Types of proof system

Peter Smith

October 13, 2010

1 Axioms, rules, and what logic is all about

1.1 Two kinds of proof system

There are at least *two* styles of proof system for propositional logic other than trees that beginners ought to know about.

The real interest here, of course, is *not* in learning yet more about classical *propositional logic* per se. For how much fun is *that*? Rather, what we are doing – as always – is illustrating some Big Ideas using propositional logic as a baby example.

The two new styles are:

1. *Axiomatic systems* The first system of formal logic in anything like the contemporary sense – Frege's system in his *Begriffsschrift* of 1879 – is an *axiomatic* one.

   What is meant by 'axiomatic' in this context? Think of Euclid's geometry for example. In such an axiomatic theory, we are given a "starter pack" of basic assumptions or *axioms* (we are often given a package of supplementary *definitions* as well, enabling us to introduce new ideas as abbreviations for constructs out of old ideas). And then the *theorems* of the theory are those claims that can deduced by allowed moves from the axioms, perhaps invoking definitions where appropriate.

   In Euclid's case, the axioms and definitions are of course geometrical, and he just helps himself to whatever logical inferences he needs to execute the deductions. But, if we are going to go on to axiomatize logic itself, we are going to need to be explicit not just about the basic logical axioms we take as given starting points for deductions, but also about the rules of inference that we are permitted to use.

   Frege's axiomatization of logic in fact goes further than familiar ('first-order'[1]) quantification theory: but its first-order fragment is equivalent to standard modern systems (though it is presented using a very idiosyncratic – and pretty unreadable – notation that never caught on). Likewise, Russell and Whitehead's system in their *Principia Mathematica* of 1910–1913 goes beyond first-order logic: it is given in another axiomatic presentation, which is *symbolically* quite a bit nicer than Frege, though a significant step backwards in terms of real rigour. Again, Hilbert and Ackermann present an axiomatic system in the first book which begins to have something like the 'look and feel' of a modern logic text, namely their rather stunning short book *Mathematical Logic* (original German edition 1928). And Hilbert and Ackermann do nicely isolate standard first-order quantificational logic, under the label 'the restricted predicate calculus'.[2]

   As we'll see, however, axiomatic systems of logic are *not* very nice to work with,

[1] First-order logic allows quantifications just over objects, and has variables that can replace names; second-order quantification allows quantifications over properties and has variables that can replace predicates.

[2] I hope your logical curiosity will prompt you to get each of those off the library shelf and browse through for just a few minutes! You can find Frege's little book reproduced in J. van Heijenoort ed. *From Frege to Gödel* and also in T. W. Bynum ed. *Frege: Conceptual Notation*, from p. 101 onwards.

David Hilbert was the greatest mathematician of his generation. His 1928 based on earlier lectures consequently had a very profound influence.

and don't reflect the natural ways of reasoning used in common-or-garden real-world deductive reasoning (e.g. in mathematical proofs).

2. *Natural deduction systems* By contrast natural deduction systems are intended to reflect much more closely those natural ways of reasoning, by encoding the sort of inference moves that we ordinarily make, and giving us a regimented framework for setting out arguments using these moves.

   As we'll see, there are various ways of setting out natural deductions. But such systems all share the core features of (i) having no axioms and lots of rules of inference (rather than having lots of axioms and just one or two rules of inference), and – crucially – (ii) allowing suppositional proofs, in which some proposition is temporarily assumed for the sake of argument and then the assumption is later 'discharged' (as, for example, in a *reductio* proof).

   A natural deduction system in pretty much the modern form was proposed by Gerhard Gentzen in his doctoral thesis of 1933.

1.2 What logic is about

The historical move from axiomatic systems to natural deduction systems is not just of formal interest: it reflects a change in the conception of what logic is fundamentally about.

Frege thinks of logic as a science, a body of truths governing a special subject matter (logical concepts such as identity, etc.). And in *Begriffsschrift* §13, he extols the general procedure of axiomatizing a science to reveal how a bunch of laws hang together: 'we obtain a small number of laws [the axioms] in which . . . is included, though in embryonic form, the content of all of them'. So it is not surprising that Frege takes it as appropriate to present logic axiomatically too.

In a significantly different way to Frege, Russell also thought of logic as science, as being in the business of systematizing the most general truths about the world. A special science like chemistry tells us truths about certain kinds of constituents of the world and certain of their properties; logic tells us absolutely general truths about *everything*. If you think like *that*, treating logic as (so to speak) the most general science, then of course you'll again be inclined to regiment logic as you do other scientific theories, ideally by laying down a few 'basic laws' and then showing that other general truths follow.

Famously, Wittgenstein in the *Tractatus* reacted radically against Russell's conception of logic. For him, truth-functional tautologies and other logical truths are indeed *tautologies* – not deep ultimate truths about the most general, logical, structure of the universe, but rather *empty* claims in the sense that they tell us nothing informative about how the world is: they merely fall out as byproducts of the meanings of the basic logical particles.

That last idea can be developed in more than one way. But one approach is Gentzen's. Think of the logical connectives as getting their meanings from how they are used in inference (so grasping their meaning involves grasping the inference rules governing their use). For example, grasping 'and' involves grasping, inter alia, that from 'A and B' you can (of course!) derive A. Similarly, grasping the conditional involves grasping, inter alia, that a derivation of the conclusion C from the temporary supposition A warrants an assertion of 'if A then C'. And now those trivial logical inference rules enable us to derive 'for free' (so to speak) various truths. For instance, suppose that *A and B*; then we can derive A (by the 'and' rule). And reflecting on that little suppositional inference, we see that a rule of inference governing 'if' entitles us to assert *if A and B, then A*. If that is right, and if the point generalizes, then we just don't have to see such logical truths as reflecting deep facts about the logical structure of the world (whatever that could mean): logical truth falls out just as a byproduct of the inference rules whose applicability is, in some sense, built into the very meaning of e.g. the connectives and the quantifiers.

However, whether or not you fully buy that story about the nature of logical truth, there surely is something odd about thinking of systematized logic as primarily aiming to regiment a special class of ultra-general *truths*. Isn't logic, rather, at bottom about good and bad reasoning practices, about what makes for a good proof? Shouldn't its prime concern be the correct styles of valid inference? So isn't it more natural for a formalized logic to highlight *rules of proof-building* rather than stressing *axiomatic truths*? (Of course, we mustn't make *too* much of this contrast: for corresponding e.g. to the inferential rule 'from A you can infer (A or B)' there is indeed the logical truth 'if A then (A or B)'. So we can, up to a point, trade off rules for corresponding truths – although of course we need to keep at least one rule of inference if we are ever to be able to infer anything from anything! Still, there is something natural about the suggestion that the study of inference should come first in logic.)

2 Logic in an axiomatic style

2.1 *M*, a sample axiomatic logic

Let's have an example of an axiomatic system to be going on with. In this system M, to be found e.g. in Mendelson's classic *Introduction to Mathematical Logic*, the only propositional connectives built into the basic language of the theory are '$\to$' and '$\neg$' ('if ... then ...' and 'not', the same choice of basic connectives as in Frege's *Begriffsschrift*). The axioms are then all wffs which are instances of the schemata[3]

Ax1. $(A \to (B \to A))$
Ax2. $((A \to (B \to C)) \to ((A \to B) \to (A \to C)))$
Ax3. $((\neg B \to \neg A) \to ((\neg B \to A) \to B))$

M's one and only rule of inference is *modus ponens*, the rule from A and $(A \to C)$ infer C.

In this axiomatic system, a proof from the given premisses $A_1, A_2, \ldots A_n$ to conclusion C is a linear sequence of wffs such that each wff is either (i) a premiss A_i, (ii) a logical axiom, i.e. an instance of Ax1, Ax2 or Ax3, or (iii) follows from two previous wffs in the sequence by modus ponens, and (iv) the final wff in the sequence if C. When there is such a proof we will write $A_1, A_2, \ldots A_n \vdash_M C$.

If we can prove C from the axioms alone without additional premisses, we'll say that C is a theorem, and write simply $\vdash_M C$.

2.2 System *M* in action

We'll first show that $(\mathsf{P} \to \mathsf{P})$ is a theorem, i.e. $\vdash_M (\mathsf{P} \to \mathsf{P})$. Here's the proof:

1.	$((\mathsf{P} \to ((\mathsf{P} \to \mathsf{P}) \to \mathsf{P})) \to ((\mathsf{P} \to (\mathsf{P} \to \mathsf{P})) \to (\mathsf{P} \to \mathsf{P})))$	by Ax2
2.	$(\mathsf{P} \to ((\mathsf{P} \to \mathsf{P}) \to \mathsf{P}))$	by Ax1
3.	$((\mathsf{P} \to (\mathsf{P} \to \mathsf{P})) \to (\mathsf{P} \to \mathsf{P}))$	from 2, 1 by MP
4.	$(\mathsf{P} \to (\mathsf{P} \to \mathsf{P}))$	Ax1
5.	$(\mathsf{P} \to \mathsf{P})$	from 4, 3 by MP

Next, let's show that $(\mathsf{P} \to \mathsf{Q}), (\mathsf{Q} \to \mathsf{R}) \vdash_{\mathsf{M}} (\mathsf{P} \to \mathsf{R})$:

1.	$(\mathsf{P} \to \mathsf{Q})$	premiss
2.	$(\mathsf{Q} \to \mathsf{R})$	premiss
3.	$((\mathsf{Q} \to \mathsf{R}) \to (\mathsf{P} \to (\mathsf{Q} \to \mathsf{R})))$	by Ax1
4.	$(\mathsf{P} \to (\mathsf{Q} \to \mathsf{R}))$	from 2, 3 by MP
5.	$((\mathsf{P} \to (\mathsf{Q} \to \mathsf{R}) \to ((\mathsf{P} \to \mathsf{Q}) \to (\mathsf{P} \to \mathsf{R})))$	by Ax2

[3] 'Instances of the schemata'? An wff is an instance of a schema (plural: schemata) if it results for systematically replacing schematic letters A, B, C etc. in the schema by particular wffs – with, in a given case, the same schematic letter always being substituted by the same wff. But you knew that!

6. $((\mathsf{P} \to \mathsf{Q}) \to (\mathsf{P} \to \mathsf{R}))$ from 4, 5 by MP
7. $(\mathsf{P} \to \mathsf{R})$ from 1, 6 by MP

I take it that neither of those proofs is wonderfully obvious or natural. For another example, consider $(\mathsf{P} \to (\mathsf{Q} \to \mathsf{R})) \vdash_M (\mathsf{Q} \to (\mathsf{P} \to \mathsf{R}))$. Again, that's an intuitively obvious validity, and can be checked by a quick tree proof (do it!). But the shortest known M-proof has twenty-one lines, starting of course

1. $(\mathsf{P} \to (\mathsf{Q} \to \mathsf{R}))$ premiss

and then going (I kid you not!) via

13. $$(((\mathsf{Q} \to ((\mathsf{P} \to \mathsf{Q}) \to (\mathsf{P} \to \mathsf{R}))) \to ((\mathsf{Q} \to (\mathsf{P} \to \mathsf{Q})) \to (\mathsf{Q} \to (\mathsf{P} \to \mathsf{R})))) \to (((\mathsf{Q} \to ((\mathsf{P} \to \mathsf{Q}) \to (\mathsf{P} \to \mathsf{R}))) \to (\mathsf{Q} \to (\mathsf{P} \to \mathsf{Q}))) \to ((\mathsf{Q} \to ((\mathsf{P} \to \mathsf{Q}) \to (\mathsf{P} \to \mathsf{R}))) \to (\mathsf{Q} \to (\mathsf{P} \to \mathsf{R})))))$$

which, if you look at it *very* hard and count brackets, is in fact an instance of the axiom schema Ax1. But even with that hint, I take it that it isn't in the slightest bit obvious how actually to complete the proof.

2.3 Augmenting the system

Let's now augment this basic system M with definitions for the other two familiar basic connectives to get the system M' where we can write $(A \ \wedge \ B)$ as an abbreviation for $\neg(A \to \neg B)$ and $(A \vee B)$ for $(\neg A \to B)$. A proof in M' is like a proof in M except that we can use the definitions to decode and to introduce abbreviations for M-wffs using the new connectives.

As a ludicrously nasty challenge, you might like to try to show that in the extended proof system, we have $(\mathsf{P} \wedge \mathsf{Q}) \vdash_{M'} \mathsf{P}$. You know that the proof is going to have to start

1. $(\mathsf{P} \wedge \mathsf{Q})$ premiss
2. $\neg(\mathsf{P} \to \neg\mathsf{Q})$ by definition of '$\wedge$'

But now where? How do you get from $\neg(\mathsf{P} \to \neg\mathsf{Q})$ to the desired conclusion P? It can be done, but as far as I know it takes well over *fifty* lines (if done from first principles, without appealing to any previously-established results about the system M).

Evidently, then, in the axiomatic systems M and M', even the most simple and intuitively primitive results about valid inference – e.g. that $(\mathsf{P} \wedge \mathsf{Q})$ entails P – can have not-at-all-simple and not-at-all-intuitive proofs which are very remote inded from "natural" ways of arguing for the validities concerned. Still, *ease-of-use* wasn't the prime concern of the founding fathers: they were much more concerned with showing in principle how the best mathematical standards of *rigour* could be brought to logical systems, and showing how the body of logical truths hangs together in virtue of flowing from a small number of self-evident assumptions.

2.4 The deduction theorem

Actually, M isn't quite as unusable as you might think at first acquaintance. True, M-proofs are ploddingly linear; in M, we can't do the natural thing of making suppositions 'for the sake of argument' and then discharging them (about which more in a moment). But there's a sort of surrogate for that. For we can show the following metalogical result – it is called the *deduction theorem*.

If $A_1, A_2, \ldots A_n, B \vdash_M C$ then $A_1, A_2, \ldots A_n \vdash_M (B \to C)$.

This says that if there is an M-proof from $A_1, A_2, \ldots A_n$ and the extra assumption B to the conclusion C, then there is also always an M-proof (differently structured but still using just the initial rules of M) from $A_1, A_2, \ldots A_n$ to $(B \to C)$. There is a parallel result for M'.

With the deduction theorem to hand, go back to the problem of showing that $(\mathsf{P} \to (\mathsf{Q} \to \mathsf{R})) \vdash_M (\mathsf{Q} \to (\mathsf{P} \to \mathsf{R}))$. It is trivially easy to show that

$$(\mathsf{P} \to (\mathsf{Q} \to \mathsf{R})), \mathsf{Q}, \mathsf{P} \vdash_M \mathsf{R}$$

since that conclusion immediately follows from the premisses by two applications of MP. Then one appeal of the deduction theorem tells us that

$$(\mathsf{P} \to (\mathsf{Q} \to \mathsf{R})), \mathsf{Q} \vdash_M (\mathsf{P} \to \mathsf{R})$$

and then another appeal tells us that

$$(\mathsf{P} \to (\mathsf{Q} \to \mathsf{R})) \vdash_M (\mathsf{Q} \to (\mathsf{P} \to \mathsf{R}))$$

That is to say, two applications of the deduction theorem tell us that there must *be* such a proof, though note it *doesn't* explicitly tell us what the proof is.

Frequent such appeals to the deduction theorem will evidently make working with a theory like M much more manageable. Standard textbooks dealing with axiomatic systems tell you how to prove the deduction theorem for a system like M.

2.5 *M* is sound and complete

The logical system M is *sound* and *complete* with respect to the usual truth-functional semantics S for '$\to$' and '$\neg$'. That is to say, if there's a proof from the premisses $A_1, A_2, \ldots A_n$ to conclusion C in system M, then the inference $A_1, A_2, \ldots A_n \therefore C$ is indeed tautologically valid using the truth-functional interpretations of the connectives given in S (the proof-system is sound, i.e. reliable). And if the inference $A_1, A_2, \ldots A_n \therefore C$ is tautologically valid there is indeed a proof from the premisses $A_1, A_2, \ldots A_n$ to conclusion C in system M (the proof-system is complete, i.e. covers all the tautologically valid inferences featuring '$\to$' and/or '$\neg$').

Actually, in advanced work, we'll for technical reasons want to consider the case which allows an infinite number of premisses to be given: but for present introductory purposes let's forget about that!

In symbols, then, soundness and completeness correspond to the two directions of the biconditional (which direction is which?):

$$A_1, A_2, \ldots A_n \vdash_M C \text{ if and only if } A_1, A_2, \ldots A_n \vDash_S C.$$

(where of course '$\vDash$' indicates tautological entailment).

Similarly, the extended axiomatic system M' is sound and complete given the truth-functional semantics S' for its four connectives: or in symbols

$$A_1, A_2, \ldots A_n \vdash_{M'} C \text{ if and only if } A_1, A_2, \ldots A_n \vDash_{S'} C.$$

Now compare: the tree-system T (i.e. the system in *IFL*) was initially introduced via the idea of 'truth-tables done backwards', so it was no surprise at all that T is sound and complete (though that doesn't mean we don't have to check that T works as advertised). But M has been – so to speak – just plonked on the table. So this time we are going to have to work a bit harder to show that the system is sound and complete.

The appendix to this handout explores the needed soundness and completeness proofs. But NB that's just for mathmos and/or logic enthusiasts.

3 Natural deduction systems

3.1 Conditional proof

Let's reinforce the point about the way that the axiomatic system M doesn't reflect natural modes of inference by thinking again about how to argue with conditionals.

Now, as we've already remarked, a *natural* way of establishing a conditional conclusion of the form 'if A then C' is to argue along the following lines: 'Suppose for the sake of argument that A is true. Then blah, blah, blah. So (still on that supposition) C. But if we can show that C given the supposition that A, then that shows that if indeed A is true, then C.' Well, we *can't* argue like that inside M.

Let's put that suppositional mode of inference to work to show that the premisses $(\mathsf{P} \to \mathsf{Q})$ and $(\mathsf{Q} \to \mathsf{R})$ entail $(\mathsf{P} \to \mathsf{R})$. We can argue: Look, suppose for the sake of argument that P *is* true. Then, from that supposition plus the first premiss, we can deduce Q by modus ponens. And from *that* plus the second premiss, we can deduce R by modus ponens again. Hence we can show that R given the supposition that P: so that shows that $(\mathsf{P} \to \mathsf{R})$.

Setting that out symbolically, we could display the argument like this, arranging things *Fitch-style* (as in a classic 1952 logic text, *Symbolic Logic: an Introduction* by Frederic B. Fitch):

$$\begin{array}{lll}
1 & (\mathsf{P} \to \mathsf{Q}) & \text{premiss} \\
2 & (\mathsf{Q} \to \mathsf{R}) & \text{premiss} \\
3 & \mid\; \mathsf{P} & \text{supposition} \\
4 & \mid\; \mathsf{Q} & \text{from 3, 1 by MP} \\
5 & \mid\; \mathsf{R} & \text{from 4, 2 by MP} \\
6 & (\mathsf{P} \to \mathsf{R}) & \text{by the subproof from 3 to 5, by CP}
\end{array}$$

Here, when we make a new temporary supposition, we indent the line of argument to the right (and use a vertical line to mark the new column, and a short horizontal line under the new supposition). When we come to the end of the 'subproof' depending on a supposition and then 'discharge' the supposition, we move back a column to the left. In this case, the little subproof from lines 3 to 5 shows that the supposition P (given the other initial premisses) implies R. We can now stop making that supposition (we discharge it), and can assert straight out that *if* P *then* R. The inference move here – the move from giving a proof of R from the supposition P to asserting $(\mathsf{P} \to \mathsf{R})$ – is an instance of *conditional proof (CP)*.

More generally, then, an inference by conditional proof has the form

$$\begin{array}{ll}
\mid\; A & \text{supposition} \\
\mid\; \vdots & \\
\mid\; C & \\
(A \to C) & \text{by CP}
\end{array}$$

Three points about this. First, subproofs can be *nested*, one inside another. Here's a demonstration, using CP, of the entailment from $(\mathsf{P} \to (\mathsf{Q} \to \mathsf{R}))$ to $(\mathsf{Q} \to (\mathsf{P} \to \mathsf{R}))$ – which we said took twenty-one pretty unobvious lines in M.

$$\begin{array}{lll}
1 & (\mathsf{P} \to (\mathsf{Q} \to \mathsf{R})) & \text{premiss} \\
2 & \mid\; \mathsf{Q} & \text{supposition} \\
3 & \mid\; \mid\; \mathsf{P} & \text{supposition} \\
4 & \mid\; \mid\; (\mathsf{Q} \to \mathsf{R}) & \text{from 3, 1 by MP} \\
5 & \mid\; \mid\; \mathsf{R} & \text{from 2, 4 by MP} \\
6 & \mid\; (\mathsf{P} \to \mathsf{R}) & \text{by the subproof from 3 to 5, by CP} \\
7 & (\mathsf{Q} \to (\mathsf{P} \to \mathsf{R})) & \text{by the subproof from 2 to 6, by CP}
\end{array}$$

The layout should make it very clear what is going on here. We first suppose Q. Then, with that supposition still in play, we also suppose P. A couple of applications of modus ponens allow us to infer R. So, given that the supposition P implies R we can now cease to make *that* supposition and infer $(\mathsf{P} \rightarrow \mathsf{R})$ by CP. But that conclusion still depends on the supposition Q which is still in play. However, since that supposition Q implies $(\mathsf{P} \rightarrow \mathsf{R})$ we can now cease to make that supposition too, and infer $(\mathsf{Q} \rightarrow (\mathsf{P} \rightarrow \mathsf{R}))$ by CP again.

A second point: we can use suppositional reasoning to establish conclusions by CP which rest on no premisses at all. Consider again the informal reasoning to the conclusion 'if A and B, then A' in §1.2. Here is a corresponding Fitch-style proof:

1	$\quad\vert\ (\mathsf{P} \land \mathsf{Q})$	supposition
2	$\quad\vert\ \mathsf{P}$	from 1 by $\land$E
3	$((\mathsf{P} \land \mathsf{Q}) \rightarrow \mathsf{P})$	by the subproof from 1 to 2, by CP

where $\land$E is the obvious 'and-elimination' inference rule which allows us to infer a conjunct from a conjunction. Note the conclusion of this proof depends on *no* initial premisses – it is a theorem, expressing a logical truth.

Thirdly, how do we prove the obvious theorem $(\mathsf{P} \rightarrow \mathsf{P})$? We want something along the following lines:

1	$\quad\vert\ \mathsf{P}$	supposition
2	$\quad\vert\ \mathsf{P}$	trivially, since we have 1!
3	$(\mathsf{P} \rightarrow \mathsf{P})$	by the subproof from 1 to 2, by CP

One way of implementing this is to add a 'reiteration' rule to allow you to repeat a wff already in play. Or we could just allow the trivial case of a subproof whose first and last wff is one and the same. But for our purposes, we needn't fuss about such very fine details.

We'll return to consider systems with other ways of laying out proofs with rules like CP in a moment.

3.2 Reductio ad absurdum

Consider next another familiar everyday rule of inference which involves making and discharging temporary suppositions, namely reductio ad absurdum (RAA). Here, we show that some proposition is false by assuming the opposite for the sake of argument, and showing that that supposition leads to contradiction. So schematically, the pattern is this:

$\quad\vert\ A$	supposition
$\quad\vert\ \vdots$	
$\quad\vert\ \bot$	
$\neg A$	by RAA

where $\bot$ stands for 'absurdity'.

And what does 'absurdity' mean here? Well, this much is sure: if we have argued ourselves to both A and $\neg A$, then we've certainly arrived at an absurd upshot. We can encapsulate that in the so-called *absurdity rule* (Abs): given A and $\neg A$ you can write down '$\bot$'.

Now, there are a number of ways of fully implementing this idea. We could treat the absurdity constant here rather as we treat '✶' on a tree – i.e. just as signal that a contradiction has indeed been reached. On this treatment '$\bot$' isn't itself a wff, and so e.g. the expression '$(\mathsf{P} \rightarrow \bot)$' would be ill-formed. And read this way, the absurdity rule, while a rule of proof-construction, isn't strictly speaking a rule-of-*inference*.

But in fact, it is actually rather neat to treat the absurdity constant '$\bot$' as built into the language as a new genuine atomic wff (the *falsum*) which always evaluates as false, and which can embed in more complex formulae, and even feature as a premiss or conclusion in an argument. And on that understanding, the absurdity rule can be treated alongside other inference rules.

Here, using the reductio rule, is a proof from the premiss $(\mathsf{P} \to \mathsf{Q})$ to the conclusion $\neg(\mathsf{P} \land \neg\mathsf{Q})$:

1	$(\mathsf{P} \to \mathsf{Q})$	premiss
2	$\mid\ (\mathsf{P} \land \neg\mathsf{Q})$	supposition
3	$\mid\ \mathsf{P}$	from 2 by $\land$E
4	$\mid\ \neg\mathsf{Q}$	from 2 by $\land$E
5	$\mid\ \mathsf{Q}$	from 1,3 by MP
6	$\mid\ \bot$	from 4, 5 by Abs
7	$\neg(\mathsf{P} \land \neg\mathsf{Q})$	from 2 to 6, by RAA

And now let's prove the reverse, that from $\neg(\mathsf{P} \land \neg\mathsf{Q})$ we can infer $(\mathsf{P} \to \mathsf{Q})$:

1	$\neg(\mathsf{P} \land \neg\mathsf{Q})$	premiss
2	$\mid\ \mathsf{P}$	supposition
3	$\mid\ \mid\ \neg\mathsf{Q}$	supposition
4	$\mid\ \mid\ (\mathsf{P} \land \neg\mathsf{Q})$	from 2, 3 by $\land$I
5	$\mid\ \mid\ \bot$	from 1, 4 by Abs
6	$\mid\ \neg\neg\mathsf{Q}$	from 3 to 5 by RAA
7	$\mid\ \mathsf{Q}$	from 6, by DN
8	$(\mathsf{P} \to \mathsf{Q})$	from 2 to 7 by CP

Here we've called upon two new, but obvious, rules. First $\land$I is the 'and-introduction' rule that allows us to infer a conjunction given both its conjuncts. And DN of course is the 'double negation' rule that allows us to strip off initial double negations.[4]

3.3 Introduction and elimination rules

3.3.1 Introduction, elimination and harmony

So far we've met a pair of rules for '$\land$' – an introduction rule $\land$I, which tells us how to argue to a wff introducing a (new) occurrence of '$\land$' as its main connective, and an elimination rule telling us how to argue from a wff with '$\land$' as the main connective to a result lacking that occurrence of the connective.

Likewise we have met a pair of rules for '$\to$'. You can think of CP as an introduction rule, which tells us how to argue to a wff introducing a (new) occurrence of '$\to$' as its main connective. And MP is an elimination rule telling us how to argue from a wff with '$\to$' as the main connective to a result lacking that occurrence of the connective.

Now, in each of those cases, the elimination rule is in a certain sense *in harmony with* the introduction rule. Gentzen puts it like this:

> The introductions represent, as it were, the 'definitions' of the symbols concerned, and the eliminations are no more, in the final analysis, than the consequences of these definitions.

[4]Note an implication of this pair of proofs! Assuming two very obvious and natural rules for the conditional, CP and MP, then – assuming some other standard rules of inference – it turns out that $(\mathsf{P} \to \mathsf{Q})$ is interderivable with $\neg(\mathsf{P} \land \neg\mathsf{Q})$. The material conditional rules, OK!

That's obvious for the pair of '$\wedge$' rules: $\wedge$E just takes out again from a conjunction what $\wedge$I puts in. For the '$\rightarrow$' rules, think of it like this. The CP introduction rule tells us that the proper warrant for $(A \rightarrow C)$ is a derivation of C from A. So to be given A and $(A \rightarrow C)$ is in effect to be given A and told that there is a derivation of C from A, so of course that entitles you to derive C – thus warranting the elimination rule MP.

It is the same for the rules for '$\vee$'. First, there is the obvious introduction rule $\vee$I, which tells us that we can infer a disjunction from one of its disjuncts. Now suppose you are given $(A \vee B)$; then you know that you are entitled to at least one of A and B but not which. So how can you proceed from there? Well you can *argue by cases*. That is to say, you can argue: I'm given $(A \vee B)$, I don't know which. But suppose A, blah, blah, blah, so C. Suppose on the other hand B, mumble, mumble, mumble, so C. So, either way, C! This seems to be warranted by the introduction rule tells you about disjunction. Laying out that reasoning Fitch-style, the rule $\vee$E that apparently harmonizes with the introduction rule (we haven't strictly shown it is the only/best candidate) has the form

$(A \vee B)$	
│ A	supposition
│ ⋮	
│ C	
│ B	supposition
│ ⋮	
│ C	
C	by $\vee$E

To illustrate, we'll use the $\vee$-rules to show $(\mathsf{P} \wedge (\mathsf{Q} \vee \mathsf{R}))$ entails $((\mathsf{P} \wedge \mathsf{Q}) \vee (\mathsf{P} \wedge \mathsf{R}))$:

1	$(\mathsf{P} \wedge (\mathsf{Q} \vee \mathsf{R}))$	premiss
2	P	from 1, by $\wedge$E
3	$(\mathsf{Q} \vee \mathsf{R})$	from 1, by $\wedge$E
4	│ Q	supposition
5	│ $(\mathsf{P} \wedge \mathsf{Q})$	from 2, 4 by $\wedge$I
6	│ $((\mathsf{P} \wedge \mathsf{Q}) \vee (\mathsf{R} \wedge \mathsf{Q}))$	from 5 by $\vee$I
7	│ P	supposition
8	│ $(\mathsf{P} \wedge \mathsf{R})$	from 2, 7 by $\wedge$I
9	│ $((\mathsf{P} \wedge \mathsf{Q}) \vee (\mathsf{R} \wedge \mathsf{Q}))$	from 8 by $\vee$I
10	$((\mathsf{P} \wedge \mathsf{Q}) \vee (\mathsf{R} \wedge \mathsf{Q}))$	from 3, 4–6, 7–9, by $\vee$E

3.3.2 The case of negation

The classical negation rules don't quite have the sort of neat, harmonious, correspondence we've just been talking about.

RAA can be said to be serve an introduction-like (and indeed is sometimes called $\neg$I). Though you might think that since the RAA rule mentions '$\bot$', and '$\bot$' is in turn introduced by a rule relating it immediately to negation, we would be going round in circles here if we tried to use RAA to introduce negation to someone. Compare, say, the introduction rule for '$\rightarrow$'. It seems that we could use *that* rule to fix the meaning of the connective for someone who hadn't encountered it before, while by contrast RAA introduces a negated conclusion ultimately by reference to a prior use of negation.

But let that first observation pass. Suppose, just suppose, we think that somehow '$\bot$', the idea of absurdity (a wildly unwanted outcome) can in fact be understood without understanding negation per se. Then RAA can be regarded as a kosher introduction rule for negation: what would be the negation-elimination rule that stands in harmony to this introduction rule? Here's RAA again:

$$
\begin{array}{ll}
\begin{array}{|l} \underline{A} \\ \vdots \\ \bot \end{array} & \text{supposition} \\
\neg A &
\end{array}
$$

Well, suppose you are given $\neg A$ Then, with the negation understood via RAA as its introduction rule, this would convey to you that there is an inference from A to $\bot$ (to warrant the assertion of $\neg A$). So, if you were *both* given $\neg A$ (so you can take it that there is such an inference from A to $\bot$) and *also* given A, then you could infer $\bot$. In other words, an 'elimination' rule co-ordinate to RAA treated as an 'introduction' rule for negation is just the absurdity rule Abs again, 'from A and $\neg A$ infer $\bot$'.

Well, so far so good. But note that leaves the double negation rule 'from $\neg\neg A$ infer A' out in the cold. Where does *that* fit into the picture? Well, arguably it doesn't – i.e. arguably DN *can't* be justified in the same way as other rules are justified. The thought is that introduction rules don't really need justifying in any substantial sense as they simply define a connective, while harmonious elimination rules are justified as in effect extracting from a complex wff what the corresponding introduction rule put in: DN has no justification of those kinds. Which is one reason you might worry about accepting DN as a law of pure logic that holds across the board. Perhaps it is a legitimate local principle in some domains of discourse but not in others, but not a fundamental law of thought that applies across the board – which is in effect the *intuitionist*'s view.

Suppose for example that you think that, in some domain, for a proposition to obtain is for there to be (at least in principle) a means of verifying it. E.g. suppose you think that for a mathematical proposition to be hold is for there to be at least in principle a way of proving it. Then, on that view, from a demonstration D that there is no means of verifying $\neg A$, i.e. no means of showing the supposition that A must lead to absurdity (so D warrants $\neg\neg A$), it evidently does not follow that there *is* a means of verifying A. So at least in this domain, from $\neg\neg A$ we are not entitled to infer A.

3.3.3 Ex falso quodlibet

Now, on the classical analysis of logical consequence, an absurdity entails any conclusion C – for it is impossible for an absurdity to be true and C false (since it is impossible for an absurdity to be true, period). We can reflect that thought in another rule of inference – often called *ex falso quodlibet* (EFQ): from $\bot$ infer anything. So once you find yourself with an absurd conclusion, you really ought to backtrack and discharge some premiss, because if you carry on then really anything goes and craziness explodes everywhere!

Now, you might hesitate over this rule. But in fact it adds almost nothing to what we already have, given the other rules. For consider the proof at the top of the next side. This shows that from the premisses $\mathsf{P}, \neg\mathsf{P}$ we can infer Q by our existing rules. And we can now generalize this form of inference, to show that we can always infer from A and $\neg A$ to any conclusion C. All that EFQ adds is that the same happens if we start from the contradiction wrapped up into a single absurd proposition, the falsum. So, in fact, adding EFQ to the other rules doesn't actually significantly add to what we can infer, and we shouldn't fuss about that if we are already prepared to accept the inference from A and $\neg A$ to any conclusion.[5]

[5]To put things the other way around: if we want to avoid the 'explosion' result that we can get

1	P	premiss
2	$\quad (\neg\mathsf{P} \wedge \neg\mathsf{Q})$	supposition
3	$\quad \neg\mathsf{P}$	from 1, by $\wedge$E
4	$\quad \bot$	from 1, 3 by Abs
5	$\neg(\neg\mathsf{P} \wedge \neg\mathsf{Q})$	from 2–4, by RAA
6	$\neg\mathsf{P}$	premiss
7	$\quad \neg\mathsf{Q}$	supposition
8	$\quad (\neg\mathsf{P} \wedge \neg\mathsf{Q})$	from 6, 7 by $\wedge$I
9	$\quad \bot$	from 5, 8 by Abs
10	$\neg\neg\mathsf{Q}$	from 7–9 by RAA
11	Q	from 10 by DN

3.4 Putting everything together

In sum, then, we now have met ten inference rules governing the connectives – $\wedge$I, $\wedge$E, $\vee$I, $\vee$E, $\to$I (a.k.a. CP), $\to$E (a.k.a. MP), $\neg$I (a.k.a. RAA), $\neg$E (a.k.a. the absurdity rule), plus DN and EFQ. And we've seen – at least in a rough and ready way – how to begin to put together inferences using these rules into a Fitch-style proof. To be sure, we haven't explained the detailed proof-building rules very carefully: but you should at least have got the basic flavour of the thing. Let's pretend we've tidied things up properly. Then, when there is a well-formed Fitch proof from the premisses $A_1, A_2, \ldots A_n$ to the conclusion S, we'll write $A_1, A_2, \ldots A_n \vdash_F C$. The usual soundness and completeness results obtain, as you'll not be surprised to be told. In symbols,

$$A_1, A_2, \ldots A_n \vdash_F C \text{ if and only if } A_1, A_2, \ldots A_n \vDash_{S''} C.$$

where S'' is the semantics you get for the language with the usual connectives plus $\bot$ by requiring that the falsum always evaluates as false.

3.5 Another way of laying out proofs

3.5.1 Gentzen-style proofs – the general idea

Fitch-style proofs are very 'natural' and easy to follow. The practice of indenting a subproof every time a new supposition is made reveals very clearly which suppositions are in play at which point in the argument (much more so than the layout that *LO* uses, which just uses a written commentary to tell us what is going on in a proof). On the other hand, Fitch's way of displaying an argument doesn't make it immediately transparent what the 'input' to each step is – which is why we have also given a step-by-step commentary to the right of a proof.

Of course, in a short proof it is easy enough to check back and see where a line must have come from. For example, take the entirely uncommented proof on the next page. It only takes a moment to see that if this is a proof it must be one whose underived initial premisses are at lines 1, 2 and 5, that line 6 is derived from line 1 (rather than any of the other preceding lines), while line 8 is derived from lines 5 and 7. But still, the layout of the proof doesn't display the relations of immediate consequence between the various steps.

---

from contradictory wffs to anything at all, it isn't enough to drop EFQ. We'll have to adjust the other principles somehow. And dropping DN won't help much as we'll still be able to get to the negation of any proposition. Now, there *are* ways of avoiding 'explosion' at greater or lesser cost – but exploring them will take us far too far afield here.

1 $(P \land Q)$
2 $(Q \to R)$
3 Q
4 R
5 $((P \land R) \to S)$
6 P
7 $(P \land R)$
8 S
9 $(S \land Q)$

Here, then, is an alternative way of laying out the same argument, *Gentzen-style*. This time, steps are separated by horizontal lines: the immediate input(s) to an inference appear above the line – inputs, plural, for rules like $\land$I and MP – and the immediate conclusion of the inference appears below the line.

$$\dfrac{\dfrac{\dfrac{\dfrac{(P \land Q)}{P} \qquad \dfrac{\dfrac{(P \land Q)}{Q} \qquad (Q \to R)}{R}}{(P \land R)} \qquad ((P \land R) \to S)}{S} \qquad \dfrac{(P \land Q)}{Q}}{(S \land Q)}$$

This layout presents visually a lot more information about the proof. It displays the premisses at the top of branches of the tree. The premiss $(P \land Q)$ appears three times, reflecting that it is used in three different threads of argument. And – most importantly – at each step, it is entirely clear what inference is being made from the inputs immediately above a line to the conclusion below.

So far so good. But how are suppositional inferences to be displayed in this type of tree layout? Well, here's an inference from the premiss P together with the two conditional premisses $(P \to Q)$, $(Q \to R)$ to the conclusion R:

$$\dfrac{\dfrac{P \qquad (P \to Q)}{Q} \qquad (Q \to R)}{R}$$

Now, instead of treating P as a premiss, we can if we want regard it as a temporary assumption waiting to be discharged, e.g. in a CP inference. So let's do that – and we'll write square brackets round an assumption at the top of a thread of argument when we go on to discharge it. Taking that step gives us:

$$\dfrac{\dfrac{\dfrac{[P] \qquad (P \to Q)}{Q} \qquad (Q \to R)}{R}}{(P \to R)}$$

And here next is the inference from $(P \to (Q \to R))$ to $(Q \to (P \to R))$ (which we previously did Fitch-style) set out in the Gentzen style. To make things clear, I'll grow the proof in stages. So we start by showing that $(P \to (Q \to R))$, along with the temporary-assumptions-waiting-to-be discharged P and Q, together entail R.

$$\dfrac{Q \qquad \dfrac{P \qquad (P \to (Q \to R))}{(Q \to R)}}{R}$$

Now we discharge the assumption P and apply CP:

$$\dfrac{\dfrac{\mathsf{Q} \qquad \dfrac{[\mathsf{P}]^{(1)} \qquad (\mathsf{P} \to (\mathsf{Q} \to \mathsf{R}))}{(\mathsf{Q} \to \mathsf{R})}}{\mathsf{R}}}{(\mathsf{P} \to \mathsf{R})}\,(1)$$

Here, for extra clarity, we label together by (1) the first assumption which is discharged and the inference line where the discharging takes place. And now we can use CP and discharge another assumption labelled (2) to get

$$\dfrac{\dfrac{\dfrac{[\mathsf{Q}]^{(2)} \qquad \dfrac{[\mathsf{P}]^{(1)} \qquad (\mathsf{P} \to (\mathsf{Q} \to \mathsf{R}))}{(\mathsf{Q} \to \mathsf{R})}}{\mathsf{R}}}{(\mathsf{P} \to \mathsf{R})}\,(1)}{(\mathsf{Q} \to (\mathsf{P} \to \mathsf{R}))}\,(2)$$

So, as we wanted, we have just the one premiss remaining, and the desired conclusion at the bottom of the proof-tree.

As another illustration of this way of laying out proofs, here is a proof warranting the inference from $\neg(\mathsf{P} \wedge \neg\mathsf{Q})$ to $(\mathsf{P} \to \mathsf{Q})$ which we earlier did Fitch-style:

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{[\mathsf{P}]^{(2)} \qquad [\neg\mathsf{Q}]^{(1)}}{(\mathsf{P} \wedge \neg\mathsf{Q})} \qquad \neg(\mathsf{P} \wedge \neg\mathsf{Q})}{\bot}}{\neg\neg\mathsf{Q}}\,(1)}{\mathsf{Q}}}{(\mathsf{P} \to \mathsf{Q})}\,(2)$$

Here at the step labelled (1) we use ¬I to discharge one of the three assumptions alive at that time which lead to a contradiction (we could, of course, discharge any of them).

3.5.2 The rules!

Let's gather our proof-building rules together again, now using this time an obvious Gentzen-style notation: so we have

$$\wedge I\colon \ \dfrac{A \qquad B}{(A \wedge B)} \qquad\qquad \wedge E_l\colon \ \dfrac{(A \wedge B)}{A} \quad \wedge E_r\colon \ \dfrac{(A \wedge B)}{B}$$

$$\vee I_l\colon \ \dfrac{A}{(A \vee B)} \quad \vee I_r\colon \ \dfrac{B}{(A \vee B)} \qquad\qquad \vee E\colon \ \dfrac{(A \vee B) \qquad \begin{matrix}[A]\\ \vdots \\ C\end{matrix} \qquad \begin{matrix}[B]\\ \vdots \\ C\end{matrix}}{C}$$

$$\to I\colon \ \dfrac{\begin{matrix}[A]\\ \vdots \\ B\end{matrix}}{(A \to B)} \qquad\qquad \to E\colon \ \dfrac{A \qquad (A \to B)}{B}$$

$$\neg I\colon \ \dfrac{\begin{matrix}[A]\\ \vdots \\ \bot\end{matrix}}{\neg A} \qquad\qquad \neg E\colon \ \dfrac{A \qquad \neg A}{\bot}$$

$$DN: \quad \frac{\neg\neg A}{A} \qquad\qquad EFQ: \quad \frac{\bot}{C}$$

We also need instructions for putting together applications of the rules thus stated into composite proofs. The tricky thing is to state carefully the principles for discharging assumptions in applications of the rules $\vee E$, $\rightarrow I$ and $\neg I$.

(i) Note, by the way, that these aren't the only set of rules we could adopt for a natural deduction logic: in fact we see some others for classical negation in just a moment (so 'natural deduction' labels a style of doing logic, rather than a unique system).

(ii) Note also that the discharge rules are permissive: in other words, discharging one or more of the relevant assumptions in applying these rules is optional – for we can't go wrong in perhaps keeping used assumptions active (we'll just end up with a proof with some unnecessary premisses still in play).

(iii) Assumptions and conclusions may be the same wff in subproofs. In other words, instead of writing

$$\begin{array}{c} [\mathsf{P}] \\ \vdots \\ \dfrac{\mathsf{P}}{(\mathsf{P} \rightarrow \mathsf{P})} \end{array}$$

(going round the houses somehow to get from P to P), we can just write

$$\frac{[\mathsf{P}]}{(\mathsf{P} \rightarrow \mathsf{P})}$$

treating P by itself as a subproof starting with P and finishing with P!

(iv) Furthermore it is convenient, in applying →I, to allow cancellation of assumptions that aren't even used, thus allowing

$$\frac{\mathsf{P}}{(\mathsf{Q} \rightarrow \mathsf{P})}$$

where we discharge the unused assumption Q. Though in the presence of the other rules, we don't really need this little convention, for consider e.g.

$$\cfrac{\cfrac{\cfrac{\cfrac{\cfrac{\mathsf{P} \qquad \cfrac{[(\neg\mathsf{P} \wedge \mathsf{Q})]^1}{\neg\mathsf{P}}}{\bot}}{\neg(\neg\mathsf{P} \wedge \mathsf{Q})}\,(1) \qquad \cfrac{[\neg\mathsf{P}]^2 \qquad [\mathsf{Q}]^3}{(\neg\mathsf{P} \wedge \mathsf{Q})}}{\bot}}{\neg\neg\mathsf{P}}\,(2)}{\mathsf{P}}}{(\mathsf{Q} \rightarrow \mathsf{P})}\,(3)$$

Here, note how the numbering helps indicate clearly which assumption is discharged at which step.

3.5.3 More on the negation rules

(a) Sometimes, instead of DN, systems use the rule 'Classical Reductio', on the right below, a companion for our previous version of RAA:

$$RAA: \quad \begin{array}{c} [A] \\ \vdots \\ \dfrac{\bot}{\neg A} \end{array} \qquad\qquad Cl.RAA: \quad \begin{array}{c} [\neg A] \\ \vdots \\ \dfrac{\bot}{A} \end{array}$$

It is plain that the other rules plus DN yields Classical Reductio: for consider

$$\dfrac{\dfrac{\begin{matrix}[\neg A]\\ \vdots \\ \bot\end{matrix}}{\neg\neg A}\;\text{RAA}}{A}\;\text{DN}$$

Conversely, given Classical Reductio plus the other rules, we get back DN: for consider

$$\dfrac{\dfrac{[\neg A] \qquad \neg\neg A}{\bot}\;\neg\text{E}}{A}\;\text{Cl.RAA}$$

(b) It is also worth noting too that, in the presence of the other rules, DN is equivalent to the law of excluded middle – i.e. to the principle that $(A \vee \neg A)$ can be proved from no premisses. Firstly, let's show how – given DN – we can show $(A \vee \neg A)$ as a theorem. And actually, this is an interesting case to prove first in Fitch-style and then in Gentzen style, to give a further comparison of the different proof-styles.

So first, here is a Fitch-style working out of a proof that it is derivable from *no* premisses) that $(\mathsf{P} \vee \neg\mathsf{P})$.

Since there are no premisses to use, and the conclusion isn't a conditional, we are going to have to kick off by assuming the opposite of what we want to prove, and aim for a reductio:

1	$\mid\; \neg(\mathsf{P} \vee \neg\mathsf{P})$	supposition

But we can't apply any rule to that, except the '$\vee$' introduction rule, so we are going to have to make another assumption to get anywhere! What shall we assume? Let's try the simplest option ...

1	$\mid\; \neg(\mathsf{P} \vee \neg\mathsf{P})$	supposition
2	$\mid\;\mid\; \mathsf{P}$	supposition

It is obvious that we can get a contradiction from *that* assumption:

1	$\mid\; \neg(\mathsf{P} \vee \neg\mathsf{P})$	supposition
2	$\mid\;\mid\; \mathsf{P}$	supposition
3	$\mid\;\mid\; (\mathsf{P} \vee \neg\mathsf{P})$	from 2 by $\vee$I
4	$\mid\;\mid\; \bot$	by 1, 3!
5	$\mid\; \neg\mathsf{P}$	from 2 to 4 by RAA

And now it is equally obvious how to finish the argument to get a contradiction from the initial supposition, and hence the desired result.

6	$\mid\; (\mathsf{P} \vee \neg\mathsf{P})$	from 5 by $\vee$I
7	$\mid\; \bot$	by 1, 6!
8	$\neg\neg(\mathsf{P} \vee \neg\mathsf{P})$	from 1 to 7 by RAA
9	$(\mathsf{P} \vee \neg\mathsf{P})$	from 8 by DN

And now that we've got the proof, we can if we want rearrange it into a Gentzen-style presentation, thus:

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{[\mathsf{P}]^{(1)}}{(\mathsf{P} \vee \neg\mathsf{P})} \qquad [\neg(\mathsf{P} \vee \neg\mathsf{P})]^{(2)}}{\bot}}{\neg\mathsf{P}}\;(1)}{(\mathsf{P} \vee \neg\mathsf{P})} \qquad [\neg(\mathsf{P} \vee \neg\mathsf{P})]^{(2)}}{\bot}}{\neg\neg(\mathsf{P} \vee \neg\mathsf{P})}\;(2)}{(\mathsf{P} \vee \neg\mathsf{P})}\;\text{DN}$$

where at the step labelled (2) we use the reductio rule →I to discharge the supposition $\neg(\mathsf{P} \lor \neg\mathsf{P})$ that we've used twice in getting to the contradiction. Which way is the proof easier to follow, Fitch-style or Gentzen-style? Why?

Obviously the proof of that instance of excluded middle generalizes. Which goes to show that DN implies excluded middle, in the presence of the other inference rules. And now let's show conversely that, if we can assume instances of excluded middle as a 'free go', then (in the presence of the other rules) we can recover DN, as promised. So consider the following pattern of inference:

$$\dfrac{(\mathsf{P} \lor \neg\mathsf{P}) \qquad \begin{matrix}\mathsf{P}\\ \vdots \\ \mathsf{P}\end{matrix} \qquad \begin{matrix}\neg\mathsf{P} \qquad \neg\neg\mathsf{P} \\ \ddots \quad \cdot^{\cdot^{\cdot}} \\ \mathsf{P}\end{matrix}}{\mathsf{P}}\ \lor\mathrm{E}$$

If we can fill in the dots, then we can use ∨E as indicated to discharge the premisses P and $\neg\mathsf{P}$, and we'll have a proof from the sole remaining assumption $\neg\neg\mathsf{P}$ to the conclusion P. Now, there is no need for any filling on the left, given we allow minimal subproofs whose initial assumption and whose conclusion is the same. And on the right we can get from the contradictory pair $\neg\mathsf{P}$ and $\neg\neg\mathsf{P}$ to anywhere we like in various ways, including by a simple appeal to EFQ. So here's a completed proof:

$$\dfrac{\dfrac{}{(\mathsf{P} \lor \neg\mathsf{P})} \qquad [\mathsf{P}]^{(1)} \qquad \dfrac{\dfrac{[\neg\mathsf{P}]^{(1)} \qquad \neg\neg\mathsf{P}}{\bot}\ \neg\mathrm{E}}{\mathsf{P}}\ \mathrm{EFQ}}{\mathsf{P}}\ (1)\ \lor\mathrm{E}$$

Again, of course, the proof generalizes.

We've shown then that in the presence of the other rules, DN is equivalent to classical reductio is equivalent to the law of excluded middle. So an intuitionist who doesn't endorse double negation won't endorse the other principles either – these are therefore the distinctively classical (i.e. non-intuitionist) laws/rules.

And that's enough to be going on with. We haven't explained entirely carefully the rules for building Gentzen-style proofs using our eight introduction and elimination rules, plus DN plus EFQ. But imagine that done. Then, when there is a well-formed Gentzen proof from the (unbracketed!) premisses $A_1, A_2, \ldots A_n$ at the top of branches to the conclusion C at the bottom, we'll write $A_1, A_2, \ldots A_n \vdash_G C$. The expected soundness and completeness results obtain, as once more will be no surprise. In symbols,

$$A_1, A_2, \ldots A_n \vdash_G C \text{ if and only if } A_1, A_2, \ldots A_n \vDash_{S''} C.$$

3.6 Natural deduction summarized

Let's summarize. An axiomatic system like M only allows ploddingly linear proofs, in which every step is either a premiss or an axiom or follows from earlier steps. We aren't allowed to make temporary suppositions 'for the sake of argument' and then discharge them, even though that is a crucial everyday mode of inference.[6] By contrast, *natural deduction* systems – and this really is perhaps their crucial differentiating feature – allow us to make temporary suppositions for the sake of argument and hence implement rules like CP and RAA. In addition, since Gentzen, it is canonical to present rules for connectives (aside from negation) – and quantifiers, when we come to them – in matching pairs of an introduction rule and an elimination rule which (as it were) undoes the effect of that introduction rule.

[6]Note then the difference between the deduction theorem which does apply to M, and CP which isn't a rule of M. CP, to repeat, is a rule of inference that operates *inside* a certain kind of proof system which allows detours via suppositional inferences. The deduction theorem is a result *about* the 'linear' system M, standing outside it, so to speak.

But we've also seen that there is more than one way of arranging a natural deduction proof, Fitch's and Gentzen's being the two nicest. The Fitch style perhaps is the more 'natural', reflecting the essentially one-thing-after-another flow of ordinary bits of argumentation: it works well as a tool to organize thoughts in the process of constructing arguments. The Gentzen mode of presentation, however, perhaps more tellingly reveals the logical structure and internal dependencies of arguments once discovered. To illustrate what I mean, check out again those two proofs of excluded middle. For more on Fitch-style proofs, there's a very clearly explained presentation in Paul Teller, *A Modern Formal Logic Primer*, Prentice-Hall, 1989 (now freely available on-line at http://tellerprimer.ucdavis.edu/). Gentzen-style proofs are used in e.g. Neil Tennant, *Natural Logic*, Edinburgh UP, 2nd edn 1990; or in the excellent but more advanced Dirk van Dalen, *Logic and Structure*, Springer, 4th edn 2004.

4 Appendix: Soundness and completeness proofs for M

We'll now return to the task that we shelved so as not to interrupt the business of comparing different styles of logical system, and show that the axiomatic system M is sound and complete. As noted before, for our purposes, you certainly don't need to master all the details of such proofs. *So this is just an appendix for enthusiasts.* Don't read on if this sort of thing fazes you: there's nothing of philosophical significance here at all. But for those with a taste for the mathematical details, it *is* worth getting some sense of how these things are proved.

4.1 The proof that all M-theorems are tautologies, more carefully done

Our soundness proof relies on two lemmata:[7]

1. Every M-axiom is a tautology. (Proof by checking that instances of Ax1, Ax2, Ax3 must indeed always be tautologies. Why is that so?)

2. If the valuation V makes both A and $(A \to C)$ true, then it makes C true too. (Proof by appeal to truth-table definition of '$\to$'. Spell that out!)

And now here's the soundness proof. Suppose $A_1, A_2, \ldots A_n \vdash_M C$. And suppose too that V is a valuation which makes all the premisses A_j true. We are assuming then that there is a proof-sequence of wffs $P_1, P_2, \ldots P_n$ which forms a proof of C, where each wff P_j is either (i) one of the A_j, or (ii) is an axiom, or (iii) follows from earlier wffs in the sequence by MP, and (iv) where $P_n = C$. Then V makes each P_j true. Why? Because as we walk along the sequence every wff P_j is either (i) one of the A_j so is true on V by hypothesis; or (ii) it is an axiom, and hence a tautology by (1), and hence true on any valuation, and hence true on V in particular; or (iii) the wff is derived by modus ponens from two earlier wffs which have already been seen to be true on valuation V as we walked past them, and so is true on V by (2). So every P_j is true on V. Whence in particular (iv) the last wff P_n, i.e. C, is true on V.

So what we've shown is that, assuming that $A_1, A_2, \ldots A_n \vdash_M C$, then any valuation V which makes the A_j all true makes C true. Which, of course, is just to say that $A_1, A_2, \ldots A_n \vDash C$.[8] QED.

An immediate corollary is that, if C is an M-theorem, then C is a tautology. In symbols,

If $\vdash_M C$ then $\vDash C$.

[7]'Lemma' (*pl.* lemmata): a subsidiary or intermediate theorem in a proof.

[8]We can drop subscripts from $\vDash$ when it is obvious which semantics is in play!

4.2 Proofs by induction

The key part of the soundness argument is essentially an argument *by induction on the length of the proof.*

Now, in general, a proof by induction has the following form. Suppose (i) $F(0)$, i.e. 0 has property F. And suppose that (ii) for every j, $F(j) \to F(j+1)$. Then since $F(0)$ and $F(0) \to F(1)$, it follows that $F(1)$. And since $F(1) \to F(2)$, it follows that $F(2)$. And since $F(2) \to F(3)$, it follows that $F(3)$. And so on through all the numbers. So, (iii) $F(n)$, whatever n is.

In this particular case, put $F(n)$ for: *in an M-proof sequence of length n, every wff is true on any valuation V which makes the premisses true.* Then, vacuously, we have (i) $F(0)$. And we also have (ii), for every j, if $F(j)$ then $F(j+1)$. Hence $F(n)$, for any n. So the last wff in any proof-sequence is true on any valuation which makes the premisses true. So it is a tautological entailment of the premisses.

And why does (ii) hold? The argument is essentially as above in §4.1. Take a proof n steps long, where by hypothesis $F(n)$. Now add another wff to the proof (premiss, axiom, deduction by MP). The next wff will still be true on any V that makes the premisses true. So we get $F(n+1)$. (Spell that out!)

4.3 Sketching a direct proof that M is complete

We want to show that M has a proof corresponding to any tautologically valid argument in '$\to$' and '$\neg$':

If $A_1, A_2, \ldots A_n \vDash C$ then $A_1, A_2, \ldots A_n \vdash_M C$.

Let's say that a set of wffs Γ (wffs from our propositional language with just '$\to$' and '$\neg$' built in) is *M-consistent* if it *isn't* the case that for some C, both $\Gamma \vdash_M C$ and $\Gamma \vdash_M \neg C$. In other words, Γ is M-consistent if you can't prove a contradiction from Γ in M. And of course, a set of wffs Γ is *semantically consistent* if there's a valuation which makes every wff in Γ true. Then it is enough to prove M's completeness that we prove

(**) If Γ is M-consistent, Γ is semantically consistent.

Why is that enough? Well, suppose $A_1, A_2, \ldots A_n \nvdash_M C$. A bit of fiddling around in M shows that this implies $A_1, A_2, \ldots A_n, \neg C$ is M-consistent (in effect, a reductio principle holds for M; if $A_1, A_2, \ldots A_n, \neg C$ is M-inconsistent and entails a contradiction, then $A_1, A_2, \ldots A_n \vdash_M C$, and vice versa). So by (**) $A_1, A_2, \ldots A_n, \neg C$ is semantically consistent. So there is a valuation which makes $A_1, A_2, \ldots A_n, \neg C$ all true, i.e. makes $A_1, A_2, \ldots A_n$ true and C false, and so $A_1, A_2, \ldots A_n \nvDash C$. In short, if $A_1, A_2, \ldots A_n \nvdash_M C$ then $A_1, A_2, \ldots A_n \nvDash C$. Contraposing gives us the completeness result. QED.

So we 'just' have to prove (**).

Here's one way to do it, informally described.

i. First we show that, for any set Γ, if it is M-consistent, it can be beefed up to a (usually) bigger but still M-consistent Γ' which includes everything in Γ but now also includes *truth-makers* for all the complex wffs in Γ. That is to say, for every wff A more complex than an atom or a negated atom we put into the set one or two other wffs which, if true, make A true. Since we are only dealing with the two connectives '$\to$' and '$\neg$', complex wffs will be of the form $\neg\neg A$, $(A \to B)$ and $\neg(A \to B)$. In the first case, the relevant truth-maker is the corresponding plain A; in the second case, $\neg A$ and B are alternative truth-makers; and in the third case, we need both A and $\neg B$ to make the complex wff true.

ii. If Γ' is M-consistent and contains truth-makers for all its complex wffs, it is semantically consistent. Why? Because there will be a valuation V which makes all the shortest wffs true – i.e. which makes true all the atoms and negated atoms in Γ' (NB no atom can appear both naked and negated in Γ', or it wouldn't be

M-consistent after all). And this valuation V which makes all the minimal truth-makers true will make the next most complex wffs true and these will make the next most complex wffs true, and so on percolating truth upwards throughout Γ'.

But of course, if V makes everything in the bigger set Γ' true, then V makes everything in the original set Γ true (since $\Gamma \subseteq \Gamma'$). So Γ is semantically consistent.

Let's now put that more carefully. Keeping to the case of a set of wffs Σ from our propositional language with just '$\rightarrow$' and '$\neg$' built in, let's say (to use fancy jargon) that Σ is *saturated* if the following conditions hold:

1. If Σ contains a wff $\neg\neg A$ it also contains A.
2. If Σ contains a wff $\neg(A \rightarrow C)$ it also contains A and $\neg C$.
3. If Σ contains a wff $(A \rightarrow C)$ it also contains $\neg A$ and/or C.

(Do those clauses look familiar? Remember tree-building: applying the unpacking rules as often as you can to the initial wffs will – if you don't hit a contradiction – give you a set of wffs Σ which will be saturated in this sense!)

Putting the argument in terms of this jargon, we can show the following:

i′. If Γ is M-consistent, it can be extended to a saturated M-consistent set Γ' such that $\Gamma \subseteq \Gamma'$.

ii′. If Γ' is saturated M-consistent, it is semantically consistent.

As we said, these two together entail (**). For if the 'little' set Γ can be extended to a 'big' set Γ' whose wffs can all be true together, obviously Γ's wffs can all be true together.

We prove (i′) by brute force. Put the wffs in Γ into some convenient order (we'll for present purposes assume Γ is finite, though that isn't actually necessary). Walk along the list. When we encounter a wff of the form $\neg\neg A$, add A to the end of the list. (Fiddling around in M, you can check that this can't change a so-far M-consistent list of wffs to an inconsistent one. Why? Because $\vdash_M \neg\neg A \rightarrow A$. So if we can prove a contradiction from stuff-plus-A, we can already prove the same contradiction from stuff-plus-$\neg\neg A$. Since $\neg\neg A$ by hypothesis didn't generate a contradiction, A won't.) When we encounter a wff of the form $\neg(A \rightarrow C)$, add A and $\neg C$ to the end of the list. (Fiddling around in M, you can again check that this can't change an M-consistent list of wffs to an inconsistent one.) If you encounter a wff of the form $(A \rightarrow C)$ add one of $\neg A$ and C, whichever doesn't make the list inconsistent (once again, it can be shown that you can always make such an addition, preserving M-consistency: for suppose otherwise; suppose both adding $\neg A$ and adding C makes the list M-inconsistent; that means the wffs already on the list, including $(A \rightarrow C)$, entail A and $\neg C$; which makes the wffs already on the list inconsistent after all, contrary to hypothesis). Keep on going, adding truth-makers for the complex wffs we encounter, while preserving M-consistency. The added wffs get simpler and simpler, and eventually you've got a saturated but still consistent set. QED.

We prove (ii′) just as we prove that there's a valuation of the wffs on a completed open branch of a tree in '$\rightarrow$' and '$\neg$' that makes every wff on the branch true (because, as we just said, the set of wffs on a completed open branch is saturated). Take the atoms that appear somewhere in the wffs in Γ. Consider the valuation that assigns the atom A the value T if it appears naked as a wff in Γ, and the value F if the atom appears negated as $\neg A$ (it can only appear one way or the other, and since all complex wffs are fully unpacked, every atom which appears in a complex wff also appears either naked or negated). Then it is easily checked that this valuation makes every wff in Γ true. QED.

So we are done, albeit with a few gaps. Still, the strategy is the thing! And note that, whatever rich-enough propositional logic L we might be using, if we can prove the corresponding version of (i) – perhaps with the notion of 'saturation' expanded to cover other connectives in the obvious ways familiar from our work with trees – we'll get a completeness proof! That's why it is worth knowing about the strategy here: it is, as it were, all-purpose.

4.4 Sketching an indirect proof that M is complete

Finally, just for masochists, let's also sketch an indirect proof of M's soundness, because that's the way *Logical Options* does things for the axiomatic system it presents.

Suppose we can show

(***) If $A_1, A_2, \ldots A_n \vdash_{T^-} C$ then $A_1, A_2, \ldots A_n \vdash_M C$.

(Here, T^- is the tree-system cut down to deal with just the connectives '$\rightarrow$' and '$\neg$' that appear in M.) Then the proof of completeness for such trees tells us that, for arguments restricted to '$\rightarrow$' and '$\neg$',

If $A_1, A_2, \ldots A_n \vDash C$ then $A_1, A_2, \ldots A_n \vdash_{T^-} C$.

And obviously these two results together entail the desired completeness result:

If $A_1, A_2, \ldots A_n \vDash C$ then $A_1, A_2, \ldots A_n \vdash_M C$.

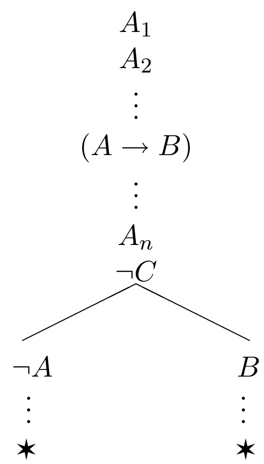
So how might we go about showing (***)?

We'll use another proof by induction, this time by induction on the complexity of tree proofs. Let's say that a tree-proof has depth d if – on the longest branch – the unpacking rules get applied d times.

(1) Then, first we show that if $A_1, A_2, \ldots A_n \vdash_{T^-} C$, then $A_1, A_2, \ldots A_n \vdash_M C$ in any case where the tree proof has depth $d = 0$. That is to say, when the tree starting $A_1, A_2, \ldots A_n, \neg C$ closes immediately, without any unpacking rules being applied. That is to say, when one of those wffs is the contradictory of the other. If A_i is the negation of A_j, then it is easy to show A_i and A_j M-entail anything you like, and so $A_1, A_2, \ldots A_n \vdash_M C$. And if A_i is the contradictory of $\neg C$, then again it is easy to show that $A_i \vdash_M C$ and hence $A_1, A_2, \ldots A_n \vdash_M C$.

(2) Suppose now that when $A_1, A_2, \ldots A_n \vdash_{T^-} C$, then $A_1, A_2, \ldots A_n \vdash_M C$, in any case where the tree proof has depth no more than $d = j$. We show on that induction hypothesis that the same thing holds when the tree proof has depth up to $d = (j + 1)$.

Suppose, then, there is a closed tree starting starting $A_1, A_2, \ldots A_n, \neg C$ which is of depth $j + 1$. How does the tree start? Either the first rule applied is the rule for splitting conditionals, or it is the rule for unpacking negated conditionals, or it is the double negation rule. Take the first case. Then the tree has the form



where, by hypothesis, eventually both branches close, and both branches have additional depth $\leq j$. And now look at those branches separately.

With trivial rearrangement, the left one forms a closed tree of depth $\leq j$, starting $A_1, A_2, \ldots, (A \rightarrow B), \ldots A_n, \neg A, \neg C$: the right one forms another closed tree of depth

$\leq j$, this time starting $A_1, A_2, \ldots, (A \rightarrow B), \ldots A_n, B, \neg C$. Since these trees *do* have depth $\leq j$, the induction hypothesis applies; so $A_1, A_2, \ldots, (A \rightarrow B), \ldots A_n, \neg A \vDash_M C$ and $A_1, A_2, \ldots, (A \rightarrow B), \ldots A_n, B \vDash_M C$. By the deduction theorem, the premisses $A_1, A_2, \ldots, (A \rightarrow B), \ldots A_n$ therefore prove $(\neg A \rightarrow C)$ and $(B \rightarrow C)$, and the premisses trivially prove $(A \rightarrow C)$. It quickly follows that those premisses therefore prove C, which was to be shown.

So much, then, for the case where the first rule applied in our tree of depth $j+1$ is the rule for splitting conditionals. The same holds similarly in the other two cases where the first rule applied is the negated-conditional rule or the double negation rule. So we've established (2) if $A_1, A_2, \ldots A_n \vdash_{T^-} C$ implies $A_1, A_2, \ldots A_n \vdash_M C$ in any case where the tree proof has depth no more than $d = j$, the implication also holds for tree-proofs of depth up to $d = (j+1)$.

Hence, from (1) and (2) by induction, the desired implication holds for trees of any depth. Phew. QED.