On Alleged Mathematical Optimality

W. Kahan, Jan. 8, 1986

We mathematicians are notoriously inclined to assign our own arcane technical meanings to ordinary words laymen use every day. Words like *Root*, *Pole*, *Ring*, *Group*, *Field* and *Order* spring to mind. Some of these words are now ambiguous because different mathematicians have redefined them differently; does the word *Field* refer to an algebraic object (a Ring with inverse), or to a smooth vector-valued function of position in space? The word *Order* has had several different technical meanings.

Mathematical games with words cause mathematicians only a little trouble; but they cause a lot of trouble when laymen become involved without realizing fully that games are being played. For instance, laymen understand the word *Optimal* in an absolute sense; nothing can surpass something that is Optimal. But we mathematicians understand that word in a relative sense; we agree first to restrict the range of possibilities under consideration; then we prove that something is Optimal relative to those prior constraints. When we boast to laymen about our accomplishments, we tend to mention the constraints only in passing, as if they were obviously inescapable, certainly not worth repeating as often as we repeat the delicious word *Optimal* . That omission can cause trouble when laymen misconstrue our claim to Optimality as mathematically rigorous justification for a choice they believe to be best possible. Should later events reveal that the choice was substantially suboptimal, the layman will feel cheated as if by a salesman who hustled him past the fine print in a contract.

Whenever we convey mathematical results to laymen, and especially when we urge them to actions motivated by our results, we bear the onus of choosing our words carefully so that laymen will not likely be misled by an imperfect understanding of what we say. To succeed in that choice, we have to control the connotations that our words will arouse in minds that neither know nor want to know about details that we have mastered. We lose that control when we choose our words so carelessly as to invite subsequent casual and careless readers to misconstrue what we have written. Whom will they blame for their misunderstandings? Current trends in custom and in law tend to blame the expert more often than the layman. When our peers misconstrue what we have written we can chide them for overlooking some technical fine point, and they will blush; but criticizing a layman for inattention to petty detail will not exculpate us from our obligation to insulate him from what he, with some justification, regards as technicalities and jargon.

As an illustration, I have chosen an example so inconsequential that I had not expected the change in terminology I advocate to excite any objection, much less passion. What makes the example interesting is that it *has* aroused passions reminiscent of mediaeval theological arguments; I don't understand why.

The term *maximum accuracy* has been defined for the result of a floating-point computation to mean that no other floating-point number exists between the computed result and what would have been obtained exactly if there were no limitations upon accuracy and range. For instance, a computation of π = 3.14159 26535 ... on a machine that carries just six significant decimal digits might yield either 3.14159 or 3.14160 , both results of "maximum accuracy" according to its definition above. But if 3.14160 has "maximum accuracy," what describes the accuracy of 3.14159 ? "Super-maximum accuracy" ? Another instance is $\sqrt{100} = 10$, for which both 9.99999 and 10.0001 are also results of "maximum accuracy." Three different results seem too many all to have "maximum accuracy." The term is encumbered by further connotations that become apparent when we consider ...

- Machines that have more than one floating-point format, say Single, Double and Extended (Quadruple) precisions.
- Languages like Common LISP that include exact rational variables as well as approximate floating-point variables.
- Calculations whose accuracies are limited more by uncertainties in the data than by roundoff.

To prevent unfortunate collisions with its diverse connotations, I have recommended that the term "maximum accuracy" be avoided in the context of rounding errors. Instead, errors due to roundoff and similar causes should be measured in ulps ; the word *ULP* stands for a *U*nit in the *L*ast *P*lace, the difference between a computed result and a neighboring value representable in the same floating-point format. For instance, 3.14159 and 3.14160 differ by an ulp, as do 9.99999 and 10 , and 10 and 10.0001 . The definition of "maximum accuracy" above can be stated thus: the error is no worse than one ulp. When a result is correctly rounded in the usual sense, its error is no worse than half an ulp. The error in 3.14159 , regarded as a six-figure approximation to π , is 0.26535... ulps.

Nowadays the term "ulp" is used by numerical analysts almost universally despite two unpleasant properties. First, it is a jargon word that has to be explained to laymen, whereas a term like "maximum accuracy" seems at first to need no explanation. Actually, it has to be explained too, as we have seen. Second, the numerical value of an ulp seems ambiguous; an ulp of 10 is 0.00001 for numbers slightly less than 10 , but 0.0001 for numbers slightly larger than 10 . This is not *ambiguity*; it is *discontinuity*, and is intrinsic in all conventional floating-point schemes used on computers nowadays. Despite discontinuity, ulps have always worked smoothly and unambiguously in practice.

Why all this fuss about so little? Most people ignore roundoff; they don't care how we describe it. But we, who work to ensure that they can ignore roundoff safely, have to be fastidious about terminology. We should avoid terms like "maximum accuracy" and "optimal arithmetic" because they impose an unfair disadvantage upon anyone who dares to suggest that very different methods might produce results with smaller errors, or in less time.

I am not the first to protest against wishful thinking and misleading terminology in mathematics. The mathematician Charles Dodgson, writing under the pseudonym "Lewis Carroll," addressed a similar topic in a more whimsical way over a century ago. Here is an extract from *Through the Looking-Glass*:

...

" ... . There's glory for you!"

"I don't know what you mean by 'glory.'" Alice said.

Humpty Dumpty smiled contemptuously. "Of course you don't -- til I tell you. I meant 'there's a nice knock-down argument for you!'"

"But 'glory' doesn't mean 'a nice knock-down argument,'" Alice objected.

"When *I* use a word," Humpty Dumpty said, in rather a scornful tone, "it means just what I choose it to mean -- neither more nor less."

"The question is," said Alice, "whether you *can* make words mean so many different things."

"The question is," said Humpty Dumpty, "which is to be master --- that's all."

Alice was too much puzzled to say anything; so after a minute Humpty Dumpty began again. "They've a temper, some of them -- particularly verbs: they're the proudest -- adjectives you can do anything with, but not verbs -- however, *I* can manage the whole lot of them! Impenetrability! That's what *I* say!"

"Would you tell me please," said Alice, "what that means?"

"Now you talk like a reasonable child," said Humpty Dumpty, looking very much pleased. "I meant by 'impenetrability' that we've had enough of that subject, and it would be just as well if you'd mention what you mean to do next, as I suppose you don't mean to stop here all the rest of your life."

"That's a great deal to make one word mean," Alice said in a thoughtful tone.

"When I make a word do a lot of work like that," said Humpty Dumpty, "I always pay it extra."

...

Enter function(s) in the Input field(s):

Input ===> A - 2.0E-9

===> B - 5.0E8

===> X - (1-A)

===> X + A*Y - 1.0

===> 3.0*X - B*Y + (B-1.0)*Z - (2.0-3.0*A)

Enter the start values for the unknowns.

A ===> 2.0E-9

B ===> 5.0E8

X ===> 1.0

Y ===> 1.0

Z ===> 2.0

Press ENTER to start Newton iteration or END KEY (PF3) to terminate PF1=Help

IDLE

---------- Nonlinear Systems - Values ------ SINGULAR JACOBIAN

COMMAND ===>

Enter function(s) in the Input field(s):

Input ===> z - 2.0E-9

===> y - 5.0E8

===> X - (1-z)

===> X + z*b - 1.0

===> 3.0*X - y*b + (y-1.0)*a - (2.0-3.0*z)

Enter the start values for the unknowns.

A ===> 2.0

B ===> 1.0

X ===> 1.0

Y ===> 5.0e8

Z ===> 2.0e-9

Press ENTER to start Newton iteration or END KEY (PF3) to terminate PF1=Help

IDLE

---------- Nonlinear Systems - Approximation ------- 4 ITERATIONS

COMMAND ===>

The command V can be used to go back to the 'Values' panel.

Unknown	Result	Last correction
A	0.100000000000000D+01	-0.3289D-25
B	0.100000000000000D+01	-0.3289D-25
X	0.999999998000000D+00	-0.1283D-16
Y	0.500000000000000D+09	0.0000D+00
Z	0.200000000000000D-08	-0.3053D-25

Unknown	Result
A	(0.999999999999999D+00 , 0.100000000000001D+01)
B	(0.999999999999999D+00 , 0.100000000000001D+01)
X	(0.999999997999999D+00 , 0.999999998000001D+00)
Y	(0.500000000000000D+09 , 0.500000000000000D+09)
Z	(0.199999999999999D-08 , 0.200000000000001D-08)