

Homework 1: CS537, Fall 2008

Due Date: 3:15pm, September 11, 2008

Please show all steps in your work.

1. (5 points) Determine the decimal number that has the IEEE 32 bits representation of $[CA3F2900]_{16}$, show all steps.
2. (5 points) Prove that a real number has a finite representation in the binary number system if and only if it is of the form $\pm m/2^n$, where n and m are positive integers.
3. (10 points) Two numbers x and y that are not machine numbers are read into an IEEE 32 bits computer. The machine computes xy^2 . What sort of relative error can be expected? Assume no underflow or overflow.
4. (10 points) Find a good strategy to compute $f(x) = \sin x + \cos x - 1$ for x near 0. In five-decimal-digit arithmetic, compute $f(0.1)$ using the standard strategy and your better strategy, and compare the difference.
5. (10 points) For what range of x is the approximation $(e^x - 1)/2x \approx 0.5$ correct to 15 decimal digits of accuracy? Using this information, *write a function procedure* for $(e^x - 1)/2x$, producing 15 decimals of accuracy throughout the interval $[-10, 10]$ and show some computed results.
6. (10 points) Let x and y be two positive normalized floating-point machine numbers in the IEEE 32 bits representation. Let $x = q \times 2^m$ and $y = r \times 2^n$ with $\frac{1}{2} \leq q, r < 1$. Show that if $m = n$, then at least one bit of significance is lost in the subtraction of $x - y$.