Exam I Name:____________________

MA 47700/MA 57700, Instructor: Nicolae Tarfulea
September 25, 2024

There are four partial credit questions. Please show all your work for full credit; displaying only the final answer will earn you very little, if anything.

Very important: throughout this exam, wherever they appear,

$$\alpha = 1 \qquad \beta = 2 \qquad \gamma = 3$$

1. (30 pts.) (a) Determine the internal representation of the decimal number $\alpha\beta\gamma.b_{-1}b_{-2}b_{-3}b_{-4}$ in the normalized floating-point number system of a 32-bit word length microcomputer.

sign of mantissa (1 bit)	exponent (7 bits)	normalized mantissa (24 bits)

Here, the fractional part $b_{-1}b_{-2}b_{-3}b_{-4}$ is one of the numbers 4375, 5625, 6250, and 8750, at your choice.

Example: $\alpha\beta\gamma.b_{-1}b_{-2}b_{-3}b_{-4} = 123.5625$

$= (1111011.1001)_2$

$= (0.11110111001)_2 \times 2^7$

$= (0.11110111001)_2 \times 2^{111_2}$

$0.4375 = (0.0111)_2$

$0.5625 = (0.1001)_2$

$0.6250 = (0.101)_2$

$0.8750 = (0.111)_2$

Answer:

0	0000111	11110111001000 ... 0

1 pt. 3 pts. 6 pts.

(b) Determine the internal representation of the **same** decimal number $\alpha\beta\gamma.b_{-1}b_{-2}b_{-3}b_{-4}$ you considered in part **(a)** in the IEEE single-precision format

sign (1 bit)	biased exp. c (8 bits)	fractional part f of the normalized mantissa (23 bits)

$$123.5625 = (1111011.1001)_2 = 1.1110111001_2 \times 2^6$$

$$c - 127 = 6 \Rightarrow c = 133 = (10000101)_2$$

1 pt. 4 pts. 5 pts.

Answer:

0	10000101	1110111001000 --- 0

(c) Convert the following 32-bit machine number to decimal. Here, s, e_5, e_4, e_2, m_2, m_3 are either 0 or 1 at your choice (there are $2^6 = 64$ possibilities; choose one).

s	$00e_5e_41e_21$	$1m_2m_3100000000000000000000$

State your choice of $s =$, $e_5 =$, $e_4 =$, $e_2 =$, $m_2 =$, and $m_3 =$

$$(-1)^s \left(0.1m_2m_31\right)_2 \times 2^{(e_5e_41e_21)_2} = \quad \text{4 pts.}$$

$$= (-1)^s\left(\frac{1}{2} + \frac{m_2}{2^2} + \frac{m_3}{2^3} + \frac{1}{2^4}\right) \cdot 2^{(e_5 \cdot 2^4 + e_4 2^3 + 1 \cdot 2^2 + e_2 \cdot 2^1 + 1)}$$

$= \ldots$ 6 pts.

Answer:

2. **(25 pts.)** Consider the **normalized** floating point number system $\mathbf{F}(10, 3, -7, 7)$.
(a) Provide the floating point equivalent for the number $\sqrt{\alpha+\beta+\gamma}$ by both chopping and rounding.

$F(10,3,-7,7) = \{\pm(0.d_1 d_2 d_3)_{10} \times 10^{e} : -7 \le e \le 7\}$ Here, $d_1 \neq 0$ unless $d_2 = d_3 = 0$.

$\sqrt{\alpha+\beta+\gamma} = \cdots$

Answer: $fl_{chop} =$ → 3 pts. $fl_{round} =$ → 3 pts.

(b) Determine the machine precision (consider both chopping and rounding), the smallest positive number, and the largest positive number.

$$u = \begin{cases} \beta^{1-k} & \text{by chopping} \\ \frac{1}{2}\beta^{1-k} & \text{by rounding} \end{cases} = \begin{cases} 10^{-2} & \text{by chopping} \\ \frac{1}{2}10^{-2} & \text{by rounding} \end{cases}$$ → 2 pts.

Smallest positive number $= (0.100)_{10} \times 10^{-7} = 10^{-8}$ → 2 pts.

Largest positive number $= (0.999)_{10} \times 10^{7} = 9{,}990{,}000$ → 2 pts.

(c) List the first three numbers greater than $1.\alpha\beta$ in $\mathbf{F}(10, 3, -7, 7)$.

$1.\alpha\beta = (0.1\alpha\beta) \times 10^{1}$

First three numbers greater than $1.\alpha\beta$:

$(0.1\alpha(\beta+1))_{10} \times 10^{1} =$, $(0.1\alpha(\beta+2))_{10} \times 10^{1} =$, $(0.1\alpha(\beta+3))_{10} \times 10^{1} =$

9 pts.

(d) How many real numbers are in this system?

$$|F(10,3,-7,7)| = 1 + 2(10-1)\cdot 10^{3-1}(7-(-7)+1)$$
$$= 1 + 2\cdot 9\cdot 100\cdot 15 = 27{,}001$$ → 4 pts.

3. (20 pts.) **(a)** Using interval arithmetic, compute a solution set for $[1+\alpha, \alpha+\beta+\gamma]\, X = [\beta,\, \beta+\gamma]$.

$$X = [\beta, \beta+\gamma] \div [1+\alpha, \alpha+\beta+\gamma]$$

$$= [\beta, \beta+\gamma] \cdot \left[\frac{1}{\alpha+\beta+\gamma}, \frac{1}{1+\alpha}\right]$$

$$= \left[\frac{\beta}{\alpha+\beta+\gamma}, \frac{\beta+\gamma}{1+\alpha}\right] \longrightarrow 8 \text{ pts.}$$

(b) Using interval arithmetic, compute bounds on the range of the function:

$$f(x,y,z) = (z-\alpha x)(\beta y + x) + \frac{z}{y} + \gamma xyz, \quad -1 \le x \le 1,\ 1 \le y \le 2,\ -2 \le z \le 2$$

$$f([-1,1],[1,2],[-2,2]) = ([-2,2] - \alpha[-1,1])(\beta[1,2] + [-1,1])$$

$$+ [-2,2] \div [1,2] + \gamma[-1,1]\cdot[1,2]\cdot[-2,2]$$

$$= [-2-\alpha, 2+\alpha]\cdot[\beta-1, 2\beta+1] + [-2,2]\cdot\left[\tfrac{1}{2}, 1\right]$$

$$+ \gamma[-2,2]\cdot[-2,2]$$

$$= [(-2-\alpha)(2\beta+1), (2+\alpha)(2\beta+1)] + [-2,2] + \gamma[-4,4]$$

$$= [(-2-\alpha)(2\beta+1) - 2 - 4\gamma, (2+\alpha)(2\beta+1) + 2 + 4\gamma]$$

$$= [-(2+\alpha)(2\beta+1) - 2 - 4\gamma, (2+\alpha)(2\beta+1) + 2 + 4\gamma]$$

$$\longrightarrow 12 \text{ pts.}$$

4. (25 pts.) **(a)** Approximate the integral $\int_0^1 e^{\sin(\gamma x)}dx$ by replacing the integrand $e^{\sin(\gamma x)}$ with its first three terms of the Taylor series expansion about 0.

$f(x) = e^{\sin(\gamma x)}$, $f'(x) = \gamma e^{\sin(\gamma x)} \cos(\gamma x)$

$f''(x) = \gamma^2 e^{\sin(\gamma x)} \cos^2(\gamma x) - \gamma^2 e^{\sin(\gamma x)} \sin(\gamma x)$

$f(x) \simeq f(0) + \frac{f'(0)}{1!} x + \frac{f''(0)}{2!} x^2 = 1 + \gamma x + \frac{\gamma^2}{2} x^2 \rightarrow$ 6 pts.

So, $\int_0^1 e^{\sin(\gamma x)} dx \simeq \int_0^1 \left(1 + \gamma x + \frac{\gamma^2}{2} x^2\right) dx$ $\rightarrow$ 6 pts.

$\simeq \left(x + \gamma \frac{x^2}{2} + \frac{\gamma^2}{2}\frac{x^3}{3}\right)\Big|_0^1 = 1 + \frac{\gamma}{2} + \frac{\gamma^2}{6}$

(b) Approximate $\sin\left(\frac{5\pi}{6}\right)$ using the first four terms of the Taylor series expansion of $\sin\left(\frac{\pi}{2} + x\right)$ about 0.

$f(x) = \sin\left(\frac{\pi}{2} + x\right)$, $f'(x) = \cos\left(\frac{\pi}{2} + x\right)$, $f''(x) = -\sin\left(\frac{\pi}{2} + x\right)$

$f'''(x) = -\cos\left(\frac{\pi}{2} + x\right)$

$f(x) \simeq f(0) + \frac{f'(0)}{1!} x + \frac{f''(0)}{2!} x^2 + \frac{f'''(0)}{3!} x^3$

$\simeq 1 + \frac{0}{1!} x + \frac{-1}{2!} x^2 + \frac{0}{3!} x^3 = 1 - \frac{1}{2} x^2 \rightarrow$ 7 pts.

$\sin\left(\frac{5\pi}{6}\right) = f\left(\frac{\pi}{3}\right) \simeq 1 - \frac{1}{2}\left(\frac{\pi}{3}\right)^2 \simeq 0.4517 \rightarrow$ 6 pts.

1	2	3	4	Total
30	25	20	25	