

Copyright Notice

The following manuscript,

EWD 1019: Another Filler of the YoP Institute

was published as

Edsger W. Dijkstra, ed., *Formal Development of Programs and Proofs*. Addison-Wesley, 1990: 209–227.

© 1989 Addison Wesley Longman Inc. Reproduced by permission of Addison Wesley Longman. All rights reserved.

Another filler of the YoP Institute

Why numbering should start at zero

When we have to characterize a set of adjacent natural numbers, say $\{2, 3, 4, 5, 6, 7, 8, 9, 10\}$, we have four options, depending on whether the bounds are included or not:

- (a) $1 < x < 11$
- (b) $1 < x \leq 10$
- (c) $2 \leq x < 11$
- (d) $2 \leq x \leq 10$

Natural numbers have the property that there is a smallest natural number. Different civilizations have made different choices for that minimum value; for the classical Greeks, for instance, it was 2, because, for them, 1 was not a number. (Consequently, Euclid had to introduce a case analysis in the justification of his algorithm for the greatest common divisor of two numbers, viz. the case that the two numbers had a common divisor versus the case that they had not, i.e. were relatively prime.)

This observation makes the convention of excluding the lower bound, as illustrated in (a) and (b) unattractive: if the range includes the smallest natural number - as the example

would for the Greeks - the convention of excluding the lower bound requires the introduction of an unnatural lower bound - as the examples (a) and (b) would do for the Greeks - . This is certainly inelegant and hence we adopt the principle of inclusion of the lower bound.

So we are left with the choice between (c) and (d). If we include the upper bound and consider the generalization $2 \leq x \leq n$, we see that for $n=2$ the range still contains the value 2. If we wish to represent the empty range by shrinking n still further, it would require $n=1$, which would be very unnatural for the Greeks. In general, the convention of including the upper bound requires the introduction of an unnatural upper bound; since this is certainly inelegant, we adopt the principle of exclusion of the upper bound.

In short, (c) is the winner: ranges of natural numbers will be indicated by

$$m \leq x < n$$

with natural bounds m and n , with $m \leq n$. Further advantages are

- the number of values in the range equals the difference of the bounds
- if the upper bound of one range equals the lower bound of another range, the ranges are

contiguous.

After these preliminary investigations we ask ourselves how we characterize the first n natural numbers. For the classical Greeks this would be

$$2 \leq x < n+2 \quad ;$$

were we to adopt the convention that the natural numbers start at 1, we would get the equally ugly

$$1 \leq x < n+1 \quad .$$

Obviously,

$$0 \leq x < n$$

is the most elegant formula; it corresponds to accepting zero as the smallest natural number. (This choice has further advantages.)

And now you know why my manuscripts start with page 0. It is really quite easy. When writing a manuscript, I have the completed pages behind me, spread out on the floor. The number on each new page equals the number of completed pages on the floor.

Austin, 1988 Jan. 25

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