Manipulating mathematical macros

A possible way of appreciating the relational calculus is to view relational expressions like R and $S;T$ as macros that, when expanded, generate xRy and $\langle \exists z :: xSz \wedge zTy \rangle$ respectively. The next step is to manipulate these abbreviations in their own right.

It seems worthwhile to point out that this technique can be applied quite generally, and to make our point we shall show an example that is <u>not</u> a standard notational device.

Suppose that we want to express that expression E is symmetric in p and q. Let $\sim E$ denote the expression that results when, in E, the parameters p and q are interchanged — formally

$$\sim E \; = \; (p, q := q, p). E \quad - \quad ;$$

the symmetry of E is then expressed by

(0) $\qquad [E = \sim E] \quad ,$

where the pair of square brackets denotes universal quantification over all global variables of E. Since for fixed p and q, $\sim E$ follows from E (and the other way round), we propose for the purpose of this discussion to abbreviate (0) to a macro

(1) $\qquad (\bullet \; E) \quad .$

And now we can list all sorts of rules about $\sim$ and $\bullet$, such as

(2) $[E = \sim\sim E]$

(3) $(\bullet\ E) \Rightarrow (\bullet\ \sim E)$

(4) $(\bullet\ E) \Rightarrow [f.E = f.(\sim E)]$

(5) $(\bullet\ \langle \underline{Q} n :: E_n \rangle) \Leftarrow (\bullet\ E_n)$ for any quantifier Q. (Note that in the antecedent universal quantification is included.)

* * *

With x and y ranging over some nonempty domain – "couples" in EWD1103 – we shall prove

(6) $\langle \forall x :: p.x = q.x \rangle$ from

(7) $\langle \uparrow y :: p.y \rangle = \langle \uparrow y :: q.y \rangle$ and

(8) $\langle \forall x,y :: p.x \downarrow q.y = q.x \downarrow p.y \rangle$.

Note that, in our new notation, demonstrandum and data are rendered:

(6a) $(\bullet\ p.x)$

(7a) $(\bullet\ \langle \uparrow y :: p.y \rangle)$

(8a) $(\bullet\ p.x \downarrow q.y)$.

The proof of (6a) $\Leftarrow$ (7a) $\wedge$ (8a) now proceeds as follows. We observe

$$
\begin{array}{ll}
 & (\bullet\ p.x) \\
= & \{\uparrow\downarrow \text{ calculus: } r \leq r\uparrow s \text{ and } a \leq b \equiv a = a\downarrow b\} \\
 & (\bullet\ p.x\downarrow\langle\uparrow y :: p.y\rangle) \\
= & \{\bullet\text{-calculus: (4) with (7a)}\} \\
 & (\bullet\ p.x\downarrow\langle\uparrow y :: q.y\rangle) \\
= & \{\uparrow\downarrow \text{ calculus: } \downarrow \text{ distributes over } \uparrow\} \\
 & (\bullet\ \langle\uparrow y :: p.x\downarrow q.y\rangle) \\
\Leftarrow & \{\bullet\text{-calculus: (5)}\} \\
 & (\bullet\ p.x\downarrow q.y) \\
= & \{(8a)\} \\
 & \text{true} \quad .
\end{array}
$$

And this concludes the example.

* * *

As said earlier, the next step is to manipulate these abbreviations "in their own right": what, up till now, were considered "rules" — like (2) through (5), and more — are partitioned into axioms and theorems — the latter being derived from the former — and thus a new universe of mathematical discourse is created.

Since this seems to be a quite general technique for mathematical innovation, we should have some insight in the items on the profit and loss account that accompanies its application.

At the profit side I can think of several items. There is firstly the abbreviation that is a direct consequence of the use of the macro; in the current example this is almost a gain of a factor of 2, but the relational calculus provides an example where this gain is bigger. Secondly, the rules of calculation can often be extended by a collection of theorems, the use of which can significantly reduce the number of steps needed: without the macros the theorems would have been too long and too elaborate to be attractive. Thirdly, the new universe of mathematical discourse can be truly more general than the original system. Subset calculus is, for instance, an overspecific model for our -pointless!- predicate calculus, and recently we saw the same macro manipulation cover both the relational calculus and the regularity calculus. In short, by discarding (notationally and conceptually) superfluous distinctions, the new universe can be truly simpler. Fourthly and finally, the new notation can suggest new ways of looking at things; for instance, it was despite EWD1103 and EWD1115 only this EWD that made me realize that a proof obligation like $\langle \forall x :: p.x = q.x \rangle$ can be seen as having to show that $p.x$ is symmetric in p and q .

But there is no such thing as a free lunch, and at the loss side are items as well. The design of a calculus requires all sorts of investments: new notations have to be designed, an axiomatization has to be chosen and a body of attractive theorems has to be constructed, and besides those formal rules we need strategies for using them. And after all these investments, the ability to exploit the available apparatus requires (as all our abilities) the maintenance of regular exercise. And finally we must report the recent experience —see WF147/148— how a proof in the relational calculus was markedly longer than a proof in the "clumsy" predicate calculus: the latter used an intermediate result that the relational calculus could not express.

Nuenen, 19 December 1991

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