Numerical Analysis Questions & Answers

Question by Student 201327118

*Proffesor, I dont understand that emax=2^8-1-1? Isnt emax=2^8-1=255? In this case IEEE single precision format In C, single precision is float float x; /*4Bytes finding Largest number*/ You teached emax=2^8-1-1=254 This is not a picture. i attached my note I drawed.*

image.jpeg

No attached pictures are allowed in the QNA thread except if they are drawings. The mathematics must be typeset within your question using LaTeX, not using pictures.

Question by Student 201327118

*Proffesor, I dont understand that emax=2^8-1-1? Isnt emax=2^8-1=255? In this case IEEE single precision format In C, single precision is float float x; /*4Bytes finding Largest number*/ You teached emax=2^8-1-1=254*

You haven't read correctly my response to your question. To give you more time to think carefully about how to formulate a question using LaTeX, you are now limited to one new question every 7 days.

Question by Student 201327103

professor, I don't know why $f_{min} = 0.00....1$ *in denormal number. I wonder why* f_{min} *is different in normal(*$f_{min} = 0.000...0$*) and denormal(*$f_{min} = 0.00....1$*)*

I don't understand the question fully. If I ask for the smallest positive number, then this can not be zero whether the number is normal or denormal because zero is not positive. I'll give you just 0.5 point bonus boost because you should put your question in more context so that I can understand better what you mean.

Question by Student 201029134

professor, When you explained ROUNDOFF ERROR, You solved
$x = -g + \sqrt{g^2 + 1}$ *with* $\varepsilon_{mach} = 10^{-8}$ *at Float*
But When you solved $x = \frac{1}{g+\sqrt{g^2+1}}$
you used $\varepsilon_{mach} = 10^{-7}$ *at Float*

What is the difference between two ε_{mach}

When using float variables in C, ϵ_{mach} should be set to 6×10^{-8}. But I won't take away points if you use a slightly more conservative value of 10^{-7}. The order of magnitude is what counts here. It was a good question, I'll give you 1.5 points bonus boost for it.

Question by Student 201327107

Professor, when you explained about float type machine precision you solved like this

$$\epsilon_{machine} = \frac{(1 * 2^{-24}) + 1}{1}$$

But I don't know how does it derived. Could you explain how it derived

No, this should read:

$$\epsilon_{\text{mach}} = \frac{1 + 2^{-24} - 1}{1}$$

The first two terms on the numerator $1 + 2^{-24}$ correspond to the sum of the smallest number 1 and the largest possible round-off error 2^{-24}. The last term on the numerator is the smallest number 1. I'll give you 0.5 point bonus boost.

Question by Student 201527145

Professor, I have a question about an assignment #1_Question #2_(d). Does the smallest possible number mean the "positive" smallest possible number? Or should I consider the "negative" smallest possible number?

Yes you are right: in Question A1Q2b and A1Q2d, we are seeking the smallest possible *positive* number. Thanks for pointing this out. I'll give you 2 points bonus boost.