Computer Architecture

Tutorial 4b – Floating Point Numbers

1) Convert –31.3 to IEEE Single Precision format.

First convert to a binary number -31.3 = -11111.01001 1001 1001

Next Normalise

$$1.11110\ 1001\ 1001\ 1001\ 1001\ 1001 \times 2^{4}$$

Significand field is **1111 0100 1100 1100 1100 110 (23 bits with 1. omitted)**

Exponent field is 4+127 = 131 = **1000 0011**
Number is -ve therefore Sign field is **1**

Sign	Exponent	Significand
1	**1000 0011**	**1111 0100 1100 1100 1100 110**

2) Convert the IEEE Single Precision format hex value C154 0000 to decimal.

C154 0000 = 1100 0001 0101 0100 0000 0000 0000 00000

Sign	Exponent	Significand
1	1000 0010	1010 1000 0000 0000 0000 000

Exponent field = 1000 0010 = 130 => Exponent = 130 - 127 = 3

Significand field = 10101 Adding Hidden Bit => 1.10101

Therefore number is 1.10101×2^{3} = 1101.01 = Decimal 13.25

Sign is 1 therefore number is **-13.25**

3) Carry out the operation 31.3 + 13.25 in IEEE single precision arithmetic

Number	Sign	Exponent	Significand
31.3	0	1000 0011	1111 0100 1100 1100 1100 110
13.25	0	1000 0010	1010 1000 0000 0000 0000 000

Significand of Larger Number = 1.1111 0100 1100 1100 1100 110

Significand of 00 000

Exponents diffe ber 1 place

Significand of Larger Number = 1.1111 0100 0

Significand of Smaller Number = 0.1101 0100 0

Significand of Sum = 10.1100 1000 1100 1100 1100 1100

Sum = $10.1100\ 1000\ 1100\ 1100\ 1100\ 1100 \times 2^{4}$

Normalise $1.01100\ 1000\ 1100\ 1100\ 1100\ 1100 \times 2^{5}$

Sign	Exponent	Significand
0	**1000 0100**	**0110 0100 0110 0110 0110 011**

4)

Bits	Binary value or special value	Decimal value or special value
0 00 00	0	0
0 00 01	0.01	0.25
0 00 10	0.10	0.50
0 00 11	0.11	0.75
0 01 00	1.00	1
0 01 01	1.01	1.25
0 01 10	1.10	1.5
0 01 11	1.11	1.75
0 10 00	10.0	2
0 10 01	10.1	2.5
0 10 10	11.0	3
0 10 11	11.1	3.5
0 11 00	∞	∞
0 11 01	NaN	NaN
0 11 10	NaN	NaN
0 11 11	NaN	NaN