**1.** [10 points] For each binary pattern please write the corresponding values of the following interpretations: Each correctly answer cell is 0.25 points. The fixed point is located 2 positions from the *LEFT* . The most significant bit is 7 while the least significant bit is 0.

*I misread this question. It should have been from the RIGHT. See the other document that solves this same question as originally stated. Either way, it’s good practice.*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| ***76543210*** | **Unsigned Int** | **Signed Int** | **Unsigned Fixed Point** | **Signed Fixed Point** |
| 1000 0000 | 128 | -128 | 2 | -2 |
| 1000 0011 | 131 | -125 |  |  |
| 1000 0001 | 129 | -127 |  |  |
| 0100 0001 | 65 | +65 |  |  |
| 0111 1111 | 127 | +127 |  |  |
| 1111 1111 | 255 | -1 |  |  |
| 1111 1100 | 252 | -4 |  |  |
| 0000 0000 | 0 | 0 | 0 | 0 |
| 0111 1110 | 126 | +126 |  |  |
| 1000 1110 | 142 | -114 |  |  |
| 0001 0000 | 16 | +16 | 0.25 | +0.25 |

***See “Workspace” for fixed point calculations.***

**2.** [10 points] What is the most negative number (largest absolute value negative) that can be represented using 16-bit signed integer representation? Please circle ALL correct ones.

Explanation:

The word “ALL” is a trick!!! There is only one answer!

**3.** [10 points] Please add the two numbers in Hex. Then convert each operand to binary and perform the same operation in binary, then repeat BASE 10. The signed integers are represented using two’s complement.

0x0E +14 0b0000 1110

+

0xFF -1 0b1111 1111

Hex: 0x0D Dec: +13 Binary: 0b0000 1101

*There is a carry of 1 that we simply ignore.*

**4.** [20 points] Determine the MINIMAL number of bits required to represent ***-511.75*** using:

4.1. [5 points] ASCII Code: 56-Bits

Each character is represented by 8-Bits of information in ASCII. Because there are 7 characters in “-127.75” (that includes the minus sign and decimal point), it requires 7x8 = 56 Bits.

4.2. [5 points] Binary Fixed Point representation: 12-Bits

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |  |  |  |  |  |  |
| 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |

4.3. [5 points] Take the result from the previous answer and shift the fixed point by 2 positions to the right and write the resulting signed decimal value.

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |  |  |  |  |  |  |
| 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |

4.4.[5 points] Write down the *unsigned* rational number stored in the 9-bit WORD below. (The right most 8 bits are fractional)

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |  |  |  |
| 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |

This represents the number . To convert this into a decimal sum up the necessary terms:

**5.** [10 points] Determine the decimal value (scientific notation) of the single precision (32-bit) floating point representations given below.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 31 | 30 | 29 | 28 | 27 | 26 | 25 | 24 | 23 | 22 | 21 | 20 | 19 | 18 | 17 | 16 | 15 | 14 | 13 | 12 | 11 | 10 | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | 1 | 0 |
| 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

The sign bit is 0, the exponent field (30 to 23) contains the value 129, and the fraction field contains . Hence, we use the following equation:

Where

**6.** [10 points] Determine the decimal value (scientific notation) of the single precision (32-bit) floating point representations given below.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 31 | 30 | 29 | 28 | 27 | 26 | 25 | 24 | 23 | 22 | 21 | 20 | 19 | 18 | 17 | 16 | 15 | 14 | 13 | 12 | 11 | 10 | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | 1 | 0 |
| 1 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

The sign bit is 1, the exponent field (30 to 23) contains the value 133, and the fraction field contains . Hence, we use the following equation:

Where

**7.** [5 points]Please write the smallest positive number that can be represented using 32-Bit floating-point representation. Give the binary representation and in decimal.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 31 | 30 | 29 | 28 | 27 | 26 | 25 | 24 | 23 | 22 | 21 | 20 | 19 | 18 | 17 | 16 | 15 | 14 | 13 | 12 | 11 | 10 | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | 1 | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

Since we are looking for the smallest positive, the sign-bit (i.e. bit 31) must be 0. Next, we set bit 23 to 1 because it is the smallest value that can be produced in the exponent field. The bits of fraction field are all set to 0. Hence,

To get the decimal representation in the form , we must solve for *x*:

Smallest positive number:

**8.** [5 points] Please write the *MOST NEGATIVE* number that can be represented using 32-Bit floating-point representation. Give the binary representation and in decimal.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 31 | 30 | 29 | 28 | 27 | 26 | 25 | 24 | 23 | 22 | 21 | 20 | 19 | 18 | 17 | 16 | 15 | 14 | 13 | 12 | 11 | 10 | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | 1 | 0 |
| 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |

This is the oppositive of finding the smallest positive so a strategy here is to invert all the bits from question 7. Hence,

To get the decimal representation in the form , we must solve for *x*:

Most negative number: -

**9.** [5 points] Please write the *LARGEST POSITIVE* number that can be represented using 32-Bit floating-point representation. Give the binary representation and in decimal.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 31 | 30 | 29 | 28 | 27 | 26 | 25 | 24 | 23 | 22 | 21 | 20 | 19 | 18 | 17 | 16 | 15 | 14 | 13 | 12 | 11 | 10 | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 | 1 | 0 |
| 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |

Since the largest positive number is a positive, the sign-bit must be 0. All other bits except bit 23 are set to 1. Hence,

To get the decimal representation in the form , we must solve for *x*:

Largest positive number: +

**10.** [20 points] In *EACH Questions 10.1-10.4* you are given SIGNED integers stored in 32-Bit registers. Indicate whether overflow occurs or not.

10.1 – (5 points)

0x0000000E +14 0b0000 0000 0000 0000 0000 0000 0000 1110

+

0xFFFFFFFF -1 0b1111 1111 1111 1111 1111 1111 1111 1111

Hex: 0x0000000D Dec:+13 Binary: 0b0000 0000 0000 0000 0000 0000 0000 1101

Overflow? NO

10.2 - (5 points)

0x7FFFFFFF 2147483647 0b0111 1111 1111 1111 1111 1111 1111 1111

-

0xFFFFFFFF -1 0b1111 1111 1111 1111 1111 1111 1111 1111

+ 0b0000 0000 0000 0000 0000 0000 0000 0001

Hex: 0x80000000 Dec: -2147483648 Binary: 0b1000 0000 0000 0000 0000 0000 0000 0000

Overflow? YES, we expected +2147483648. This overflow occurs because the maximum value represented in a 32-bit WORD is +2147483647.

10.3 - (5 points)

0x80000000 -2147483648 0b1000 0000 0000 0000 0000 0000 0000 0000

-

0xFFFFFFFF -1 0b1111 1111 1111 1111 1111 1111 1111 1111

+ 0b0000 0000 0000 0000 0000 0000 0000 0001

Hex: 0x80000001 Dec: -2147483647 Binary: 0b1000 0000 0000 0000 0000 0000 0000 0001

Overflow? NO

10.4 - (5 points)

0x7FFFFFFF 2147483647 0b0111 1111 1111 1111 1111 1111 1111 1111

+

0xFFFFFFFF -1 0b1111 1111 1111 1111 1111 1111 1111 1111

Hex: 0x7FFFFFFE Dec: +2147483646 Binary: 0b0111 1111 1111 1111 1111 1111 1111 1110

Overflow? NO

Workspace

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |  |  |
| 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

Unsigned Fixed Point:

Signed Fixed Point:

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |  |  |
| 1 | 0 | 0 | 0 | 0 | 0 | 1 | 1 |

Unsigned Fixed Point:

Signed Fixed Point:

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |  |  |
| 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |

Unsigned Fixed Point:

Signed Fixed Point:

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |  |  |
| 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 |

Unsigned Fixed Point:

Signed Fixed Point: Same as unsigned

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |  |  |
| 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |

Unsigned Fixed Point:

Signed Fixed Point: Same as unsigned

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |  |  |
| 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |

Unsigned Fixed Point:

Signed Fixed Point: :

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |  |  |
| 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 |

Unsigned Fixed Point:

Signed Fixed Point: :

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |  |  |
| 0 | 1 | 1 | 1 | 1 | 1 | 1 | 0 |

Unsigned Fixed Point:

Signed Fixed Point: : Same as unsigned

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |  |  |
| 1 | 0 | 0 | 0 | 1 | 1 | 1 | 0 |

Unsigned Fixed Point:

Signed Fixed Point: :

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |  |  |
| 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |

Unsigned Fixed Point:

Signed Fixed Point: Same as unsigned