# Floating Point Addition

Addition: $(5.38 * 10^5) + (4.99 * 10^5)$

$(5.38 \times 10^5 + 4.99 \times 10^5)$

$(9.99 * 10^4) + (-1.0 * 10^5)$

$(-1.0 \times 10^5) + (9.99 \times 10^4)$

$(-1.0 \times 10^5) + (0.999 \times 10^5)$

Sign: (+)

(−)

Exponent: "stay the same" → from first / largest value.

(5)

(5)

Significand:

$10.37 \rightarrow \boxed{1.037 \times 10^6}$

$0.001$

$\boxed{1.0 \times 10^2}$

# Floating Point Representation

Floating Point (Float) = $(-1)^s * (1.\text{significand})_2 * 2^{(\text{exponent}-127)}$

| 31 | 30 29 28 27 26 25 24 23 | 22 21 20 19 18 17 16 15 14 13 12 11 10 09 08 07 06 05 04 03 02 01 00 |
|---|---|---|
| s | exponent | significand |

1.037

1.04

$(0.75)_{10} = \frac{1}{2} + \frac{1}{4}$

-0.75 in Floating Point?

sign: 1

sig: 0.11

exp: -1

$(.5)_{10} + (.25)_{10} = 0.75$

sign: 1

exp = exp - 127 = -1 = 126

sig: 1.sig = "1." 1

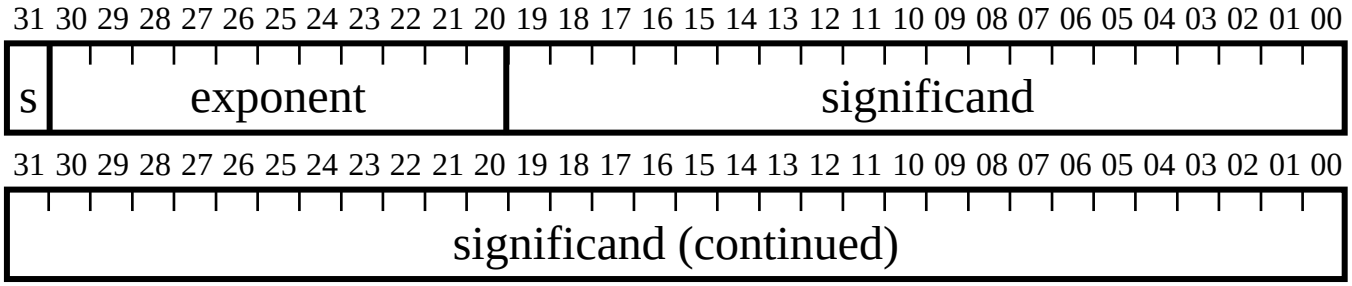
| 31 | 30 29 28 27 26 25 24 23 | 22 21 20 19 18 17 16 15 14 13 12 11 10 09 08 07 06 05 04 03 02 01 00 |
|---|---|---|
| 1 | 01111110 | 1000 |

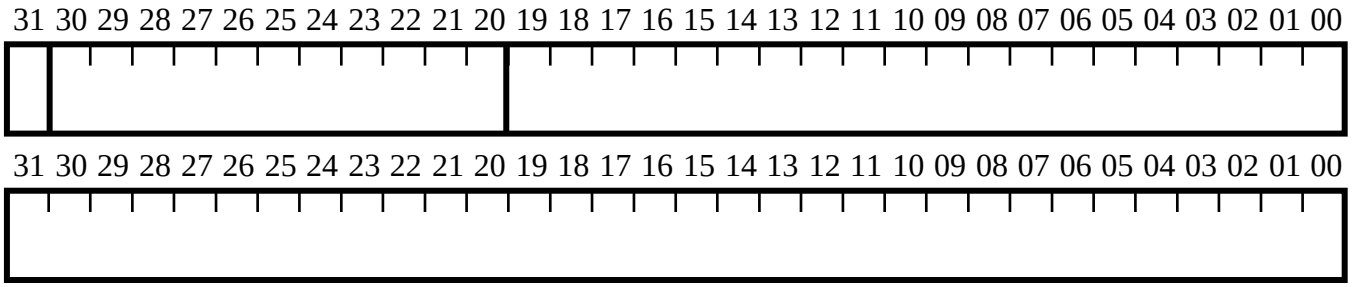
exponent: -1 = exp - 127

(1.)1001

# Double Precision Representation

Double Precision (double) = $(-1)^s * (1.\text{significand}) * 2^{(\text{exponent}-1023)}$




-0.75 in Double Precision?