CSC 336F ASSIGNMENT #1 Due: 12 October 2018.

This assignment is due at the **start** of your lecture on Friday, 12 October 2018.

For the questions that require you to write a MatLab program, hand-in the program and its output as well as any written answers requested in the question. Your program and its output, as well as your written answers, will be marked. Your program should conform to the usual CS standards for comments, good programming style, etc.

When first learning to program in MatLab, students often produce long, messy output. Try to format the output from your program so that it is easy for your TAs to read and to understand your results. For example, if you are asked to print a table of values, as in Question 5, print them as a table that fits on one page. To this end, you might find it helpful to read "A short description of fprintf" on the course webpage. Marks will be awarded for well-formatted, easy-to-read output.

Also, your TAs will appreciate your using a word processor to write the answers to questions (or parts of questions) that do not require a program. If you do write those answers by hand, make sure that they are easy to read.

1. [6 marks: 2 marks for each part]

   What are the approximate absolute and relative errors in approximating π by the following values?

   (a) 3.14

   (b) 3.14159

   (c) 3.141592654

   For the purposes of this question, you can assume that the "true" value of π is 3.14159265358979.

   Correctly round each error to three significant digits (using the *round-to-nearest* rounding rule.)

   You might find it helpful to use a computer or a calculator to assist you with the required arithmetic for this question.

2. [10 marks: 1 mark for each part]

   In a floating-point number system with parameters $\beta = 10$, $p = 3$, $L = -10$ and $U = +10$ that uses the *round-to-nearest* rounding rule and allows gradual underflow with subnormal numbers, what is the result of each of the following floating-point arithmetic operations?

   (a) $5.35 \cdot 10^0 + 2.46 \cdot 10^{-2}$

   (b) $4.53 \cdot 10^1 - 6.38 \cdot 10^{-1}$

   (c) $5.65 \cdot 10^1 + 5.23 \cdot 10^{-4}$

(d) $6.54 \cdot 10^4 - 8.73 \cdot 10^6$

(e) $5.21 \cdot 10^8 \times 4.25 \cdot 10^{-5}$

(f) $-4.32 \cdot 10^7 \times 3.25 \cdot 10^3$

(g) $5.41 \cdot 10^{-5} \times 4.27 \cdot 10^{-5}$

(h) $-6.52 \cdot 10^{-6} \times 4.75 \cdot 10^{-6}$

(i) $-6.46 \cdot 10^{-7} \times 1.32 \cdot 10^{-6}$

(j) $3.82 \cdot 10^{-6} \times 1.25 \cdot 10^{-7}$

Write each answer as a normalized 3-decimal-digit floating-point number, if possible. If that is not possible, write your answer as either a subnormal 3-decimal-digit floating-point number or zero, if that is the most accurate representation. If that is not possible either, then write your answer as +Inf, −Inf or NAN, whichever best represents your answer.

You might find it helpful to use a computer or a calculator to assist you with the required arithmetic for this question.

3. [10 marks: 5 marks for each part]

Consider the function $f(x) = \tan(x)$ for $x \in (-\pi/2, \pi/2)$.

(a) Is this function well-conditioned or ill-conditioned in a relative sense with respect to small relative changes in the value of the input argument x for x close to 0?

Is this function well-conditioned or ill-conditioned in a relative sense with respect to small relative changes in the value of the input argument x for x close to $\pi/2$? (Note: $x < \pi/2$, but close to $\pi/2$.)

Justify your answer for each case.

(b) Write a little MatLab program to verify your predictions from part (a).

Hand in your MatLab program and its output.

Also include a brief explanation of why you believe your computational results from part (b) support your theoretical predictions from part (a).

4. [10 marks: 5 marks for each part]

(a) Do question 2 on the 2017 final exam.

(b) Do question 3 on the 2017 final exam.

You can find the 2017 final exam on the webpage

`http://www.cs.toronto.edu/~krj/courses/336/exam.2017.pdf`

5. [15 marks: 5 marks for each part]

(a) Write a MatLab function exp1 to approximate e^x by summing the series

$$1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots$$

from left to right until the accumulated sum stops changing.

Test your program by computing exp1(x) for $x = -25, -24, -23, \ldots, +25$ (i.e., x = $-25 : +25$ in MatLab).

For each value of x, compute the relative error

$$\frac{\text{exp1}(x) - \exp(x)}{\exp(x)}$$

where $\exp(x)$ is the MatLab function that approximates e^x, and print both x and the relative error.

For the purpose of this question, assume $\exp(x) = e^x$.

Format your output neatly.

(b) For what values of x does your function produce accurate approximations to e^x and for what values of x does your function produce poor approximations to e^x?

Explain why your function performs well in the cases where it produces accurate approximations to e^x and also explain why your function performs poorly in the cases where it produces poor approximations to e^x.

Don't just say that it performs poorly because there is rounding error. There is rounding error in your computations for all values of x (except possibly $x = 0$). However, in some cases the rounding errors are insignificant and you obtain a good approximation to e^x, while in other cases the rounding errors are significant and you obtain a poor approximation to e^x. Explain why.

(c) Make a small change to your function exp1 so that it produces accurate approximations to e^x for all $x = -25, -24, -23, \ldots, +25$. Call your new function exp2. (If you find it helpful, you can call exp1 from within exp2.)

For each value of x, compute the relative error

$$\frac{\text{exp2}(x) - \exp(x)}{\exp(x)}$$

where $\exp(x)$ is the MatLab function that approximates e^x, and print both x and the relative error.

Format your output neatly.

Hint: note $e^x = 1/e^{-x}$.