Owais Kazi

Assignment 1 - Floating Point Arithmetic

1a. 8.125
- Binary value = 1000.001
- Single Precision representation = 1.000001 * 2^3

1b. 1 ulp = change in the last bit
1.00000100000000000000001 * 2^3
Change = $2^{-23} * 2^{20}$ = 2^{-20}

2a. 1/7
- Binary Representation ⅛+0/16+0/32+1/64+.... = $0.\overline{001}$

2b. 1.00100100100100100100100 * 2^{-3}

2c. $0.\overline{001}$ * $2^{-23} * 2^{-3}$ = $0.\overline{001} * 2^{-26}$
= 0.1*(1+ $2^{-3} + 2^{-6}$ +...)* 2^{-26}
= 0.1*(1/1- 2^{-3})* 2^{-26}
= 0.1 * (8/7) * 2^{-26}
Converting it to decimal, we get
= ½ * 8/7* 2^{-26} = 0.6* 2^{-26}
$\hat{x}$ -x = (1-0.6)* 2^{-26} = 0.4* 2^{-26}

3. 0.135$\overline{135}$ = 0.135 * (1+ $10^{-3} + 10^{-6}$ +...)
= 0.135 * (1/ (1- 10^{-3}))
= 0.135*(1000/999)
= 135/999 = 15/111

4. False.
Reason - 2 does not have the same factor as 10. Both can't terminate the same way.

5a. $x + y = a$
$x + (1 + 2^{-n})y = b$
Subtracting both the equations
$y - (1 + 2^{-n})y = a - b$
$y = b - a/2^{-n}$
$x = (a(2^{-n} + 1) - b)/ 2^{-n}$

5b. Let's consider change in b by 1 ulp $b^{`} = b + 2^{-23}$

$y^{`} = (b + 2^{-23} - a)/2^{-n}$

Change = $y^{`} - y$ = 2^{n-23}

Similarly change in x = $x^{`} - x$ = 2^{n-23}

Even if n is relatively modest, then b is subject to roundoff error.

6a. 6*10*10*10*10*10 = 600,000

6b. $2^{23} - 2^{21}$

6c. 6* 10^5 will be less than $2^{23} - 2^{21}$. One hole will shared between multiple pigeons.

6d. 6* 10^6 will be less than $2^{23} - 2^{21}$. One hole will shared between multiple pigeons.

7a. (+,1.5,0) = ± 1.5, (-,1.5,1) = $-2^{1.5}$, (+,1.5,2) = $2^{2^{1.5}}$, (-,1.5,-1) = $-2^{-1.5}$

7b. $0 \rightarrow \pm s$, $\pm 1 \rightarrow \pm 2^s$, $\pm 2 \rightarrow \pm 2^{2^s}$, $\pm 3 \rightarrow \pm 2^{2^{2^s}}$, $\pm 4 \rightarrow \pm 2^{2^{2^{2^s}}}$

7c. (+,s,4) = $2^{2^{2^{2^s}}}$ and the max value of s = 2^{27}

(+,s,4) = $2^{1.0531 * 10^{65}}$

To find value of x in $2^{1.0531 * 10^{65}}$ = G^x

$1.0531 * 10^{65} \, log\, 2 = x \, log\, G = x \, log\, 10^{100} = 100x$

$x \approx 3.1701 * 10^{62}$

$G^{3.1701 * 10^{65}} \approx (+, s, 4)$

7d. (+,s,5) = $2^{2^{2^{2^{2^s}}}}$ and max value of s = 2^{27}

(+,s,5) = $2^{2^{1.0531 * 10^{65}}}$

Relation between $2^{2^{1.0531 * 10^{65}}}$ and 10^G

Let x = $2^{1.0531 * 10^{65}}$

$2^x = 10^y$

$x \, log\, 2 = y \, log\, 10$

$y = x \, log\, 2$

The largest representable value of (+,s,5) is greater than googolplex.

7e. (+,s,6) = $2^{2^{2^{2^{2^{2^s}}}}}$ can be call as 'saturn

(+,s,7) = $2^{2^{2^{2^{2^{2^{2^{s}}}}}}}$ can be call as ‘uranus’

Collaborated with Amitabh Das for the assignment.