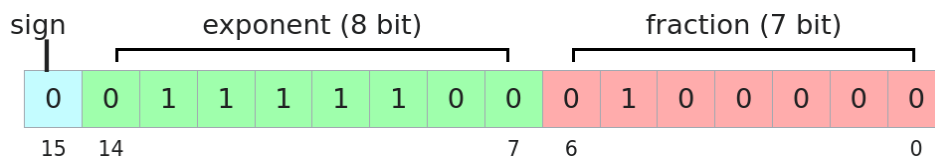


1. The **bf16** brain floating-point format consists of a sign bit, an 8-bit exponent followed by a 7-bit mantissa for a total of 16 bits.



https://en.wikipedia.org/wiki/Bfloat16_floating-point_format

The sign bit is set for a negative number, the exponent is stored as a biased integer with the offset being 127 and the mantissa is stored with a leading implicit bit similar to classical floating-point formats.

Find the decimal equivalents of

(i) 0 10000000 1001001

(ii) 0 01111011 1001101

Compute the error $\varepsilon = x^* - x$ where x^* is the **bf16** closest to

(iii) $x = 2.718$

(iv) $x = 12345.6$

2. The Julia function

```
function f(x,y)
    return (abs(x)^3+abs(y)^3)^(1/3)
end
```

implements the calculation of

$$f(x, y) = (|x|^3 + |y|^3)^{1/3}.$$

Rewrite this function in a way that avoids overflow. Use your routine to compute $f(x, y)$ when

(i) $x = 5.0 \times 10^{120}$ and $y = 3.0 \times 10^{130}$

(ii) $x = 1.0$ and $y = 3.0 \times 10^{130}$

(iii) $x = 5.0 \times 10^{120}$ and $y = 1.0$

(iv) $x = 0.0$ and $y = 0.0$

Note the number 5.0×10^{120} may be expressed in Julia as **5.0e120**.

3. Perform a forward and backward error analysis for $f(x) = \sqrt[3]{x}$. You should find the relative error is reduced by a third when applying f .