

* در محل با محدودیت بیت مواضع هستیم ، یا باید خطا نداشته باشیم

یا خطای مطلق بعضی را تعیین کنیم که در این صورت خطاهای

نسبی متفاوتتر در وجودش آید . مثلاً غره : $4,8 \rightarrow 4,77$

$17,8 \rightarrow 17,77$

← خطای مطلق مثل هم است ولی خطاهای نسبی کاملاً متفاوت می‌شود.

و منطقاً هم هست در مورد غره .

ولی در بعضی مسائل ، خطای نسبی است که مهم است .

(مثلاً اندازه گیری فاصله بین دو شهر و طول کمانه قطعه خطای مطلق

متفاوت و خطای نسبی است که مهم است .)

* اگر در محاسباتی $3,294$

$$\begin{cases} 3,294 \times 10^2 \\ 3,294 \times 10^{34} \\ 3,294 \times 10^{-18} \end{cases}$$

با $\pm 0,0001$ خطا باشد ، خطای

مطلق همان ضریب متفاوت

است ولی خطای نسبی یکسان است .

* در بعضی مسائل لازم داریم خطای نسبی کنترل شود و در بعضی
مسائل لازم داریم خطای مطلق کنترل شود.

* در این حالت $\pm M \times B^E$ \rightarrow unique نیست

اگر $\begin{matrix} 1 < 14 < 3 \\ 1 < 14 < 8 \end{matrix}$ \leftarrow unique نیست

* اگر B چون میان بانه، ابتدا برای جمع و ضرب رفتار
مشکل داریم، به و قیاس لازم که دو عدد هم توان نسبی و مبر
باید بتواند شماره باشد.

* binary $\rightarrow \pm M \times 2^E$, $1 < 14 < 2$

* مثلا: $\begin{matrix} 3,294 \times 10^3 \\ 2,445 \times 10^3 \end{matrix} \rightarrow$ راحت هم توان میشن

باید به 10^3 تقسیم شه و اینجا،
هم توان کردن ممکنه.

* سیستم اعداد علمی مائری : $1, 14 \leq 2$, $\pm (14)_2 \times 2^E$

- حجم بیت بستر برای M هرت گنم ، دقت بالا مرد . (ارقام باستان)
- حجم بیت بستر برای E هرت گنم ، range دما صلی مرد مرد

Δ $\pm 14 \times (2)^E$ $1 \leq 14 \leq 2$ *

* لعل اولی سے

ما این غائی را
(نگارنده است)

4= ① تعداد احصاء فی F

② کوه غاس اعداد منفی $(\pm M)$

(3) مقدار ست غاٹس =

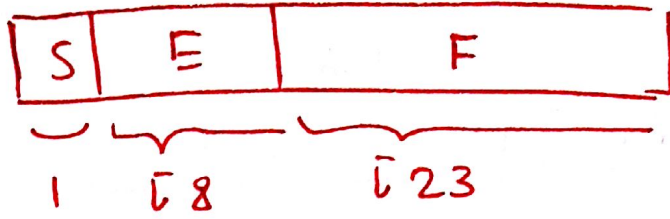
(4) کوہ غائب اعداد منفی E

← شماره‌ها برای اعداد منق؟ ← (بچه‌ها نفس، زورما درست نیست)

E: 2's complement biased

14/F : 2's complement, چون غرض shift داد

* استاندارد IEEE 754 :



32 بیت :

از راس عدد به سطر

$$(-1)^S 1.F \times 2^{E-127}$$

$$\begin{cases} E \neq \phi \\ E \neq 255 \end{cases}$$

* فرض کنیم S بیت ، هم ترین بیت ، بیت بالای E هست

البته فرض خواستیم 2's comp. باشد ، فقط توان رو منفی در کردیم .

رابطه دیکه هم این که انتظوری برای مقایسه کردن ضریب راصه ، اول

S و E اول کنترل می شه (اول S) و بعد به مقایسه integer

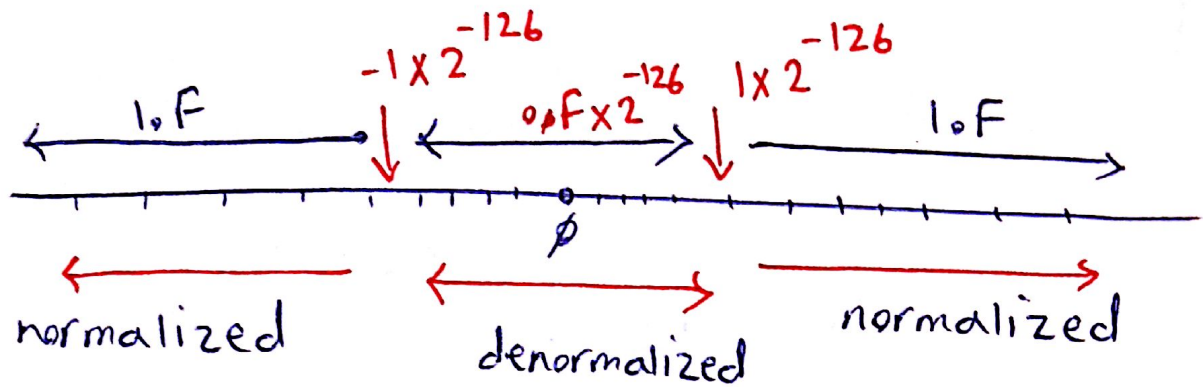
بین F ها (اگر S و E برابر بود بین دو عدد)

(سیستم positional است .)

در پس چسبندگی E-127 ، ضریب برابر رو 2's comp.

برای E هست .

$$* E = \phi : (-1)^S 0.F \times 2^{-126}$$



$$* E = 255 \quad +\infty \quad -\infty \quad NaN \text{ (Not a Number)}$$

کدها و اعداد خاص

$$1.11 \dots 11 \times 2^{127} : \text{بزرگ ترین عدد مثبت}$$

$$\approx 2^{128} \approx 10^{38}$$

$$\approx -10^{38} : \text{کوچک ترین عدد}$$

* single precision (دقت عادی) 32 bits

7، 24، 8 (صحیح، اعشاری، علامت)

* double precision (دقت مضاعف) 64 bits

$$\begin{array}{lcl} 8 \rightarrow 11 & 2^{E-127} \rightarrow 2^{E-1023} \\ 23 \rightarrow 52 & 2^{-126} \rightarrow 2^{-1022} \end{array}$$

$$E = 255 \rightarrow E = 2047$$

مسالہ: نمائش 10^{-3} نہ صورت دت جاری IEEE-754

$$10^{-3} \times 16 = 0,016$$

$$0,016 \times 16 = 0,256$$

$$0,256 \times 16 = 4,096$$

$$0,096 \times 16 = 1,536$$

$$0,536 \times 16 = 8,576$$

$$0,576 \times 16 = 9,216$$

$$0,216 \times 16 = 3,456$$

$$0,456 \times 16 = 7,296$$

$$0,296 \times 16 = 4,736$$

$0,0000 \ 0000 \ 0100 \ 1000 \ 1000 \ 0011 \ 0111 \ 0100$
 (درند بیکت مانا)
 $\times 2^{-10}$
 کہ در فرم ب را و با 23 رقم اعشار رست.

$$\Rightarrow \begin{cases} S = 0 \\ E = 127 - 10 = 117 \rightarrow 011101010101010101010101 \end{cases}$$

* است با در است با در است / کوثر است optional
 است 9 pm