

Largely on nomenclature.

Before embarking on a new book - something that is under consideration - I would very much like to decide (wisely!) with respect to notation and nomenclature all those conventions that will permeate all through the text. It seems wise not only to write such conventions down, but to include them for the sake of feedback in the EWD-series. Your comments are solicited!

For the pronunciation of the relations I propose the following:

for " $x = y$ "	read "x equals y"
for " $x \neq y$ "	read "x differs from y"
for " $x > y$ "	read "x exceeds y"
for " $x \geq y$ "	read "x is at least y"
for " $x < y$ "	read "x is less than y"
for " $x \leq y$ "	read "x is at most y"

The above reflects my current habit. I adopted "at least" and "at most" about 20 years ago; they are much shorter than "greater than or equal to" and "less than or equal to" (not to mention the horribly ungrammatical "greater or equal to"! ). An equally conscious, but much more recent, choice was "differs from". This latter choice is probably more fundamental; the trouble with "not equal to" and "unequal to" is that they treat equality as the basic concept and difference as its absence, an asymmetry

without justification. When introducing complementary concepts, we should find names for both and not arbitrarily call the one the negation of the other. The term "exceeds" was adopted quite unconsciously, but I like it so much that I can only regret having been unable to find as nice a verb to express "is less than".

For ordered sequences - a common topic in computing - we also need terms. I propose

$K(0) > K(1) > \dots$  is a "decreasing" sequence,  
 $K(0) \geq K(1) \geq \dots$  is a "descending" sequence,  
 $K(0) < K(1) < \dots$  is an "increasing" sequence,  
 $K(0) \leq K(1) \leq \dots$  is an "ascending" sequence.

The term "ascending" has been introduced to replace terms like "non-decreasing" and "weakly increasing". The term "non-decreasing" is even worse than "unequal": it suggests that a sequence is either decreasing or non-decreasing. Quod non. (It is only true for monotonic sequences; so we find "monotonically increasing" and "monotonically non-decreasing", but that I can only regard as a verbose patch: what would an increasing sequence be that isn't monotonic?) The adverb "weakly" as used in "weakly increasing" should be avoided for its hidden negation: a weakly increasing sequence is not necessarily an increasing sequence. Let a desperately angry animal always be an angry animal, independently of the meaning of "desperately". (From this point of view there would be no objection

to "strongly ascending" for "increasing".)

For 1, 2, 3, ..... I propose "the positive integers" and for 0, 1, 2, 3, ..... I propose "the natural numbers". There are two reasons to include zero in the set of natural numbers. Firstly there is the linguistic reason that, if zero were excluded, the term "natural number" would be superfluous: "positive integer" would do. Secondly there is the mathematical reason that in many an argument zero is a most "natural" number. (More natural, as a matter of fact, than many mathematicians are aware of. I have often seen the case analysis "either  $n$  is not divisible by  $p$ , or  $n$  has  $k$  factors  $p$ " while the argument for the second case was perfectly valid for  $k=0$ .)

Whenever a range of natural numbers has to be indicated, I shall do so lower bound included and upper bound excluded, as in the existential quantification in the following description of mathematical induction:

"In order to prove for some predicate  $P$   
 $(\underline{A} n: n \geq 0: P(n))$

one proves

$(\underline{A} n: n \geq 0: P'(n))$

with  $P'$  defined by

$P'(n) = P(n) \text{ or } (\underline{E} i: 0 \leq i < n: \underline{\text{non}} P(i)).$ "

Exclusion of the lower bound, as well as inclusion of the upper bound, would have forced bounds

outside the realm of natural numbers. I intend to let a lower bound never exceed an upper bound; then we have the (minor?) advantage that two ranges can be joined to form a single one if the upper bound of the one equals the lower bound of the other. Finally, the number of elements in the range equals the difference of the bounds. It now stands to reason to identify the elements of a sequence of length  $M$  - rows of a matrix or characters of a string - by a subscript in the range  $0 \leq \text{subscript} < M$ .

Minor remark. All this deviates from established mathematical tradition, and, consequently, from most of the programming tradition. Since the traditional conventions just "grew", it is not amazing that they leave room for improvement. What may amaze my reader, however, is the amount of importance I by now attach to such a detail. (It almost amazes myself.) But I cannot help it: the evidence is too overwhelming.

Let me just mention the most recent example I encountered. From the University of St. Andrews I received the manual for a programming language they had designed for educational purposes. I remember one example of its string handling facilities. The little program was supposed to strip off the leading zeros from a string of decimal digits. The text contained quite a few additions and subtractions of 1, which (except one addition) would have been superfluous if subscription of string characters had

started at zero. More seriously, the authors had written the manual under the illusion that the program could produce the error message "the input string is empty" without realizing that this case was indeed subsumed by the other error message "the input string consists of zeros only"! A very telling bug.

\* \* \*

The above text remained -not fully without reason- for a long time in my file of unfinished manuscripts. I now declare it finished.

4 March 1981  
Plataanstraat 5  
5671 AL NUENEN  
The Netherlands

prof. dr. Edsger W. Dijkstra  
Burroughs Research Fellow