: so don't get bogged down in tackling Exercises if you find them hard going. Just read on and make sure you at least grasp the general *principles* at stake in these chapters about proofs.

## 20.6 'Availability'

(a) Suppose we are working in a language where 'P' means *Hillary is a woman* and 'Q' means *Donald is a woman*. Consider this inference:

**G** ¬(P ∧ Q) ∴ ¬P

This has a true premiss and false conclusion, so it is of course invalid! Yet here is a 'proof' which purports to derive **G**'s conclusion from its premiss:

| | | |
|---|---|---|
| (1) | │ ¬(P ∧ Q) | (Prem) |
| (2) | │ │ P | (Supp) |
| (3) | │ │ │ Q | (Supp) |
| (4) | │ │ │ (P ∧ Q) | (∧I 2, 3) |
| (5) | │ │ │ ⊥ | (Abs 4, 1) |
| (6) | │ │ ¬Q | (RAA 3–5) |
| (7) | │ │ Q | (∧E 4) |
| (8) | │ │ ⊥ | (Abs 7, 6) |
| (9) | │ ¬P | (RAA 2–8) |

Since **G** is invalid, something must be badly amiss with this putative 'proof'. But what?

Everything preceding line (7) is fine. We make a first temporary supposition at line (2), indenting the line of proof; and then we make another supposition at (3), so we indent the proof a second time. The inner proof-within-a-proof then tells us that the supposition 'Q' at (3) leads to absurdity; so we can discharge that supposition and correctly appeal to (RAA) to derive '¬Q' at line (6). So far so good.

However, note that by the time we get to line (6), we have discharged the supposition at line (3). Hence, from (6) onwards, *the wff at line (3) and the wffs it implies are no longer being assumed to be true*. Once a subproof is finished or closed off, its innards are, so to speak, all 'packed away'. In particular, by the time we get to line (7), the wff '$(P \wedge Q)$' from line (4) is no longer available to be used as an input to a further inference. So we *can't* now use it to infer '$Q$'.

(b) *What happens in subproofs, stays in subproofs* – we could call this the Vegas rule! And for another example of what can go wrong if you flout the rule, here is a second 'proof' for **G**:

| | | |
|---|---|---|
| (1′) | $\neg(P \wedge Q)$ | (Prem) |
| (2′) | $Q$ | (Supp) |
| (3′) | $P$ | (Supp) |
| (4′) | $(P \wedge Q)$ | (∧I 3′, 2′) |
| (5′) | $\bot$ | (Abs 4′, 1′) |
| (6′) | $(Q \wedge Q)$ | (∧I 2′, 2′) |
| (7′) | $\neg P$ | (RAA, 3′–5′) |

This time the howler is at line (7′). Why? Everything is unproblematic up to and including line (5′). Then we terminate the inner subproof, and basically ignore it at the next step: but that's fine – we don't go *wrong* by not immediately applying reductio and by applying another rule instead. We then terminate the subproof from (2′) to (6′). We are allowed to do this too. We can't go wrong by closing off a subproof whenever we want – so long as we no longer rely on its initial supposition or anything else we deduced while that supposition was in force.

So, when we close off the subproof from (2′) to (6′), by the Vegas rule all its contents are to be considered as 'packed away', *including the inner sequence of wffs from (3′) to (5′)*. That's why we can't *now* apply (RAA) to that subproof to infer (7′).

(c) With these two examples vividly in mind, here is a key definition (there seems to be no universally-agreed terminology, but the idea is a standard one):

> After a subproof is finished and its initial temporary supposition is discharged, its wffs and any subproofs nested inside it become *unavailable* for future use.

Wffs/subproofs earlier in a proof are, of course, deemed available until they become unavailable. Our Fitch-style structuring of proofs makes it quite beautifully clear when subproofs are finished, and hence when their contents become unavailable.

We can now state more carefully the crucial constraint which will govern the application of all the rules of inference in our proof system, to avoid howlers like those in our last two 'proofs':

> When extending a proof by applying a rule of inference, the rule can only invoke as its inputs some previous wffs/subproofs that are *available* at that stage in the proof.

In other words, when we talk of the 'given' inputs to a rule of inference, these inputs must be available, not-packed-away-in-a-finished-subproof, at the point when the rule is applied.

## 20.7 Explosion and absurdity again

The inference rules which we have introduced for conjunction and negation seem intuitively compelling, and the way we have been putting them together in building up formal proofs seems entirely natural. But they lead to a perhaps unexpected consequence.

(a) Consider, however, the following two inferences:

**C″** $\neg P \therefore \neg(P \land \neg Q)$,
**E′** $P, \neg(P \land \neg Q) \therefore Q$.

Both are valid. The first is a minor variant of **C′**, and is proved in the same way. The second we have met before. So here are (entirely natural-seeming) Fitch-style proofs for both:

| | | |
|---|---|---|
| (1) | $\neg P$ | |
| (2) | | $(P \land \neg Q)$ |
| (3) | | $P$ |
| (4) | | $\bot$ |
| (5) | $\neg(P \land \neg Q)$ | |

| | | |
|---|---|---|
| (1) | $P$ | |
| (2) | $\neg(P \land \neg Q)$ | |
| (3) | | $\neg Q$ |
| (4) | | $(P \land \neg Q)$ |
| (5) | | $\bot$ |
| (6) | $\neg\neg Q$ | |
| (7) | $Q$ | |

Now, we can chain together correctly constructed proofs so that the conclusion of one becomes a premiss for the next, and the result will be a longer, equally correct, proof. Or at least, this is what we normally assume we can do, e.g. in mathematics. And surely, this is how logical inference can contribute massively to knowledge. We chain together relatively obvious pieces of reasoning into something longer, less obvious, but still reliably truth-preserving; and this way we get to warrant some new and perhaps surprising conclusion.

Our formal proof system will allow us to combine proofs in this way. So let's in particular chain together our two compelling proofs above (just moving the right-hand proof's initial premiss to the top, since – merely as a matter of style – we insist on listing premisses at the outset). We then get the proof **H**:

| | | | |
|---|---|---|---|
| (1) | $P$ | | (Prem) |
| (2) | $\neg P$ | | (Prem) |
| (3) | | $(P \land \neg Q)$ | (Supp) |
| (4) | | $P$ | ($\land$E 3) |
| (5) | | $\bot$ | (Abs 4, 2) |
| (6) | $\neg(P \land \neg Q)$ | | (RAA 3–5) |
| (7) | | $\neg Q$ | (Supp) |
| (8) | | $(P \land \neg Q)$ | ($\land$I 1, 7) |
| (9) | | $\bot$ | (Abs 8, 6) |
| (10) | $\neg\neg Q$ | | (RAA 7–9) |
| (11) | $Q$ | | (DN 10) |

So from the contradictory pair ‘P’ and ‘P’, we can derive ‘Q’.

And the proof idea easily generalizes. Replacing ‘P’ throughout by $\alpha$ and ‘Q’ throughout by $\gamma$, we will get a proof from any contradictory pair $\alpha$, $\neg\alpha$ to an arbitrary conclusion $\gamma$. Explosion! (Compare §17.1.)

(b) When we first met the explosive inference in §6.5, it was as a probably unwelcome upshot of our informal definition of validity. We then asked: should we revise the definition because of this consequence? Or should we learn to live with the idea that a contradiction entails anything?

In that early discussion, we said – then without explanation – that rejecting explosion comes at a high price. We can now see one reason why. In order to block our particular formal proof **H**, we have just three options:

(1) We can reject the proof which we took to validate the inference **C**″.
(2) We can reject the proof which we took to validate the inference **E**′.
(3) We can drop the assumption that we can unrestrictedly chain correct inferences together to get another correct inference.

None of these options looks very attractive at all. Which is why the great majority of logicians do bite the bullet and accept explosion.

Philosophers being the contentious lot that they are, there exists a vocal minority who resist, still insisting that explosion is a fallacy of irrelevance. We can't here investigate the various ways that have been suggested for escaping **H** and similar explosive arguments – though, if forced to choose, I'd tinker with (3). It is fair to say, though, that no escape route has proved very popular, particularly among mathematicians who want a *practicable* logical system.

(c) Now for another explosive inference. We have officially added the absurdity sign to our PL languages as a special wff, and it has featured in our formal (RAA) arguments as the conclusion of some subproofs. But, like any other wff, it can also in principle appear e.g. as a premiss in an argument.

(An aside here. We could – with only relatively minor adjustments – have arranged things so that the absurdity sign *only* gets used to conclude subproofs. And then we could downgrade the sign to being something like an exclamation mark signalling that you have hit a contradiction. There are some philosophical attractions in doing things this way. But we will continue along the more well-trodden path of treating ‘$\bot$’ as a wff.)

Since ‘$\bot$’ is like a generalized contradiction, we will now expect to be able to argue from it to any conclusion we want. And we can: consider for example this proof, **I**:

| | | | |
|---|---|---|---|
| (1) | $\bot$ | | (Prem) |
| (2) | | $\neg$P | (Supp) |
| (3) | | ($\neg$P $\wedge$ $\bot$) | ($\wedge$I 2, 1) |
| (4) | | $\bot$ | ($\wedge$I E) |
| (5) | $\neg\neg$P | | (RAA 2–4) |
| (6) | P | | (DN, 5) |

Needless to say, this line of argument also generalizes. Replace ‘P’ by any wff $\alpha$, and we will get a derivation from '⊥' to $\alpha$. This corresponds to the explosive tautologically valid inference we tagged *ex falso quodlibet* in §17.3.

(d) The availability of proofs like **I** means that it is strictly speaking redundant to add the following as separate rule of inference, given the other rules already in play:

*Ex Falso Quodlibet*: Given '⊥', we can infer any wff $\alpha$.

However, despite the redundancy, it *is* usual to include this (EFQ) rule in the basic package of inference rules for a Fitch-style proof system for a propositional logic which uses the absurdity sign. There are a number of reasons for making this addition. The shallow one is that it makes some proofs neater. A deeper reason is that we can see (EFQ) as the rule which tells us just what it is to be an all-purpose absurdity – it is something such that if *that* is true, then really *anything* goes!

So we will now adopt (EFQ) as a further rule for our proof system. Using this rule we get a *much* snappier proof than **H** taking us from 'P' and '¬P' to 'Q', namely **J**:

| | | |
|---|---|---|
| (1) | P | (Prem) |
| (2) | ¬P | (Prem) |
| (3) | ⊥ | (Abs 1, 2) |
| (4) | Q | (EFQ 3) |

Quite often, the (EFQ) rule is presented right at the very start of an exposition of a proof system for propositional logic. And then the beginner is immediately faced with proofs like our last one which can seem far too much like suspicious trickery. That's why we have left adding the (EFQ) rule until the end of this chapter, *after* the explosive consequence 'P, ¬P ∴ Q' has *already* been proved. The derivation **J** which uses (EFQ) just allows us to get more quickly to a destination that we can equally well get to the long way round, invoking only our initial package of rules.

## 20.8 Summary

We have adopted five basic rules for arguing with conjunctions, negations, and the absurdity sign. (∧I) and (∧E) tell us that we can argue from conjuncts to conjunctions and back again. (Abs) tells us how to mark that we have got entangled in the absurdity of blatant contradiction. (RAA) tells us that if a supposition leads to absurdity, then we can infer its negation. (DN) allows us to eliminate (initial) double negations.

In general terms, these basic rules seem entirely natural and compelling. So we indeed ought to be able to use them to construct genuine *proofs* which show that the eventual conclusion will be true if the premisses are. But we have choices about how to implement the rules in detail.

We have adopted a *Fitch-style* vertical layout, whose distinctive feature is that we indent the line of reasoning when we make a new temporary supposition, and cancel the indent when we drop that supposition again.

It is crucial that the inputs invoked in applying a rule of inference are still available to be used, i.e. they are not buried inside a subproof which is already finished. What happens in a subproof, stays in a subproof.

We noted that, if we are allowed to chain proofs in the standard sort of way, then our rules – natural though they are – warrant the explosive inference from contradictory wffs to any conclusion we like. We also in effect get for free the further rule (EFQ) which tells us that we can infer anything from absurdity – though we will officially add this as a rule in its own right.

## Exercises 20

# 21 PL proofs: disjunction

There are a number of everyday modes of inference which involve making temporary suppositions for the sake of argument. So far, we have met *reductio* proofs. In this chapter we introduce the important idea of a *proof by cases*.

A formal proof system which smoothly handles temporary suppositions – and so copes with reductio and proof by cases in a natural way – is standardly called a *natural deduction* system. There is more than one type of natural deduction system on the market. But we have to concentrate on just one. So we continue to present our Fitch-style system, extending it now to deal with disjunctions. (The discussion, with lots of proof examples, is inevitably rather dense in places. Do take things slowly.)

## 21.1 The iteration rule

First, however, consider the following little inference which involves none of the propositional connectives:

**A** Popper was born in 1902. Quine was born in 1908. Russell was born in 1872. Hence, Quine was born in 1908.

Given our understanding of what makes for validity, this counts as trivially valid, for there is certainly no possible situation in which our premisses are all true and the conclusion false – see §3.2(b).

Correspondingly, in a suitable PL language,

**A′** P, Q, R ∴ Q

is also trivially valid. We will therefore want to be able to construct a proof of the conclusion from the premisses in our PL proof system.

One way of getting a proof ending in the right conclusion is simply to adopt a new rule:

> *Iteration*: At any inference step, we can reiterate any available wff.

This easy rule cannot lead us astray! If a wff is already available at a given step, repeating it doesn't involve any new commitment. (In fact, we will almost always be iterating premisses or 'live' – i.e. undischarged – suppositions.)

Using our iteration rule, call it '(Iter)' for short, we can construct the required miniature proof for the inference in **A′**:

| | | |
|---|---|---|
| (1) | P | (Prem) |
| (2) | Q | (Prem) |
| (3) | R | (Prem) |
| (4) | Q | (Iter 2) |

Now, we could have managed perfectly well without using any iteration rule. For – assuming that we can use the usual rules for conjunction – our rule is actually redundant. In the present case, we could argue as follows:

| | | |
|---|---|---|
| (1) | P | (Prem) |
| (2) | Q | (Prem) |
| (3) | R | (Prem) |
| (4) | $(Q \land Q)$ | ($\land$I 2, 2) |
| (5) | Q | ($\land$E 4) |

We said in §20.1(a) that we can appeal to the same input twice when applying ($\land$I), as in (4). And note, the trick in this proof can be applied quite generally: in other words, whenever we are tempted to use (Iter), we could use the two conjunction rules instead.

Still, it does seem rather perverse to invoke the conjunction rules in a little two-step dance in order to derive a conclusion which contains no conjunction from premisses which equally contain no conjunctions. So, just on the grounds that it occasionally makes for more natural proofs, we will adopt the trivially safe additional rule (Iter) – though it is worth stressing that there is no right or wrong policy here.

## 21.2 Introducing and eliminating disjunctions

(a) We move on, then, to introduce two principles for reasoning with disjunctions (take all disjunctions in this chapter to be inclusive).

Consider first the following little inference:

**B** Quine is a logician. Therefore either Popper or Quine is a logician, or Russell is one.

This is of course trivially valid. You can infer an inclusive disjunction from either disjunct. So from the premiss that Quine is a logician you can infer that Popper or Quine is a logician. And from that interim conclusion you can now infer that either Popper or Quine is a logician, or Russell is one. It is as easy as that.

Going formal, the inferential principle for arguing *to* a disjunction becomes simply this:

> $\lor$-*Introduction*: Given a wff $\alpha$, we can infer $(\alpha \lor \beta)$ for any wff $\beta$. Equally given $\beta$, we can infer $(\alpha \lor \beta)$ for any $\alpha$.

With 'P' meaning *Popper is a logician*, etc., our everyday inference **B** can be rendered into a suitable PL language as

**B′** $Q \therefore ((P \lor Q) \lor R)$.

And we can then mirror our informal proof which shows that **B** is valid by a direct – and truth-preserving! – formal PL proof:

| | | |
|---|---|---|
| (1) | Q | (Prem) |
| (2) | (P ∨ Q) | (∨I 1) |
| (3) | ((P ∨ Q) ∨ R) | (∨I 2) |

(We use the natural label for our new rule in the commentary, and indicate the line number of the wff which an application of the rule appeals to.)

(b) To work towards the corresponding ∨-elimination rule, let's start by considering a couple of informal one-premiss inferences. First:

**C** Either Popper and Quine are both logicians, or Quine and Russell are both logicians. Therefore Quine is a logician.

This is valid. Why? Here's an informal proof, spelt out in laborious detail:

> Suppose for the sake of argument that the first disjunct of our premiss is true, so Popper and Quine are both logicians. Then, on that assumption, Quine in particular is a logician.
>
> Suppose alternatively that the second disjunct of our premiss is true, so Quine and Russell are both logicians. It follows on that assumption too that Quine is a logician.
>
> So either way – whichever disjunct of our initial premiss holds – we get our desired conclusion.

The inferential principle in play here is a version of *proof by cases*. We are given that (at least) one of two cases $A$ and $B$ holds. We show that in the first case $C$ follows. We show that, equally, in the second case $C$ follows. Since $C$ follows in either case, and we are given that (at least) one of the cases holds, we can infer $C$ outright.

Here's a second inference:

**D** Polly and Quentin are both married. So it isn't the case that either Polly is unmarried or that Quentin is.

This too is valid. How can we informally show that? We could argue like this.

> Suppose for a moment the opposite of the conclusion holds, i.e. suppose that either Polly is unmarried or Quentin is unmarried.
>
> But the first case leads to absurdity, since it is flatly contradicted by our initial premiss which tells us that Polly *is* married. And likewise the second case leads to absurdity, since it is flatly contradicted by our initial premiss which tells us that Quentin *is* married.
>
> So either way, whichever disjunct of our initial supposition holds, we get absurdity. Therefore that supposition leads to absurdity. Hence it has to be rejected.

Again, in the middle of this informal proof we use a version of proof by cases. We suppose that (at least) one of two cases $A$ and $B$ holds. We show that in the first case absurdity follows. We show that, equally, in the second case absurdity follows. Since we can infer an absurdity in either case, the supposition that one of the cases $A$ and $B$ holds is itself absurd.

The key idea, then, is this: if we are given a disjunction, and each disjunct leads to the same conclusion (whether an ordinary proposition or absurdity), then we can infer that conclusion outright. Going formal, this natural principle is captured by the following rule of inference in our Fitch-style system. Suppose $\alpha$, $\beta$ and $\gamma$ are any wffs (possibly the absurd wff '$\bot$'), then:

> $\lor$-*Elimination*: Given $(\alpha \lor \beta)$, a finished subproof from the temporary supposition $\alpha$ to $\gamma$, and also a finished subproof from the temporary supposition $\beta$ to $\gamma$, we can infer $\gamma$.

So let's put this to work.

First, **C** can be rendered into a suitable PL language as

**C′** $((\mathsf{P} \land \mathsf{Q}) \lor (\mathsf{Q} \land \mathsf{R})) \therefore \mathsf{Q}$.

Then we can closely mirror our informal proof for **C** with the following formal proof for **C′**:

| | | |
|---|---|---|
| (1) | $((\mathsf{P} \land \mathsf{Q}) \lor (\mathsf{Q} \land \mathsf{R})$ | (Prem) |
| (2) | │ $(\mathsf{P} \land \mathsf{Q})$ | (Supp) |
| (3) | │ $\mathsf{Q}$ | ($\land$E 2) |
| (4) | │ $(\mathsf{Q} \land \mathsf{R})$ | (Supp) |
| (5) | │ $\mathsf{Q}$ | ($\land$E 4) |
| (6) | $\mathsf{Q}$ | ($\lor$E 1, 2–3, 4–5) |

Three explanatory comments about the layout here:

(i) At line (2) we suppose for the sake of argument that the first disjunct of (1) is true, indenting our proof a column rightwards as we do so. When we finish this first subproof after line (3), we head back a column leftwards. However, we *immediately* make the alternative supposition that the second disjunct of (1) is true, giving us a new supposition at line (4) to start the second subproof. And therefore we indent our column of reasoning rightwards again. Hence the proof layout as displayed.

(ii) There is therefore a little *gap* between the vertical lines marking the extents of the two different finished subproofs – but it helps the eye if we put a very small bar in the gap to emphasize the break between the subproofs.

(iii) Our short-form annotation of the final step records the *three* inputs to the final $\lor$-elimination inference; we give the line number of the relevant disjunction, and then give the extents of the two relevant subproofs.

Putting '$(\mathsf{P} \lor \mathsf{Q})$' for $\alpha$, '$(\mathsf{Q} \lor \mathsf{R})$' for $\beta$, and '$\mathsf{Q}$' for $\gamma$, the final line is indeed an application of the ($\lor$E) rule.

Likewise, using the obvious glossary, here is a formal version of **D** (giving us an instance of one of De Morgan's laws – see §13.4):

**D′** $(\mathsf{P} \land \mathsf{Q}) \therefore \neg(\neg\mathsf{P} \lor \neg\mathsf{Q})$.

And we can replicate our informal line of reasoning for **D** as follows:

| | | |
|---|---|---|
| (1) | $(\mathsf{P} \land \mathsf{Q})$ | (Prem) |
| (2) | $(\neg\mathsf{P} \lor \neg\mathsf{Q})$ | (Supp) |
| (3) | $\neg\mathsf{P}$ | (Supp) |
| (4) | $\mathsf{P}$ | ($\land$E 1) |
| (5) | $\bot$ | (Abs 4, 3) |
| (6) | $\neg\mathsf{Q}$ | (Supp) |
| (7) | $\mathsf{Q}$ | ($\land$E 1) |
| (8) | $\bot$ | (Abs 7,6) |
| (9) | $\bot$ | ($\lor$E 2, 3–5, 6–8) |
| (10) | $\neg(\neg\mathsf{P} \lor \neg\mathsf{Q})$ | (RAA 4,10) |

This time, putting '$\neg\mathsf{P}$' for $\alpha$, '$\neg\mathsf{Q}$' for $\beta$ and '$\bot$' for $\gamma$, line (9) is again an application of the ($\lor$E) rule.

(c) Let's take another example. So consider the informal argument

**E** Popper is a clear writer. Either Quine or Russell is a clear writer. Hence either Popper and Quine are clear writers, or both Popper and Russell are.

Valid again, as is shown by considering cases.

> Take the case where Quine is a clear writer. Then Popper and Quine are clear writers – so, it will true that either Popper and Quine are clear writers or Popper and Russell are. Similarly in the other case, where Russell is a clear writer. Either way, our desired conclusion follows.

Going formal, here is a corresponding PL inference:

**E′** $\mathsf{P}, (\mathsf{Q} \lor \mathsf{R}) \therefore ((\mathsf{P} \land \mathsf{Q}) \lor (\mathsf{P} \land \mathsf{R}))$.

This can be shown to be valid by the following proof:

| | | |
|---|---|---|
| (1) | $\mathsf{P}$ | (Prem) |
| (2) | $(\mathsf{Q} \lor \mathsf{R})$ | (Prem) |
| (3) | $\mathsf{Q}$ | (Supp) |
| (4) | $(\mathsf{P} \land \mathsf{Q})$ | ($\land$I 1, 3) |
| (5) | $((\mathsf{P} \land \mathsf{Q}) \lor (\mathsf{P} \land \mathsf{R}))$ | ($\lor$I 4) |
| (6) | $\mathsf{R}$ | (Supp) |
| (7) | $(\mathsf{P} \land \mathsf{R})$ | ($\land$I 1, 6) |
| (8) | $((\mathsf{P} \land \mathsf{Q}) \lor (\mathsf{P} \land \mathsf{R}))$ | ($\lor$I 7) |
| (9) | $((\mathsf{P} \land \mathsf{Q}) \lor (\mathsf{P} \land \mathsf{R}))$ | ($\lor$E 2, 3–5, 6–8) |

(d) Now take the trivial inference

**F** $(\mathsf{P} \lor \mathsf{P}) \therefore \mathsf{P}$

We can prove the conclusion from the premisses here by another proof by cases:

| | | |
|---|---|---|
| (1) | $(\mathsf{P} \lor \mathsf{P})$ | (Prem) |
| (2) | $\quad\quad \mathsf{P}$ | (Supp) |
| (3) | $\quad\quad \mathsf{P}$ | (Iter 2) |
| (4) | $\quad\quad \mathsf{P}$ | (Supp) |
| (5) | $\quad\quad \mathsf{P}$ | (Iter 4) |
| (6) | $\mathsf{P}$ | ($\lor$E 1, 2–3, 4–5) |

(Minor point: we could also allow the same subproof to be invoked twice by ($\lor$E), just as we allow the same wff to be invoked twice e.g. in applying ($\land$I).)

Note our use of the iteration rule (Iter) to get us from the supposition '$\mathsf{P}$' at (2) to the subproof's conclusion '$\mathsf{P}$' at (3). As we remarked in the previous section, we could equally derive the same conclusion by introducing and then eliminating a conjunction. But this would again seem rather unnatural. We want to be able to warrant the conjunction-free **F** without invoking conjunction rules, and (Iter) enables us to do that.

(e) Having introduced our $\lor$-rules and seen some easy examples of the rules in use, let's pause for a diagrammatic summary:

*Rules for disjunction*

($\lor$I)

$$\begin{array}{cc} \alpha & \beta \\ \vdots & \vdots \\ (\alpha \lor \beta) & (\alpha \lor \beta) \end{array}$$

($\lor$E)

$$\begin{array}{l} (\alpha \lor \beta) \\ \vdots \\ \quad | \ \alpha \\ \quad | \ \vdots \\ \quad | \ \gamma \\ \quad | \ \beta \\ \quad | \ \vdots \\ \quad | \ \gamma \\ \gamma \end{array}$$

Strictly speaking, we do not *require* the three elements for an application of ($\lor$E) to appear in the order diagrammed, nor do we require one subproof to appear immediately after the other. But in practice, that's the usual layout.

We noted in §20.1(a) how ($\land$E) allows us to recover from a conjunction $(\alpha \land \beta)$ the inputs that ($\land$I) requires in arguing to that conjunction. So ($\land$E) can be used to 'reverse' an application of ($\land$I). Now, we obviously can't reverse an application of ($\lor$I) in quite the same way. For of course we cannot argue *backwards* from a disjunction $(\alpha \lor \beta)$ to recover whichever input ($\lor$I) might use in arguing to that disjunction.

But ($\lor$E) can be used to undo an application of ($\lor$I) in a closely related sense. For ($\lor$E) allows us to argue *forwards* from $(\alpha \lor \beta)$ to any conclusion we could already have derived equally from $\alpha$ and $\beta$ before we used ($\lor$I). If we can already infer $\gamma$ from $\alpha$, and also infer $\gamma$ from $\beta$, then ($\lor$E) allows us *still* to get to $\gamma$, now from $(\alpha \lor \beta)$.

So our ($\lor$I) and ($\lor$E) rules also do seem to fit together very nicely – as it is often put, they are in *harmony*.

## 21.3 Two more proofs

(a) For another example, consider the following (interpret the atoms however you like):

**G** ¬(¬P ∨ ¬Q) ∴ (P ∧ Q)

This is valid, being an instance of another of De Morgan's Laws.

How can we show this to be a valid inference by a PL proof in our system? Given that our conclusion is a conjunction, a natural strategy is first to derive one conjunct from our given premiss, and then to derive the other conjunct, before conjoining these two interim conclusions. Which breaks down the proof into two simpler tasks.

How then can we prove 'P' from our premiss? There's no rule we can directly apply to the premiss. (Reality check: why would it be a howler to try to infer 'P' from '¬(¬P ∨ ¬Q)' by a straight application of (DN)?) What else can we do, then, but suppose '¬P', and aim for a contradiction?

```
(1)   | ¬(¬P ∨ ¬Q)             (Prem)
      |---
(2)   |    | ¬P                (Supp)
      |    |---
(3)   |    | (¬P ∨ ¬Q)         (∨I 2)
(4)   |    | ⊥                 (Abs 3, 1)
(5)   | ¬¬P                    (RAA 2–4)
(6)   | P                      (DN 5)
(7)   |    | ¬Q                (Supp)
      |    |---
(8)   |    | (¬P ∨ ¬Q)         (∨I 7)
(9)   |    | ⊥                 (Abs 8, 1)
(10)  | ¬¬Q                    (RAA 7–9)
(11)  | Q                      (DN 10)
(12)  | (P ∧ Q)                (∧I 6, 11)
```

(b) Now let's show the following is valid by another Fitch-style derivation of the conclusion from the premiss:

**H** (P ∨ (Q ∨ R)) ∴ ((P ∨ Q) ∨ R)

We have a disjunctive premiss. We need to use it in a proof by cases. So we can already sketch out the shape of our hoped-for proof:

```
| (P ∨ (Q ∨ R))
|---
|    | P                 (Suppose first disjunct of initial premiss is true)
|    |---
|    | ⋮
|    | ((P ∨ Q) ∨ R)     (Target conclusion)
|    
|    | (Q ∨ R)           (Suppose second disjunct is true)
|    |---
|    | ⋮
|    | ((P ∨ Q) ∨ R)     (The same target conclusion)
| ((P ∨ Q) ∨ R)          (∨E from initial premiss and two subproofs)
```

How is the first subproof to go? Plainly, we can use (∨I) twice to arrive at '((P ∨ Q) ∨ R)'. How is the second subproof to go? This time we have a disjunctive temporary supposition '(Q ∨ R)'. Hence – and this is the important point to see – we will now have to embed *another* proof by cases (*inside* the main one), with two new subproofs starting 'Q' and 'R' and each aiming at the same conclusion '((P ∨ Q) ∨ R)'.

With these guiding thoughts in mind, a finished proof is now quite easily found:

```
(1)   | (P ∨ (Q ∨ R))                  (Prem)
      |----
(2)   |   | P                          (Supp)
      |   |----
(3)   |   | (P ∨ Q)                    (∨I 3)
(4)   |   | ((P ∨ Q) ∨ R)              (∨I 4)
      |
(5)   |   | (Q ∨ R)                    (Supp)
      |   |----
(6)   |   |   | Q                      (Supp)
      |   |   |----
(7)   |   |   | (P ∨ Q)                (∨I 6)
(8)   |   |   | ((P ∨ Q) ∨ R)          (∨I 7)
      |   |
(9)   |   |   | R                      (Supp)
      |   |   |----
(10)  |   |   | ((P ∨ Q) ∨ R)          (∨I 9)
(11)  |   | ((P ∨ Q) ∨ R)              (∨E 5, 6–8, 9–10)
(12)  | ((P ∨ Q) ∨ R)                  (∨E 1, 2–4, 5–11)
```

## 21.4 Disjunctive syllogisms

(a) We turn next to discuss so-called *disjunctive syllogism* inferences. Informally, these are inferences where we take a disjunction and the negation of one disjunct, and then infer the other disjunct. Thus from *A or B* and *Not-A* we can infer *B*; and equally, from *A or B* and *Not-B* we can infer *A*. Given our standard understanding of negation and disjunction, such inferences are certainly valid (see argument **B** in §1.3).

So consider a formal instance, e.g.

**I** (P ∨ Q), ¬P ∴ Q

How can we derive the conclusion from the premisses by a PL proof, *using just our current rules of inference?* We will need a proof by cases, shaped like this:

```
| (P ∨ Q)          (Prem)
| ¬P               (Prem)
|----
|   | P            (Supp)
|   |----
|   | ⋮
|   | Q
|
|   | Q            (Supp)
|   |----
|   | ⋮
|   | Q
| Q                (∨E, from the first premiss and the two subproofs)
```

Filling in the *second* subproof takes no work – it is just an iteration step. But what about the *first* subproof? In that subproof, the initial premiss '¬P' and the temporary assumption 'P' are both available. We will then have to appeal to some explosive argument to get us to 'Q'.

There are various options. We could for example borrow the long-way-round proof **H** from the last chapter. But now we see the point of having the fast-track explosive rule (EFQ) to hand. For that enables us to give the following much snappier proof:

| | | | |
|---|---|---|---|
| (1) | (P ∨ Q) | | (Prem) |
| (2) | ¬P | | (Prem) |
| (3) | | P | (Supp) |
| (4) | | ⊥ | (Abs 3, 2) |
| (5) | | Q | (EFQ 4) |
| (6) | | Q | (Supp) |
| (7) | | Q | (Iter) |
| (8) | Q | | (∨E) |

And the proof-strategy generalizes as you would now expect: we can warrant any PL disjunctive syllogism of the form $(\alpha \vee \beta), \neg\alpha \therefore \beta$, and we can similarly warrant any inference of the form $(\alpha \vee \beta), \neg\beta \therefore \alpha$.

(b) What we have just given is indeed the kind of derivation for a disjunctive syllogism that you standardly meet in Fitch-style proof systems. But on reflection, isn't it rather peculiar?

Return to informal argumentation for a moment. Assume you believe that *A or B* holds. Then suppose you realize that the first disjunct is inconsistent with something else you already accept, *X*. In this case, don't you now simply *rule out the first option A*, and immediately conclude that *B* without any further fuss – and then argue on from there to some desired conclusion *C*? You certainly don't ordinarily think that you have to pause to show, by some explosive inference, that *A* plus *X* together entail *B*!

If we want to stay closer to ordinary ways of reasoning, then – following on from what we've just said – one option is to liberalize our (∨E) rule by *adding* a further clause:

> Given $(\alpha \vee \beta)$, a subproof from one of the disjuncts as temporary supposition to the wff $\gamma$ and also a subproof from the other disjunct to absurdity, we can infer $\gamma$.

And then our proof for **I** could be rather more naturally completed by appeal to this additional clause, without any unnatural play with explosion:

| | | | |
|---|---|---|---|
| (1) | (P ∨ Q) | | (Prem) |
| (2) | ¬P | | (Prem) |
| (3) | | P | (Supp) |
| (4) | | ⊥ | (Abs 3, 2) |
| (5) | | Q | (Supp) |
| (6) | | Q | (Iter 6) |
| (7) | Q | | (Liberalized ∨E 1, 3–4, 5–6) |

Is it worth liberalizing our (∨E) rule like this, though? The shortening of a proof by just one line is neither here nor there. So it comes down to a question of trading off economy of rules against intuitive naturalness of proofs, and you pays your money and you takes your choice. We *did* introduce the rule (Iter) even though it is strictly speaking redundant, because it made some proofs more natural. We could equally well liberalize our (∨E) rule on just the same grounds. But – like it or not – the usual line is to stick to our original version of (∨E) and then use (EFQ) to derive a disjunctive syllogism.

And on balance it is probably unwise to depart significantly from the conventional wisdom in an introductory text like this: for a start, we want to our story to integrate smoothly with more advanced texts. So – with some regrets – we will leave our (∨E) rule in its original standard form.

(c) Let's have another example of an argument where we need to get use (EFQ) in order to massage our formal proof to conform to our standard (∨E) rule. So consider:

**J** Either Popper and Quine were both born in the twentieth century or Quine and Russell were both born in the twentieth century. But Russell wasn't born in the twentieth century. So Popper was.

Why is this valid? Informally,

> Consider cases from the disjunctive premiss. The second case – Quine and Russell both being born in the twentieth century – is quickly ruled out by the premiss that Russell *wasn't* born in the twentieth century. So that leaves only the first case in play: and in that case, Popper was indeed born in the twentieth century.

Using a language with the obvious glossary, inference **J** can be formally rendered as

**J′** $((\mathsf{P} \land \mathsf{Q}) \lor (\mathsf{Q} \land \mathsf{R})),\ \neg\mathsf{R}\ \therefore\ \mathsf{P}$.

And then here is a formal approximation of our sketched proof:

| | | |
|---|---|---|
| (1) | $((\mathsf{P} \land \mathsf{Q}) \lor (\mathsf{Q} \land \mathsf{R}))$ | (Prem) |
| (2) | $\neg\mathsf{R}$ | (Prem) |
| (3) | \| $(\mathsf{P} \land \mathsf{Q})$ | (Supp) |
| (4) | \| $\mathsf{P}$ | (∧E, 3) |
| (5) | \| $(\mathsf{Q} \land \mathsf{R})$ | (Supp) |
| (6) | \| $\mathsf{R}$ | (∧E, 5) |
| (7) | \| $\bot$ | (Abs 6, 2) |
| (8) | \| $\mathsf{P}$ | (EFQ 7) |
| (9) | $\mathsf{P}$ | (∨E 1, 3–4, 5–8) |

Here, the second supposition (5) leads to absurdity at (7), as in the intuitive proof. But to shoehorn our formal proof into the standard form of a (∨E) argument, we need to get (5) to lead eventually to 'P'. We bridge the gap by appealing to (EFQ).

(d) Just for the fun of the ride, let's finish this section with a more complicated example, where our ∨-elimination rule is appealed to twice over (with one occurrence inside another, as in the proof of **H**). So take the following inference:

**K** Either Polly is unmarried or Roland is married. But at least one of Polly and Quentin is married. And it's not true that Quentin is married while Sebastian isn't. So either Roland or Sebastian is married.

This is valid. Informally,

Consider cases from the first premiss.

Suppose Polly is unmarried – then, by disjunctive syllogism using the second premiss, Quentin is married. And then by the third premiss Sebastian is married; and hence Roland or Sebastian is married.

Suppose alternatively that Roland is married – it is immediate that Roland or Sebastian is married.

In both cases we get our conclusion.

See if you can construct a formal version of this informal proof before reading on.

(e) **K** can be regimented into a PL language like this:

**K′** $(\neg P \lor R)$, $(P \lor Q)$, $\neg(Q \land \neg S) \therefore (R \lor S)$.

Note, we have in fact already met this argument – it was **E** in §15.5 where we showed it to be valid by the truth-table test. But now we can turn our informal proof by cases for **K**'s validity into a Fitch-style proof warranting **K′** as follows:

| | | |
|---|---|---|
| (1) | $(\neg P \lor R)$ | (Prem) |
| (2) | $(P \lor Q)$ | (Prem) |
| (3) | $\neg(Q \land \neg S)$ | (Prem) |
| (4) |     $\neg P$ | (Supp) |
| (5) |         $P$ | (Supp) |
| (6) |         $\bot$ | (Abs 5,4) |
| (7) |         $Q$ | (EFQ 6) |
| (8) |         $Q$ | (Supp) |
| (9) |         $Q$ | (Iter 8) |
| (10) |     $Q$ | (∨E 1, 5–7, 8–9) |
| (11) |         $\neg S$ | (Supp) |
| (12) |         $(Q \land \neg S)$ | (∧I 10, 11) |
| (13) |         $\bot$ | (Abs 12, 3) |
| (14) |     $\neg\neg S$ | (RAA 11–13) |
| (15) |     $S$ | (DN 14) |
| (16) |     $(R \lor S)$ | (∨I 15) |
| (17) |     $R$ | (Supp) |
| (18) |     $(R \lor S)$ | (∨I 17) |
| (19) | $(R \lor S)$ | (∨E 2, 4–16, 17–18) |

Getting from (4) to (15) just involves chaining two now familiar patterns of inference that you should now recognize. So the whole proof almost writes itself.

## 21.5 Summary

In this chapter, we have added to our Fitch-style proof system the introduction and elimination rules ($\vee$I) and ($\vee$E) for dealing with disjunctions.

These new rules again seem entirely compelling. However, if we restrict ourselves to just the standard ($\vee$E) rule, some proofs – for disjunctive syllogisms, for example – can look rather unnatural, needing to appeal to (EFQ). But this is usually taken to be a price worth paying for keeping our rules simple.

Note: we have yet to summarize the permissible ways of assembling inferences together into structured proofs (so far we have been relying on examples): we will give the official story at the beginning of Chapter 24.

## Exercises 21

# 22 PL proofs: conditionals

We started Chapter 18 by saying (in effect) 'It would be nice if we can treat 'if' as another truth-functional connective. For then we'll be able to carry over our familiar apparatus of truth-table testing, etc.' However, in Chapter 19 (and in its Exercises), we noted real problems with simply identifying 'if' and the truth-functional '→'. Though we did end that chapter by suggesting that, all the same, '→' will serve as a cleaned-up *substitute* for the ordinary conditional for many purposes.

Let's now start from a different thought: 'It would be nice if we can treat the conditional as another connective governed by intuitively compelling inference rules, and can extend our now familiar Fitch-style proof system to cover arguments involving conditionals. Perhaps this will give us a more appealing way of handling the logic of 'if'!'

So set aside for a moment the official truth-functional semantics for the arrow symbol in PL languages. Keep the same syntax, but regard the symbol '→' just as standing in for a conditional governed by the two natural-seeming rules of inference we are about to introduce. Where does this take us?

## 22.1 Rules for the conditional

(a) As we saw in Chapter 18, the fundamental principle for arguing *from* a conditional is (MP), modus ponens. So here is one inference rule we'll certainly want for '→':

> *Modus ponens*: Given $\alpha$ and $(\alpha \to \gamma)$, we can infer $\gamma$.

What about arguing *to* a conditional proposition? Well, here is Jo considering a particular given triangle $\Delta$. She wants to prove the conditional *if* $\Delta$ *is isosceles, then* $\Delta$*'s base angles are equal*. What does Jo do? She supposes for the sake of argument that $\Delta$ is isosceles ('*Let* $\Delta$ *be isosceles*, . . . '). Then, on the basis of that supposition, she establishes that $\Delta$*'s base angles are equal*. And she takes it that doing this warrants the desired conditional assertion.

The inferential principle that Jo is relying on is this:

> Suppose that – given some background assumptions – we can argue from the additional temporary supposition $A$ to the conclusion $C$. Then – given the same background – we can infer *if A then C*.

In fact, in ordinary reasoning, we often hardly notice the difference between (i) making a supposition $A$ and drawing a conclusion $C$ and (ii) asserting a conditional *if A then C*.

So now for a formal version of this inferential principle:

> *Conditional proof*: Given a finished subproof starting with the temporary supposition $\alpha$ and ending $\gamma$, we can infer $(\alpha \to \gamma)$.

Immediately putting these two rules into diagrammatic form, we have:

> *Rules for the conditional*
>
> (MP) $$\begin{array}{c} \alpha \\ \vdots \\ (\alpha \to \gamma) \\ \vdots \\ \gamma \end{array}$$
>
> (CP) $$\begin{array}{l} \quad\left|\begin{array}{l} \alpha \\ \hline \vdots \\ \gamma \end{array}\right. \\ (\alpha \to \gamma) \end{array}$$

We can think of (CP) as our →-introduction rule. What then is the matching →-elimination rule that extracts from a conditional wff just what the introduction rule requires us to put in? (CP) grounds $(\alpha \to \gamma)$ in a proof from $\alpha$ to $\gamma$; so the corresponding elimination rule should say that, given $(\alpha \to \gamma)$, then if we do have $\alpha$ we should be able to conclude $\gamma$. Which is exactly what (MP) says.

(b) The various intuitively valid inferences involving conditionals that we met in §18.4 can now be warranted by proofs using our new conditional rules plus our existing rules for conjunction, disjunction, and negation.

Start with the modus tollens inference

**A** $(\mathsf{P} \to \mathsf{Q}), \neg\mathsf{Q} \therefore \neg\mathsf{P}$.

The proof strategy should be obvious – we want to prove the negation of something, and so we appeal to (RAA), as in:

| | | | |
|---|---|---|---|
| (1) | $(\mathsf{P} \to \mathsf{Q})$ | | (Prem) |
| (2) | $\neg\mathsf{Q}$ | | (Prem) |
| (3) | | $\mathsf{P}$ | (Supp) |
| (4) | | $\mathsf{Q}$ | (MP 3, 1) |
| (5) | | $\bot$ | (Abs 4, 2) |
| (6) | $\neg\mathsf{P}$ | | (RAA 3–5) |

Here we annotate the application of (MP) in the predictable way.

And note that the proof-strategy here generalizes: whenever we have $(\alpha \to \gamma)$ and $\neg\gamma$, we can of course use a proof of the same shape to derive $\neg\alpha$. It would therefore be redundant to add a separate modus tollens rule to our repertoire of rules for the conditional.

Next, we noted that the informal counterpart of the inference

**B** $(\mathsf{P} \to \mathsf{Q}), (\mathsf{Q} \to \mathsf{R}) \therefore (\mathsf{P} \to \mathsf{R})$

is intuitively correct for simple conditionals. And here is how to derive the conclusion from the premisses, using our conditional rules. The proof strategy is the default one when the target conclusion is a conditional. We temporarily assume the antecedent for the sake of argument and then aim for the consequent, so we can use the introduction rule (CP).

| | | |
|---|---|---|
| (1) | $(\mathsf{P} \to \mathsf{Q})$ | (Prem) |
| (2) | $(\mathsf{Q} \to \mathsf{R})$ | (Prem) |
| (3) | $\quad \mathsf{P}$ | (Supp) |
| (4) | $\quad \mathsf{Q}$ | (MP 3, 1) |
| (5) | $\quad \mathsf{R}$ | (MP 4, 2) |
| (6) | $(\mathsf{P} \to \mathsf{R})$ | (CP 3–5) |

The way we annotate the application of the conditional proof rule (CP) by giving the extent of the relevant subproof is also just as you would expect.

Next consider the contraposition inference (i.e. the inference from a conditional to its contrapositive) in

**C** $(\mathsf{P} \to \mathsf{Q}) \therefore (\neg\mathsf{Q} \to \neg\mathsf{P})$.

Again, we are aiming to prove a conditional. So again the strategy is to temporarily suppose the antecedent '$\neg\mathsf{Q}$' is true and aim to derive to consequent '$\neg\mathsf{P}$' with a view to using (CP). But this just sets us off on the same modus tollens inference as in **A** after line (2):

| | | |
|---|---|---|
| (1) | $(\mathsf{P} \to \mathsf{Q})$ | (Prem) |
| (2) | $\quad \neg\mathsf{Q}$ | (Supp) |
| (3) | $\quad\quad \mathsf{P}$ | (Supp) |
| (4) | $\quad\quad \mathsf{Q}$ | (MP 3, 1) |
| (5) | $\quad\quad \bot$ | (Abs 4, 1) |
| (6) | $\quad \neg\mathsf{P}$ | (RAA 2–5) |
| (7) | $(\neg\mathsf{Q} \to \neg\mathsf{P})$ | (CP 2–6) |

And now for the version of proof by cases that we met in §18.4:

**D** $(\mathsf{P} \lor \mathsf{Q}),\ (\mathsf{P} \to \mathsf{R}),\ (\mathsf{Q} \to \mathsf{R}) \therefore \mathsf{R}$.

To argue from the disjunctive first premiss, we have to use ($\lor$E). And then the proof goes in the predictable way.

| | | |
|---|---|---|
| (1) | $(\mathsf{P} \lor \mathsf{Q})$ | (Prem) |
| (2) | $(\mathsf{P} \to \mathsf{R})$ | (Prem) |
| (3) | $(\mathsf{Q} \to \mathsf{R})$ | (Prem) |
| (4) | $\quad \mathsf{P}$ | (Supp) |
| (5) | $\quad \mathsf{R}$ | (MP 4, 2) |
| (6) | $\quad \mathsf{Q}$ | (Supp) |
| (7) | $\quad \mathsf{R}$ | (MP 6, 3) |
| (8) | $\mathsf{R}$ | ($\lor$E 1, 4–5, 6–7) |

Let's have one more example in this initial group. So consider the inference

**E** $(P \to (Q \to R))$ ∴ $(Q \to (P \to R))$.

This inference intuitively *ought* to be valid (why?). So how do we derive the conclusion from the premiss?

The target conclusion is a conditional, so the obvious thing to do is assume the antecedent 'Q' and aim for the consequent '$(P \to R)$', preparing to use (CP). Our new target then is *another* conditional, and to prove this we again assume its antecedent 'P' and aim for its consequent 'R', preparing to invoke (CP) again.

With that plan in mind, the proof is then easily completed:

| | | |
|---|---|---|
| (1) | $(P \to (Q \to R)$ | (Prem) |
| (2) | $\quad Q$ | (Supp) |
| (3) | $\quad\quad P$ | (Supp) |
| (4) | $\quad\quad (Q \to R)$ | (MP 3, 1) |
| (5) | $\quad\quad R$ | (MP 2, 4) |
| (6) | $\quad (P \to R)$ | (CP 3–5) |
| (7) | $(Q \to (P \to R))$ | (CP 2–6) |

## 22.2 More proofs with conditionals

Let's work through a few more proofs – starting with two desirable results, but ending in a perhaps unexpected place.

(a) In §19.2, we met the following two informal arguments (keeping the same labels):

**X**(i) The claim *if A then C* rules out having *A* true and *C* false. So *if A then C* implies *it isn't the case that both A and not-C*.

**Y**(i) Suppose *if A then C*. So we either have *not-A*, or we have *A* and hence *C*. So *if A then C* implies *either not-A or C*.

Using our new derivation rules for conditionals, we can now provide proofs for corresponding formal inferences, as in:

**F** $(P \to Q)$ ∴ $\neg(P \land \neg Q)$
**G** $(P \to Q)$ ∴ $(\neg P \lor Q)$.

The strategy for the first proof is straightforward; we just follow the implied reductio in the informal argument **X**(i):

| | | |
|---|---|---|
| (1) | $(P \to Q)$ | (Prem) |
| (2) | $\quad (P \land \neg Q)$ | (Supp) |
| (3) | $\quad P$ | (∧E 2) |
| (4) | $\quad Q$ | (MP 3, 1) |
| (5) | $\quad \neg Q$ | (∧E 2) |
| (6) | $\quad \bot$ | (Abs 4, 5) |
| (7) | $\neg(P \land \neg Q)$ | (RAA 2–6) |

The second proof is a little more interesting. The informal argument in **Y**(i) quietly appeals to the law of excluded middle – but we won't meet the formal version of that until the next chapter. So let's instead proceed as follows: we temporarily suppose the opposite of the conclusion, and aim to reduce that to absurdity. So then our proof should have the shape

| | |
|---|---|
| $(\mathsf{P} \to \mathsf{Q})$ | (Prem) |
|     $\neg(\neg\mathsf{P} \lor \mathsf{Q})$ | (Supp) |
|     $\vdots$ | |
|     $\bot$ | |
| $\neg\neg(\neg\mathsf{P} \lor \mathsf{Q})$ | (RAA) |
| $(\neg\mathsf{P} \lor \mathsf{Q})$ | (DN) |

How do we join up the dots? We can't yet apply a conditional rule to the initial premiss, nor can we apply a negation rule to the supposition. And it would be pointless to conjoin them or to disjoin either with some other wff. There's only one sensible thing to do! Namely, make *another* supposition. The obvious thing to try is assuming '$\mathsf{P}$', which sets up a modus ponens inference. Things then go easily enough:

| | | |
|---|---|---|
| (1) | $(\mathsf{P} \to \mathsf{Q})$ | (Prem) |
| (2) |     $\neg(\neg\mathsf{P} \lor \mathsf{Q})$ | (Supp) |
| (3) |         $\mathsf{P}$ | (Supp) |
| (4) |         $\mathsf{Q}$ | (MP 3, 1) |
| (5) |         $(\neg\mathsf{P} \lor \mathsf{Q})$ | ($\lor$I 4) |
| (6) |         $\bot$ | (Abs 5, 2) |
| (7) |     $\neg\mathsf{P}$ | (RAA 2–6) |
| (8) |     $(\neg\mathsf{P} \lor \mathsf{Q})$ | ($\lor$I 7) |
| (9) |     $\bot$ | (Abs 8, 2) |
| (10) | $\neg\neg(\neg\mathsf{P} \lor \mathsf{Q})$ | (RAA 2–9) |
| (11) | $(\neg\mathsf{P} \lor \mathsf{Q})$ | (DN 10) |

(b) As we noted before, **X**(i) is just a version of the basic falsehood condition (FC) from §18.2. Hence any good rules for arguing with a conditional should certainly allow us to show that '$(\mathsf{P} \to \mathsf{Q})$' entails '$\neg(\mathsf{P} \land \neg\mathsf{Q})$' (and should also allow us to show that it entails the equivalent '$(\neg\mathsf{P} \lor \mathsf{Q})$'). So far, then, so good.

But what about the other two informal arguments we met in §19.2, namely

**X**(ii) Suppose we are given that *it isn't the case that both A and not-C*. Then we can infer that if *A* is actually true we can't have *not-C* as well: in other words *if A then C*.

**Y**(ii) Suppose we are given *either not-A or C*. Then if not the first, then the second. So we can infer *if A then C*.

Well, these two arguments can also be mirrored using our formal derivation rules for conditionals.

For example, we can warrant the following inferences:

**H** $\neg(\mathsf{P} \wedge \neg\mathsf{Q}) \therefore (\mathsf{P} \to \mathsf{Q})$

**I** $(\neg\mathsf{P} \vee \mathsf{Q}) \therefore (\mathsf{P} \to \mathsf{Q})$.

For the first of these we can argue simply as follows:

```
(1)  │ ¬(P ∧ ¬Q)                 (Prem)
     ├──
(2)  │    │ P                    (Supp)
     │    ├──
(3)  │    │    │ ¬Q              (Supp)
     │    │    ├──
(4)  │    │    │ (P ∧ ¬Q)        (∧I 2, 3)
(5)  │    │    │ ⊥               (Abs 4, 1)
(6)  │    │ ¬¬Q                  (RAA 2–6)
(7)  │    │ Q                    (DN 7)
(8)  │ (P → Q)                   (CP 2–8)
```

And for the second we just echo the informal disjunctive syllogism in **Y**(ii):

```
(1)  │ (¬P ∨ Q)                  (Prem)
     ├──
(2)  │    │ P                    (Supp)
     │    ├──
(3)  │    │    │ ¬P              (Supp)
     │    │    ├──
(4)  │    │    │ ⊥               (Abs 2, 3)
(5)  │    │    │ Q               (EFQ 4)
     │    │
(6)  │    │    │ Q               (Supp)
     │    │    ├──
(7)  │    │    │ Q               (Iter 5)
(8)  │    │ Q                    (∨E 1, 3–5, 6–7)
(9)  │ (P → Q)                   (CP 2–8)
```

Do note that the application of (∨E) here is quite correct, even though the subproofs are indented more than one column in from the original disjunction at line (1). All that matters is that, at line (8), '$(\neg\mathsf{P} \vee \mathsf{Q})$' plus the two subproofs are indeed all available.

(c) Put together the trivial proof from '$\neg\mathsf{P}$' to '$(\neg\mathsf{P} \vee \mathsf{Q})$' with a proof for **I** and we get a proof warranting the inference

**J** $\neg\mathsf{P} \therefore (\mathsf{P} \to \mathsf{Q})$.

Compare §19.4.

Two comments on this. First, note that there is a much quicker proof for this inference, again using explosion:

```
(1)  │ ¬P                        (Prem)
     ├──
(2)  │    │ P                    (Supp)
     │    ├──
(3)  │    │ ⊥                    (Abs 2, 1)
(4)  │    │ Q                    (EFQ 3)
(5)  │ (P → Q)                   (CP 2–4)
```

But second, note that explosion is not really of the essence here. Suppose we had adopted the arguably more natural liberalized (∨E) rule suggested in passing in §21.4(b). We could then have proved **I** – and hence proved **J** – without relying on explosion.

## 22.3 The material conditional again

(a) In §18.4, we saw that the material conditional satisfies the modus ponens principle (MP). In §18.7, we saw that, if $\alpha$ together with some background premisses entails $\gamma$, then those premisses entail $(\alpha \to \gamma)$, where the arrow is specifically the material conditional. So the material conditional satisfies the conditional proof principle (CP).

It is no surprise that the material conditional obeys (MP) and (CP). What is perhaps more surprising is a converse point. Suppose we add a conditional (symbolized by the arrow) to our Fitch-style PL logical system for the other connectives. Suppose we initially require only that this conditional is governed by the two intuitively compelling principles (MP) and (CP). Then, generalizing from the examples we met in the last section, a wff of the form $(\alpha \to \gamma)$ will imply and be implied by the corresponding truth-functional wff $\neg(\alpha \land \neg\gamma)$. Similarly, $(\alpha \to \gamma)$ will imply and be implied by $(\neg\alpha \lor \gamma)$.

We wondered in the preamble to this chapter whether we could avoid the problems of identifying 'if' with the material conditional by introducing a conditional into our formal logic via intuitively appealing inference rules rather than via a truth-table definition. But it now seems that this is no help, at least if we start from a standard propositional logic for the other connectives. Add to that a conditional '→' obeying the natural rules (MP) and (CP), and '→' will end up behaving exactly like the material conditional again.

(b) We can return, then, to the remarks in §19.6 about adopting the material conditional. Nearly all mathematicians cheerfully accept the usual logic for the other connectives. And mathematicians reason all the time using a conditional which unrestrictedly obeys (CP) as well as (MP) – in particular, it is absolutely standard in informal mathematical reasoning to prove a conditional by assuming the antecedent and aiming for the consequent. Therefore, these mathematicians are reasoning with a conditional which behaves like the material conditional. That's why, as we said, mathematics textbooks will often *explicitly* adopt the material conditional, announcing that this is indeed what is officially meant by (their symbol for) 'if'.

## 22.4 Summary

We add now to our PL proof apparatus inference rules for the conditional – namely (MP) and (CP), i.e. modus ponens and conditional proof. Once more, the inference rules we are adding to our system seem intuitively very natural.

However, it quickly turns out that these additional rules make a wff of the form $(\alpha \to \gamma)$ interderivable with the corresponding wffs $\neg(\alpha \land \neg\gamma)$ and $(\neg\alpha \lor \gamma)$. In other words, at least when added to our core PL system, '→' will behave like the material conditional.

## Exercises 22

# 23 PL proofs: theorems

A proof in our Fitch-style system has to start from *something*. But it doesn't have to start from a permanent premiss. A mere temporary supposition is good enough. This short chapter illustrates this simple but important new idea.

## 23.1 Theorems

(a) Consider the following little proof. We start with *no* premisses, but make a temporary supposition for the sake of argument, and then continue to get proof **A**:

| | | |
|---|---|---|
| (1) | | |
| (2) | │ │ $(\mathsf{P} \land \neg\mathsf{P})$ | (Supp) |
| (3) | │ │ $\mathsf{P}$ | ($\land$E) |
| (4) | │ │ $\neg\mathsf{P}$ | ($\land$E) |
| (5) | │ │ $\bot$ | (Abs ) |
| (6) | │ $\neg(\mathsf{P} \land \neg\mathsf{P})$ | (RAA 2–5) |

Let's adopt a standard bit of terminology:

> A wff which is provable in our Fitch-style proof system from *no* premisses is said to be a *theorem* of the system.

Then we have just shown that one formal instance of Aristotle's Law of Non-Contradiction is a theorem. And the proof strategy can be generalized: *any* wff of the form $\neg(\alpha \land \neg\alpha)$ can similarly be shown to be a theorem. Put any wff $\alpha$ for '$\mathsf{P}$' throughout the proof, and we will get a derivation of the corresponding $\neg(\alpha \land \neg\alpha)$ from no premisses.

Is this some dubious sleight of hand, as we apparently pull truths out of an empty logical hat? No. In fact, on reflection, our result is exactly what we would hope for! We can intuitively establish that any instance of Aristotle's Law is true just by reflecting on how the connectives appear in it. So the rules governing the connectives *ought* to be enough by themselves to generate instances of the Law.

(b) We can also show that an instance of the Law of Excluded Middle like '$(\mathsf{P} \lor \neg\mathsf{P})$' is a theorem.

How will the proof go? We can't get the disjunctive conclusion using ($\lor$I), as we plainly can't get either '$\mathsf{P}$' or '$\neg\mathsf{P}$' from no premisses. So strategically our only hope is to suppose the opposite, '¬(P ∨ ¬P)', and aim for a contradiction. But now note that we can't usefully apply our negation or disjunction rules to this supposition either: so it seems that we will need to make *another* supposition in order sensibly to move on.

In fact, the simplest further supposition will do the trick. Here is how, **B**:

| | | |
|---|---|---|
| (1) | │ | |
| (2) | │ │ ¬(P ∨ ¬P) | (Supp) |
| (3) | │ │ │ P | (Supp) |
| (4) | │ │ │ (P ∨ ¬P) | (∨I 3) |
| (5) | │ │ │ ⊥ | (Abs 4, 2) |
| (6) | │ │ ¬P | (RAA 3–5) |
| (7) | │ │ (P ∨ ¬P) | (∨I 6) |
| (8) | │ │ ⊥ | (Abs 7, 2) |
| (9) | │ ¬¬(P ∨ ¬P) | (RAA 2–8) |
| (10) | │ (P ∨ ¬P) | (DN 9) |

Check that you understand how the proof almost writes itself. Once again, the proof-idea generalizes. Any wff of the form $(\alpha \vee \neg\alpha)$ is a theorem.

(c) Now for a couple of examples involving conditionals. Firstly, we will show that '(P → (Q → (P ∧ Q)))' is a theorem – as you would expect, since it is in general a trivial logical truth that if *A* is true, then if *B* is true as well, we will have *A and B*! Here's the easy proof, **C**:

| | | |
|---|---|---|
| (1) | │ | |
| (2) | │ │ P | (Supp) |
| (3) | │ │ │ Q | (Supp) |
| (4) | │ │ │ (P ∧ Q) | (∧I 2, 3) |
| (5) | │ │ (Q → (P ∧ Q)) | (CP 3–4) |
| (6) | │ (P → (Q → (P ∧ Q))) | (CP 2–5) |

That was easy! But our next example is trickier. Take the wff

**D** (((P → Q) → P) → P)

This is an instance of what is known as Peirce's Law. If we interpret '→' as the material conditional then, as can be easily checked, it is a tautology.

In our proof system, then, we will evidently need a proof of the following shape:

| | |
|---|---|
| │ | |
| │ │ ((P → Q) → P) | (Supp) |
| │ │ ⋮ | |
| │ │ P | |
| │ (((P → Q) → P) → P) | (CP) |

But how do we fill in the dots? We will need to apply (MP) if we are to make any use of the temporary supposition. But that means we will need to first derive the antecedent of our supposition, i.e. '$(\mathsf{P} \to \mathsf{Q})$'. And how are we going to arrive at *that*?

Well, we can make another supposition '$\neg\mathsf{P}$', and then derive '$(\mathsf{P} \to \mathsf{Q})$' as in the short proof for **J** in the last chapter. We then get the following (non-obvious!) proof:

| | | |
|---|---|---|
| (1) | | |
| (2) | $((\mathsf{P} \to \mathsf{Q}) \to \mathsf{P})$ | (Supp) |
| (3) | $\neg\mathsf{P}$ | (Supp) |
| (4) | $\mathsf{P}$ | (Supp) |
| (5) | $\bot$ | (Abs 4, 3) |
| (6) | $\mathsf{Q}$ | (EFQ 5) |
| (7) | $(\mathsf{P} \to \mathsf{Q})$ | (CP 4, 6) |
| (8) | $\mathsf{P}$ | (MP 7, 2) |
| (9) | $\bot$ | (Abs 8, 3) |
| (10) | $\neg\neg\mathsf{P}$ | (RAA 3–9) |
| (11) | $\mathsf{P}$ | (DN 101) |
| (12) | $(((\mathsf{P} \to \mathsf{Q}) \to \mathsf{P}) \to \mathsf{P})$ | (CP 2–11) |

It is interesting to note that this theorem of our system only involves the conditional; yet we do need to use more than our two conditional rules in order to prove it. (In particular, we have invoked (DN) at line (9): and it can be shown that we *have* to use (DN) or some equivalent to derive instances of Peirce's Law in our proof system.)

## 23.2 Derived rules

(a) Let's take one more example of a theorem. Back in §14.1 we saw that

**E** $((\mathsf{P} \land \mathsf{Q}) \lor (\neg\mathsf{P} \lor \neg\mathsf{Q}))$

is a PL logical truth. So we will want to be able to prove it from no premisses. Here is an outline proof:

| | | |
|---|---|---|
| (1) | | |
| (2) | $((\neg\mathsf{P} \lor \neg\mathsf{Q}) \lor \neg(\neg\mathsf{P} \lor \neg\mathsf{Q}))$ | (Excluded middle, as in **B**) |
| (3) | $(\neg\mathsf{P} \lor \neg\mathsf{Q})$ | (Supp) |
| (4) | $((\mathsf{P} \land \mathsf{Q}) \lor (\neg\mathsf{P} \lor \neg\mathsf{Q}))$ | ($\lor$I 3) |
| (5) | $\neg(\neg\mathsf{P} \lor \neg\mathsf{Q})$ | (Supp) |
| (6) | $(\mathsf{P} \land \mathsf{Q})$ | (De Morgan's Law from 5, as in **G** §21.3) |
| (7) | $((\mathsf{P} \land \mathsf{Q}) \lor (\neg\mathsf{P} \lor \neg\mathsf{Q}))$ | ($\lor$I 6) |
| (8) | $((\mathsf{P} \land \mathsf{Q}) \lor (\neg\mathsf{P} \lor \neg\mathsf{Q}))$ | ($\lor$E 2, 3–4, 6–8) |

Filling in the details, that's a twenty-five line proof. But we have kept things a lot shorter by allowing ourselves to 'plug in' derivations from two earlier proofs.

(b) Picking up that last point, here is a general observation. Suppose we want to cope with increasingly messy PL proofs. Then:

(i) We can generalize from particular proofs that we have already produced, and build up a library of *derived rules* of inference (rules that must be truth-preserving if our original rules are).

For example, we could officially adopt the Law of Excluded Middle, as described in the next section. Or we could adopt versions of De Morgan's Law like *From* $\neg(\neg\alpha \vee \neg\beta)$ *you can infer* $(\alpha \wedge \beta)$. Or here's a Disjunctive Syllogism rule: *From* $(\alpha \vee \beta)$ *and* $\neg\alpha$ *you can infer* $\beta$.

(ii) We can then invoke these derived rules in order to speed up new proofs.

And that indeed is how introductory logic texts often proceed.

But not us. For we are not going to make a fetish of the ultimately tedious business of producing lots of complex PL proofs. Once you have got the hang of how to use our formal logics here and later in the book, and understand their role as ideal models of explicit rigour, then there is little point in making it easier to produce more and yet more proofs inside our various systems. As stressed before in §20.5, understanding the basic motivations and guiding principles of our logical systems is *much* more important.

## 23.3 Excluded middle and double negation

Having mentioned theorems and derived rules, we finish the chapter with a simple observation. Consider the following two rules of inference, one old, one new:

*Double negation* (DN): Given $\neg\neg\alpha$, we can infer $\alpha$.

*Law of excluded middle* (LEM): For any wff $\alpha$, we can add $(\alpha \vee \neg\alpha)$ at any step in a proof.

Now, (DN) is built into our proof system as a basic rule. (LEM) isn't. But we have just seen how any instance of $(\alpha \vee \neg\alpha)$ *can* already be proved at any point by using (DN) (along with other rules). In other words, we could add (LEM) to our system as a derived rule, and it would make no difference to what we can establish.

Suppose, however, that we *hadn't* adopted the rule (DN), but had *instead* adopted the rule (LEM) as basic alongside the other rules. Then, in this revised system, we have the following short proof, by disjunctive syllogism:

| | | |
|---|---|---|
| (1) | $\neg\neg$P | (Prem) |
| (2) | (P ∨ ¬P) | (LEM) |
| (3) |     P | (Supp) |
| (4) |     P | (Iter 3) |
| (5) |     ¬P | (Supp) |
| (6) |     ⊥ | (Abs 5, 1) |
| (7) |     P | (EFQ 6) |
| (8) | P | (∨E 2, 3–4, 5–7) |

And of course this proof idea generalizes too. In our revised system, we can still always get from $\neg\neg\alpha$ (whether as a premiss or as a wff in the middle of a proof) to $\alpha$ by using (LEM) and other background rules. So in our new system (DN) would be a derived rule.

In short, taking the other rules – i.e. (Iter), (EFQ) and the pairs of introduction and elimination rules – as fixed background, *it doesn't matter whether we adopt (DN) as our additional negation-involving rule or adopt (LEM) instead*. We will say something about the significance of this observation at the end of the next chapter.

## 23.4 Summary

A theorem of our proof system is a wff which can be proved from no premisses at all. Intuitively, theorems will be logical truths.

We saw that, given the other inferences, having the rule (DN) in place is equivalent to having the rule (LEM) which allows us to invoke an instance of the law of excluded middle at any point in a proof.

## Exercises 23

# 24 PL proofs: metatheory

We have now worked through a range of proofs *inside* our Fitch-style natural deduction system. In this chapter we mostly stand *outside* this system in order to note some important results about it. In a word, we do some . . .

## 24.1 Metatheory

We have already remarked on some small metatheoretic results. For example, having shown that we could argue from '(P ∧ Q)' to '(Q ∧ P)' in our initial proof system, we then noted that – by replacing 'P' and 'Q' with $\alpha$ and $\beta$ throughout the little proof – we can similarly get a proof from any wff $(\alpha \wedge \beta)$ to $(\beta \wedge \alpha)$. So we concluded that it would be redundant to add to our system a further rule (∧C) allowing us directly to swap the order of conjuncts in a conjunction – see §20.1(d). This redundancy result is, of course, not itself a *theorem* in the sense of a wff derived inside our PL proof system: it is a *metatheorem* about our system.

For another example, we have just proved the metatheorem that, keeping the other rules fixed, we could replace the rule (DN) with the rule (LEM) without changing what can be proved in the system.

As we will see, however, the most basic metatheorems about our Fitch-style proof system are these:

> Our proof system is *sound* in the sense that, if we can get from (zero or more) PL premisses to a PL conclusion by a derivation following our rules of inference, then the conclusion really *is* tautologically entailed by the premisses. In other words, a 'proof' works as advertised, as a genuine proof.
>
> Our proof system is also *complete* in the sense that we don't need to add any more inference rules for it to capture all tautological entailments. If a PL wff is tautologically entailed by some PL premisses, then we can get to that conclusion via a derivation from the premisses, using our proof system as it stands without further augmentation.

Our Fitch-style system is certainly not the only sound and complete one for arguing in PL languages with their truth-functional connectives. As we have stressed before, there are other frameworks for proof-building on the market. And even within a natural deduction framework, there are various choices for the fundamental rules. (Still, for brevity's sake, we will continue to refer to a derivation in our particular Fitch-style natural deduction system as simply a 'PL proof'.)

## 24.2 Putting everything together

Before discussing soundness and completeness further, our first task must be to put together a specification for our proof system in a rather tidier way. We bring together in one place the rules of inference that have been scattered through the last four chapters. And then we will officially specify the ways in which we are allowed to chain inference steps together to form proofs. (We are putting our logical house in order: but like other housekeeping, it is unavoidably a bit tedious!)

(a) First, then, the rules of inference again. We start with the rule which allows us to simply repeat ourselves when it is useful to do so:

(Iter) At any inference step, we can reiterate any available wff.

Next we have a rule governing the absurdity sign, plus a group of eight rules that come in introduction/elimination pairs. (We explained at the end of §20.2(d) why we can think of the negation rules as forming a pair of this kind. And we explained at the end of §22.1(a) why we can also think of the conditional rules as forming another introduction/elimination pair.) As usual, $\alpha$, $\beta$ and $\gamma$ are arbitrary wffs, and then:

(EFQ) Given '$\bot$', we can infer any wff $\alpha$.
(RAA) Given a finished subproof starting with the temporary supposition $\alpha$ and concluding '$\bot$', we can infer $\neg\alpha$.
(Abs) Given $\alpha$ and $\neg\alpha$, we can infer '$\bot$'.
($\wedge$I) Given $\alpha$ and $\beta$, we can infer $(\alpha \wedge \beta)$.
($\wedge$E) Given $(\alpha \wedge \beta)$, we can infer $\alpha$. Given $(\alpha \wedge \beta)$, we can infer $\beta$.
($\vee$I) Given $\alpha$, we can infer $(\alpha \vee \beta)$ for any $\beta$. Given $\beta$, we can infer $(\alpha \vee \beta)$ for any $\alpha$.
($\vee$E) Given $(\alpha \vee \beta)$, a finished subproof from $\alpha$ as supposition to $\gamma$ and also a finished subproof from $\beta$ as supposition to $\gamma$, then we can infer $\gamma$.
(CP) Given a finished subproof starting with the temporary supposition $\alpha$ and ending $\gamma$, we can infer $(\alpha \to \gamma)$.
(MP) Given $\alpha$ and $(\alpha \to \gamma)$, we can infer $\gamma$.

Finally, we have the additional negation rule – an outlier, which is independent of the basic introduction/elimination pairs:

(DN) Given $\neg\neg\alpha$, we can infer $\alpha$.

In each case, the given inputs for the application of a rule of inference must be *available* – i.e. occur earlier in the proof, but not inside an already finished subproof.

(b) Now we review how to put together inference steps in order to construct complete PL proofs. We need to record some preliminary decisions (in order to pin down a few details).

(i) Line numbers on the left of a derivation, and commentary (in abbreviated English) on the right, are helpful optional extras, but won't count as part of the proof itself. But what about the vertical lines marking columns of reasoning? And what about the horizontal bars separating premisses and temporary suppositions from what follows? These do carry important information, so let's retain them. So when we talk about 'columns' we will take these as marked as before by a vertical line, one for the whole 'home' column, and others lasting for the duration of a subproof.

(ii) After we start a proof by giving some premisses, or start a subproof with a temporary supposition, we must make at least one further step!

(iii) We can only finish a proof overall when we are in its *home* column, i.e. the column which started with the original premisses, if any. (In other words, we don't get a properly completed proof if we stop in the middle of a subproof, with some undischarged temporary supposition left dangling.)

(iv) We now repeat a point that we have made before. Consider the following:

```
(1)  | P                  (Prem)
(2)  | Q                  (Prem)
     |---
(3)  |      | R           (Supp)
     |      |---
(4)  |      | (P ∧ R)     (∧I 1, 3)
(5)  | (P ∧ Q)            (∧I 1, 2)
```

Here we have a correctly formed subproof which, however, turns out to be an irrelevant digression. We finish the subproof, but then make no later use of it in reaching our final conclusion.

Now, we *could* require proofs to be written tidily and economically, so that the *only* subproofs which appear along the way are ones which are actually used as inputs to later rules of inference. And there are indeed natural deduction systems which in effect rule out such redundant detours. However, we can afford to be relaxed about this. After all, digressions may be inefficient: but we do not ordinarily regard them as fallacious. Hence we will allow subproofs to be completed and then, in effect, abandoned. So our example here will count as a correctly formed (even if digressive) proof. For us, being a proof is one thing: being an elegantly minimal proof is something else!

With these (very minor) clarifications in place, we can now give an official summary of the principles for building PL proofs. We take it as understood that a proof in our system is a snaking vertical line of wffs, arranged in marked columns and jumping at most one column left or right at each step. Then:

A PL proof begins with zero or more wffs as premisses, entered at the top of the *home* (left-most) column of the proof, followed by a horizontal bar.

Then at each further step we do one of the following (keeping going until we eventually end the proof):

(1) We remain in the current column, and apply a PL rule of inference to add a wff. In applying a rule, *the 'given' inputs to the rule must be available*, i.e. must occur earlier but *not* inside an already finished subproof.

(2) We move one column to the right and add a wff (which becomes a new temporary supposition) followed by a horizontal bar. This starts a new subproof.

(3) We move one column to the left (only if we are *not* in the home column). This ends the current subproof.

(4) We end the whole proof (only if we *are* in the home column).

These moves are subject to the following constraints:

(i) After stating the premisses or stating a temporary supposition, and drawing a horizontal line, we must next do (1) or (2).

(ii) After (3), moving a column left, we must also next do (1) or (2).

The first constraint simply stops us putting down premisses or a temporary supposition and doing *nothing* with them. The second constraint stops us having proofs or subproofs which end with no final wff.

Just check through some of our earlier examples, to convince yourself that they do indeed conform to our now official proof-building rules.

(c) One important remark. Look again at our rules of inference and at our rules for putting them together in forming PL proofs. These rules are of course *motivated* by thoughts about truth-preservation (we do want a sound proof system). But the rules nowhere *mention* anything semantic; they just describe permissible symbol-shuffling moves. A computer can be programmed to check whether an array of symbols constitutes a PL proof, simply by checking the syntactic form of the expressions involved. In short:

It is a purely syntactic matter whether an array of wffs and absurdity signs is put together in such a way as to form a properly constructed PL proof.

## 24.3 Vacuous discharge

The general guiding ideas behind our PL proof system are, we can agree, very natural and attractive. This section is about one bothersome point of detail.

(a) Consider, for example, the following derivation, **A**:

| | | | |
|---|---|---|---|
| (1) | P | | (Prem) |
| (2) | ¬P | | (Prem) |
| (3) | | ¬Q | (Supp) |
| (4) | | ⊥ | (Abs 1, 2) |
| (5) | ¬¬Q | | (RAA 3–4) |
| (6) | Q | | (DN 5) |

Our (Abs) rule tells us that given ‘P’ and ‘¬P’ we can declare an absurdity. At line (4), both ‘P’ and ‘¬P’ *are* available. So we can indeed correctly put down ‘⊥’ at that line. Our (RAA) rule tells us that, given an available subproof starting with the temporary supposition ‘¬Q’ and concluding ‘⊥’, we can discharge the supposition and infer ‘¬¬Q. At line (5) just such a subproof *is* available, so (RAA) allows us to infer ‘¬¬Q’ at that line. Hence, with a final (DN) step, we get another proof in our system warranting the explosive inference ‘P, ¬P ∴ Q’.

You might, however, be rather suspicious about *this* derivation. For note that the supposition ‘¬Q’ at line (3) is not actually *used* in deriving the absurdity at line (4). But isn’t the idea behind a reductio proof that, if an assumption *A leads to* or *generates* a contradiction, we can infer *not-A*? By contrast, our official (RAA) rule allows what is known by the unprepossessing name of ‘vacuous discharge’. Which means that we are allowed to discharge the temporary supposition $\alpha$ at the top of a subproof ending in absurdity and infer $\neg\alpha$, *even if $\alpha$ is not involved in deriving the absurdity*.

(b) Here is an even simpler derivation, **B**:

| | | |
|---|---|---|
| (1) | Q | (Prem) |
| (2) | \| P | (Supp) |
| (3) | \| Q | (Iter 1) |
| (4) | (P → Q) | (CP 2–3) |

To infer a wff of the form $(\alpha \to \gamma)$ by (CP), all we need is an available finished subproof starting $\alpha$ and ending $\gamma$; and that’s what we have here at line (4).

Again, you might be suspicious. The supposition ‘P’ at line (2) in **B** is not actually used in deriving (3), the final wff ‘Q’ of the subproof. But isn’t the idea behind a ‘conditional proof’ that, if an assumption *A* leads to or generates a derived conclusion *C*, then we are entitled to infer i*f A then C*. By contrast, our official (CP) rule also allows vacuous discharge. Which means that we are allowed to discharge the temporary supposition $\alpha$ at the top of a subproof ending in $\gamma$ and derive $(\alpha \to \gamma)$, *even if $\alpha$ is not involved in deriving $\gamma$*.

(c) So the question arises: is allowing vacuous discharge as in these cases a bug in our proof system? Or is it an acceptable feature?

Concentrating on the simpler second example, it might be suggested that we should make the conditional proof rule stricter, as follows:

($CP_2$) Given an available subproof starting with the temporary supposition $\alpha$ and ending with $\gamma$, *where $\alpha$ is actually used in deriving $\gamma$*, then we can infer $(\alpha \to \gamma)$.

Consider, however, the following proof, **C**:

| | | |
|---|---|---|
| (1) | Q | (Prem) |
| (2) | \| P | (Supp) |
| (3) | \| (P ∧ Q) | (∧I 2,1) |
| (4) | \| Q | (∧E, 3) |
| (5) | (P → Q) | ($CP_2$ 2–4) |

Here the wff officially appealed to in deriving (4), namely (3), is itself derived from (2). So we do still get a proof even with our revised rule ($CP_2$).

In sum, a proof such as **B** using (CP) and vacuous discharge can only too easily be massaged into a proof such as **C** using ($CP_2$) without vacuous discharge.

(d) To avoid the revised rule ($CP_2$) in effect collapsing back into our original rule (CP), should we *also* try to block the sort of two-step dance involved at lines (3) and (4) in **C**? Shall we ban using an introduction rule and then simply 'undoing' the result by using the corresponding elimination rule? The trouble with this suggestion is that such little detours strike us as redundant rather than wrong, so an outright ban would initially seem to be hard to motivate. Moreover it turns out that, in general, the business of 'normalizing' proofs to iron out all such detours is non-trivial.

But we will not try to tangle any further with complications like these. For the headline news is that other standard natural deduction systems also permit vacuous discharge when applying rules using temporary assumptions like (RAA), (∨E) and (CP), so this isn't just a feature of our system. And this feature doesn't enable us to prove anything that we can't prove otherwise, so it isn't a vicious bug. (For example, as an alternative to **B**, we can of course derive '$(\mathsf{P} \to \mathsf{Q})$' from '$\mathsf{Q}$' by going via '$(\neg\mathsf{P} \lor \mathsf{Q})$' – how?)

The path of least resistance, then, is simply to continue to allow vacuous discharge of 'unused' temporary suppositions as a special case.

## 24.4 Generalizing proofs

Now back to more important general issues. We pick up again an important point which we have already frquently noted: PL proofs generalize.

Suppose we take an array of PL expressions which are interrelated by our rules of inference so as to form a correctly formed proof. And suppose we systematically swap out the atoms in this proof for other wffs (obeying the crucial rule – same atom, same replacement throughout). Then we will get an array of new wffs.

But now note that nothing in our rules for PL proof-building gives a special role to atomic wffs. So given that the old wffs in the old array were originally related to each other in the right way to make them a proof, then – because the pattern of relationships stays fixed – the new wffs in the new array will still be related in the right way to make them a proof. In simpler words, any derivation built up the same way as our original proof will also be a correctly formed proof.

As an aside, note that we can put the same point like this, highlighting a parallel with §16.4 on generalizing tautological entailments. Take a correctly formed PL proof. Systematically replace its atoms with schematic variables (same atom, same replacement; different atoms, different replacements). Call the result a *proof schema*. Then

> Any substitution instance of a proof schema is also a proof.

We get a substitution instance, of course, by systematically replacing schematic variables with wffs again.

## 24.5 '⊨' and '⊢'

(a) Recall some standard metalinguistic notation from §14.2 and §16.3. Assuming the $\alpha$s and $\gamma$ are all wffs of some PL language, then:

> We abbreviate the claim that the premisses $\alpha_1, \alpha_2, \ldots, \alpha_n$ tautologically entail the conclusion $\gamma$ as follows: $\alpha_1, \alpha_2, \ldots, \alpha_n \vDash \gamma$.
>
> We abbreviate the claim that $\gamma$ is a tautology as follows: $\vDash \gamma$.

We now add some new, equally standard, notation:

> We abbreviate the claim that there is a PL proof from the premisses $\alpha_1, \alpha_2, \ldots, \alpha_n$ to the conclusion $\gamma$ as follows: $\alpha_1, \alpha_2, \ldots, \alpha_n \vdash \gamma$.
>
> We abbreviate the claim that $\gamma$ is a PL theorem – i.e. the claim that $\gamma$ can be derived from zero premisses – as follows: $\vdash \gamma$.

The old double turnstile is the *semantic* turnstile, defined in terms of the semantic property of being a truth-preserving inference. Correspondingly, the new single turnstile is the *syntactic* turnstile, defined in terms of the syntactic property of being a well-constructed PL proof. So the turnstiles have quite different definitions.

(b) Still, despite their different definitions, it is easy to see that the semantic and syntactic turnstiles march in step in many particular cases. For example, we know from §§14.2, 16.4 and 17.1 respectively that for any PL wffs $\alpha, \beta, \gamma$, we have the semantic facts

(1) $\vDash (\alpha \lor \neg\alpha)$,
(2) $(\alpha \lor \beta), \neg\alpha \vDash \beta$,
(3) $\alpha, \neg\alpha \vDash \gamma$.

And we now know from §§23.1, 21.4, and 20.7 respectively that these are exactly matched by the syntactic facts that for any $\alpha, \beta, \gamma$,

(1′) $\vdash (\alpha \lor \neg\alpha)$,
(2′) $(\alpha \lor \beta), \neg\alpha \vdash \beta$,
(3′) $\alpha, \neg\alpha \vdash \gamma$.

Likewise, for example, we saw in Exercises 16 that the following two principles hold for tautological entailment:

(4) If $\alpha \vDash \gamma$, then $\alpha, \beta \vDash \gamma$,
(5) If $\alpha \vDash \beta$ and $\beta \vDash \gamma$, then $\alpha \vDash \gamma$.

We have results about provability:

(4′) If $\alpha \vdash \gamma$, then $\alpha, \beta \vdash \gamma$,
(5′) If $\alpha \vdash \beta$ and $\beta \vdash \gamma$, then $\alpha \vdash \gamma$.

For the first, just note that adding an extra premiss to a correctly formed proof still gives us a correctly formed proof (inserting an unused premiss doesn't ruin things). For the second, we just take a derivation from $\alpha$ to $\beta$ and splice that together with a proof from $\beta$ to $\gamma$ to get a proof from $\alpha$ from $\gamma$.

Again in §18.7(c), we showed that

(6) $\alpha \vDash \gamma$ if and only if $\vDash (\alpha \to \gamma)$.

Similarly, we have

(6′) $\alpha \vdash \gamma$ if and only if $\vdash (\alpha \to \gamma)$.

We can leave it as an exercise to check that last claim.

## 24.6 Soundness and completeness

(a) We won't consider any more examples of how facts about provability match facts about tautological entailment case by case, because we can show that the two sorts of fact must *always* march in step.

As announced at the beginning of the chapter, we have the following two metatheorems. First, and crucially,

> Our PL proof system is *sound*: if $\alpha_1, \alpha_2, \ldots, \alpha_n \vdash \gamma$, then $\alpha_1, \alpha_2, \ldots, \alpha_n \vDash \gamma$.
>
> As a special case, if $\vdash \gamma$ then $\vDash \gamma$.

So our proof system is indeed trustworthy: a completed derivation of some wff from given premisses is indeed always a *proof*, showing that the conclusion really does follow from the premisses. And if a wff is a theorem, then it is indeed always logically true. (It is only a very mild annoyance that 'sound' is used in elementary logic in two different ways – to say of a particular argument that it has true premisses and makes a valid inference move, and to say of a formal proof system that its demonstrations are reliably truth-preserving. This double use of the word is unlikely to cause any confusion.)

Secondly, we have the converse result:

> Our PL proof system is *complete*: if $\alpha_1, \alpha_2, \ldots, \alpha_n \vDash \gamma$, then $\alpha_1, \alpha_2, \ldots, \alpha_n \vdash \gamma$.
>
> As a special case, if $\vDash \gamma$ then $\vdash \gamma$.

In other words, there are no 'missing rules' in our PL system. Given a tautologically valid PL inference, we can always derive its conclusion from the premisses by a derivation in the system as it stands.

(b) Let's comment on these results. First, then, on the idea of soundness.

So far we have only introduced *one* family of formal languages, PL languages. And we defined a suitable notion of validity-in-virtue-of-logical-structure for inferences framed in a PL language, namely tautological validity. We have since introduced a Fitch-style framework for proving conclusions from premisses in a PL language. The aim of this PL proof system all along has been to enable us to establish that various inferences are truth preserving without going through the palaver of a truth-table test. And certainly, we have chosen proof-building rules that intuitively *seem* reliable. So it would be Very Bad

News if our proof system can after all lead us astray and purport to validate inferences that aren't valid! Hence we do essentially require our PL proof system to be sound.

We are soon going to introduce another family of formal languages, QL languages, for regimenting quantificational inferences. We will eventually define a suitable notion of validity-in-virtue-of-logical-structure for inferences framed in a QL language, namely q-validity. And we will extend our Fitch-style framework to enable us to prove conclusions from premisses in a QL language. Again, we will want to be able to rely on the so-called proofs – i.e. we will require every derivation of a conclusion from premisses to be q-valid, making our QL proof system sound in the relevant sense.

And so it goes. Beyond the scope of this book, there are other families of logical languages with different types of semantics, and there will be correspondingly different definitions of validity-in-virtue-of-logical-structure for inferences in such languages. We will usually want to design deductive systems for arguing formally in such languages. There will be many options other than Fitch-style systems. But we will *always* require our deductive systems to be sound, i.e. we require a 'proved' conclusion to validly follow from the given premisses (in the appropriate sense of validity).

(c) Completeness, however, is another matter. Of course, it is great if we *can* find a nice tidy formal proof-system for a given class of logical languages in which it is possible to derive all the valid inferences. We can do this for tautologically valid inferences in PL languages. We can, it turns out, also do this for q-valid inferences in QL languages.

But we can't do it across the board. Roughly speaking, some logical languages are too 'infinitary' for the relevant notion of validity to be fully captured by a finitely specifiable formal deductive system. However, that has to be a story for another time, in some more advanced discussion. For the moment, we will just say: completeness in a proof system is highly desirable, but is not always obtainable once we move beyond the elementary logics which are our topic in this book.

(d) The main thing you should take away from this chapter is an *understanding* of the claims that our PL proof system is sound and is complete, in the senses just explained. You do not need to know *how we establish* these claims to be true.

In fact, a soundness proof is relatively easy. Say that a line on a proof is *good* if the wff on that line is tautologically entailed by the premisses and suppositions available at that line. Then we can show that a proof starts with a good line and that every proof-building rule takes us from some previous good line(s) to another good line. So the last line of a proof must be good too. But at the last line there can be no undischarged suppositions. So this means that at the last line of a proof, the initial premisses must indeed tautologically entail the concluding wff. To fill out this argument-sketch, we have to check case-by-case that applications of the various proof-building rules preserves goodness – which is tedious but not difficult. We give details in an Appendix.

A completeness proof is harder. But we give details of two interestingly different ways of proving the completeness of our PL system in the Appendix.

Still, soundness and completeness proofs arguably sit exactly on the borderline between introductory and more advanced logic. So to repeat, for our purposes in the rest of this book, you can very safely skip the proofs.

## 24.7 Excluded middle again

We finish this chapter – and our discussion of propositional logic – by revisiting the double negation rule and its equivalent, the law of excluded middle. As we will see, these have a distinctive status that sets them apart from the other PL rules, an observation which has considerable philosophical interest.

(a) Start from the PL proof system which we codified in §24.2. Now simply omit the rule (DN). Call the resulting cut-down system IPL – the reason for the label will become clear in a moment.

In the light of §23.3, adding the rule (DN) to IPL is equivalent to adding the rule (LEM) which allows us to invoke an instance of the law of excluded middle at any stage in a proof. And the point we want to emphasize now is that (DN) or (LEM) *are* substantive additions to IPL, enabling us to derive conclusions that do not follow if we can only use the rules of the cut-down system. To put it another way, (DN) – or equivalently, (LEM) – is independent of the other rules of the original system PL.

(b) How do we *prove* the independence result? Here is a strategy. We find some new way of interpreting IPL and its connectives, an interpretation on which the rules of our cut-down deductive system would still be acceptable, but on which (DN) or (LEM) would plainly *not* be acceptable. It then follows that buying the rules of IPL can't commit us to (DN) or (LEM).

Now, as we said when we first introduced it, the (DN) rule reflects the classical understanding of negation as truth-value-flipping. So an interpretation on which (DN) *doesn't* hold will have to revise, in particular, the interpretation of negation. How can we do that?

It is natural, perhaps, to think of a correct assertion as one that corresponds to some realm of facts (whatever that means exactly). But suppose that we instead think of correctness as a matter of being informally provable or being fully justified or (for short) being *warranted*. And suppose that we think of a correctness-preserving inference as one that preserves warranted assertibility. Then here is a quite natural four-part story about how to characterize the connectives in this new framework:

(i) $(\alpha \wedge \beta)$ is correct, i.e. warranted, iff $\alpha$ and $\beta$ are both correct, i.e. are both warranted.

(ii) $(\alpha \vee \beta)$ is correct, i.e. warranted, iff at least one disjunct is correct, i.e. there is a warrant for $\alpha$ or a warrant for $\beta$.

(iii) A correct conditional $(\alpha \to \beta)$ must be one that, together with the warranted assertion $\alpha$, will enable us to derive another warranted assertion $\beta$ by using modus ponens. Hence $(\alpha \to \beta)$ is correct iff there is a way of converting a warrant for $\alpha$ into a warrant for $\beta$.

(iv) Finally, $\neg\alpha$ is correct iff we have a warrant for ruling out $\alpha$ as leading to something absurd.

We can still accept (EFQ) as essentially defining the absurdity constant – and then it will follow that the absurd can't be warrantedly assertible (or else everything would be assertible).

With the connectives understood this way, the familiar *introduction* rules for the connectives will *still* be acceptable, i.e. they will be warrant-preserving (think through why is this so). But – as we saw – the various elimination rules in effect just 'undo' the effects of the introduction rules: so they should come along for free once we have the introduction rules. Hence, thought of as a warrant-preserving system of rules, all our IPL rules can remain in place.

However (DN) will *not* be acceptable in this framework. We might have a good reason for ruling out being able to rule out $\alpha$, so we can warrantedly assert $\neg\neg\alpha$. But that doesn't put us in a position to warrantedly assert $\alpha$. We might just have to remain neutral.

Likewise (LEM) will *not* be acceptable. On the present understanding, $(\alpha \lor \neg\alpha)$ would be correct, i.e. warranted, just if there is a warrant for $\alpha$ or a warrant for $\neg\alpha$, i.e. a warrant for ruling out $\alpha$. But again, why should there always be a way of justifiably deciding a conjecture $\alpha$ in the relevant area of enquiry one way or the other? Some things may be beyond our ken.

Hence, if we are thinking of our proof system as a framework for arguments preserving warranted assertability, we shouldn't endorse (DN) or (LEM).

(Be *very* careful here! It is one thing to stand back from endorsing (LEM). It would be something else to actually accept some instances of the 'opposite' schema $\neg(\alpha \lor \neg\alpha)$. In fact, in IPL, any instance of $\neg(\alpha \lor \neg\alpha)$ leads to outright contradiction!)

(c) So now we see the significance of the fact that the (DN) rule is an outlier, not one of the introduction/elimination rules for the connectives. Its special status leaves room for an interpretation on which the remaining rules – the rules of IPL – hold good, but (DN) doesn't. Which is what we wanted to show.

True, our version of the argument might seem a bit rough-and-ready; after all, the notions of warrant or informal proof are not ideally clear. But let's not fuss about that. For we can in fact develop a rigorous story inspired by these notions which gives us an entirely uncontroversial technical proof that (DN) and its equivalents are, as claimed, independent of the other rules of PL.

Things *do* get controversial, when it is claimed that in some particular domain of enquiry (DN) and (LEM) really don't apply, because in this domain there can indeed be no more to correctness than warrant or informal provability. For example, so-called *intuitionists* hold that mathematics is a case in point. Mathematical truth, they say, doesn't consist in correspondence with facts about objects laid out in some Platonic heaven (after all, there are familiar worries: what kind of objects could these ideal mathematical entities be? how could we possibly know about them?). Rather, the story goes, being mathematically correct is a matter of being assertible on the basis of a proof – meaning not a proof in this or that formal system but any proof satisfying informal mathematical standards.

This view is radical and highly contentious – taken seriously, it involves having to reject swathes of standard mathematics, because it changes the logic we can use. For an intuitionist, the appropriate propositional logic for arguing with the connectives is not the full *classical* two-valued logic PL where (LEM) holds, but rather the cut-down *intuitionist* logic we have already labelled IPL because this is the right logic for correctness-as-informal-provability.

(d) In this introductory book, we obviously can't even begin to discuss those intriguing issues about the nature of truth and provability in mathematics. But we have said enough to raise some horribly difficult questions. Are there really domains where we can't apply the full classical PL apparatus (even when we continue to set aside issues of vagueness)? More generally, is there One True Logic, or should we be pluralists allowing different principles to govern reasoning in different areas of enquiry? Is a pluralist approach consistent with whatever is true in the thought that logic should be topic-neutral?

Here, however, we must set such issues aside, and we will just have to follow absolutely standard practice and concentrate on the *classical* logic which includes the law of excluded middle, reflecting the classical understanding of negation as truth-value-flipping. For this is the logic which is routinely deployed by mathematicians – along with the rest of us, at least most of the time, when we can set aside issues about vagueness or troublesome arguments about paradoxical sentences (like the Liar, see §10.3). Classical logic is, so to speak, the baseline against which alternative logics are usually measured: so it is – on almost everyone's view – the logic you need to understand first.

## 24.8 Summary

We have reviewed our Fitch-style proof system, restating the inference rules, and now specifying the permissible ways of assembling inferences together into structured proofs.

In fixing the details of our system, various choices have had to be made. For example, we allow 'stray' unused subproofs to clutter up a proof. We noted in particular that we allow 'vacuous discharge'.

We defined the syntactic turnstile '$\vdash$' (indicating provability) to go alongside the semantic turnstile '$\vDash$' (indicating entailment). We noted particular cases where properties of '$\vdash$' match properties of '$\vDash$'.

In fact, our PL proof system is sound and complete. A conclusion derived from some premisses by a PL proof is indeed tautologically entailed by those premisses. And conversely, a conclusion tautologically entailed by some premisses can be derived from them by a PL proof.

Detailed proofs of these claims are in an Appendix on 'Soundness and completeness for PL proofs'.

## Exercises 24