Biased Exponents, Why?

Floating point representations were designed for ease of comparisons.

Sign of number comes first

This makes it easy to determine if one is larger when signs are different. Also, compares with Ø are easy.

Exponent comes next

Bigger exponent means bigger number when signs are the same.

Single precision numbers use a bias of 127 (Ø111 1111). Add exponent to 127 to get bias.
ex. $-1 + 127 = Ø111\ 111Ø$. Store 126 to represent -1. So this forces all negative exponents to start with Ø.

Notice That for positive exponents we have $1 + 127 = 128$ or $1000\ 0000$. Notice That all Positive will now start with 1. $0 < 1$ reflects negative < positive.

So, previously we wished to store $1 0 1.1 1 0 1 \times 2^{37}$

Step 1 - Normalize, so 1.011101×2^{39}

Step 2 - Sign $= 0$

Step 3 - Exponent $= 127 + 39 = 166 = \underbrace{1\overbrace{010\ 0110}^{38}}$

So, Think of This as $1 + 38 = 39$

The left most exponent bit represents 128, so, The others come from There.

Finally,

$$\underbrace{0}_{S}\ \underbrace{101\ 0011\ 0}_{Exp}\ \underbrace{011\ 1010\ 0000 \ldots 0000}_{Fraction}$$

In Hex we have $533A0000_{16}$

Let's go the other way.
Which floating point value
is represented by:
C120 0000$_{16}$?

$$\underbrace{1}_{s}\underbrace{100\,0001\,0}_{Exp}\underbrace{010\,0000\,\ldots\,0}_{fraction}$$

$$(-1)^s \times (1 + fraction) \times 2^{(exp)}$$

130 - 127

$-1 \times 1.010\ldots \times 2^3$

$1 + \frac{0}{2} + \frac{1}{4}$

1.25

-1.25×2^3

-1.25×8

-10.0_{10}

exp = 1000 0010

128 + 2

$$\begin{array}{r} 130 \\ -127 \\ \hline 3 \end{array}$$

Friday, February 23, 2024 2:28 PM

Given $\phi 2140000_{16}$

which base 1ϕ floating point value?

$\underbrace{\phi}_{S}\overbrace{\phi\phi\phi\ \phi\phi 1\phi\ \phi}\phi\phi 1\ \phi 1\phi\phi\ \phi\phi\phi\phi \ldots \phi\phi\phi\phi$

positive

$exp = 4 - 127 = -123$

frac 1.00101×2^{-123}

$1 + \frac{1}{8} + \frac{1}{32}$

$1 + 0.125 + 0.03125 = 1.15625$

1.15625×2^{-123}

-139.90625 ?

To add $9.999 \times 10^{1} + 1.610 \times 10^{-1}$

1. Convert to common exponent, usually smaller number.

$$\begin{array}{r} 9.999 \times 10^{1} \\ + \ 0.016 \times 10^{1} \\ \hline 10.015 \times 10^{1} \end{array}$$

2. Add

3. Normalize 1.0015×10^{2}

Note, at this time we check for over/under flow.

4. Round

1.002×10^{2}

Since we only had 3 digits to the right in both original numbers.

Now, in binary!

$$1.\emptyset\emptyset\emptyset \times 2^{-1}$$
$$+ (-1.11\emptyset) \times 2^{-2}$$

$(\emptyset.5_{1\emptyset} + -\emptyset.4375)$

$$\begin{array}{r} 0.5000 \\ -\,0.4375 \\ \hline 0.0625 \end{array}$$

1. Shift smaller and add

$$\begin{array}{r} 1.\emptyset\emptyset\emptyset \times 2^{-1} \\ + \; -0.111 \times 2^{-1} \\ \hline 0.001_2 \times 2^{-1} \end{array}$$

2. Normalize

$1.000_2 \times 2^{-4}$ (no under/over flow)

3. Round and renormalize if needed

$1.000_2 \times 2^{-4}$

$= \emptyset.\emptyset625$

$1/2 = \emptyset.5$, $1/4 = \emptyset.25$, $1/8 = \emptyset.125$

$1/16 = \emptyset.\emptyset625$, $1/32 = \emptyset.\emptyset3125$ . . .