UNIVERSITY OF VICTORIA
MIDTERM EXAM 1 OCTOBER 6 2020
COMPUTER SCIENCE 349A

NAME: Zheng Yin STUDENT NO. V00915261
INSTRUCTOR: Rich Little

DURATION: 50 minutes

YOU MAY ANSWER ON THE EXAM PAPER, ON YOUR OWN PAPER OR TYPESET YOUR ANSWERS AND RETURN A SINGLE PDF FILE TO THE ASSIGNMENT SUBMISSION WINDOW

STUDENTS MUST COUNT THE NUMBER OF PAGES IN THIS EXAMINATION PAPER BEFORE BEGINNING TO WRITE, AND REPORT ANY DISCREPANCY IMMEDIATELY TO ME.

THIS QUESTION PAPER HAS 4 PAGES INCLUDING THIS COVER PAGE.

NOTES: (0) OPEN BOOK EXAM; ONLY BASIC CALCULATORS ARE ALLOWED,
(1) THERE ARE A TOTAL OF 18 MARKS,
(2) YOU MAY NOT USE THE INTERNET FOR ANY PURPOSE OTHER THAN THE COURSE WEBSITE ON CONNEX WITH THE EXCEPTION OF AN ONLINE PDF TOOL IF NECESSARY

Question	Possible marks	Actual marks
1	6	
2	6	
3	6	
Total	18	

1. Consider an alien floating-point number system that is base-8 (octal) (maybe they look like octopuses and live in an ocean-world). Answer the following questions using a $k = 2$ precision normalized floating-point representation with $b = 8$:

   (a) [2 points] What is the floating-point representation in this alien system of the decimal number 128.6 using rounding?

   (b) [2 points] What is the largest number (in decimal) that can be represented without loss of accuracy in this system if the exponent contains 2 digits in octal?

   (c) [2 points] How many distinct, non-zero positive numbers can be represented without loss of accuracy in this system if the exponent contains only 1 digit in octal and is signed, i.e. is positive or negative?

(a) $k=2$

$128.6_{10} =$

$128/8 = 0$
$16/8 = 0$
$2/8 \quad 2$

$128_{10} = 200_8$

$0.6\times8 = 4.8 = 4$
$0.8\times8 = 6.4 \quad 6$
$0.4\times8 = 3.2 \quad 3$
$0.2\times8 = 1.6 \quad 1$
$0.6\times8 = 4.8 \quad 4$
$0.8\times8 = 6.4 \quad 6$
$0.4\times8 = 3.2 \quad 3$
$0.2\times8 = 1.6 \quad 1$

$0.6_{10} = 0.463463$

$128.6_{10} \approx 200.463463_8$

$k=2$ precision normalize

$200.463463_8 = 0.200463463\times8^3 = 0.20\times8^3$

$\therefore 0.20 \times 8^3$ is the floating-point representation in this alien system

(b) $\boxed{S_m \,\vdots\, b \,|\, e \,|\, e \,|\, e \,|\, d_1 \,|\, d_2}$

The largest $b^e = 8^{77_8}$ in this number system

$77_8 = 63$

$7\cdot8 + 7\cdot1 - 1 = 63$

$8^{77_8} = 8^{63} = 0.78\times10^{57}$ The largest number is 0.78×10^{57}

(c) $k=2$ precision

$\pm0.d_1d_2 \times 8^{e_m}$

For e_m is negative $-1 \sim -7$

$7\cdot8\cdot7 = 392$

For e_m is postive

$7\cdot8\cdot8 = 448$

$392 + 448 = 840$

Total 840 non-zero Positive number.

2. (a) [4 points] Evaluate

$$g(x) = \frac{x - \sin x}{x^3}$$

using $b = 10$, $k = 4$ floating-point arithmetic with *chopping* where $x = 0.02468$ radians.

(b) [2 points] Given that the actual value is $g(0.02468) = 0.1666616$, what is the relative error of your approxiamtion in (a)?

(a) with $b=10$ $k=4$ chopping

$fl(\sin(x)) = fl(\sin(0.02468)) = fl(0.0246774941) = 0.02467$

$fl(x - \sin(x)) = fl(0.02468 - 0.02467) = fl(0.00001) = 0.00001$

$fl(x^3) = fl(0.00001503\,2) = 0.00001503$

$fl\left(\frac{x - \sin x}{x^3}\right) = fl\left(\frac{0.00001}{0.00001503}\right) = fl(0.665335994) = 0.6653$

(b) $V_A = 0.1666616 \times 10^0$

$V_E = 0.6653$

$\delta = \left|\frac{0.1666616 - 0.6653}{0.1666616}\right| = \left|\frac{-0.4986384}{0.1666616}\right| = 2.991921354 \approx 299.1\%$

The relative error is 299.1%

3. [6 points] The evaluation of $f(x) = 1 - \cos x$ in floating-point arithmetic suffers from subtractive cancellation when x is close to $2\pi = \underline{6.2831853}...$ radians. Illustrate this by evaluating $f(6.27)$ using $b = 10$, $k = 4$ floating-point arithmetic with rounding and comparing to the actual value. Determine an alternate way of evaluating $f(x)$ that does not suffer from subtractive cancellation for this same value.

$fl(f(x)) = fl(1-\cos(x)) = fl(fl(1) - fl(\cos(x)))$

Subtractive cancellation of x. keep 8 significant digit

$x = 6.2831853$

$fl(1) - fl(\cos(6.2831853))$

$= fl(1) - fl(1)$

$= fl(1-1)$

$= 0.000 \times 10^0$

$fl(f(6.27)) = fl(fl(1) - fl(\cos(6.27))) = fl(fl(1) - fl(0.999913075))$

$= fl(1. - 0.9999)$

$= fl(0.0001)$

$= 0.0001$

true Value $= 1 - \cos 6.27$

$= 0.00008692$

$\left| \frac{0.00008692 - 0.0001}{0.00008692} \right| = 0.1504$

$\approx 15\%$

Too large

If we use rounding

$(1-\cos x) = (1-\cos x) \cdot \left(\frac{1+\cos(x)}{1+\cos(x)} \right) = \frac{1-\cos^2(x)}{1+\cos(x)} = \frac{\sin^2(x)}{1+\cos(x)}$

with $x = 6.27$

$fl(\sin(x)) = fl(-0.013184925) = -0.01318$

$fl(\sin^2(x)) = fl(0.000173712) = 0.0001737$

$fl(\cos(x)) = fl(0.999913075) = 0.9999$

$fl(1+\cos(x)) = fl(1.9999) = 2.000$

$fl(\sin^2(x)/1+\cos(x)) = fl(0.0001737/2) = fl(0.00008685) = 0.00008685$

RE: $\left| \frac{0.00008692 - 0.00008685}{0.00008682} \right| = 0.000805338 \approx 0.805338\%$

Smallerr