Name _____

Student Number _____

Instructor: S. Qiao

SFWR ENG 3X03/COMP SCI 4X03

Day Class
Duration of examination: 50 minutes
McMaster University Midterm Examination October 2012

This examination paper includes **5** pages and **6** questions. You are responsible for ensuring that your copy of the paper is complete. Bring any discrepancy to the attention of your invigilator.

SPECIAL INSTRUCTIONS: This paper must be returned with your answers. Use of McMaster standard (Casio-FX991) calculator only is allowed.

1. Give the IEEE single precision binary representation of each of the following decimal numbers. For example, the IEEE single precision binary representation of +2.0 is

```
0 10000000 00000000000000000000000
```

(a) (1 mark) -3

Answer: `1 10000000 10000000000000000000000`

(b) (1 mark) +0.25

Answer: `0 01111101 00000000000000000000000`

(c) (1 mark) +3.5

Answer: `0 10000000 11000000000000000000000`

2. Give the IEEE single precision binary representation of each of the following floating-point numbers.

(a) (1 mark) NaN (not a number)

Answer: `0 11111111 10000000000000000000000`

(b) (1 mark) The machine precision ϵ_M

Answer: `0 01101000 00000000000000000000000`

(c) (1 mark) $+0.10000000000000000000000 \times 2^{-126}$

Answer: `0 00000000 10000000000000000000000`

3. Give the IEEE single precision floating-point approximation of each of the following real numbers. For example,

$$\text{fl}(0.1) = 1.10011001100110011001101 \times 2^{-4}$$

(a) (1 mark) $\text{fl}(0.3) =$

Answer: $1.00110011001100110011010 \times 2^{-2}$

(b) (1 mark) $\text{fl}(1/3) =$

Answer: $1.01010101010101010101011 \times 2^{-2}$

(c) (1 mark) $\text{fl}(3.3) =$

Answer: $1.10100110011001100110011 \times 2^{1}$

4. Assuming a floating-point number system where $\beta = 10$, $t = 4$, $e_{\min} = -15$, and $e_{\max} = +16$, consider the evaluation of the expression

$$(7.654 + 5.432) - (7.651 + 5.423).$$

(a) (1 mark) What is the *exact* result?

Answer: 0.012

(b) (1 mark) What is the computed result? (Do not remove parentheses.)

Answer: 0.02

(c) (1 mark) How would you rearrange the order of the operations to improve the computed result?

Answer: $(7.654 - 7.651) + (5.432 - 5.423) = 0.012$

(d) (1 mark) Explain your results.

Change catastrophic cancellation $(7.654+5.432)-(7.651+5.423)$ to begign cancellations $(7.654-7.651)$ and $(5.432 - 5.423)$.

5. Let the upper triangular matrix

$$A = \begin{bmatrix} 1 & -2 & -2 & -2 \\ & 1 & -2 & -2 \\ & & 1 & -2 \\ & & & 1 \end{bmatrix}.$$

(a) (2 marks) Find the last column of A^{-1}:

Answer: Soving the triangular system

$$\begin{bmatrix} 1 & -2 & -2 & -2 \\ & 1 & -2 & -2 \\ & & 1 & -2 \\ & & & 1 \end{bmatrix} x = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix},$$

we get the last of column of A^{-1}:

$$\begin{bmatrix} 18 \\ 6 \\ 2 \\ 1 \end{bmatrix}$$

(b) (1 mark) Find A^{-1}:

Answer:

$$\begin{bmatrix} 1 & 2 & 6 & 18 \\ & 1 & 2 & 6 \\ & & 1 & 2 \\ & & & 1 \end{bmatrix}.$$

6. (8 marks) Consider a system where $t = 3$, $\beta = 10$, $e_{\min} = -15$, and $e_{\max} = +16$. Apply the Gaussian elimination with pivoting

```
[A, p] = decomp(A)
```

(without condition number estimation) to the matrix:

$$A = \begin{bmatrix} 4.00 & 5.00 & 6.00 \\ 1.00 & 2.00 & 3.00 \\ 3.00 & 2.00 & 0.00 \end{bmatrix}.$$

Show the steps and the changes of the entries of A and the pivoting vector p, which is initialized as

$$\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}.$$

Step 1. A and p:

$$A \to \begin{bmatrix} 4.00 & 5.00 & 6.00 \\ 0.250 & 0.750 & 1.50 \\ 0.750 & -1.75 & -4.50 \end{bmatrix} \qquad p \to \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$$

Step 2. A and p:

$$A \rightarrow \begin{bmatrix} 4.00 & 5.00 & 6.00 \\ 0.750 & -1.75 & -4.50 \\ 0.250 & -0.429 & -0.430 \end{bmatrix} \qquad p \rightarrow \begin{bmatrix} 1 \\ 3 \\ -1 \end{bmatrix}$$