

Midterm Examination

Problems:

1. [50 points] Consider the two numbers in radix decimal: $A = 123$ and $B = -1023$.

1.1. Explicitly convert each decimal number to binary (radix two) in 2s complement

(a) $A = 123$ is positive.

$$123 / 2 = 61 + 1$$

$$61 / 2 = 30 + 1$$

$$30 / 2 = 15 + 0$$

$$15 / 2 = 7 + 1$$

$$7 / 2 = 3 + 1$$

$$3 / 2 = 1 + 1$$

$$1 / 2 = 0 + 1$$

so binary of A in 2's complement is 01111011.

(b) $B = -1023$.

According to Yale, Sanjay(2004, Page 26) *Example 2.1*. I followed steps from this example.

1. Let B be +1023 and find the binary for +1023.

$$1023 / 2 = 511 + 1$$

$$511 / 2 = 255 + 1$$

$$255 / 2 = 127 + 1$$

$$127 / 2 = 63 + 1$$

$$63 / 2 = 31 + 1$$

$$31 / 2 = 15 + 1$$

$$15 / 2 = 7 + 1$$

$$7 / 2 = 3 + 1$$

$$3 / 2 = 1 + 1$$

$$1 / 2 = 0 + 1$$

So, the representation for +1023 is 0111111111

2. The complement of +1023 is 10000000000

3. Adding 1 to 1000000000 gives us 10000000001 and it's 2's complement binary.

1.2. Explicitly convert each decimal number to hexadecimal (radix sixteen) in 1s complement.

According to Yale, Sanjay(2004, Page 42, ¶ 5) "If we had first broken the string at four-bit boundaries and then converted each four-bit string to its equivalent hex digit.

(a) For A. We have 01111011 in 2's complement.

Convert it to 1's complement: 10000100.

Break it to four-bit boundaries from left to right: 0001 0000 0100.

Convert to hex for each part: 1 0 4.

So, the representation of 1's complement hexadecimal number for A is x104.

(b) For B, we have 10000000001 in 2's complement.

Convert it to 1's complement: 0111111110.

Break it to four-bit Boundaries from left to right: 0011 1111 1110.

Convert it to hex for each part: 3 F E.

So, the representation of 1's complement hexadecimal number for B is x3FE.

1.3. Using 2s complement integer arithmetic, explicitly find the result A + B in binary.

According to Yale, Sanjay(2004, Page 30) *Example 2.3*.

The decimal value $A = 123$ is represented as 01111011.

The decimal value $B = -1023$ is represented as 1000000001.

$A+B$ is

$$\begin{array}{r} 00001111011 \\ + 10000000001 \\ \hline 10001111100 \end{array}$$

So, $A+B$ represented 10001111100.

1.4. Using 2s complement integer arithmetic, explicitly find the result $A - B$ in binary.

According to Yale, Sanjay(2004, Page 30) *Example 2.4*.

$A - B$ we can change it to $A + (-B)$

$-B$ is +1023 and the representation of it is 0111111111.

$A + (-B)$ is

$$\begin{array}{r} 00001111011 \\ + 01111111111 \\ \hline 10001111010 \end{array}$$

So, $A - B$ represented 10001111010.

1.5. Using 2s complement integer arithmetic, explicitly find the result A/B (long division) in binary.

The decimal value $A = 123$ is represented as 01111011.

The decimal value $B = -1023$ is represented as 1000000001.

Find A / B by long division:

$$\begin{array}{r}
 10000000001 \overline{) 0.000111101} \\
 \underline{01111011} \\
 11110110 \\
 \underline{0} \\
 111101100 \\
 \underline{0} \\
 1111011000 \\
 \underline{0} \\
 1111011000 \\
 \underline{10000000001} \\
 11101011110 \\
 \underline{10000000001} \\
 11010111010 \\
 \underline{10000000001} \\
 10101110010 \\
 \underline{10000000001} \\
 1011100010 \\
 \underline{0} \\
 10111000100 \\
 \underline{10000000001} \\
 1110000110
 \end{array}$$

1.6. Repeat the division in (1.5), but now show the result including any explicit binary fractions, and express the result with the explicit binary point on display

I don't understand this question very well. This is all I can do.

$$A / B =$$

$$0.00011110101110000101000111101011100001010001111010111000010100...$$

2. [50 points] Consider the real number $X = -1024$ (not an integer), and the real number $Y = 2$ (not an integer).

2.1. Encode X in IEEE 754 single precision floating point, showing explicitly your construction.

Consider the decimal number $X = -1024$.

(a) the binary value for 1024 is

$$1024 / 2 = 512 + 0$$

$$512 / 2 = 256 + 0$$

$$256 / 2 = 128 + 0$$

$$125 / 2 = 62 + 1$$

$$62 / 2 = 31 + 1$$

$$31 / 2 = 15 + 1$$

$$15 / 2 = 7 + 1$$

$$7 / 2 = 3 + 1$$

$$3 / 2 = 1 + 1$$

$$1 / 2 = 0 + 1$$

the binary value for 1024 is 1111111100

(b) the sign bit 1 reflect the fact -1024 is negative.

(c) Normalizing the binary value the result obtained is -1.1111111×2^9

(d) The exponent is unsigned number 132, represented as 10001000 in binary, because the real exponent is +9 ($136 - 127 = +9$).

(e) The fraction is 23 bits of precision, removing the leading 1. So, the fraction is

111111100000000000000000

(f) The floating point representation of -1024 is as follows:

sign	Exponent	Fraction
1	10001000	111111100000000000000000

so, the answer is 1 10001000 111111100000000000000000

2.2. Encode Y in IEEE 754 single precision floating point, showing explicitly your construction.

Consider the decimal number $Y = 2$.

(a) the binary value for 2 is $(10)_2$

$$2 / 2 = 1 + 0$$

$$1 / 2 = 0 + 1$$

(b) the sign bit 0 reflect the fact 2 is positive.

(c) Normalizing the binary value the result obtained is 1.0×2^1

(d) The exponent is unsigned number 128, represented as 10000000 in binary, because the real exponent is +1 ($128 - 127 = +1$).

(e) The fraction is 23 bits of precision, removing the leading 1. So, the fraction is

00000000000000000000000

(f) The floating point representation of 2 is as follows:

sign	Exponent	Fraction
0	10000000	00000000000000000000000

so, the answer is 0 10000000 00000000000000000000000

2.3. Explicitly show the steps required in IEEE 754 single precision floating point to form $X + Y$.

On the Website, *Floating Point Tutorial* (2012, ¶ 8) states the following steps to adding floating point.

	S1	E1	M1
X =	1	10001000	11111110000000000000000

	S2	E2	M2
Y =	0	10000000	00000000000000000000000

(a) Absolute value of X is greater than absolute value of Y.

(b) Result of initial exponent $E = E1 = (10001000)_2 = (136)_{10}$

(c) $E1 - E2 = (100100 - 1000000) = 136 - 128 = 6$

(d) Shift the mantissa M2 by (E1-E2) so that the exponents are same for both numbers.

M2 = 1.00000000000000000000000
= 0.00000100000000000000000

(e) Sign is not equal. So, replace S with 1 and subtract M1 and shifted M2 for M.

$$\begin{array}{r}
 1.111111100000000000000000 \\
 - 0.000001000000000000000000 \\
 \hline
 1.111110100000000000000000
 \end{array}$$

$X + Y =$

S	E	M
1	10001000	111110100000000000000000

so, the representation for $X + Y$ is 1 10001000 111110100000000000000000

References

Floating Point Tutorial. Retrieved 2012, from the World Wide Web:

<http://www.rfwireless-world.com/Tutorials/floating-point-tutorial.html>

Yale P. and Sanjay P. (2004). *introduction to computing systems*, New York: The McGraw-Hill companies, Inc.