Ovn cf dsf qsf t f oubujpo

98/11/29

بسم الله الرحمن الرحیم

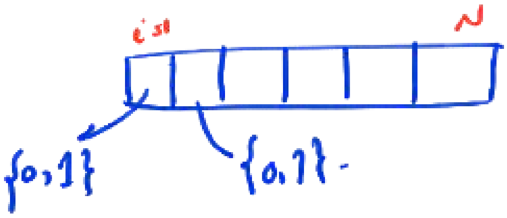
* Number Reresentation

نمایش اعداد

0 , 1

مبتنی بر خواص مغناطیسی

N - Bits → 2^N



IEEE : Institute of Electrical and Electronic Engineering

$$\begin{cases} 1 \text{ Byte} = 8 \text{ Bits} \\ 1 \text{ K Byte} = 2^{10} \text{ Bytes} = 1024 \text{ Bytes} \end{cases}$$

Number Representation

- class 1 : Fixed Point Rep.
- class 2 : Floating Rep.

Ex 1: FPR

$$R \equiv \widetilde{\text{Sign}} \times (\alpha\, 2^n + \alpha\, 2^{n-1} + \dots \alpha\, 2^1 + \widetilde{\alpha\, 2^0} +$$

$\underbrace{\quad}_{\text{علامت}}$ $\quad n \quad n-1$

$\alpha_{-1}2^{-1} + \alpha_{-2}2^{-2} + \dots \alpha_{-m}2^{-m})$

$$N-1-1 = n+m$$

$$N-2 = n+m$$

ذخیره‌سازی هر خانه (0, 1)

Ex2: $R = (1-2S)\, 2^{e-Bias}\, (1.f)$

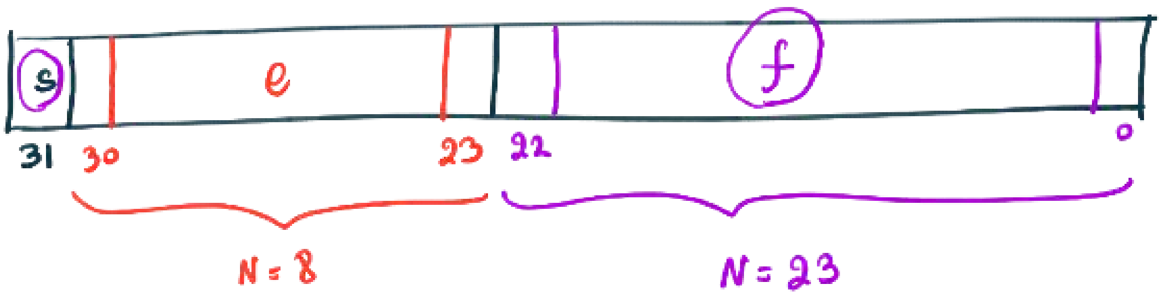
$$\boxed{R = \underbrace{(1-2S)}_{\text{علامت}}\, \underbrace{2^{e-Bias}}_{\text{نما}}\, \underbrace{(1.f)}_{\text{اعشار}}}$$

Precision دقت

بازه نمایش اعداد؟

1. Single Precision → Bytes = 2^5 Bits, $\boxed{32 \text{ Bits}}$ تعداد خانه
2. Double Precision → $\boxed{64 \text{ Bits}}$ → $[2^{-1022} - 2^{+1023}]$
3. Extended Precision ← $[2^{-32766} - 2^{+32767}]$

تخصیص خانه در نمایش floating برای N=32



$e-Bias$

Bias = ...

$R = (1-2s)\,2^{} \quad (1.f)$

Bias = …

$2^8 \longrightarrow [0, 255] \xrightarrow{\text{IEEE}} [1, 254]$

$$\boxed{1 \leqslant e \leqslant 254 \rightarrow 2^e}$$

$$\boxed{\text{Bias} = 127} \longrightarrow 2^{e-\text{Bias}} = 2^{e-127}$$

$$2^{e-\text{Bias}} = [2^{-126} - 2^{+127}]$$

بازه عملی

$$1.f = 1 + f = 1 + (m_{22}2^{-1} + m_{21}2^{-2} + m_{20}2^{-3} + \dots + m_0 2^{-23})$$

$$R_{max} = \begin{cases} s = 0 \\ e = 254 \\ f: \{\underbrace{1, 1, \dots 1}_{23}\} \end{cases} = +1 \times 2^{+127} \times \Big[\overbrace{1. \sum_{m=0}^{22} 2^{-(m+1)}}^{1+1}\Big]$$

$\underbrace{\sum_{m=0}^{22} 2^{-(m+1)}}_{1}$

$$= 2^{127} \times 2$$

$$\boxed{R_{max} = 2^{128}}$$

مثبت

$$R_{min} = \begin{cases} s = 0 \\ e = 1 \\ f: \{\underbrace{0, 0 \dots 1}_{23}\} \end{cases} = +1 \times 2^{-126} \times (0 + 2^{-23})$$

$$= 2^{-149}$$

Ex 3 : $K = 639.6875 \rightarrow$

$$+\underbrace{639}.\underbrace{6875} = (1-2S)\, 2^{e-Bias}\, (1.f)$$

① $S = 0$ ✓

② Bias = 127 → $e = ?$ ✓

③ $f = ?$ ✓

$639 \,|\, 2 \quad 319 \,|\, 2 \qquad 1 \quad 1$

$\Rightarrow 639 = 1\times 2^9 + 0\times 2^8 + 0\times 2^7 + 1\times 2^6 + 1\times 2^5 + 1\times 2^4 + 1\times 2^3 + 1\times 2^2 + 1\times 2^1 + 1\times 2^0$

$= 1001111111$

$0.6875 \,|\, 2^{-1} = 0.5 \quad 1 \longrightarrow m_{22}$

0.5000

$0.1875 \,|\, 2^{-2} = 0.25 \quad 0 \longrightarrow m_{21}$

$0.1875 \,|\, 2^{-3} = 0.125 \quad 1 \longrightarrow m_{20}$

0.1250

$0.0625 \,|\, 2^{-4} = 0.0625 \quad 1 \longrightarrow m_{19}$

0.0625

0.0000

$2^6 \quad 2^5 \quad 2^2 \quad 2^1 \quad 2^0 \quad 2^{-1} \quad 2^{-2} \quad 2^{-3} \quad 2^{-4}$

$639.6875 = 1001111111.1011$

$2^9 \quad 2^8 \quad 2^7 \quad 2^4 \quad 2^3$

$$= 1.\overbrace{0011111111\,1011}^{f} \times 2^9$$

$$= (1.f) \times 2^9 \times (+1)$$

$$(1-2s)\,2^{e-127}\,(1.f) =$$

$$e - 127 = 9$$

$$e = 136$$

$$136 = 1000\,1000$$

0	1	0	0	0	1	0	0	0	0	0	1	1	1	1	1	1	1	1	0	1	1	0	.	.	

$$639.6875$$

Ex 4:

$$7 + 10^{-7} = ?$$

$$\downarrow \qquad \downarrow$$

$$\mathbb{R} \qquad \mathbb{R}$$

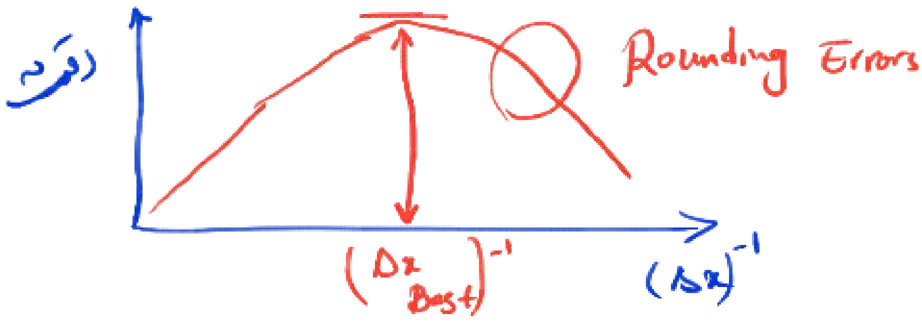
$$0 \quad 10000010 \quad 111 \quad 000\ldots$$

$$7 = 111 = \underbrace{1.11}_{(1.f)} \times \underbrace{2^2}_{2^{e-Bias}} \qquad e = 129$$

$2^2 \quad 2^1 \quad 2^0$

$$7 + 10^{-7} = 7$$

$$\underline{f(x+\Delta x)} = f(x) + \Delta x\, f' + \underbrace{\frac{\Delta x^2}{2!} f'' + \cdots}_{O(\Delta x^2)}$$



بسم الله

Number Representation theory

در کامپیوتر چون از خواص مغناطیسی استفاده می شود برای ضبط داده ها از حالت های مغناطیسی حوزه ی مغناطیسی یعنی ۰ و ۱ استفاده می شود. اگر N بیت داشته باشیم پس تعداد داده ی قابل ذخیره سازی برابر با 2^N خواهد بود. حال اگر یکی از بیت ها را به علامت نسبت دهیم در آن صورت بیشینه عدد قابل ذخیره 2^{N-1} خواهد بود.

بر اساس قرارداد

$$1 \text{ Byte} = 8 \text{ Bits}$$

$$1 KB = 2^{10} B = 1024 \text{ Bytes}$$

- به طور کلی دو نوع نمایش داریم
  1. fixed Point Representation
  2. Floating Point Representation

- در خصوص دقت ذخیره سازی نیز داریم:

1. Single Precision
2. Double Precision
3. Extended Precision

- نمایش Fixed Point داریم:

$$R = \text{Sign} \times (\alpha_n 2^n + \alpha_{n-1} 2^{n-1} + \dots + \alpha_0 2^0 + \alpha_{-1} 2^{-1} + \dots \alpha_{-m} 2^{-m})$$

اگر N بیت داشته باشیم یک بیت برای علامت پس $n+m=N-2$ چون صفر داشتیم شد $N-2$.

- نمایش Floating

$$R = (1-2S)\, 2^{e-\text{Bias}} \times (1+f)$$

برای Single Precision تعداد 32 بیت داریم

s	e	f

31 30 23 22 0

برای نمایش e در واقع 2^8 حالت داریم یعنی در بازه $[0,255]$ که مورد 255 رزرو شده است

برای اعداد خاص پس $1 \leq e \leq 254$. عبارت Bias به این دلیل آمده است که از

254 نصف برای توان منفی و نصفی برای توان مثبت در نظر گرفته شود پس 2^{e-125}

بخش مربوط به اعشار هم $1.f$ است یعنی

$$f = m_{22} 2^{-1} + m_{21} 2^{-2} + \dots \quad m_0 2^{-23}$$

پس x_{Max} عبارت است

$$S=0$$

$$e=254 \longrightarrow 2^{127}$$

$$1.f = 1 + 2^{-1} + 2^{-2} + \dots \; 2^{-23} = 1+1 = 2$$

پس

$$x_{Max} = + 2^{127} \, 2$$

برای x_{min} البته منظورم عدد مثبت است پس

$$e=1 \longrightarrow e\text{-Bias} = -126$$

$$1.f \rightarrow 0 + 0 + \dots \; 2^{-23} = 2^{-23}$$

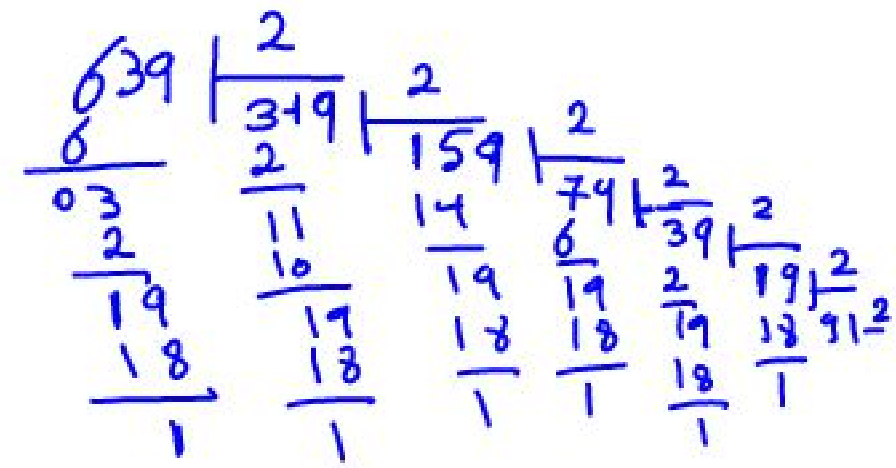
پس

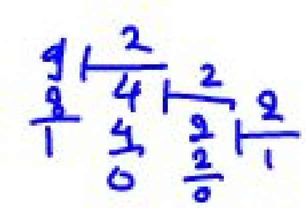
$$x_{min} = 2^{-126} \, 2^{-23} = 2^{-149}$$

توجه شود در این حالت بخش $1+f$ تبدیل شده به $0+f$

مثال عدد 639.6875 در چارچوب Single Precision چگونه نمایش داده می‌شود؟

$$\underbrace{639}_{\text{در مبنای 2 البته با توانهای مثبت}}.\underbrace{6875}_{\text{در مبنای 2 البته با توانهای منفی}} \Rightarrow$$





یعنی $639 = 1001111111$

و $0.6875 = \underbrace{\frac{1}{2}+\frac{1}{8}+\frac{1}{16}}_{1\times2^{-1}+0\times2^{-2}+1\times2^{-3}+1\times2^{-4}} = 1011$

پس

$$639.6875 = 1001111111.1011 = 1.0011111111011 \times 2^{9}$$

چون در پایه دو است پس به تعداد اعشار توان 2 خواهیم داشت. این وضعیت برای Decimal یعنی پایه 10 خواهیم داشت.

پس تا اینجا معلوم شد که $e-127=9$ یعنی $e=136$

و $1.f = 1 + 0011111111011$

پس 136 را نیز بر پایه 2 می‌نویسیم پس $136 = 10001000$

0 10001000 0011111111011000 . . .

مثال: عبارت زیر چه عددی را نشان می‌دهد؟

1 01111100 11000 . . . . . .

$S = 1 \longrightarrow (-)$

$01111100 = 124 \longrightarrow e = 124 \Rightarrow e - Bias = -3$

$11000\ldots = 0.75$

ج:

$(1-2S)\,2^{-3}(1+0.75) = -1.75 \times 2^{-3} = -0.21875$

همین جا می توان روش معکوس شن آنچه در مثال قبل گفته بودم