CPE 431/531 Just Enough Floating-Point Fall 2022

**The University of Alabama in Huntsville**

**ECE Department**

**CPE 431 01/01R, CPE 531 01/91 Fall 2022**

**Due September 13, 2022 – You must show your work to get full credit. Use online calculators to check your answers.**

**1.0 (5), 2.0 (5), 3.0 (10), 4.0 (10), 5.0 (10), 6.0 (10)**

**1.0 <3.2>** Assume 12 and 155 are unsigned 8-bit decimal integers. Calculate 12 – 155 in binary. Is there overflow, why or why not?

12 – 155 = -143

1\_0111\_ 0001

**Answer:**  Yes because at 8 bits signed you only get 113

**2.0 <3.5>** What decimal number does the bit pattern **0xCB9A\_0A89** represent if it is a two’s complement integer? An unsigned integer?

1100\_1011\_1001\_1010\_0000\_1010\_1000\_1001

31,30,27,25,24,23,20,19,17,11,9,7,3,0

**Unsigned Decimal: 3,415, 870, 089**

**Two’s Comp Decimal: -1,268,386,441**

**3.0 <3.5>**What decimal number does the bit pattern **0x6DB8\_0000** represent if it is a floating point number? Use the IEEE 754 standard and express in decimal scientific representation.

32 bit (single precision)

Binary: 0|110\_1101\_1|011\_1000-0000\_0000\_0000\_0000

23\_4-----

1101\_1011 (219)

1000\_0001 (bias)

0101\_1100 (92) = exp

2^-2, 2^-3, 2^-4 = 0.4375 (fraction)

**Answer: 1.4375 \* 2^92 = 7.1182 E27**

**4.0 <3.5>** Write down the hexadecimal representation of the decimal number 57812.59375 assuming the IEEE 754 single precision format.

Binary: 1110\_0001\_1101\_0100.10011

Scientific Notation: 1.1100\_0011\_1010\_1001\_0011 \* 2^15

Calculate Exp:

0000\_1111 (15)

0111\_1111 (127)

1000\_1110 (142) – exp

Put it together: 0|100\_0111\_0|110\_0001\_1101\_0100\_1001\_1000

0100\_0111\_0110\_0001\_1101\_0100\_1001\_1000

**Answer: 0x4761\_D498**

**5.0 <3.5>**Write down the hexadecimal representation of the decimal number -5932.515625 assuming the IEEE 754 double precision format.

Ignore sign when changing to binary

Binary: 1\_0111\_0010\_1100.100001

Scientific: 1.011100101100100001 \* 2^12 (sign bit = 1)

Calculate Exp: 000\_0000\_1100 (12)

111\_1111\_1111 ( 1023)

100\_0000\_1011 (1035)

Put it together: 1|100\_0000\_1011|011100101100100001

1100\_0000\_1011\_0111\_0010\_1100\_1000\_0100\_0000\_0000\_0000\_0000\_0000\_0000\_0000\_0000

**Answer: C0B7\_2C84\_0000\_0000**

**6.0 <3.5>** Write down the hexadecimal representation of the decimal number -1947.75 assuming it was stored using the single precision IBM format (base 16, instead of base 2, with 7 bits of exponent, bias = 64, no implicit numbers)

Sign = 1

Hex: 79B.C

Scientific Notation: 0.79BC \* 16^3

Calculate Exp: 3 + 40 (bias: 64) = 43 ( 67 decimal)

Put it together: 1100\_0011 (43 and the sign bit in binary) = C3

Add the rest….

**Answer: 0xC379\_BC00**