Peter Smith, *Introduction to Formal Logic* (CUP, 2nd edition)

Exercises 19: More on 'if's and '$\to$'s

(a) *Show that the following hold for* PL *languages with the relevant atoms:*

(1) $\vDash ((\mathsf{P} \to \mathsf{Q}) \lor (\mathsf{Q} \to \mathsf{P}))$

(2) $((\mathsf{P} \land \mathsf{Q}) \to \mathsf{R}) \vDash ((\mathsf{P} \to \mathsf{R}) \lor (\mathsf{Q} \to \mathsf{R}))$

(3) $((\mathsf{P} \to \mathsf{Q}) \to \mathsf{R}), \neg\mathsf{P} \vDash \mathsf{R}$

Do these results raise more problems for equating '$\to$' with 'if'?

More generally, the pattern of inference $\gamma \mathrel{.\,\dot{}\,.} (\alpha \to \gamma)$ *is valid. What about a corresponding ordinary-language inference of the form* C; so if A then C*?*

The pattern of inference $(\beta \to \gamma) \mathrel{.\,\dot{}\,.} ((\alpha \land \beta) \to \gamma)$ *is also valid. What about a corresponding ordinary-language inference of the form* C, if B; so C, if A and B*?*

It is easy to check the claims (1) to (3) – use truth-tables. Similarly, we can check that for any α, β, γ, we have (*) $\gamma \vDash (\alpha \to \gamma)$ and (**) $(\beta \to \gamma) \vDash ((\alpha \land \beta) \to \gamma)$.

Before commenting on the vexed issue of how these easy technical results about the validity of some arguments involving the *material* conditional relate to issues about the *ordinary-language* conditional, let's recall the following:

- A conditional like e.g. 'If you heat an iron bar, it expands' is in fact a generalization, equivalent to 'Any iron bar is such that, if you heat it, it expands'. Likewise too for some uses of e.g. 'If Fido is a dog, Fido has four legs', if 'Fido' is here acting not as a proper name for a specific hound, but as a dummy name (in effect, a free variable) doing generalizing work. Let's set aside conditionals like these. They plainly aren't singular material conditionals!
- A conditional like 'If Oswald hadn't shot Kennedy in Dallas, someone else would have', or 'If I had put money into Bitcoin at the beginning, I would have now been richer' is a counterfactual, possible-world conditional. We evaluate the actual value of a conditional of this kind – *if P had been the case, Q would have been the case* – by considering possible worlds different from the actual world, 'nearby possible worlds' where P is true, and consider whether Q would be true there. Plainly such conditionals aren't material conditionals (the actual truth-value of a material conditional depends on the actual values of its antecedent and consequent).
- The use of some past tenses like 'had' and 'would have' are in context fairly reliable indicators that a conditional is a possible-world conditional. Other cases may not be so clear. Take the future tensed 'if the Captain turns his key tomorrow, then the missile will fire'. This arguably invites a possible world reading: go to the nearest possible future world where the Captain turns his key tomorrow – which may be the actual future world – and then consider what happens in such a world. And what, after the event, about the past-tense counterpart? Consider 'if the Captain turned his key, then the missile fired': it can in some contexts seem rather natural to take this too as having possible-world truth-conditions – as saying, in effect, that if we take a world like this one where the Captain turned his key (the actual past world if the Captain did turn his key there, or if he didn't then a world as close as possible to the actual world but where he did turn his key), then the missile fired in that world. [Or maybe those last claims are wrong – the story about ordinary-language conditionals does indeed get murky!!]

When we compare inferences with the material conditional with vernacular inferences, therefore, we need to keep in mind that the material conditional is, at best, a candidate for capturing the core content of *singular indicative conditionals.*

So now let's turn to (*). An inference of the form $\gamma \mathrel{.\,\dot{}\,.} (\alpha \to \gamma)$ with the material conditional is valid. What about a corresponding ordinary-language inference of the form *C; so if A then C*? We could argue like this:

> Suppose that C is straight-out true. Then, C is part of the story about how things are. Maybe A then correctly tells us another part of the story, maybe it doesn't. But given that C really is true, then it is a truth however things stand in the rest of the story about the world; it is true whichever way things happen to have panned out with A. So in particular, C is true (even) if A is true.

"You mean that given Fido has four legs, it follows that if he had a leg amputated yesterday, he has four legs? Really?" Well, yes: given Fido actually has four legs, then he actually has four legs – and he still actually has four legs, even if one was amputated yesterday (there must have been a leg transplant or a prosthetic leg fitted, even miraculous regrowth, but anyway there are four legs! . . .).

Of course it is no doubt very much more likely that if Fido had had a leg amputated yesterday, he would now have only three legs. So faced with Fido in the actual world, four-leggedly bounding about happily, we naturally hear the conditional 'if he had a leg amputated yesterday, then he now has three legs' as being a true (or very likely true) *counterfactual* conditional – a conditional that tells us what happens at a possible world similar to this one except that poor Fido loses a leg yesterday. But still, given that Fido has four legs, then read as an indicative, strictly this-worldly conditional, 'Fido has four legs, [even] if P' will come out true, even if P is an unlikely tale about amputation!

Moving on to (**), take an informal correlate of $(\beta \to \gamma) \therefore ((\alpha \land \beta) \to \gamma)$, i.e. an ordinary-language inference of the form *C, if B; so C, if A and B*? We could argue like this:

> Given C really does hold if B is true, then it holds if B is true – whatever the rest of the story about the world. So C holds if B is true, however things stand with A. So in particular C holds if B is true and in fact A happens to be true as well.

"You mean given that (a) Ben has four legs today, if he's a cat, then (b) Ben has four legs today, if he's a cat and had a leg amputated yesterday?" Well, yes, we can similarly dig in here and insist that given that the two conditionals here are read as a singular, this-worldly conditionals (not as implicitly general or counterfactual) then (b) follows from (a).

Let's have some scene-setting. Suppose Jill has a pet cat and a pet parrot; I recall one of them is called 'Ben', but I can't remember which. Still, reasonably enough, I say (a) 'If Ben is a cat, he has four legs'. My grounds for this might be general thoughts about cats, but (a) is a singular indicative conditional.

You now chip in 'Ahah, but Ben – whichever pet he is – could have had a leg amputated yesterday!'. I could now go jump either way. On the one hand, now bearing in mind the possibility you have raised, I could withdraw my confident assent to the unqualified conditional. On the other hand, I could dismiss the possibility as too outlandish – especially perhaps because Jill has been listening and hasn't dissented from what I said. So I stick to (a) 'If Ben is a cat, he has four legs'. And for long as I stick to (a), I am committed to agreeing that Ben has four legs, if a cat, whatever else is true in the world. Though of course, I accept the probable truth of the *counterfactual* conditional that if the cat had had a leg amputated yesterday, he wouldn't now have four legs! Or so the story might go.

OK, we have so far given some considerations that suggest that the truth of (*) and (**) doesn't count against the idea that the core content of some indicative singular conditionals is reflected well enough by the material conditional.

Now what about (1)? Both the conditional *if the weather is too hot, then the weather is too cold* and the conditional *if the weather is too cold, then the weather is too hot* look false (indeed, you might wonder whether they are not just false but *necessarily* false). So their disjunction should surely be false. But – putting P for *the weather is too hot* and Q for *the weather is too cold* – the formal rendition of that false disjunction comes out as a tautology! Awkward!

Or perhaps not so awkward after all. For arguably our inclination to read *if the weather is too hot, then the weather is too cold* and *if the weather is too cold, then the weather is too hot* as obviously false arises from reading them as general and/or counterfactual conditionals, and such conditionals are apt for regimentation as in (1).

Re (2): Richard Jeffrey offered the following example. Suppose a dual key system was in place. Then it seems that it could be true that *if the Captain turned his key and the First Officer turned his key, then the missile fired* without it being true either that *if the Captain turned his key, then the missile fired* or that *if the First Officer turned his key, then the missile fired.* But – putting P for *The Captain turned his key*, Q for *The First Officer turned* his *key*, R for the missile fired – we get a valid argument from the truth of $((\mathsf{P} \wedge \mathsf{Q}) \to \mathsf{R})$ to the conclusion that at least one of $(\mathsf{P} \to \mathsf{R})$ and $(\mathsf{P} \to \mathsf{S})$ is true! Can we really logically rule out dual key systems?!

Of course not. But again, in the light of our preliminary remarks, we can perhaps naturally the key-turning conditionals as counterfactual conditionals too. And we know that such conditionals don't behave like the material conditionals in (2).

Re (3): Again Richard Jeffrey offers a nice example.The situation is this: *if* the enemy retreats if we advance, then we will win the battle. But we are going to play safe and not advance. Never mind; we still win. Because – putting P for *we advance*, Q for *the enemy retreats*, R for *we will win* – we know that $((\mathsf{P} \to \mathsf{Q}) \to \mathsf{R})$ and $\neg\mathsf{P}$ entail victory, R! That's zapping the enemy with logic . . . !

But once more, it seems open to us to interpret the relevant conditional 'if the enemy retreats if we advance, then we will win the battle' as one of those future-tense possible-world conditionals which is not a plain this-worldly conditional and so isn't a candidate for being regimented by the material conditional in (3).

Still, there's a lot of wriggling going on in those comments! If we keep ruling out different classes of conditionals from being apt for regimentation using the arrow conditional of formal logic, just which ordinary-language conditionals are essentially material conditionals? We'll have to leave the question hanging.

(b) *Assess the following arguments for validity:*

(1) *Beth is a logician, provided that Amy is one too. And given that Beth is a logician, Chloe is a philosopher. But, hold on! – Chloe* isn't *a philosopher. So Amy isn't a logician.*

Intuitively valid. From the second and third premisses, we can infer that Beth isn't a logician (that's a version of modus tollens, treating the 'given' as expressing a conditional). And from *that* and the first premiss, we can infer that Amy isn't a logician (another modus tollens, treating 'provided that' too as expressing a conditional).

For a formal version, let's allow ourselves to use letters from the beginning of the alphabet as propositional letters. Then we can render the argument like this, with the obvious glossary:

$$(\mathsf{A} \to \mathsf{B}), (\mathsf{B} \to \mathsf{C}), \neg\mathsf{C} \ \therefore \neg\mathsf{A}$$

It is easy to check with a truth table that this formalized version is tautologically valid, and hence valid. [We can ignore all the lines where A is false – those can't be 'bad' lines with true premisses and a false conclusion. We can then ignore the two of the four remaining lines where C is true – those also can't be 'bad' lines with true premisses and a false conclusion. So that leaves just two lines to check . . .] And we can carry back this verdict to the original argument because 'provided that' and 'given' don't here seem to require more than a truth-functional connection.

(2) *Jo went up the hill. Why so? Because Jo went up the hill if Jack and Jill did. While Jack went up the hill if Jane did, and either Jane didn't go or Jill did. But Jane* did *go up the hill!*

Intuitively valid. Jane went up the hill. So from the second premiss we know that (i) Jack did too, and also (by disjunctive syllogism) that (ii) Jill went up the hill. But from (i) and (ii) together with the first premiss we can derive that Jo went up the hill.

To render the argument formally, use the glossary

> P: Jack went up the hill,
> Q: Jill went up the hill,
> R: Jane went up the hill,
> S: Jo went up the hill.

We then get

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

And running up a truth table to test this inference is entirely straightforward! First, (1) evaluate the conclusion at each line, and then (2) – forgetting the lines where the conclusion is true, and which therefore can't be 'bad' lines – evaluate the third premiss, which rules out another four lines as possibly 'bad'. Next (3) look at the first premiss, which rules out one more line. And that leaves just three lines of work to do at step (4) . . .

P	Q	R	S	$((\mathsf{P} \land \mathsf{Q}) \to \mathsf{S})$	$((\mathsf{R} \to \mathsf{P})$	$\land$	$(\neg\mathsf{R} \lor \mathsf{Q}))$	R	S
T	T	T	T						T
T	T	T	F	F				T	F
T	T	F	T						T
T	T	F	F					F	F
T	F	T	T						T
T	F	T	F	T		F	F	T	F
T	F	F	T						T
T	F	F	F					F	F
F	T	T	T						T
F	T	T	F	T	F	F		T	F
F	T	F	T						T
F	T	F	F					F	F
F	F	T	T						T
F	F	T	F	T	F	F		T	F
F	F	F	T						T
F	F	F	F					F	F
				3		4		2	1

So our argument is tautologically valid and hence valid; and since the truth-functional rendition doesn't seem to misrepresent the core content of the original argument, we can conclude that that too is valid (as we knew perfectly well!).

(3) *If Angharad speaks Welsh, so does either Bryn or Carys. But Bryn doesn't speak Welsh. So if Carys doesn't speak Welsh, neither does Angharad.*

This time let's start with a formal rendition, with the obvious glossary:

$$(\mathsf{A} \to (\mathsf{B} \lor \mathsf{C})), \neg\mathsf{B}\ \therefore\ (\neg\mathsf{C} \to \neg\mathsf{A})$$

A quick truth table shows it is tautologically valid and so valid. And we can carry back the verdict to the original argument.

Or informally we could argue like this. Suppose Carys doesn't speak Welsh. Then, by the second premiss, neither Bryn nor she speaks Welsh. So by the first premiss, Angharad doesn't speak Welsh. In sum, if Carys doesn't speak Welsh, neither does Angharad.

(4) *Veronika is a violinist only if Peter plays the cello. Marek is a violinist if and only if Jiří plays the viola. Veronika doesn't play the violin if Marek doesn't. Jiří doesn't play the viola only if Peter doesn't play the cello. Hence either Veronika isn't a violinist or Marek isn't.*

In real life, Jiří has now left the quartet, so the example isn't as nice! But pressing on .... Let's adopt the glossary

V: Veronika plays the violin,
M: Marek plays the violin,
J: Jiří plays the viola,
P: Peter plays the cello.

And we can render the argument like this:

$$(\mathsf{V} \to \mathsf{P}), ((\mathsf{M} \to \mathsf{J}) \land (\mathsf{J} \to \mathsf{M})), (\neg\mathsf{M} \to \neg\mathsf{V}), (\neg\mathsf{J} \to \neg\mathsf{P}) \therefore (\neg\mathsf{V} \lor \neg\mathsf{M})$$

We can then quickly see that the valuation which makes the four propositional variables all true will make the premisses true and conclusion false. So the inference is not tautologically valid. But it isn't valid for any other reason, so the inference fails.

(5) *It isn't the case that Popper is a good philosopher and Quine isn't. But in fact, Quine isn't a good philosopher, unless Russell is one too. And if Russell is merely a wordy fool, then he isn't a good philosopher. Hence if Popper is a good philosopher, then it isn't the case that either Russell is a wordy fool or isn't himself a good philosopher.*

For a reminder about how to handle 'unless' see Exercises 18 (a) 8. Using the obvious glossary, we can then render the argument

$$\neg(\mathsf{P} \land \neg\mathsf{Q}), (\mathsf{Q} \to \mathsf{R}), (\mathsf{W} \to \neg\mathsf{R}) \therefore (\mathsf{P} \to \neg(\mathsf{W} \lor \neg\mathsf{R}))$$

A truth table shows that this is tautologically valid and hence valid. Can we transfer this verdict to the original argument? It seems that that need involve no more than a thin truth-functional conditional and the rendition is faithful enough: so that argument is valid too.

Here's a direct argument for the validity of the original argument. Suppose Popper is a good philosopher. Then by the first premiss Quine must be too. So by the second premiss so Russell must be a good philosopher – and hence by the third premiss he can't be a wordy fool. So if Popper is a good philosopher, then Russell is a good philosopher and not a wordy fool. In other words, if Popper is a good philosopher, then Russell is neither a wordy fool nor a not a good philosopher.

(6) *If the switch had been flipped, the light would have been on. If the light had been on, the burglar would have been discovered. So, if the switch had been flipped, the burglar would have been discovered.*

Suppose we render this argument as having the shape $(\mathsf{P} \mathbin{\Box\!\!\to} \mathsf{Q}), (\mathsf{Q} \mathbin{\Box\!\!\to} \mathsf{R}) \therefore (\mathsf{P} \mathbin{\Box\!\!\to} \mathsf{R})$, using '$\Box\!\!\to$' for the counterfactual conditional. Now, a corresponding argument of the shape $(\mathsf{P} \to \mathsf{Q}), (\mathsf{P} \to \mathsf{R}) \therefore (\mathsf{P} \to \mathsf{R})$, with the familiar material conditional is valid: the plain arrow connective is transitive. But is $\Box\!\!\to$ transitive?

While transitivity holds for simple indicative conditionals it arguably doesn't hold for possible world conditionals of the kind that feature here. Take another example. I say, reasonably enough,

(i) If there had been snow in the valley yesterday, I would have gone skiing yesterday.

And yes, in nearby possible worlds in which there was enough snow yesterday, I'd have gone skiing. Now, it isn't the avalanche season – they are a rarity this time of year. But

(ii) If an avalanche had been taking place yesterday, there would have been snow in the valley yesterday.

In that remote possible world in which there is an avalanche, there would have been snow in the valley. But that remote circumstance isn't what we were considering in (i), we were considering the ordinary snow as forecast!

Now, I'm not suicidal. This is surely false:

(iii) If an avalanche had then been taking place yesterday, I would have gone skiing yesterday.

In that remote possible world in which there is an avalanche, I certainly don't go skiing! So here we have an inference of the shape $(\mathsf{Q} \;\Box\!\!\rightarrow \mathsf{R}), (\mathsf{P} \;\Box\!\!\rightarrow \mathsf{Q}) \therefore (\mathsf{P} \;\Box\!\!\rightarrow \mathsf{R})$ with – it seems – true premisses and a false conclusion.

To put the point abstractly, we evaluate the premiss (ii) *if P had been true, Q would have been true* and the conclusion (iii) *if P had been true, R would have been true* by looking at worlds as similar as possible to the actual world except that P is true there – and those are [in our example] very remote worlds. But we evaluate the premiss (i) *if Q had been true, R would have been true* by looking at worlds as similar as possible to the actual world except that Q is true there. And in our example that's a world much closer to the actual world. In sum, the worlds that are relevant to evaluating (ii) and (iii) aren't the worlds that are relevant to evaluating (i). That's why we can't smoothly combine our verdicts about (i) and (ii) together to get a verdict on (iii), and why arguments of the form $(\mathsf{P} \;\Box\!\!\rightarrow \mathsf{Q}), (\mathsf{Q} \;\Box\!\!\rightarrow \mathsf{R}) \therefore (\mathsf{P} \;\Box\!\!\rightarrow \mathsf{R})$ are not in general valid. [Or so goes a standard story.]

Back then to the burglar scenario. There's a video camera in the room where the valuables are stored. Recording is triggered by sound, which also turns the light on. Unfortunately the system hasn't been set up very well, and the sound trigger isn't very sensitive. And during the break-in last night, the burglar was unusually quiet and didn't trigger the system. *If the light had been on, the burglar would have been discovered*, but the light didn't go on, and he got away scot-free. Now, being a hyper-modern place, everything is controlled by voice-activated apps – the only switch in the building is behind glass, only to be broken in extreme emergency: that switch activates a klaxon alarm system. There was no unlikely emergency. But *if the switch had been flipped, the burglar would have been scared away and not discovered.* Though, in the extremely unlikely scenario of that alarm system being triggered all the lights go on, so in particular *If the switch had been flipped, the [relevant] light would have been on.* So (arguably!) in that scenario, the counterfactual premisses are true but the conclusion isn't.

But this claim, that counterfactual conditionals are not transitive, i.e. that arguments of the form $(\mathsf{P} \;\Box\!\!\rightarrow \mathsf{Q}), (\mathsf{Q} \;\Box\!\!\rightarrow \mathsf{R}) \therefore (\mathsf{P} \;\Box\!\!\rightarrow \mathsf{R})$ are not valid, is one more contestable and contested claim about conditionals, another issue we will just have to leave hanging.