IFL: Logicbite 20

Introducing truth trees

Peter Smith

One way of determining whether an inference is tautologically valid – as you've rather painfully found out! – is by using a laborious brute-force truth-table test. But there are *other* ways of testing for tautological validity. In particular, there is the method of truth trees. Sadly, this is not guaranteed to be much faster in *every* case: but it is elegant and usually a *lot* nicer to use. Moreover, as you'll soon find out, there is no hope of running something like a brute-force truth-table test in quantificational logic; but we *can* use truth trees to prove validity for arguments involving quantifiers. So truth trees are good to know about!

You can find out about truth trees for propositional logic in the supplementary chapters here. Alternatively, this Logicbite gives an introductory taster – enough perhaps for you to get the general idea – and then you can decide whether you want to find out more.

❦

One preliminary point before we get going. In what follows, I will usually prefer to say of a wff, not that it is *false* on a given valuation of its atoms, but rather that its *negation is true.* For example, instead of saying that a (tautologically) invalid inference is one where there's a valuation which makes the premisses true and conclusion false, I will now prefer to say that *an inference is (tautologically) invalid just if there is a relevant valuation which makes the premisses and the negation of the conclusion all true.*

Obviously nothing can turn on this alternative way of putting things; in our two-valued logical framework, a wff is false if and only if its negation is true. But there is a reason for this apparently quirky preference for writing things of the form $\neg\alpha \coloneqq \mathrm{T}$ rather than $\alpha \coloneqq \mathrm{F}$. I reveal all towards the end.

❦

Considered the following daft inference with fifty premisses – all negated atoms, all different – and an unrelated conclusion:

A $\quad \neg\mathsf{P}, \neg\mathsf{P}', \neg\mathsf{P}'', \neg\mathsf{P}''', \ldots, \neg\mathsf{P}''^{\cdots}{}'' \therefore \mathsf{Q}$

It would be quite crazy to try to write down a full truth table to test the status of this argument (which would take millennia). It is enough to remark that **A** is invalid if and only if there is a valuation of the fifty-one atoms which makes the premisses and negated conclusion all true – i.e. a valuation such that

$$\neg\mathsf{P} \coloneqq \mathrm{T},\ \neg\mathsf{P}' \coloneqq \mathrm{T},\ \neg\mathsf{P}'' \coloneqq \mathrm{T},\ \neg\mathsf{P}''' \coloneqq \mathrm{T},\ \ldots,\ \neg\mathsf{P}''^{\cdots}{}'' \coloneqq \mathrm{T},\ \neg\mathsf{Q} \coloneqq \mathrm{T}.$$

Then we merely have to note that we can immediately read off just such a valuation, namely

$$\mathsf{P} \coloneqq \mathrm{F},\ \mathsf{P}' \coloneqq \mathrm{F},\ \mathsf{P}'' \coloneqq \mathrm{F},\ \mathsf{P}''' \coloneqq \mathrm{F},\ \ldots, \mathsf{P}''^{\cdots}{}'' \coloneqq \mathrm{F},\ \mathsf{Q} \coloneqq \mathrm{F}.$$

Which settles it. **A** is invalid. (In this Logicbite, by the way, take 'valid'/'invalid' always to mean *tautologically* valid/invalid.)

❦

Take next the argument

B $\quad (\mathsf{P} \land \neg\mathsf{Q}), (\mathsf{R} \land \neg\mathsf{S}) \therefore (\mathsf{Q} \lor \mathsf{S}).$

This too is obviously invalid, as we can confirm with a sixteen-line truth table. But we can also argue as follows, this time laying out our reasoning step by step.

The inference **B** is invalid if and only if there is a valuation of the four atoms which makes the premisses true and the negation of the conclusion true too, so

(1)	$(\mathsf{P} \land \neg\mathsf{Q}) \coloneqq \mathrm{T}$
(2)	$(\mathsf{R} \land \neg\mathsf{S}) \coloneqq \mathrm{T}$
(3)	$\neg(\mathsf{Q} \lor \mathsf{S}) \coloneqq \mathrm{T}$

This time, we can't instantly read off a valuation which makes (1) to (3) hold. But note that by the truth table for conjunction, (1) holds if and only if these claims hold too:

(4)	$\mathsf{P} \coloneqq \mathrm{T}$
(5)	$\neg\mathsf{Q} \coloneqq \mathrm{T}$

Similarly, (2) holds if and only if these claims do too:

(6)	$\mathsf{R} \coloneqq \mathrm{T}$
(7)	$\neg\mathsf{S} \coloneqq \mathrm{T}$

Finally, (3) holds if and only if the plain disjunction '$(\mathsf{Q} \lor \mathsf{S})$' is false, which requires that both disjuncts are false. Avoiding talk of falsehood, this means that the negation of each disjunct has to be true:

(8)	$\neg\mathsf{Q} \coloneqq \mathrm{T}$
(9)	$\neg\mathsf{S} \coloneqq \mathrm{T}$

So, putting everything together, we can ensure that (1) to (3) hold – i.e. we can make the premisses and negated conclusion of **B** all true – by taking the valuation where $\mathsf{P} \coloneqq \mathrm{T}$, $\neg\mathsf{Q} \coloneqq \mathrm{T}$ (i.e. $\mathsf{Q} \coloneqq \mathrm{F}$), $\mathsf{R} \coloneqq \mathrm{T}$, and $\neg\mathsf{S} \coloneqq \mathrm{T}$ (i.e. $\mathsf{S} \coloneqq \mathrm{F}$). Hence **B** is indeed invalid.

❦

For a contrasting third example, consider

C $(\mathsf{P} \land \neg\mathsf{Q}) \therefore \neg(\mathsf{Q} \land \mathsf{R})$.

Again, the inference is invalid just if there is a valuation of the three atoms here which makes the premiss and negated conclusion both true, i.e. which satisfies

(1)	$(\mathsf{P} \land \neg\mathsf{Q}) \coloneqq \mathrm{T}$
(2)	$\neg\neg(\mathsf{Q} \land \mathsf{R}) \coloneqq \mathrm{T}$

For (1) to obtain, we need

(3)	$\mathsf{P} \coloneqq \mathrm{T}$
(4)	$\neg\mathsf{Q} \coloneqq \mathrm{T}$

By the truth table for negation applied twice, for (2) to obtain, we need

(5)	$(\mathsf{Q} \land \mathsf{R}) \coloneqq \mathrm{T}$

and hence we need both

(6)	$\mathsf{Q} \coloneqq \mathrm{T}$
(7)	$\mathsf{R} \coloneqq \mathrm{T}$

But hold on! We have now hit a contradiction. *No valuation can give us (4) and (6), i.e. no valuation can make* $\neg$Q *and* Q *simultaneously true together.* So there can not possibly be a valuation satisfying both (1) and (2) and making the premiss of **C** true and conclusion false.

Which shows that **C** is valid.

❦

In these three examples we have been – so to speak – *working backwards.*

On the brute-force truth-table method, in order to decide whether a given argument is tautologically valid, we start by laying out all the combinatorially possible valuations for the relevant atoms. Then we plod forwards through each valuation, calculating the truth values of the premisses and conclusion as we go, looking for a 'bad' line. But now we are starting on the very same decision task from the other end. We suppose that there *is* a bad valuation, i.e. one where the premisses are true and the negation of the conclusion is true too. Then we try to unravel the implications of this supposition, working backwards towards uncovering a valuation of the relevant atoms that is indeed bad.

If we *can* uncover a bad valuation, then obviously the argument being tested is *not* valid. On the other hand, if we get inextricably entangled in contradiction as we work backwards, then (by an informal reductio ad absurdum inference!) that shows that there is no bad valuation after all, and the argument in question is *valid.*

❦

So far, so straightforward. But now here is an obviously invalid little argument:

D $(\mathsf{P} \lor \mathsf{Q})\ \therefore\ \mathsf{P}.$

What happens if we try working backwards to expose a valuation which makes the premiss true and negated conclusion true too?

Start, as before, with the assumption that there *is* a bad valuation, one which makes the following hold:

(1)	$(\mathsf{P} \lor \mathsf{Q}) \coloneqq \mathrm{T}$
(2)	$\neg\mathsf{P} \coloneqq \mathrm{T}$

Given the truth table for disjunction, (1) tells us that either (3a) $\mathsf{P} \coloneqq \mathrm{T}$ or (3b) $\mathsf{Q} \coloneqq \mathrm{T}$ or perhaps both – and either alternative suffices to verify the disjunction. *But (1) doesn't tell us which alternative holds.* We will have to consider the two cases in turn.

In fact, the alternative (3a) where $\mathsf{P} \coloneqq \mathrm{T}$ need not detail us long, for it is immediately inconsistent with (2), i.e. with $\neg\mathsf{P} \coloneqq \mathrm{T}$.

The alternative (3b) where $\mathsf{Q} \coloneqq \mathrm{T}$ fares better; there is no contradiction to be found this time. The resulting valuation on which $\neg\mathsf{P} \coloneqq \mathrm{T}$ (i.e. $\mathsf{P} \coloneqq \mathrm{F}$), $\mathsf{Q} \coloneqq \mathrm{T}$ makes the premiss of **D** true and the negated conclusion true too. As we knew all along, then, the inference **D** is indeed invalid.

❦

We want a perspicuous way of setting out reasoning like this when our line of argument forks, and we have to explore alternatives. The obvious device is to draw forking branches in our chains of reasoning. We then get (upside down, trunk-at-the-top) *truth trees*, and can display our reasoning about **D** like this:

(1)	$(\mathsf{P} \lor \mathsf{Q}) \coloneqq \mathrm{T}$	(premiss assumed true)
(2)	$\neg\mathsf{P} \coloneqq \mathrm{T}$	(negated conclusion assumed true)
(3)	$\mathsf{P} \coloneqq \mathrm{T}$ $\mathsf{Q} \coloneqq \mathrm{T}$	(alternatives from 1)
	✻	(contradiction on left branch!)

This mode of presentation should be self-explanatory. The main novelty compared with our informal reasoning a moment ago is that, instead of considering options sequentially, we consider them side-by-side; we split our reasoning into left and right branches as we consider the alternative ways of making (1) come out true.

The other novelty is that we now begin to use '✻' as an *absurdity sign*. We could alternatively use '$\bot$' – but officially that's a wff, and it is conventional to treat '✻' here just as quasi-punctuation – like a bold exclamation mark – signalling that we have hit a contradiction on a path down the truth tree. In other words, '✻' at the end of a path indicates that it features a pair of wffs α and $\neg\alpha$ which are both assigned T (note that the wff α here doesn't have to be an atom).

Our miniature downward branching tree for argument **D**, then, has two paths through it running from top to bottom, as we explore alternative potential ways of making true the initial claims at the top of its trunk. The left-hand path is, as we will say, *closed* with an absurdity marker; but the other path remains *open*. And from the open path we can read off a perfectly consistent assignment of values to the relevant atoms which satisfies (1) and (2), making **D** invalid.

❦

Take next the following argument:

E $\quad \mathsf{P}, \neg(\mathsf{P} \land \neg\mathsf{Q}) \;\therefore\; \mathsf{Q}.$

This is invalid if and only if we can consistently make the premisses and negated conclusion all true, as follows:

(1)	$\mathsf{P} \coloneqq \mathrm{T}$	(premiss assumed true)
(2)	$\neg(\mathsf{P} \land \neg\mathsf{Q}) \coloneqq \mathrm{T}$	(premiss assumed true)
(3)	$\neg\mathsf{Q} \coloneqq \mathrm{T}$	(negated conclusion assumed true)

Now, a conjunction is false when at least one conjunct is false. Equivalently, a negated conjunction is true just when at least one negated conjunct is true. In other words, given $\neg(\alpha \land \beta) \coloneqq \mathrm{T}$ we must have at least one of $\neg\alpha \coloneqq \mathrm{T}$ or $\neg\beta \coloneqq \mathrm{T}$. This is the principle we can apply in moving from (2) to the alternatives at (4):

(4)	$\neg\mathsf{P} \coloneqq \mathrm{T}$ $\neg\neg\mathsf{Q} \coloneqq \mathrm{T}$	(alternatives from 2)
	✻ ✻	

This time *both* alternatives immediately produce contradictory assignments of truth values ($\mathsf{P} \coloneqq \mathrm{T}$ and $\neg\mathsf{P} \coloneqq \mathrm{T}$ on the left-branching path from top to bottom; $\neg\mathsf{Q} \coloneqq \mathrm{T}$ and $\neg\neg\mathsf{Q} \coloneqq \mathrm{T}$ on the right-branching path). So both paths get closed off with the absurdity sign. Which shows that neither of the potential ways of satisfying conditions (1) to (3) is consistent. Therefore the inference **E** has to be valid.

❦

As you'll have noticed, the wffs on our trees so far have been *all* be assigned the value T. That's because, whenever you might have been tempted to say a wff was false, you were asked instead to say that its negation is true. The pay-off from this little bit of cheap trickery is that we can now move so very easily to so-called *unsigned* truth trees, where wffs appear 'naked'. The move couldn't be simpler! *We just delete* ':= T' *wherever it appears*! So, we tidy our last 'working backwards' tree to look like this:

(1)	P		(premiss)
(2)	$\neg(\mathsf{P} \wedge \neg\mathsf{Q})$		(premiss)
(3)	$\neg\mathsf{Q}$		(negated conclusion)
(4)	$\neg\mathsf{P}$	$\neg\neg\mathsf{Q}$	(alternatives from 2)
	✻	✻	

And we could if we want suppress the line numbers on the left and the commentary on the right to get an even snappier presentation of the tree-form demonstration that argument **E** is tautologically valid.

❦

Let's leave things at this point. You've now seen the most basic headline news about truth trees. The leading idea is that, instead of hacking forward through all the lines of a truth table for an argument, searching for a 'bad line', we work backwards. We suppose the argument up for assessment is invalid, i.e. suppose that there *is* a bad line on its table, and work out what a corresponding valuation must look like (if one exists). As we work backwards, we have to consider branching alternatives – and that's why trees get into the picture. Our next examples would involve us tracing out trees that branch multiple times (but that doesn't involve essentially new ideas). It you want to see how the story develops in detail, though, now is the time to look at these supplementary chapters, revised from the first edition of *IFL*.