IEEE 754 Standards for Single Precision Representation

http://numericalmethods.eng.usf.edu

IEEE-754 Floating Point Standard

- Standardizes representation of floating point numbers on different computers in single and double precision.
- Standardizes representation of floating point operations on different computers.

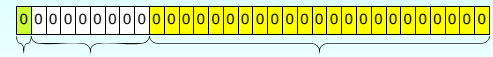
One Great Reference

What every computer scientist (and even if you are not) should know about floating point arithmetic!

http://www.validlab.com/goldberg/paper.pdf

IEEE-754 Format Single Precision

32 bits for single precision

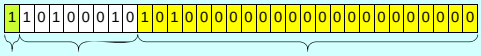


Sign (s) | Biased Exponent (e') | Mantissa (m)

$$\text{Value} = (-1)^s \times (1.m)_2 \times 2^{e'-127}$$

4

Example#1



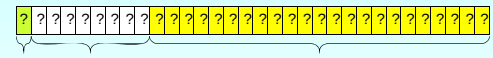
Sign (s) | Biased Exponent (e') | Mantissa (m)

$$\begin{aligned}\text{Value} &= (-1)^s \times (1.m)_2 \times 2^{e'-127} \\ &= (-1)^1 \times (1.10100000)_2 \times 2^{(10100010)_2-127} \\ &= (-1)\times(1.625)\times 2^{162-127} \\ &= (-1)\times(1.625)\times 2^{35} = -5.5834\times 10^{10}\end{aligned}$$

5

Example#2

Represent -5.5834x10^10 as a single precision floating point number.



Sign (s) | Biased Exponent (e') | Mantissa (m)

$$-5.5834\times 10^{10} = (-1)^1 \times (1.?)\times 2^{\pm ?}$$

6

Exponent for 32 Bit IEEE-754

8 bits would represent

$$0 \le e' \le 255$$

Bias is 127; so subtract 127 from representation

$$-127 \le e \le 128$$

7

Exponent for Special Cases

Actual range of e'

$$1 \le e' \le 254$$

$e' = 0$ and $e' = 255$ are reserved for special numbers

Actual range of e

$$-126 \le e \le 127$$

Special Exponents and Numbers

$e' = 0$ — all zeros

$e' = 255$ — all ones

s	e'	m	Represents
0	all zeros	all zeros	0
1	all zeros	all zeros	-0
0	all ones	all zeros	∞
1	all ones	all zeros	$-\infty$
0 or 1	all ones	non-zero	NaN

IEEE-754 Format

The largest number by magnitude

$$(1.1.....1)_2 \times 2^{127} = 3.40 \times 10^{38}$$

The smallest number by magnitude

$$(1.00.....0)_2 \times 2^{-126} = 2.18 \times 10^{-38}$$

Machine epsilon

$$\varepsilon_{mach} = 2^{-23} = 1.19 \times 10^{-7}$$

10

Additional Resources

For all resources on this topic such as digital audiovisual lectures, primers, textbook chapters, multiple-choice tests, worksheets in MATLAB, MATHEMATICA, MathCad and MAPLE, blogs, related physical problems, please visit

http://numericalmethods.eng.usf.edu/topics/floatingpoint_representation.html

THE END

http://numericalmethods.eng.usf.edu