ICS 233 – Computer Architecture & Assembly Language

Assignment 3: Floating-Point Representation and Arithmetic

1. **(4 pts)** What is the decimal value of the following single-precision floating-point numbers?

   a) `1 01011010 001 0100 0000 0000 0000 0000` (binary)
   b) `0 10001101 100 1000 0000 0000 0000 0000` (binary)

2. **(3 pts)** Show the IEEE 754 binary representation for: `-75.4` in …

   a) Single Precision
   b) Double precision

3. **(6 pts)** x = `1 10001101 101 1000 0000 0000 0000 0000` (binary)
   and y = `0 01111101 110 0000 0000 0000 0000 0000` (binary)
   are single-precision floating-point numbers. Perform the following operations showing all work:

   a) $x + y$
   b) $x * y$

4. **(4 pts)** x = `0 10111111 011 1110 0100 0000 0000 0000` (in binary)
   and y = `0 01111111 111 1000 0000 0000 0000 0000` (in binary)
   and z = `1 10111111 011 1110 0100 0000 0000 0000` (in binary)
   represent single precision IEEE 754 floating-point numbers. Perform the following operations showing all work:

   a) $x + y$
   b) Result of (**a**) + z
   c) Why is the result of (**b**) counterintuitive?

5. **(3 pts)** IA-32 offers an 80-bit extended precision option with a 1 bit sign, 16-bit exponent, and 63-bit fraction (64-bit significand including the implied 1 before the binary point). Assume that extended precision is similar to single and double precision.

   a) What is the bias in the exponent?

   b) What is the range (in absolute value) of normalized numbers that can be represented by the extended precision option?