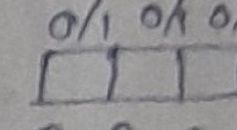
-N/L A1- Assignment-1 - Problem-1 | D. Pradeep Chandra - 25173

Given, and $n = (1.f)_2 \times 2^{(exponent-1)}$ as toy system and here we get exponent has 2 bits & mantissa f has 5 bits So, As per the question exponent - 1 = exp where exponent is an actual exponent and the exp which is exponent field can be taken as (-1, 0, 1). so, $n = (1.f)_2 \times 2^{exp}$

(a) Soli: So, we know that mantissa has 3 bits 

$\rightarrow 2^3$ possibilities where as exponent field exp can take. (-1, 0, 1) which is 3 values.

As per multiplication rule in combinatorics, we can have $2^3 \times 3 = 24$ possibilities.

i.e., Total number of numbers formed in toy system = 24

(b) Soli: we know list out all the numbers represented by the toy system. where 3 tables formed and each table has 8 numbers with their respective exponent field (-1, 0, 1)

exp = -1 (Table-1)

Normalized form (1.f)	Binary representation	Decimal (Base-10)
1.000×2^{-1}	$(0.1000)_2$	0.5
1.001×2^{-1}	$(0.1001)_2$	0.5625
1.010×2^{-1}	$(0.1010)_2$	0.625
1.011×2^{-1}	$(0.1011)_2$	0.6875
1.100×2^{-1}	$(0.1100)_2$	0.75
1.101×2^{-1}	$(0.1101)_2$	0.8125
1.110×2^{-1}	$(0.1110)_2$	0.875
1.111×2^{-1}	$(0.1111)_2$	0.9375

exp = 0 (Table-2)

Normalized form (1.f)	Binary representation	Decimal Base-10
1.000×2^{0}	$(1.000)_2$	1
1.001×2^{0}	$(1.001)_2$	1.125
1.010×2^{0}	$(1.010)_2$	1.25
1.011×2^{0}	$(1.011)_2$	1.375
1.100×2^{0}	$(1.100)_2$	1.5
1.101×2^{0}	$(1.101)_2$	1.625
1.110×2^{0}	$(1.110)_2$	1.75
1.111×2^{0}	$(1.111)_2$	1.875

exp = 1 (Table-3)

Normalized form (1.f)	Binary Representation	Decimal (Base-10)
1.000×2^{1}	$(10.000)_2$	2
1.001×2^{1}	$(10.010)_2$	2.25
1.010×2^{1}	$(10.100)_2$	2.5
1.011×2^{1}	$(10.110)_2$	2.75
1.100×2^{1}	$(11.000)_2$	3
1.101×2^{1}	$(11.010)_2$	3.25
1.110×2^{1}	$(11.100)_2$	3.5
1.111×2^{1}	$(11.110)_2$	3.75

(c) Soli: The minimum value in the system = 0.5
The maximum value in the system = 3.75

(d) Soli: Actually what I found that absolute gaps are not constan
they increase as the numbers get larger. This was the characteristi
of floating point systems: the precision decreases (i.e, the gap
increase) as the magnitude of numbers increases same happens
in toy floating point system.

If we observes, the absolute gap blw two consecutive
numbers in table-1 = 0.0625 $\longrightarrow$ exp = -1
table-2 = 0.125 $\longrightarrow$ exp = 0
table-3 = 0.25 $\longrightarrow$ exp = 1

Sol:- The machine epsilon ($\epsilon_{machine}$) is defined as the smalles value ϵ such that

$$1 + \epsilon \neq 1$$

so, the relative error $= \frac{|x - x'|}{|x|} \leq \epsilon_{machine}$ is only done by minimizing by doing 1st two consecutive numbers.

i.e., $\epsilon_{machine} = 0.5625 - 0.5 = 0.0625$

it's where we get minimum value.