Homework Assignment #1 – Number Representations

The purpose of this assignment is to let you be familiar and become comfortable with binary representations which are used heavily in computer organization.

Problem 1 (30 points, 10 points each) Convert the following decimal numbers into (a) 8-bit, (b) 16-bit, and (c) 32-bit binary numbers. For negative numbers, use the 2's complement. State "overflow" if a number cannot be represented correctly. Hint: You may want to use the sign extension rule.

1) 45_{ten}.
2) -81_{ten}.
3) $-3{,}000_{ten}$.

Solution:

Decimal number	8-bit two's complement binary number	16-bit two's complement binary number	32-bit two's complement binary number
45	0010 1101	0000 0000 0010 1101	0000 0000 0000 0000 0000 0000 0010 1101
-81	1010 1111	1111 1111 1010 1111	1111 1111 1111 1111 1111 1111 1010 1111
-3,000	Overflow	1111 0100 0100 1000	1111 1111 1111 1111 1111 0100 0100 1000

Problem 2 (30 points) What decimal number does each of the following two's complement binary number represent?

1) 1111 1111 1111 1111 1111 1111 1011 1101 $_{two}$.
2) 1111 1111 1111 1111 1111 1111 1101 1001 $_{two}$.
3) 0111 1111 1111 1111 1111 1111 1011 1111 $_{two}$.

Solution:

1. **-67** $_{ten}$
2. **-39** $_{ten}$
3. 2147483583 $_{ten}$ (2^{31}-1-2^{6})

Problem 3 (40 points, sub-problem 3 is 20 points, others 10 points each) Show the IEEE 754 binary representation for the following floating-point numbers in single and double precision. Give your results in hexadecimal format.

1) $19_{\ ten}$.
2) $2.375_{\ ten}$.
3) -0.3_{ten}.

Solution:

Decimal no.	Single precision floating point no	Double precision floating point no
1) 19	0x41980000	0x4033000000000000
2) 2.375	0x40180000	0x4003000000000000
3) -0.3	0xBE999999	0xBFD3333333333333

1)

- Single Precision:

$19_{ten} = 10011_{two} = 1.0011 * 2^4 = 1.0011 * 2^{131-127}$

=> So we have: F = 001100…000 (23 bits)

E = 131_{ten} = 1000 0011 (8 bits)at

Therefore the presentation is

0100 0001 1001 1000 0000 0000 0000 0000

0x41980000

- Double Precision:

$19_{ten} = 10011_{two} = 1.0011 * 2^4 = 1.0011 * 2^{1027-1023}$

=> F = 00110…000 (52 bits)

E = 1027_{ten} = 10000000011 (11 bits)

Therefore the presentation is

0100 0000 0011 0011 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000

0x4033000000000000

2)

single precision:

$2.375_{ten} = 10.011_{two} = 1.0011 * 2^1 = 1.0011 * 2^{128-127}$

0100 0000 0001 1000 0000 0000 0000 0000

0x40180000

double precision:

$2.375_{ten} = 1.0011_{two} = 1.0011 * 2^1 = 1.0011 * 2^{1024-1023}$

0100 0000 0000 0011 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000 0000

0x4003000000000000

3)

Single precision:

We find the presentation of its complement 0.3 and then change the sign bit

$0.3_{ten} = 0.0100110011\ldots._{two} = 1.00110011\ldots * 2^{-2} = 1.00110011\ldots * 2^{125-127}$

1011 1110 1001 1001 1001 1001 1001 1001

Double precision:

$0.3_{ten} = 0.0100110011\ldots._{two} = 1.00110011\ldots * 2^{-2} = 1.00110011\ldots * 2^{1021-1023}$

1011 1111 1101 0011 0011 0011 0011 0011 0011 0011 0011 0011 0011 0011 0011 0011