A calculational example

In this note x and y are variables of some type E and P, Q and R are predicates on E, i.e. functions of type $E \to \mathbb{B}$ (where $\mathbb{B}$ denotes the boolean domain {true, false}).

We introduce square brackets [...] to denote what is known as "the everywhere operator", which is a function of type $(E \to \mathbb{B}) \to \mathbb{B}$. For any predicate P it is defined by

$$[P] \equiv \langle \forall x :: P.x \rangle \quad ,$$

the right-hand side to be read as "for all x (of type E), $P.x$ (= true)". The right-hand side has the problem that writing it down forces us to make an irrelevant choice, as we also could have written $\langle \forall y :: P.y \rangle$ instead. The left-hand side $[P]$ - read as "everywhere P" is freed from that clutter.

In order to render a universal quantification by the everywhere operator as given by (0), it is essential that the term be of the form $P.x$, i.e. a func-

tion -here P- applied to arbitrarily chosen dummy. As it stands, universal quantifications like $\langle \forall x :: \neg P.x \rangle$ or $\langle \forall x :: P.x \Rightarrow Q.x \rangle$ cannot be rendered by the everywhere operator because the terms $\neg P.x$ and $P.x \Rightarrow Q.x$ respectively do not have the required shape.

There are two ways out of this dilemma, viz. the introduction of new names and of "function expressions". The introduction of new names is mathematically very simple: if R is defined by

$$\langle \forall x :: R.x \equiv \neg P.x \rangle \quad , \quad \text{then}$$

$$[R] \equiv \langle \forall x :: \neg P.x \,] \quad .$$

Similarly, if R is defined by

$$\langle \forall x :: R.x \equiv (P.x \Rightarrow Q.x) \rangle \quad \text{then}$$

$$[R] \equiv \langle \forall x :: P.x \Rightarrow Q.x \rangle \quad .$$

It is very simple, but does not help much either: there is little we can do with $[R]$ without referring back to R's definition.

The way of introducing "function expressions" that I shall mention here is called "lifting". It is the notational

convention that application to an argument distributes to the left over the operators on the type of the function value. That is, we introduce for example the new expressions $\neg P$ and $P \Rightarrow Q$ given by

$$\langle \forall x :: (\neg P).x \equiv \neg(P.x) \rangle \qquad \text{and}$$

$$\langle \forall x :: (P \Rightarrow Q).x \equiv (P.x \Rightarrow Q.x) \rangle \ .$$

Notice that this convention implies "overloading" of the operators: at the right-hand side of $\equiv$, the boolean operators —here $\neg$ and $\Rightarrow$— have arguments and yield values of type $\mathbb{B}$, while at the left-hand side they have arguments and yield values of type $\mathbb{E} \to \mathbb{B}$.

It is the convention of lifting that allows us to apply the everywhere operator to a whole class of function expressions, e.g.

$$[\neg P] \equiv \langle \forall x :: \neg P.x \rangle \qquad \text{and}$$

$$[P \Rightarrow Q] \equiv \langle \forall x :: P.x \Rightarrow Q.x \rangle \ .$$

Finally, notice that $[P \equiv Q]$ expresses that P and Q are different names for the same predicate.

* * *

In the following, the only logical connectives we shall use are — listed in the order of increasing syntactic binding power — $\equiv$, $\Rightarrow$, $\wedge$. From predicate calculus we shall use that for all P, Q, R we have

(0) $[P \wedge Q \Rightarrow R] \equiv [Q \Rightarrow (P \Rightarrow R)]$

(1) $[P \Rightarrow Q] \equiv [P \wedge Q \equiv P]$;

without reference we shall use that $\wedge$ is

associative: $[P \wedge (Q \wedge R) \equiv (P \wedge Q) \wedge R]$
symmetric: $[P \wedge Q \equiv Q \wedge P]$
idempotent: $[P \wedge P \equiv P]$.

All the above we are going to use in our study of the "priming operator" given by the assertion that for all P, predicate P' is given by

(2) $[P' \equiv f.P \Rightarrow P]$

for some function f from predicates to predicates, i.e. of type $(\mathbb{E} \to \mathbb{B}) \to (\mathbb{E} \to \mathbb{B})$.

Our first lemma puts no further requirements on function f.

Lemma 0 : $[P \Rightarrow P']$.

Proof We observe for any P

$$
\begin{array}{ll}
 & [P \Rightarrow P'] \\
\equiv & \quad \{(2), \text{ the definition of priming}\} \\
 & [P \Rightarrow (f.P \Rightarrow P)] \\
\equiv & \quad \{(0) \text{ with } P, Q, R := f.P, P, R\} \\
 & [f.P \wedge P \Rightarrow P] \\
\equiv & \quad \{(1) \text{ with } P, Q := f.P \wedge P, P\} \\
 & [f.P \wedge P \wedge P \equiv f.P \wedge P] \\
\equiv & \quad \{\text{properties of } \wedge\} \\
 & \text{true}
\end{array}
$$

(End of Proof.)

Our next investigation explores what we may have to postulate about f, so that we can show that the priming operation is idempotent, i.e. that P' is a fixpoint of the priming operation, that is, in formula, that

$$[P' \equiv P''] \quad . \tag{3}$$

So we start a proof of the above. We observe for any P

$$
\begin{array}{ll}
 & [P' \equiv P''] \\
\equiv & \quad \{(2) \text{ with } P := P'\} \\
 & [P' \equiv f.P' \Rightarrow P'] \\
\equiv & \quad \{(2)\} \\
 & [P' \equiv f.P' \Rightarrow (f.P \Rightarrow P)] \\
\equiv & \quad \{(0) \text{ with } P, Q, R := f.P, f.P', P\}
\end{array}
$$

$$
\begin{array}{ll}
 & [P' \equiv f.P \wedge f.P' \Rightarrow P] \\
\equiv & \quad \{ (4), \text{ see below} \} \\
 & [P' \equiv f.P \Rightarrow P] \\
\equiv & \quad \{(2)\} \\
 & \text{true}
\end{array}
$$

and this completes the demonstration of (3), conditional on the demonstration of of (4).

Remark The above proof appeals 3 times to (2), the definition of the priming operation. This is no accident : (3) contains 3 primes. (End of Remark.)

The lemma (4) we need above is

$$(4) \qquad [f.P \wedge f.P' \equiv f.P] \quad .$$

To demonstrate this we observe

$$
\begin{array}{ll}
 & [f.P \wedge f.P' \equiv f.P] \\
\equiv & \quad \{ (1) \text{ with } P, Q := f.P, f.P' \} \\
 & [f.P \Rightarrow f.P'] \\
\Leftarrow & \quad \{ f \text{ is monotonic, see (5) below with } Q := P' \} \\
 & [P \Rightarrow P'] \\
\equiv & \quad \{ \text{Lemma 0} \} \\
 & \text{true}
\end{array}
$$

where "f is monotonic means" that we have for any P, Q that

(5) $\quad [P \Rightarrow Q] \;\Rightarrow\; [f.P \Rightarrow f.Q]$

and thus we have proved

Lemma 1 For monotonic f, the priming operation given by $[P' \equiv f.P \Rightarrow P]$ is idempotent.

Monotonic functions are fairly common. One way of establishing monotonicity is by

Lemma 2 A function that distributes over $\wedge$ is monotonic.

Proof Given that $f.$ distributes over $\wedge$ we observe for any P, Q

$$
\begin{array}{ll}
 & [\, f.P \Rightarrow f.Q \,] \\
\equiv & \quad \{ (1) \text{ with } P, Q := f.P, f.Q \} \\
 & [\, f.P \wedge f.Q \equiv f.P \,] \\
\equiv & \quad \{ f. \text{ distributes over } \wedge \} \\
 & [\, f.(P \wedge Q) \equiv f.P \,] \\
\Leftarrow & \quad \{ \text{Leibniz: substituting equals for equals} \} \\
 & [\, P \wedge Q \equiv P \,] \\
\equiv & \quad \{ (1) \} \\
 & [\, P \Rightarrow Q \,]
\end{array}
$$

with which (5) has been established.

(End of Proof.)

Because universal quantification distributes over $\wedge$ in a sense to be shown

below, we now consider an f satisfying for all R, x

$$(7) \quad (f.R).x \equiv \langle \forall y : y<x : R.y \rangle$$

where the right-hand side should be read as "for all y such that $y<x$, $R.y$ (holds)". Then f distributes over $\wedge$.

Proof We show $[f.(P \wedge Q) \equiv f.P \wedge f.Q]$ by observing for any P, Q, x

$$\begin{aligned}
& (f.(P \wedge Q)).x \\
\equiv & \quad \{ (7) \text{ with } R := P \wedge Q \} \\
& \langle \forall y : y<x : (P \wedge Q).y \rangle \\
\equiv & \quad \{ \text{lifting} \} \\
& \langle \forall y : y<x : P.y \wedge Q.y \rangle \\
\equiv & \quad \{ \text{distribution of } \forall \text{ over } \wedge \} \\
& \langle \forall y : y<x : P.y \rangle \wedge \langle \forall y : y<x : Q.y \rangle \\
\equiv & \quad \{ (7) \text{ with } R := P \text{ and with } R := Q \} \\
& (f.P).x \wedge (f.Q).x \\
\equiv & \quad \{ \text{lifting} \} \\
& (f.P \wedge f.Q).x
\end{aligned}$$

and as this holds for any x, we have proved $[f.(P \wedge Q) \equiv f.P \wedge f.Q]$.

(End of Proof)

The combination $(E, <)$ is called a <u>well-founded</u> set if the following form of proof by mathematical induction

is valid: in order to show $P.x$ for any x, it suffices to show that $P.x$ is true under the assumption that P holds for all smaller values, i.e.

(8) $\langle \forall y : y < x : P.y \rangle \Rightarrow P.x$ for all x.

With f as defined in (7), proof obligation (8) takes the form

$$[\, f.P \Rightarrow P \,] \quad ,$$

or, with priming defined as in (2),

$$[\, P' \,] \quad .$$

In other words: "$(E, <)$ is well-founded" means "for all predicates P on E, $[P'] \Rightarrow [P]$". That is, if $(E, <)$ is well-founded, we can prove $[P]$ by proving $[P']$ instead, which is in general easier since - see Lemma 0 - $[P \Rightarrow P']$. But it makes no sense to try to prove $[P']$ in turn by mathematical induction since the priming operation in question guarantees that $[P' \equiv P'']$.

Concluding remarks

This note is not about the (meta)theorem that the decision to prove something by

mathematical induction does not need to be repeated because it is idempotent (though some would consider this an unusual theorem and we have proved it simply) but about calculation. We should try to get an appreciation for which elements have made this calculation so effective.

There is to begin with "the everywhere operator": by doing away with the dummy of the universal quantifier $\forall$ it certainly contributes to the brevity of our formulae. A second issue is the decision to denote this operator by a parenthesis pair surrounding its argument. The observation that this argument is usually an expression of some complexity supports this notational decision. Writing $[P \equiv Q]$ instead of $P = Q$ deviates, however, from the tradition of denoting a binary relation by an infix operator, and some think this deviation too high a price to pay.

There is no controversy about the benefits of introducing the notational possibility of manipulating unapplied functions, nor about borrowing by "lifting" the operators needed to form function-valued expressions.

unapplied functions – P, $P \Rightarrow Q$, $f.P$, P'', etc. – occur all over the text.

I would like to point out that, though the introduction of the name R as on EWD 1278-1 did not buy us much, we used the same techique at great advantage in formula (7) for the definition of f.

Remark For the sake of completeness I mention that (7) can be viewed as two hidden λ-abstractions: we could have defined f by

$$f = \langle \lambda R :: \langle \lambda x :: \langle \forall y : y < x : R.y \rangle \rangle \rangle$$

but I don't consider that an improvement over (7). In my limited experience, the λ-calculus is better avoided when you don't need it. (End of Remark.)

The calculation further derives its crispness from the degree in which it has been disentangled. Our final goal of the calculation was a (meta)theorem about mathematical induction, but none of the specifics of mathematical induction enter the calculation:

- $[P \Rightarrow P']$ is demonstrated without any assumption about f,

- using the above, $[P' \equiv P'']$ is demonstrated assuming only f to be monotonic,
- independent of the above it is shown that f is monotonic if it distributes over $\wedge$,
- independent of the above it is shown that universal quantification distributes over $\wedge$,
- finally the link to mathematical induction is made: implication (8) has a universally quantified antecedent.

In short: "Ain't it a beauty?"

Austin, 10 September 1998

prof. dr Edsger W. Dykstra
Department of Computer Sciences
The University of Texas at Austin
Austin, TX 78712 - 1188
USA