<u>Essays on the nature and role of mathematical elegance</u> (1)

<u>Introduction</u>.

I would not try to write these essays, if I did not have the feeling that they are urgently needed. I observe that the vast majority of my contemporaries cannot think very effectively --and in this respect our educational systems have something to answer for-- because they have not learned how to do so. I also observe a widespread underestimation of what can be achieved by thinking properly, and sometimes even an unwillingness to admit that, for certain goals, proper thinking is an indispensable prerequisite, an unwillingness of which a sad waste of effort, and ultimately, undeniable failure are the certain results. Finally, I observe the widespread assumption that effective thinking can be neither taught, nor learned. This last assumption is alarmingly common among those engaged in education. (Some will say that this is obvious because, if it were true, it would relieve our teachers from the responsibility to try to teach their students how to think effectively, but I find it hard to accept that explanation, for, if it were true, it would make the teacher's job rather futile.)

I regard the obvious consequences of these attitudes and assumptions as regrettable. We have surrounded ourselves by artefacts of needlessly clumsy designs, which make them cumbersome to develop and to maintain, and dangerous to rely upon; nevertheless they are relied upon, partly because we have the feeling that we have no choice, partly --and that is where I protest-- their imperfections have been raised to the status of a Law of Nature. Secondly we observe --mainly, but not exclusively, under the auspices of that branch of human activity and aspiration called "Artificial Intelligence"-- a great amount of research and development effort spent on trying to reduce the need for (what I am now forced to call) Natural Intelligence; I have reasons to expect these efforts to be vain. Thirdly, we see our educational systems affected: they are getting geared to "Otto Normalverbraucher" to the extent that it is now regarded as an indecency to require from a university graduate that he expresses himself adequately.

The purposes of these essays are threefold. Firstly, to show what effective thinking is all about, secondly to show the circumstances under which it is indispensable, and thirdly, to try to convince my readers that it can be taught. (Here a caveat has to be inserted: stating that something can be taught is not stating that

anybody, regardless of gifts, previous experience and personal objectives, can learn
it: it means that the ability to learn it is by no means restricted to a (statistic-
ally insignificant) minority of "natural geniuses".)

<div align="center">*   *   *</div>

I undertake the writing of these essays with great hesitation. From experience
I have learned that "thinking" and, even more so, the effectiveness with which we do
so, is a very touchy subject. It is so now, but it must always have been so: think-
ing is one of our most elusive, but also one of our most intimate activities, and
just considering the possibility that we are not doing it well enough might very
well hurt our ego. (As one of my former bosses once remarked: "Apparently, common
sense is the most equally distributed commodity, for no one complains that he has
not enough of it.".) From equally sad experiences I have learned that thinking
(and talking!) about ways of thinking and trying to evaluate their adequacy or in-
adequacy is spontaneously regarded by many (one's best friends and closest colleagues
included) as unacceptably presumptuous, and I have come to the conclusion that it is
a topic as unacceptable as sex among the Victorians. The mechanism is obvious. "To
be able to think, one must be a Genius. To be willing to talk about how to think,
one must not only be a Genius, but even know about oneself that one is a Genius."
But the whole point of these essays is that, although elusive and intimate, thinking
--like sex!-- is a most natural human activity: we all do it (and, as with sex: we
all do it as well, as we have learned to do it.)

Besides those emotional obstacles, further hurdles may be provided by current
views about what society is or should be. The teaching of effective thinking is
certainly the teaching of an ability, of a methodology, and the first effect of
teaching a methodology --rather than disseminating knowledge-- is that of enhancing
the capacities of the already capable, thus magnifying the difference in intelligence.
In a society in which the educational system is used as an instrument for the es-
tablishment of a homogenized culture, in which the cream is prevented from rising
to the top, the teaching of thinking could be politically unpalatable. This hurdle
can, indeed, be observed in many countries with socialist governments. Ironically
enough a view of management that is usually associated with undiluted capitalism,
may provide the same hurdle, namely the view that for the sake of stability and
continuity the industrial organization should be based on the employment of the
mediocre, because of those we have an unlimited supply. Carried to its extreme

this view results in the opinion that the real training of intellects is pointless, because no organization can take the risk of employing them and, thereby, making itself dependent on them.

In short, my topic is in many ways an impopular one. I know that that should not prevent me from tackling it, if I feel it as one of my most urgent tasks --and I feel that way-- . Yet I thought it wise to explain my "great hesitation": in doing so I may have removed (or at least: lowered) some of the obstacles.

*   *   *

I am a programmer and my direct concern is how to avoid the complexity that makes larger programs intellectually unmanageable and, as a result, not a safe tool to use. The attitude of the computing community towards this complexity is one of great ambivalence, and because this ambivalence seems greater than the computing community is eager to admit, it may clarify the situation when I describe its nature. On the one hand it is now quite openly admitted that in the whole process of using computers the limits of our programming ability present a more serious bottleneck than the intrinsic constraints of the hardware, and that efforts to trespass those limits result in systems of uncontrollable complexity. The attitude of many programmers, on the other hand, resembles that of the (later) saint who prayed "Dear Lord, please make me virtuous, but not yet.". Deep in their hearts the idea of complete intellectual control over their designs does not really attract them: many, I found, derive a great part of their professional excitement from not quite understanding what they are doing and from the glorious risks they take in their daring irresponsibility. Besides that it is very questionable, whether simplicity has any sales value. He who regularly addresses Western academic audiences quickly discovers that, on the average, his audience is impressed to the extent it has not understood him: by a perfectly understandable lecture many people in the audience feel somewhat cheated, and they leave the lecture hall afterwards, complaining to each other "Well, that was all rather trivial, wasn't it?". As a result, most audiences exert on most speakers a pressure --subconscious at both sides-- to be occasionally unnecessarily obscure. In analogy, I would not be amazed at all, if the customer exerted a similar pressure on manufacurers of hard- and software, and market analysis had shown that complexity sells better. I would not be amazed at all. (The recent IBM slogan "Management of Complexity" leads me to suspect that market analysis has, indeed, made that discovery; the slogan itself seems

IBM's answer to the 1975 Los Angeles Conference on Software Reliability, where
C.A.R.Hoare stated emphatically, that for reliability simplicity is an absolute
prerequisite.)   *   *   *
                                                 *

Most computer systems are ugly, and in general: the larger the uglier. One
of the purposes of these essays is to convince my colleagues that in (the practice
of!) automatic computing, elegance is not a dispensable luxury, but a must, a
"matter of life and death" so to speak.

I turned my attention to the role and nature of elegance in the wider context
of mathematics in general; I did so in the hope of extracting for my young and con-
fused field guidance from more than twenty-five centuries of mathematical experience.
Exploring the garden of mathematics I found, indeed, absolute beauties of simplicity!
As part of those explorations I asked all sorts of mathematical colleagues for con-
tributions: could they tell me a really elegant solution? When I had explained my
little project, most of them expressed grave doubts as to whether there existed such
a thing as --more or less objective-- mathematical elegance. Most of them felt
--because the beauty is supposed to be in the eye of the beholder-- that mathematical
elegance was too much dependent on personal taste and experience to have any guiding
value at all. I was pleased to discover that there was a much greater consensus
about what is mathematically elegant, than they themselves had suspected. The con-
sensus was, indeed, overwhelming. For me that was very encouraging, for it was a
strong suggestion that mathematical elegance, after all, might be not such an
elusive topic as it originally seemed to be.

Apparently, the lacking awareness of the general consensus about mathematical
elegance has prevented "elegance" from becoming a generally accepted criterion for
the quality of mathematical work. To my surprise, and somewhat to my disappointment,
I found a lot of mathematicians with respect to their own work not very "elegance-
conscious". I found them doing all sorts of clumsy things, clumsy things that I,
as a programmer, had already learned to avoid many years ago.

The reason why, so often, mathematicians can come away with a rather inelegant
way of working is probably that, at least in comparison to software projects, ma-
thematical projects are relatively "small". Yet I regret this general lack of
elegance: it makes mathematical texts less attractive to write and harder to enjoy,

and seems to impose an unnecessary limit on what mathematics can eventually achieve.
It is, therefore, my fervent hope that these essays will serve a dual purpose,
namely, that not only the practice of computing may profit from our mathematical
culture, but that also the practice of doing mathematics may profit from our com-
putational experience. It is the possibility of such a cross-fertilization that
makes the writing of these essays such an exciting challenge.

Essays on the nature and role of mathematical elegance (2)

On reasoning and pondering.

Of all our thinking activities, one class stands out very clearly, viz.
all manipulations that are formalized --or could readily be so-- by techniques
such as arithmetic, formula manipulation, symbolic logic, etc.. I shall denote
these activities by the term "reasoning", a term of which I hope that it is suf-
ficiently descriptive. A single name seems justified, because they have a few
common characteristics.

First of all, whenever they have been applied, in principle there never
needs to be an argument, whether they have been applied correctly or not, for each
step (which is always from a finite repertoire) can be checked.

Secondly, as soon as it has been decided in sufficient detail, what has to
be achieved by them, it is no longer a problem how to achieve it.

Thirdly --and this is not independent of the first two-- we know how to
teach them: arithmetic at primary schools, formula manipulation at secondary
schools, and symbolic logic at universities.

Fourthly, we are very good at doing modest amounts of reasoning. When
large amounts of it are needed, however, we are powerless without mechanical aids.
Multiplying two two-digit numbers is something we all can do; for the multiplication
of two five-digit numbers most of us would prefer the assistance of pencil and
paper; the multiplication of two hundred-digit numbers is a task that, even with

the aid of pencil and paper, most of us would not care to undertake.

These various instances of reasoning are emotionally appreciated quite differently, so differently as a matter of fact, that quite a few people (I discovered) wondered whether arithmetic should be classified as a thinking activity. But let us be careful: less than 500 years ago a professor of mathematics at one of Europe's universities stated that for the more gifted and industrious student it was not impossible to master long divisions: his students had to do it with Roman numerals! The majority of us might feel that the introduction of the decimal number system has reduced arithmetic to a boring routine, but that is probably only, because most of us are so poor at arithmetic, and can do no more than laboriously applying the routine steps. Calculating prodigies, however, have many different ways of arriving at a result, and, as a consequence, can get quite excited by their mental gyrations, excited because they can search for a shorter way.

As said: we are very good at modest amounts of reasoning. In many arguments, however, the amount of reasoning used often becomes the stumbling block, and I therefore relate the effectiveness of the way in which we have arranged our thoughts, to the extent in which we have been able to reduce the demands on our reasoning powers.

Let me give now just one example. In the late 18th century a German schoolmaster gave --with the intention of keeping his pupils busy for another hour-- the task of adding hundred terms of an arithmetic progression to a class of little boys who, of course, had never heard of arithmetic progressions. (To quote E.T.Bell: "The problem was of the following sort: 81297 + 81495 + ... + 100899 , where the step from one number to the next is the same all along (here 198) and a given number of terms (here 100) are to be added.") The youngest pupil, however, wrote down the answer instantaneously and waited, gloriously, with his arms folded, for the next hour while his classmates toiled: at the end of the hour it turned out that little Johann Friederich Carl Gauss had been the only one to hand in the correct answer. Young Gauss had seen instantaneously how to sum such a series analytically: the sum equals the number of terms, multiplied by the average of the first and the last term. In two respects this is a classical example: firstly young Gauss produced his answer about a thousand

times as fast as his classmates, secondly he was the only one to produce the correct answer. So much for the effective ordering of one's thoughts!

Prior to the existence of automatic computers the need to reduce, whenever possible, the demands on our limited reasoning powers was obvious. With today's possibility of mechanizing not only numerical computations, but also the other formal manipulations that we have classified as "reasoning", this need may seem less urgent. Some seem even to believe that the need no longer exists, but I believe that the latter are mistaken. Firstly, laboriously carrying out great numbers of avoidable steps of reasoning remains a waste, whether mechanized or not. Secondly, the discovery of a way of avoiding a large amount of reasoning often transcends the specific instance: the formula for the summation of arithmetic progressions did not only save young Gauss on that specific occasion an hour of hard labour. And thirdly, unless we can control --nearly in the sense of: "prevent"-- the growth of the amount of reasoning needed, mechanization will not help very long: without special precautions, that growth tends to be exponential. (The fact that, already now, so-called Artificial Intelligence projects tend to be most demanding as far as computing power is concerned should be a warning.) My conclusion is that the need to reduce, whenever possible, the amount of reasoning needed, is undiminished. In view of the fact that we are now trying to accomplish so much more difficult things than in the pre-computer age, I am even willing to argue that that need has increased.

<p style="text-align:center">*　　　*　　　*</p>

In order to learn as much as I could about how the amount of reasoning needed is reduced most effectively, I decided to have a closer look at (known) mathematical solutions of acknowledged elegance. In the beautiful arguments I encountered, the amount of reasoning has, indeed, been reduced very effectively: they are <u>very</u> <u>short</u>.

We should bear in mind,however, that just the criterion of length is often too crude. Sometimes, arguments have been shortened by omission: a big step is made, the writer has just been too lazy to spell it all out, and the reader has to supply the missing reasoning himself; that is clearly not the elegance we are looking for. Other arguments have been kept short by an appeal to knowledge that is often so extensive, that one cannot escape the feeling that

an egg is being cracked with a sledgehammer; and that isn't the kind of elegance
we are looking for either. An active mathematician commands a great manipulative
agility, an experienced mathematician has an extensive knowledge, but I think
that a wise one tries to use eventually as little as possible of either.

*   *   *

      I did some experiments with the following simple problem in classical
mechanics. Given a flat earth, a homogenous acceleration of gravity g , and
neglecting air resistance, we are asked to determine at what distance a cannon
ball hits the ground again, if it is shot away with an initial velocity v under
and angle phi with the horizontal direction. I have posed this problem to a
number of grown-up professional mathematicians. A surprisingly large fraction
could not really solve it --"It was too long ago that they had done problems
like that."-- , a few applied mathematicians solved it in a straightforward
manner by writing down the differential equations of motion, solving them,
adjusting the constants so as to satisfy the initial conditions, and finally
setting $y = 0$ in the equation for the trajectory. (The story doesn't tell
whether the eventual simplicity of the answer made them suspicious that they
could have reached the answer in a simpler manner.) Those that could not solve
it and those that solved it in the above manner, essentially reacted in the same
way: they tried to solve the problem by means of the reasoning techniques that
were their daily routine; for the majority this routine pattern was insufficient,
only the routine pattern of the applied mathematicians enabled those to solve
the problem (but only after a fashion, as we shall see in a moment).

      Two of them reacted quite differently: they gave the answer instantaneously
because they knew it. And they knew it because they had solved the problem at
the age of 14 --when both had posed it to themselves-- by a very simple argument,
so simple that it is impossible to forget it once you have seen it. Their argument
had to be simple because at that time 14-year old boys had not had analytical
geometry --and so they did not know what a parabola was--, and they did not know
what differential equations were, let alone that they knew how to solve them.
They did not know much more than the definition of velocity --"the rate of change
of position"-- and the definition of acceleration --"the rate of change of ve-
locity"-- and a rather informal introduction to Newtonian mechanics. The boys'
argument was as follows.

Because the horizontal speed is constant -- $v_x = v * \cos(\text{phi})$ -- the required distance is equal to $t * v_x$, where t is the time of travel through the air. During the time of travel, the vertical speed has changed from $v_y = v * \sin(\text{phi})$ --for reasons of symmetry!-- to $-v_y$, i.e. the total decrement of the vertical speed equals $2 * v_y$. Because the rate of change of vertical velocity is given to be g , we derive $t = 2 * v_y / g$, and the required answer is therefore

$$\frac{2 * v_x * v_y}{g}$$

(an answer of which all except the factor 2 can be derived by dimensional analysis, a derivation which essentially repeats the boys' argument.)

The little experiment is interesting in two ways. Firstly we can try to appreciate why the boys' argument is so much more attractive than the applied mathematicians' derivation of the parabolic trajectory. It is, because in the boys's argument the reasoning consists of two different parts: a --very simple-- argument about the horizontal motion and an --also very simple-- argument about the vertical motion, the two being combined by the very "narrow" interface of t , the time of travel through the air. This simplicity is lost as soon as one writes down the equation of the trajectory. Secondly I don't think it very noteworthy that the little boys each found the argument: they used about the only tools they had. But it is worth observing that the grown-up mathematicians immediately tried to solve the problem with the familiar reasoning apparatus of their daily routines! The ones that did not solve the problem should have realized the unsuitability of their daily used equipment --and should, to my taste, have found the boys' argument-- the applied mathematicians that solved the differential equations should, to my taste, have realized that they were cracking an egg with a sledgehammer --an activity I don't like, no matter how effective a sledgehammer may be for cracking eggs-- .

From this experiment I concluded (perhaps a little bit rapidly!) that while many mathematicians acquire a great manipulative agility in certain areas, and become very familiar with certain patterns of reasoning, relatively few tend to challenge the "adequacy" of those methods when faced with a "new" problem --where "inadequacy" may present itself in two ways: either by the inability to solve the problem at all, or by the inability to solve the problem

with an amount of reasoning near the minimum-- . Questions like "Is all that
reasoning going to be sufficient?" and "Is all that reasoning really necessary?"
seem to be ignored.

This is a very sad observation, because it is quite clear that a versatile
and effective problem solver must ask himself these questions all the time. And
when, evidently, many otherwise quite capable mathematicians do not ask them-
selves these questions all the time, I can only suspect that this is because
they haven't been taught to do so. Two observations, indeed, point in that
direction, the observation of our universities and the observation of our
scientific journals.

For centuries we know two ways of transmitting knowledge, insights, and
abilities to the next generation. The one is the tradition of the guilds, in
which the young apprentice works for seven meagre years with a master, absorbing
the skills implicitly, by "osmosis" so to speak. The other is the prevailing
tradition at the universities, where via explicit --and recorded!-- formulation
teachers try to bring knowledge and insights into the public domain. Till this
very moment, both teaching techniques are applied. Whereas, for instance, the
physicists adhere very strongly to the tradition of the universities, the
medical training takes to a large extent place as in the guilds. Mathematics
is in a somewhat in-between position: mathematical results are taught quite
openly, on the question how to do mathematics, however, most teachers are re-
markably silent. (And we can hardly blame them for that. Most mathematicians
quite understandably shun the "soft" subject of heuristics, they feel that they
can apply it, but that they lack the ability to teach it. Add to that a cul-
tural environment in which "thinking" is a taboo subject, and their silence on
how to do mathematics becomes quite understandable.)

Secondly, the same taboo controls the style of most publications. The
final results are published, but seldomly the reasons why the author looked
for them, and seldomly the way in which the author arrived at them. Such
"background noise" is regarded by many editors as irrelevant, and, as a result,
by many authors as "no one's business". (To sketch the heuristics and to disclose
some of the blind alleys one went into is really unusual: in a number of cases
where I did so, the referees' reports drew attention to my refreshing "courage".

The mere fact that referees use such terminology in their reports tells quite a tale: describing, or even indicating, how one has arrived at one's results is evidently appreciated as "exposing oneself".) It is one of my purposes to break this taboo: even if we cannot teach the skill of how to do mathematics as solidly as we can teach "mathematics itself", the very least we can do is not to hide the heuristics we used ourselves.

* * *

Besides the explicit and formal thinking activity that we have called "reasoning", a second thinking activity has now become apparent: all the considerations that help the experienced problem solver in choosing the most effective argument. I think that this second kind of thinking differs so much from the first one, that it deserves a special name; let us call it "pondering".

At least in the early days of Artificial Intelligence, the validity of the distinction between what we have called "reasoning" and "pondering" respectively, has been challenged violently, but that need not worry us. The Artificial Intelligentsia argued --and perhaps they still do!-- that the difference between reasoning and pondering was not essential because, ultimately, both would be reduced to the same activity: symbol manipulation. But the objections of the Artificial Intelligentsia don't worry me --and should not worry you-- for the following reasons. Firstly, if they think it profitable to view all mental activity as symbol manipulation, that is their business, but no one is forced to share their belief. I, for my part, don't think it very profitable (unable to see, for instance, how it can help me to view an important mental activity like falling in love --either with a sweet girl or with a sweet argument-- as an act of symbol manipulation). Secondly I have learned to view the dichotomy between "essential" and "gradual" differences --like most philosophical terms-- with the gravest suspicion, for it is usually fruitless and misleading: to present something --even if it seems technically allowed-- as a "gradual" and, therefore, "inessential" difference becomes a distortion each time the "gradual" difference is large enough. (Which computing scientist has not learned that a difference of only several orders of magnitude usually makes already "all the difference"!)

To maintain the distinction between "reasoning" and "pondering" seems

meaningful for the following reasons. Firstly, they serve different purposes:
it is the purpose of reasoning to present a solution in a convincing way, it
is the purpose of pondering to discover how the amount of reasoning needed
can be reduced. (Note that it is usually not the purpose of pondering and
reasoning together to solve a problem as quickly as possible. Suppose that
by one day of pondering we see a way of solving a problem via an argument that
takes another day to write down --and requires a similar amount of time to be
read, checked and understood-- . Compare this with the situation where four
days of pondering shows us a way of writing down the solution in fifteen
minutes --on the proverbial "backside of an envelope"-- . In the latter case
it took us twice as long to solve the problem, but our second solution is vastly
superior to the first one!) The second reason why it seems meaningful to me
to maintain the distinction between reasoning and pondering is that, their
techniques being so different from each other, they have to be taught in quite
different ways.

Eventually pondering should become one of the main topics of these essays.
Before we can approach that subject, however, we should try to get a clearer
idea of what makes an argument a really nice argument. For that purpose I
propose to turn my attention to some quantitative aspects of reasonings.

<div align="center">* * *</div>

Illustration. At primary school we were taught how to solve what we called
"encountering and overtaking problems". A typical specimen would be:

"At 12 o'clock the two hands of the clock coincide. What are the
other moments of the day at which this is the case?"

Another beloved form in which such an implicit linear equation would be posed
to us was the question where two walkers, starting at a given distance from
each other with given speeds, would meet. When we had made hundreds of such
sums, someone would pose us the following problem:

"At a given moment Piet starts walking from A to B ; at the same
moment a fly departs (with a greater speed) from his nose towards Jan,
who leaves at the same moment B on a walk to A . As soon as the fly
has reached the nose of Jan, it returns and flies back to the nose of
Piet, from where it returns again to the nose of Jan, etc. until it is

killed between the two colliding noses. If the distance between A and B is 20 km, Piet walks with a speed of 4 km/h, the fly flies with a speed of 13 km/h, and Jan walks with a speed of 5 km/h, how far has the fly flown before it is killed?"

Conditioned as we were , we first went into the foreseeable blind alley. When you have found a simple way to derive the answer (28.888... km), you will observe that you have arranged your reasoning in a pattern that is very similar to "the boys' argument" referred to above. (End of illustration.)

Plataanstraat 5   prof.dr.Edsger W.Dijkstra

5671 AL NUENEN   Burroughs Research Fellow

The Netherlands