# Heuristics for a calculational proof

Edsger W. Dijkstra[†]

Abstract Using a small, yet sufficiently sophisticated example, we show the heuristics that lead to a compact and clear solution. An implicit message is that the calculational style has the advantage that the formalism used provides major heuristic guidance for the design of the proof.

Keywords heuristics, calculational proofs, formal proofs, proof design, proof derivation, avoiding rabbits, formula manipulation.

[†] Department of Computer Sciences, The University of Texas at Austin, Austin TX 78712-1188, USA

The purpose of this note is to show a motivated design of a calculational proof and to give the reader an appreciation for the kind of heuristic guidance that (the properties of) the formalism used can provide. Our example stems from the Nineteenth International Mathematical Olympiad, 1977 (see [0]) and has the advantage that many people have solved it differently. The solution derived here is essentially the one by van de Snepscheut and Bird [1]; it has independently been designed by Misra [2].

With the natural numbers IN defined as all integers n such that $0 \leq n$ , and using an infix dot to denote function application, we can state the problem as follows:

"A total function f that maps natural numbers to natural numbers has the property that

$$f.(f.n) < f.(n+1) \qquad (0)$$

for all natural n . Show that f is the identity function."

* * *

It is only natural, though superfluous, to check that the identity function of type IN→ IN satisfies (0) ; it does since $n < n+1$ .

It is more instructive to try to check that all givens are needed. Certainly we need something like (0), for not any function of type IN→ IN is the identity function, and (0) cannot be weakened much either because

$$f.(f.n) \leq f.(n+1)$$

would admit for f any constant function. Also the constraint that f maps naturals to naturals is not void: had it been "integers to integers", $f.n = n-1$ would have been admissible.

In other words, our proof has to contain a step that is valid for the natural domain, but not the integer one. Since the naturals are well-founded whereas the integers are not, it is sweetly reasonable to propose

(α) In our proof of

$$f.n = n \qquad (1)$$

for all natural $n$ , we shall try to use mathematical induction over the natural numbers.

In following (α), it would be rash to conclude that (1) has to be our induction hypothesis: in view of our obligation to conclude equalities (1) where only inequalities (0) are given, it is sweetly reasonable to propose

(β) We shall try to construct a ping-pong argument in which

$$n \leq f.n \qquad \text{and} \qquad (2)$$

$$f.n \leq n \qquad , \qquad (3)$$

both for all natural $n$ , are dealt with separately.

In choosing which of the above two to prove by mathematical induction, the choice immediately falls on (2) since the base

$$0 \le f.0$$

is an immediate consequence of f's type $\mathbb{N} \to \mathbb{N}$. For the induction step we proceed by observing

$$
\begin{array}{cl}
 & n+1 \le f.(n+1) \\
= & \quad \{\text{arithmetic, heading for } (0)\} \\
 & n < f.(n+1) \\
\Leftarrow & \quad \{(0)\} \\
 & n \le f.(f.n)
\end{array}
$$

and here we are stuck, because $n \le f.n$ is too weak an induction hypothesis to conclude the last result from. So we had better backtrack and look for a stronger induction hypothesis.

Can we conclude a stronger base from the fact that $f$ is of type $\mathbb{N} \to \mathbb{N}$ ? Well, we can take that fact itself; in order to make it amenable to manipulation, we eliminate the symbol $\mathbb{N}$ by considering the naturals as a subset of the integers and formulate the fact that $f$ is of type $\mathbb{N} \to \mathbb{N}$ as

$$\langle \forall n :: \ 0 \le n \ \Rightarrow \ 0 \le f.n \rangle \qquad (4)$$

Asked for which induction hypothesis (4) acts as a proper base, any computing scientist that has designed invariants by replacing constants by variables will come up with the induction hypothesis

$$\langle \forall n :: j \leq n \Rightarrow j \leq f.n \rangle \qquad (5)$$

Concerning the decision to replace <u>both</u> 0's by the induction variable $j$, we point out that

- induction hypothesis $\langle \forall n :: j \leq n \Rightarrow 0 \leq f.n \rangle$ would lead to a trivial step and thus prevent the required exploitation of (0)
- induction hypothesis $\langle \forall n :: 0 \leq n \Rightarrow j \leq f.n \rangle$ would lead to a step that cannot be proved
- (5) does the job since

$$\langle \forall n :: n \leq f.n \rangle \Leftarrow \langle \forall j,n :: j \leq n \Rightarrow j \leq f.n \rangle \quad . (6)$$

There is a totally different reason why it is more attractive to prove (5) inductively for all $j$ than it is to prove (2) inductively for all $n$. The reason is that (2) contains the induction variable as argument of (the unknown) function $f$, whereas (5) contains the induction variable $j$ in perfectly manageable positions.

The base having been taken care of by (4), we now turn to the induction step to prove (5) by mathematical induction over $j$. To this end we observe for any natural $n$ and $j$

$j+1 \leq f.n$

$=$ {arithmetic, heading for (0)}

$j < f.n$

$\Leftarrow$ { (0) with $n := n-1$, i.e.
$f.(f.(n-1)) < f.n$ for $1 \leq n$ }

$j \leq f.(f.(n-1)) \wedge 1 \leq n$

$\Leftarrow$ {ex hypothese (5) with $n := f.(n-1)$ }

$j \leq f.(n-1) \wedge 1 \leq n$

$\Leftarrow$ {ex hypothese (5) with $n := n-1$}

$j \leq n-1 \wedge 1 \leq n$

$=$ { $0 \leq j$ }

$j+1 \leq n$ .

In view of (6) we have thus dealt with ping, i.e. (2) .

* * *

For pong – i.e. $f.n \leq n$ – we observe that mathematical induction over $n$ is (as yet) not indicated because the base is not obvious. To relate (3) to (0)

$$f.(f.n) < f.(n+1) \quad ,$$

we rewrite (3) as

$$f.n < n+1 \quad ,$$

i.e. the given (0) has at both sides of < an f-application more than the demonstrandum (3) . Now this looks very similar to monotonicity! In fact it is.

A usual way of expressing that $f$ is monotonic is that for any $x, y$

$$x \geq y \;\Rightarrow\; f.x \geq f.y \qquad ;$$

by taking the contrapositive, we get for total orders the alternative

$$x < y \;\Leftarrow\; f.x < f.y \qquad (7)$$

Under the assumption of $f$'s monotonicity, the demonstration of (3) is a walkover: we observe for any natural $n$

$$\begin{array}{ll}
 & f.n \leq n \\
= & \quad \{\text{arithmetic, heading for (7)}\} \\
 & f.n < n+1 \\
\Leftarrow & \quad \{\text{(7) with } x,y := f.n,\, n+1\} \\
 & f.(f.n) < f.(n+1) \\
= & \quad \{(0)\} \\
 & true \qquad ,
\end{array}$$

but this still leaves us with the obligation to demonstrate that $f$ is monotonic. (Note that the assumption was safe in the sense that the identity function is, indeed, monotonic.)

Monotonicity of a function $f$ on naturals can be expressed by an expression like (7) with 2 universally quantified dummies, or by

$$\langle \forall x :: f.x \leq f.(x+1) \rangle \qquad (8)$$

which quantifies over 1 dummy. The latter is usually the most convenient form to demonstrate monotonicity; the former, which includes the consequences of transitivity, is the most convenient characterization for the exploitation of monotonicity.

Remark The above paragraph covers a standard ingredient of the intellectual baggage of professional reasoners about sorting. (End of Remark.)

In order to demonstrate the monotonicity of $f$ , we prove (8) by observing for any natural $x$

$$\begin{array}{ll} & f.x \\ \leq & \quad \{ (2) \text{ with } n := f.x \} \\ & f.(f.x) \\ \leq & \quad \{ \text{from } (0) \text{ with } n := x \} \\ & f.(x+1) \qquad ; \end{array}$$

which concludes pong, and thus the whole proof.

* * *

A few final remarks. Firstly, had we written

$$\langle \forall n :: n \in \mathbb{N} \Rightarrow f.n \in \mathbb{N} \rangle$$

instead of (4), a simple mechanical routine would no longer have sufficed for the generalization to (5). This illustrates the advantage of avoiding special terminology, notations, or symbols (such as $\mathbb{N}$), i.e. of bandwidth reduction in general.

Secondly we would like to point out that the occurrence of a reductio ad absurdum often indicates that the author has failed to formulate a contrapositive, such as (7).

Acknowledgements I thank van de Snepscheut, Bird, and Misra for their inspiration, and the Tuesday Afternoon Clubs of Austin and of Eindhoven for their comments.

[0] Greitzer, Samuel L., "International Mathematical Olympiads, 1959 - 1977", Mathematical Association of America, 1978

[1] Snepscheut, Jan L.A. van de, JAN 161 "A little problem posed by R.S. Bird" d.d. 1989.12.11, California Institute of Technology

[2] Misra, Jayadev, Private Communication

Nuenen 11 July 1994