Exercise 5 from WF122

Today's exercise is the transitivity of the implication, in symbols:

(0) $[\ (X \Rightarrow Y) \wedge (Y \Rightarrow Z) \quad \Rightarrow \quad (X \Rightarrow Z)\ ]$.

(I should have said "punctual transitivity" .)

My first concern in tackling such a large calculandum is proof shape. However, many narrow calculations depend implicitly on transitivity. For example:

$[\![$ Context: $(X \Rightarrow Y) \wedge (Y \Rightarrow Z)$

X

$\Rightarrow$ { $X \Rightarrow Y$ }

Y

$\Rightarrow$ { $Y \Rightarrow Z$ }

Z

$]\!]$

is a nice exercise in being orderly, but it hardly constitutes a proof of (0) , as the viability of the calculation depends on the transitivity of $\Rightarrow$.

What about the proof shape:

$(X \Rightarrow Y) \wedge (Y \Rightarrow Z)$

${}^{\Rightarrow}/_{\equiv}$ { ... }

$X \Rightarrow Z$?

I think this should only involve the normal (ie non-punctual) transitivity of $\Rightarrow$. First of all, if our calculation has only one weakening step, then we at most use Leibniz. For example:

EX5–1

$$A$$

$$\equiv \quad \{ \; ... \; \}$$

$$B$$

$$\Rightarrow \quad \{ \; ... \; \}$$

$$C$$

stands for:

$$[\; A \equiv B \;] \quad \wedge \quad [\; B \Rightarrow C \;] \qquad ,$$

which implies $[\; A \Rightarrow C \;]$ by Leibniz. If we have more than one weakening step, then this requires only normal transitivity. For example:

$$A$$

$$\Rightarrow \quad \{ \; ... \; \}$$

$$B$$

$$\Rightarrow \quad \{ \; ... \; \}$$

$$C$$

stands for:

$$[\; A \Rightarrow B \;] \quad \wedge \quad [\; B \Rightarrow C \;] \qquad ,$$

which implies $[\; A \Rightarrow C \;]$ by transitivity of $\Rightarrow$.

I propose to not use the transitivity of $\Rightarrow$. It is a stronger result if we don't use it, and besides, standard transitivity follows from punctual transitivity. Therefore, in our calculation, we will permit at most one implicational step.

The next question is what proof shape to use. To me, the most reasonable choices are:

$$(X \Rightarrow Y) \;\wedge\; (Y \Rightarrow Z)$$

$$^{\Rightarrow}/_{\equiv} \quad \{ \; ... \; \}$$

$$X \Rightarrow Z$$

or:

$[\![$ Context: $X \Rightarrow Y$

$Y \Rightarrow Z$

$^{\Rightarrow}/_{\equiv}$ { ... }

$X \Rightarrow Z$

$]\!]$

or:

$[\![$ Context: $Y \Rightarrow Z$

$X \Rightarrow Y$

$^{\Rightarrow}/_{\equiv}$ { ... }

$X \Rightarrow Z$

$]\!]$.

The contextual calculations are narrower, but destroy symmetry. Between the two of them, the latter seems to be the simpler, as the context holds $Y \Rightarrow Z$, and the syntactic goal is to change Y to Z in a weakening calculation. (In more sophisticated language, we are observing the monotonicity of $\Rightarrow$ in its right argument.)

Though I feel the wider calculation may be the more fruitful one here, I am eager to try the latter narrow calculation first. Remember that we are allowed at most one weakening step. Here we go!

$[\![$ Context: $Y \Rightarrow Z$

$X \Rightarrow Y$

$\equiv$ { From the start, I aim for an equivalence preserving step, knowing that we are allowed at most one weakening step. There are many ways to equivalently rewrite $X \Rightarrow Y$, but we should only do this with a goal in mind. Our goal is to replace Y with Z , and there is no non-trivial way to rewrite $X \Rightarrow Y$ which accomplishes this goal. Rather, we ought to use the context $Y \Rightarrow Z$, which has occurrences of the crucial symbols Y and Z . But how to use $Y \Rightarrow Z$, when we cannot assume monotonicity? The only option I can think of is to use Leibniz, which means we should use $Y \Rightarrow Z$ in one of the equivalent forms $Y \equiv Y \wedge Z$ or $Y \vee Z \equiv Z$. Of these two, the latter would require the introduction of "$\vee Z$". Weakening rule $[\, P \Rightarrow P \vee Q \,]$ comes to mind, but as I am not eager to use a weakening step immediately, I vote for using $Y \equiv Y \wedge Z$, which can be used obviously and immediately. Here we use punctuality of $\Rightarrow$ in its right argument. }

$X \Rightarrow Y \wedge Z$

$\Rightarrow$ { We have introduced Z : now we have to eliminate Y . Since we know we have to weaken somewhere, I propose to use $[\, Y \wedge P \Rightarrow P \,]$, a proposal further supported by the presence of "$Y \wedge$" already. To use this, we need to write $X \Rightarrow Y \wedge Z$ equivalently (!) in the form $Y \wedge P$ for some P . I really have no idea how to do this, as our only rewrites for $\Rightarrow$ introduce $\vee$ or $\equiv$. So we will have to investigate this step more widely. }

$X \Rightarrow Z$

$]\!]$.

That was rough, and we're not yet done. We still have to prove:

$$(1) \qquad [\ (X \Rightarrow Y \wedge Z) \ \Rightarrow \ (X \Rightarrow Z)\ ] \quad .$$

(I'm hoping to not use context $Y \Rightarrow Z$ again, since we already used it.) A narrow calculation will not help us in proving (1) —that's how we got stuck!— so we have to try for a wide calculation. The shape of (1) should make us wonder if $\Rightarrow$ enjoys distributivity with respect to $\Rightarrow$.

Towards this end, we explore proving:

$$(2) \qquad [\ P \Rightarrow (Q \Rightarrow R) \ \equiv \ (P \Rightarrow Q) \Rightarrow (P \Rightarrow R)\ ] \quad .$$

Implication can be rewritten as $\equiv$ and $\wedge/\vee$, and also as $\neg$ and $\vee$. Since $\vee$ distributes over $\equiv$ and $\wedge/\vee$, the proof should go through. (I am giving this motivation as a sketch because it is not my primary aim. For more on this design, see JAW46 .) Here is a proof of (2) :

$$
\begin{array}{ll}
 & P \Rightarrow (Q \Rightarrow R) \\
\equiv & \{ \ \Rightarrow \text{ into } \neg/\vee \ ; \ \Rightarrow \text{ into } \equiv/\wedge \ \} \\
 & \neg P \ \vee \ (Q \equiv Q \wedge R) \\
\equiv & \{ \ \vee \text{ over } \equiv \text{ and over } \wedge \ \} \\
 & \neg P \vee Q \ \equiv \ (\neg P \vee Q) \wedge (\neg P \vee R) \\
\equiv & \{ \ \equiv/\wedge \text{ into } \Rightarrow \ \} \\
 & \neg P \vee Q \ \Rightarrow \ \neg P \vee R \\
\equiv & \{ \ \neg/\vee \text{ into } \Rightarrow \ \} \\
 & (P \Rightarrow Q) \Rightarrow (P \Rightarrow R) \qquad ,
\end{array}
$$

so that we can now prove (1) as follows:

$$
\begin{array}{ll}
 & (X \Rightarrow Y \wedge Z) \ \Rightarrow \ (X \Rightarrow Z) \\
\equiv & \{ \ \Rightarrow \text{ over } \Rightarrow \ \} \\
 & X \ \Rightarrow \ (Y \wedge Z \Rightarrow Z) \\
\equiv & \{ \text{ weakening } \} \\
 & X \ \Rightarrow \ \mathbf{true} \\
\equiv & \{ \ \Rightarrow \text{ into } \neg/\vee \ \} \\
 & \neg X \vee \mathbf{true} \\
\equiv & \{ \text{ predicate calculus } \} \\
 & \mathbf{true} \qquad .
\end{array}
$$

Proving (1) settles (0) .

Now that we have gone this route, and developed some tools, other approaches present themselves. For example, now that we know that $\Rightarrow$ distributes over itself, we might massage (0) as follows:

$$
\begin{array}{ll}
 & (X \Rightarrow Y) \wedge (Y \Rightarrow Z) \ \Rightarrow \ (X \Rightarrow Z) \\
\equiv & \{ \text{ shunting } \} \\
 & (Y \Rightarrow Z) \ \Rightarrow \ ((X \Rightarrow Y) \Rightarrow (X \Rightarrow Z))
\end{array}
$$

and then calculate:

$[\![$ Context: $Y \Rightarrow Z$

$(X \Rightarrow Y) \Rightarrow (X \Rightarrow Z)$

$\equiv$ { $\Rightarrow$ over $\Rightarrow$ }

$X \Rightarrow (Y \Rightarrow Z)$

$\equiv$ { context: $Y \Rightarrow Z$ }

$X \Rightarrow$ **true**

$\equiv$ { see above }

true

$]\!]$.

The reader could fairly ask why we did not shunt (0) into the form:

$$(X \Rightarrow Y) \Rightarrow ((Y \Rightarrow Z) \Rightarrow (X \Rightarrow Z)) \quad ,$$

and explore the distributivity of $(\Rightarrow Z)$ over $\Rightarrow$. We certainly could have, but I imagine this distributivity is not as nice. (I tried it once before and will try it again in an ancillary EX , I think.)

Before we abandon the narrow approach to (0) , let us try the other narrow proof shape:

$[\![$ Context: $X \Rightarrow Y$

$Y \Rightarrow Z$

$\equiv$ { as before, using $Y \equiv X \vee Y$ and the punctuality of $\Rightarrow$ }

$X \vee Y \Rightarrow Z$

$\equiv$ { a predicate calculus theorem I know — see below }

$(X \Rightarrow Z) \wedge (Y \Rightarrow Z)$

$\Rightarrow$ { weakening }

$X \Rightarrow Z$

$]\!]$.

(Actually, the result of the first three lines of the calculation equivale the calculandum $(Y \Rightarrow Z) \ \Rightarrow \ (X \Rightarrow Z)$.)

Wow! That was really lovely. Of course I now see I missed the opportunity to use a similar move in the other narrow proof shape:

$[\![$ Context: $\quad Y \Rightarrow Z$

$\quad X \Rightarrow Y$

$\equiv \quad \{$ context: $\ Y \equiv Y \wedge Z$ and the punctuality of $\Rightarrow \ \}$

$\quad X \ \Rightarrow Y \wedge Z$

$\equiv \quad \{$ predicate calculus $\}$

$\quad (X \Rightarrow Y) \ \wedge \ (X \Rightarrow Z)$

$\Rightarrow \quad \{$ weakening $\}$

$\quad X \Rightarrow Z$

$]\!]$.

It still remains to prove:

(3a) $\quad [\quad P \vee Q \Rightarrow R \quad \equiv \quad (P \Rightarrow R) \wedge (Q \Rightarrow R) \quad]$

(3b) $\quad [\quad P \Rightarrow Q \wedge R \quad \equiv \quad (P \Rightarrow Q) \wedge (P \Rightarrow R) \quad]$.

These are standard lattice theory results. Rather than just prove them, we ought to design them.

Formulae (3) embody a sort of distributivity. Actually, (3b) is literally distributivity of $(P \Rightarrow)$ over $\wedge$, so it makes sense to write $(P \Rightarrow)$ as $(\neg P \vee)$, which we already know distributes over $\wedge$:

EX5–7

$$
\begin{array}{ll}
& P \Rightarrow Q \wedge R \\
\equiv & \{ \Rightarrow \text{ into } \neg/\vee \} \\
& \neg P \vee (Q \wedge R) \\
\equiv & \{ \vee \text{ over } \wedge \} \\
& (\neg P \vee Q) \wedge (\neg P \vee R) \\
\equiv & \{ \neg/\vee \text{ into } \Rightarrow , \text{ twice} \} \\
& (P \Rightarrow Q) \wedge (P \Rightarrow R) \qquad .
\end{array}
$$

Now, I know that the same approach, with de Morgan's rule, will prove (3a) . But how can that proof be designed?

If we are hesitant to apply negation to $P \vee Q$, we can be more non-committal, using different unfoldings of $\Rightarrow$ which still allow us to distribute R to both P and Q :

$$
\begin{array}{ll}
& P \vee Q \Rightarrow R \\
\equiv & \{ \text{predicate calculus} \} \\
& P \vee Q \equiv (P \vee Q) \wedge R \\
\equiv & \{ \wedge \text{ over } \vee \} \\
& P \vee Q \equiv (P \wedge R) \wedge (Q \wedge R)
\end{array}
$$

or:

$$
\begin{array}{ll}
& P \vee Q \Rightarrow R \\
\equiv & \{ \text{predicate calculus} \} \\
& (P \vee Q) \vee R \equiv R \\
\equiv & \{ \vee \text{ over } \vee \} \\
& (P \vee R) \vee (Q \vee R) \equiv R \qquad .
\end{array}
$$

But now I am stuck, as I don't know how to turn these into conjunctions. So I content myself with:

$P \vee Q \Rightarrow R$

$\equiv \quad \{ \ \Rightarrow \text{ into } \neg/\vee \ \}$

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

$\equiv \quad \{ \text{ de Morgan } \}$

$(\neg P \ \wedge \ \neg Q) \ \vee \ R$

$\equiv \quad \{ \ \vee \text{ over } \wedge \ \}$

$(\neg P \ \vee \ R) \ \wedge \ (\neg Q \ \vee \ R)$

$\equiv \quad \{ \ \neg/\vee \text{ into } \Rightarrow \text{ , twice } \}$

$(P \Rightarrow R) \ \wedge \ (Q \Rightarrow R)$,

even though the heuristics are not satisfying.

Before attempting a wide calculation of (0) , I want to observe that I missed the opportunity to use (3b) in my first proof, because I envisioned getting from $X \Rightarrow Y \wedge Z$ to $X \Rightarrow Z$ using weakening in the form $[\ Y \wedge Z \ \Rightarrow \ Z\ ]$. And in that proof, we did. However to prove (0) using (3b) , we never use weakening with these instantiations, but instead use $[\ (X \Rightarrow Y) \wedge (X \Rightarrow Z) \ \Rightarrow \ (X \Rightarrow Z)\ ]$.

* *
*

Finally, let's explore the proof shape:

$(X \Rightarrow Y) \ \wedge \ (Y \Rightarrow Z)$

$\Rightarrow/\equiv \quad \{ \ ... \ \}$

$X \Rightarrow Z$.

We know that a weakening step is necessary, hence that precisely one such step will occur. Moreover, that step will consist of dropping a conjunct. So we may refine our proof shape as follows:

$$
\begin{array}{ll}
 & (X \Rightarrow Y) \wedge (Y \Rightarrow Z) \\
\equiv & \quad \{ \text{ ... } \} \\
 & P \wedge Q \\
\Rightarrow & \quad \{ \text{ weakening } \} \\
 & P \\
\equiv & \quad \{ \text{ ... } \} \\
 & X \Rightarrow Z \qquad .
\end{array}
$$

The question now is how to refine P and Q . Unfortunately, the naive choice $P, Q := (X \Rightarrow Y), (Y \Rightarrow Z)$ does not work because the middle step would then eliminate Z from the manipulandum. To simplify the design, we might take $P := X \Rightarrow Z$, making the last step superfluous. The question then becomes how to design the following step:

$$
\begin{array}{ll}
 & (X \Rightarrow Y) \wedge (Y \Rightarrow Z) \\
\equiv & \quad \{ \text{ ... } \} \\
 & (X \Rightarrow Z) \wedge Q \qquad ,
\end{array}
$$

for some Q , which we may freely choose.

Syntactically, we are guided by the goal of joining X and Z together in the same $\Rightarrow$-expression , whereas in the top line of our calculation, they are each contained in $\Rightarrow$-expressions with Y , which are themselves linked by $\wedge$.

Since we are trying to regroup the subexpressions, it would be wonderful if we could rewrite $\Rightarrow$ as a conjunction, but we can only do this as a disjunction, viz $[\ P \Rightarrow Q \ \equiv \ \neg P \vee Q\ ]$. Distributivity of $\wedge$ over $\vee$ would allow us to group X and Z , but it is not immediately clear if this approach would work. So, let's explore:

$$(X \Rightarrow Y) \wedge (Y \Rightarrow Z)$$
$$\equiv \quad \{ \ \Rightarrow \text{ into } \neg/\vee \ \}$$
$$(\neg X \vee Y) \wedge (\neg Y \vee Z)$$
$$\equiv \quad \{ \ \wedge \text{ over } \vee \ \}$$
$$(\neg X \wedge \neg Y) \vee (\neg X \wedge Z) \vee (Y \wedge \neg Y) \vee (Y \wedge Z)$$
$$\equiv \quad \{ \ Y \wedge \neg Y \text{ is the identity of } \vee \ \}$$
$$(\neg X \wedge \neg Y) \vee (\neg X \wedge Z) \vee (Y \wedge Z) \qquad .$$

I honestly don't know how to proceed!

* *
*

At this point I have really exhausted what I can do with this problem. My thoughts have gestated for more than a week now, and I feel I've done the best I can. I'm quite proud of my three narrow proofs which use (2) , (3a) , and (3b) , respectively. It was an exhausting investigation, and I would like to see others' proofs!

The Hungarian Pastry Shop, 3 October 2009

Commentary: The one shape I did not consider was the widest one, that is, manipulating all of (0) . In this way, I failed to find the most traditional of all proofs of (0) , a version of which incidentally is found in Edsger Dijkstra's EWD1123 . Here is my version of that proof:

$$(X \Rightarrow Y) \wedge (Y \Rightarrow Z) \quad \Rightarrow \quad (X \Rightarrow Z)$$
$$\equiv \quad \{ \text{ shunting } \}$$
$$X \wedge (X \Rightarrow Y) \wedge (Y \Rightarrow Z) \quad \Rightarrow \quad Z \qquad ,$$

followed by:

EX5–11

$$X \;\wedge\; (X \Rightarrow Y) \;\wedge\; (Y \Rightarrow Z)$$

$\equiv$ { Modus Ponens: $[\;\; P \wedge (P \equiv Q) \;\;\equiv\;\; P \wedge Q \;\;]$ }

$$X \;\wedge\; Y \;\wedge\; (Y \Rightarrow Z)$$

$\equiv$ { Modus Ponens }

$$X \;\wedge\; Y \;\wedge\; Z$$

$\Rightarrow$ { weakening }

$$Z \qquad .$$

Not bad, but I far prefer my lattice theory proofs.

While we're on the subject, here's another proof, borrowed from Dijkstra:

$[\![$ Context: $(X \Rightarrow Y) \;\wedge\; (Y \Rightarrow Z)$

$$X$$

$\equiv$ { context: $X \Rightarrow Y$, equivalently $X \;\equiv\; X \wedge Y$ }

$$X \wedge Y$$

$\equiv$ { context: $Y \Rightarrow Z$, equivalently $Y \;\equiv\; Y \wedge Z$ }

$$X \;\wedge\; (Y \wedge Z)$$

$\equiv$ { associativity of $\wedge$ }

$$(X \wedge Y) \;\wedge\; Z$$

$\equiv$ { context: $X \Rightarrow Y$ }

$$X \;\wedge\; Z$$

$\Rightarrow$ { weakening }

$$Z$$

$]\!]$.

This is a lovely proof, but I've never found good heuristics for it!

NYC, 6 November 2009