

Homework 1: CS321-003, Spring 2007

Due Date: 1:50pm, January 29, 2007

Please show all steps in your work.

1. (10 points)
 - Convert $(45653.127664)_8$ to binary and to decimal, separately, show all steps.
 - Determine the decimal number that has the IEEE 32 bits representation of $[C553E000]_{16}$, show all steps.
2. (10 points) Let x, y and z be three machine numbers in the 32-bit machine. By analyzing the relative error in the worst case determine how much roundoff error should be expected in computing $(xy)z$. (Hint: $\text{fl}(xy) = xy(1 + \delta)$ with $|\delta| \leq 2^{-24}$.)
3. (10 points) Write a computer code in Fortran or C to compute the unit roundoff error (machine epsilon) on a given machine using double precision. You need to compute the largest positive machine number ϵ such that $1 + \epsilon = 1$.
4. (10 points) Show 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.
5. (10 points) Write a routine in Fortran or C that computes e^x by summing n terms of the Taylor series until the $(n + 1)$ st term t is such that $|t| < \epsilon = 10^{-6}$. Use the reciprocal of e^x for negative values of x . Test on the following data: 0, 1, -1, 0.5, -0.123, -25.5, -1776, 3.14159. Compute the relative error, the absolute error, and n for each case, using the exponential function on your computer system for the exact value. Sum no more than 25 terms.