

Essays on the nature and role of mathematical elegance (3)

On notation.

To be good a notation should satisfy at least three requirements: it should be unambiguous, it should be short, and it should be geared to our manipulative needs. Obvious as these requirements are, they are violated by wide-spread notational habits.

A rich deposit of poor notations is provided by the way in which telephone numbers are pronounced. In many countries telephone numbers are pronounced in pairs of digits: "83 18 27" becomes "eighty-three, eighteen, twenty-seven". Before we had a six-digit telephone number, we had "18 27", which, when pronounced according to this convention, cannot be distinguished from "18 20 07". In the Netherlands it is not uncommon to parse five-digit telephone numbers differently: "12 3 48" becomes "twelve, three, forty-eight" which is then audibly undistinguishable from "12 03 48". (In French it is even worse, because "72" is pronounced as "sixty-twelve" which makes it hard to distinguish between "11 60 12" and "11 72". As in French, "80" is pronounced as "four-twenty" and "91" as "four-twenty-eleven" the possibilities for ambiguity are in French still much richer.)

But apart from the ambiguity, the Dutch pronunciation is definitely not geared to our manipulative needs: telephone numbers are sequences of digits to be dialed in the order from left to right, but in Dutch --as with the old four-and-twenty blackbirds-- two-digit numbers are pronounced from right to left, so "83 18 27" becomes "three-and-eighty, eighteen, seven-and-twenty". These three inversions certainly don't facilitate dialing. (For me the only way out was to train myself to pronounce --in whatever language I am speaking-- telephone numbers as what they are, viz. sequences of digits, and I know my current telephone number as "eight, three, one, eight, two, seven".)

The most common way of sinning against shortness is repetition. Although the notation of music has a perfect way of indicating that, say, 16 bars have to be played twice, another music printing convention is that the printed text of a piece of music should end at the bottom of a page, an aesthetic convention which is sometimes met by "padding out" and printing the 16 bars twice. As a young boy trying to learn to play the piano I was always very annoyed

when I discovered that a music printer had played that trick upon me: with my untrained memory it was at that stage a laborious process to verify that the two printed sequences of 16 bars were, indeed, equal, a process that needed the forefingers of both hands for pointing at the bars to be compared.

Shortness is a virtue because lack of brevity can be harmful in many ways. Repetition forces the reader --see above-- to verify that long sequences of symbols are indeed identical; the repetitive style makes mathematical texts insipid, gives the writing of them the flavour of manual labour, and, when adopted, has a deteriorating influence on the quality of people's handwriting. (People write as well as they think necessary, and the redundancy of repetitive texts invites sloppiness.) Worse of all: it hides structure.

We have seen how in music printing the available structure "those 16 bars have to be played twice" can be hidden by printing them twice, but that was only a simple example of structure hiding. In a puzzle book I found the following problem: "Given twelve coins and the knowledge that one of them is false and of incorrect weight, determine the identity of the false coin and whether it is too heavy or too light by using a simple scale not more than three times." In the upside-down part of the book that contains the solutions, the solution of this puzzle occupies three pages of fine print. The author starts by numbering the coins from 1 through 12. Then he describes the first weighing as (1 2 3 4) against (5 6 7 8). For each of the three possible outcomes he describes the second weighing, each of which has again three possible outcomes. For these nine possible outcomes of the second weighing he prescribes a third weighing: only 24 of the 27 outcomes can occur and each of them miraculously identifies the false coin and establishes whether it is too heavy or too light. (Even the structure of the tree has been hidden by traversing it "depth first".) The structure of the solution was completely buried under the poor notation. Here is my effort.

Let us represent each coin by "0", "±", "+", or "-", depending on our knowledge about it, according to the following table:

0 = true
± = possibly false
+ = possibly too heavy
- = possibly too light

introduced "y" to stand for "+" or "-" (and "ȳ" to stand for the other!) . I could then have characterized case 1a as "y y x" . I haven't done so, because the introduction of the abbreviation would have cost more than it would have saved. Please note that the price for not introducing y and ȳ is paid:

in case 2a) where I am overspecific: comparing (- - +) with (- - +) would also have been right, (this is now left to the intelligence of the reader); in case 2b) where I have to enumerate the two forms "+ + -" and "- - +". The symmetry between "+" and "-" is exploited, however, by never separating explicitly, when the scale tips, the case that the one side goes down from the case that the other does so: in case 2a) we refer to "the heavy side" and "the light side", which, of course, should suffice because the starting situation of that weighing experiment was symmetrical.

The main gain from this descriptive technique is that it provides a catalogue of the possible situations with still 1, 2, or 3 weighings to go. If we were to extend the problem to, say, 36 coins and four usages of the scale, a description of the solution in our style would be one paragraph longer, in the style of the author of the puzzle book it would require nine pages of fine print.

(In the above remarks I have confined myself to the descriptive technique for the solution in the puzzle book. An alternative would have been to characterize all sets for which the problem can be solved by using the scale n times, by giving explicitly the general strategy for the first weighing of a sequence of at most n usages of the scale, and proving the adequacy of that strategy by induction over n . But then we would have taken the problem to a next level of sophistication that is beyond the scope of the puzzle book. Try to solve this more general problem: you will discover that its solution can be described clearly in about the same space as I needed for the problem of the 12 coins!)

(Just for the record: the description of the solution given above is the result of my second effort, polished twice.)

(To be continued)

Plataanstraat 5
5671 AL NUENEN
The Netherlands

prof.dr.Edsger W.Dijkstra
Burroughs Research Fellow