

2-1/

$$UZ(B) = \sum_{i=0}^n 2^i \times B_i$$



$$FZ(B) = \sum_{i=-n}^m 2^i \times B_i$$

$$\begin{aligned} &0010.1011 \\ &\quad \downarrow \quad \downarrow \quad \downarrow \quad \searrow \\ &2^1 + 2^{-1} + 2^{-3} + 2^{-4} \\ &2 + 1/2 + 1/8 + 1/16 \\ &= 2.6875 \end{aligned}$$

fixed-point binary number

2.5
2.0

$$\begin{aligned} 10111 &\xrightarrow{\text{mod}} 0111 \\ 10111 &\xrightarrow{\text{sat}} 1111 \end{aligned}$$

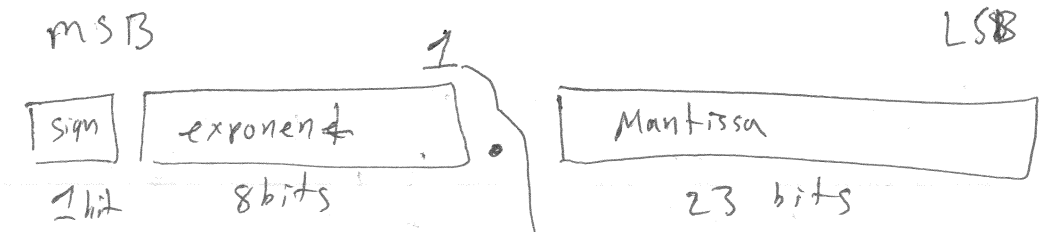
$$\begin{aligned} 01111 &\xrightarrow{\text{mod}} 1000 \\ 01111 &\xrightarrow{\text{trunc}} 0111 \end{aligned}$$

$N =$ sign \rightarrow magnitude \rightarrow exponent
 s m e
 $(-1)^s \times m \times 2^e$

$$\begin{aligned} &\overset{(0,1)}{\rightarrow} \overset{a}{\rightarrow} \overset{b}{\rightarrow} \\ &(s_1, m_1, e_1) * (s_2, m_2, e_2) \\ &= (s_1 \times s_2, \underbrace{m_1 \times m_2}_{\downarrow 2^a}, \underbrace{e_1 + e_2}_{\downarrow b+1}) \\ &(-1)^{s_1} (m_1) (2^{e_1}) \times (-1)^{s_2} (m_2) (2^{e_2}) \end{aligned}$$

32-bit \sim 18 trillion
29 trillion positive
4.5 tril $\in [0, 1]$

L-2/



$$(-1)^s \times (1.M_0 \dots M_{22}) \times 2^{(E)-127}$$

0 1000 0001 1100 0

+ 129 - 127 = 2 1 + 1/2 + 1/4 = 1.75

$$2^2 = 4$$

$$2^7 + 2^0 = 128 + 1 = 129$$

exp ≠ 0 & ≠ 255 ⇒ normalized
 exp = 0 ⇒ de-normalized
 no hidden bit, E = 1 - field value

exp = 255 (all 1s) → mantissa = 0 ⇒ infinity
 → ≠ 0 ⇒ NaN

+ 88.00 1, 2, 4, 8, 16, 32, 64 88 - 64 = 24 - 16 = 8

~~01011000~~ $88 = (-1)^s \times (1.M) \times 2^E = 6$

01011000 1. × 2^E $\frac{88}{64}$

$$1.375 = 1.011$$

$$\begin{array}{r} 64 \sqrt{88} \\ - 64 \\ \hline 240 \\ - 192 \\ \hline 480 \\ - 448 \\ \hline 320 \end{array}$$

+ 6 0.375
 0 1000 0100 0110

0 1000 0111 001 1000 0001 0

304.125