

Name:

Id#

COE 301/ICS 233, Term 171

Computer Architecture & Assembly Language

Quiz# 4 Solution

Date: Thursday, Nov. 16, 2017

- (i) Q1. What is the decimal value of the following single-precision floating-point number?

1100 0010 1110 1101 1000 0000 0000 0000. (2 Points)

$$\begin{aligned}
 &= - (1.1101101100000000...0)_2 * 2^{(133-127)} \\
 &= - (1.1101101100000000...0)_2 * 2^6 \\
 &= - (1110110.1100000000...0)_2 = - 118.75
 \end{aligned}$$

- (ii) Show the single-precision floating-point binary representation for: **120.125**. (2 Points)

$$\begin{aligned}
 120.125 &= (1111000.001)_2 = (1.111000001)_2 * 2^6 \\
 \text{Exp.} &= 6 + 127 = 133 \\
 &\text{Single precision binary representation:}
 \end{aligned}$$

0100 0010 1111 0000 0100 0000 0000 0000

- (iii) Perform the following floating-point operation rounding the result to the **nearest even**. Perform the operation using **guard, round** and **sticky** bits.

(6 Points)

$$\begin{array}{r}
 1.000\ 1000\ 0000\ 0000\ 0000\ 0000\ 000\ 000 \quad \times 2^8 \\
 - 1.000\ 0000\ 0000\ 0000\ 0000\ 0100\ 000 \quad \times 2^4
 \end{array}$$

$$\begin{array}{r}
 1.000\ 1000\ 0000\ 0000\ 0000\ 0000\ 000\ 000 \quad \times 2^8 \\
 - 1.000\ 0000\ 0000\ 0000\ 0000\ 0100\ 000 \quad \times 2^4 \\
 \hline
 = 1.000\ 1000\ 0000\ 0000\ 0000\ 0000\ 000\ 000 \quad \times 2^8 \\
 - 0.000\ 1000\ 0000\ 0000\ 0000\ 0000\ 010 \quad \times 2^8 \text{ (align)} \\
 \hline
 = 01.000\ 1000\ 0000\ 0000\ 0000\ 0000\ 000\ 000 \quad \times 2^8 \\
 + 11.111\ 0111\ 1111\ 1111\ 1111\ 1111\ 110 \quad \times 2^8 \text{ (2's complement)} \\
 \hline
 \end{array}$$

$$\begin{aligned}
&= 00.111\ 1111\ 1111\ 1111\ 1111\ 1111\ 110 \times 2^8 \\
&= +0.111\ 1111\ 1111\ 1111\ 1111\ 1111\ 110 \times 2^8 \\
&= +1.111\ 1111\ 1111\ 1111\ 1111\ 1111\ 100 \times 2^7 \text{ (normalize)}
\end{aligned}$$

Next, we round to the nearest even by adding 1 and the result becomes:

$$= +10.000\ 0000\ 0000\ 0000\ 0000\ 0000 \times 2^7 \text{ (round)}$$

Next, we renormalize the result and the result becomes:

$$= +1.000\ 0000\ 0000\ 0000\ 0000\ 0000 \times 2^8 \text{ (renormalize)}$$