CS & IT ENGINEERING

Data Structures

Hashing

Lecture No.- 02

By- Pankaj Sharma Sir

Recap of Previous Lecture

Topic Hashing Part 01

Hash func.

CRT

1. Linear probing
2. Quadratic probing

Topics to be Covered

Topic Hashing Part 02

Double hashing
Chaining

Topic : Hashing

Double Hashing

let $h(k) = k \bmod m$ is the hash func.

Collission occurs

Collission Resolution

(i) $H(k,i) = (h(k)+i) \bmod m$ L.P

(ii) $H(k,i) = (h(k)+i^2) \bmod m$ Q.P

(iii) $H(k,i) = (h(k)+i\,h'(k)) \bmod m$ D.H

$h(k)$ → Primary hash function

$h'(k)$ → Secondary hash function

what if $h'(k) = 0$?

$$\boxed{H(k,i) = h(k)}$$

$\Rightarrow$ $\boxed{h'(k) \text{ can never be } 0}$

$2 \rightarrow 2, 7$

$57 \rightarrow 2, 8$

$m = 11$ Keys: 13, 17, 21, 2, 57, 28, 30, 27

Given

$$h(x) = x \bmod 11$$

$$h'(x) = 7 - (x \bmod 7)$$

(i) $h(13) = 2$

(ii) $h(17) = 17 \bmod 11 = 6$

(iii) $h(21) = 21 \bmod 11 = 10$

(iv) $h(2) = 2$ coll.

$$H(2,1) = (h(2) + 1 \cdot h'(2)) \bmod 11$$
$$= (2 + 1 \cdot 5) \bmod 11 = 7$$

(v) $h(57) = 57 \bmod 11 = 2$ coll.

$$H(57,1) = (h(57) + 1 \cdot h'(57)) \bmod 11$$
$$= (2 + 1 \cdot 6) \bmod 11 = 8$$

(vi) $h(28) = 28 \bmod 11 = 6$ coll.

$$H(28,1) = (h(28) + 1 \cdot h'(28)) \bmod 11$$
$$= (6 + 1 \cdot 7) \bmod 11 = 2 \text{ coll.}$$

$$H(28,2) = (h(28) + 2 \cdot h'(28)) \bmod 11$$
$$= (6 + 2 \cdot 7) \bmod 11 = 9$$

(vii) $h(30) = 8$ coll

$$H(30,1) = (h(30) + 1 \cdot h'(30)) \bmod 11$$
$$= (8 + 1 \cdot 5) \bmod 11 = 2 \text{ coll}$$

$$H(30,2) = (h(30) + 2 \cdot h'(30)) \bmod 11$$
$$= (8 + 2 \cdot 5) \bmod 11 = 7 \text{ coll}$$

$$H(30,3) = (h(30) + 3 \cdot h'(30)) \bmod 11$$
$$= (8 + 3 \cdot 5) \bmod 11 = 1$$

(viii) $h(27) = 27 \bmod 11 = 5$

0	
1	30
2	13
3	
4	
5	27
6	17
7	2
8	57
9	28
10	21

Two hash function

$\downarrow$

$\boxed{\text{Time consuming}}$

λ : load factor

$$\lambda = \frac{\text{no. of keys}}{\text{Table size}} = \frac{n}{m}$$

$n = 20$

$m = 30$

$$\lambda = \frac{20}{30} = \frac{2}{3}$$

$n = 10$

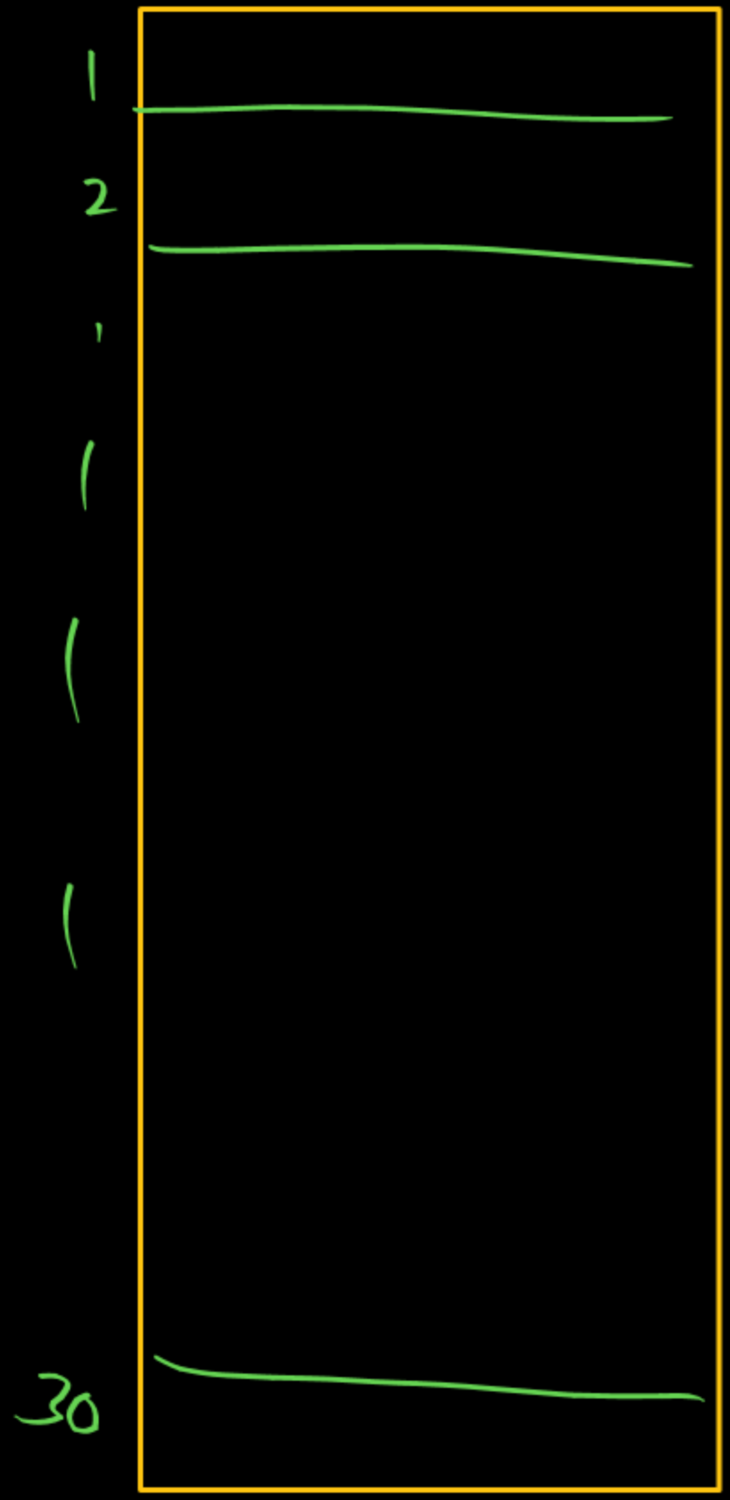
$m = 30$

$$\lambda = \frac{1}{3}$$

$n = 40$

$m = 30$

$$\lambda = \frac{4}{3}$$



$\boxed{\lambda \leq 1}$

Linear Probing

Quadratic Probing

Double hashing

$\boxed{\lambda > 1}$

Separate chaining

$m=10$ $h(k) = k \bmod 10$

Keys: 400, 300, 125, 605, 6, 36, 16, 96

$h(400) = 0$

$h(300) = 0$] Collission

Woost case

100, 200, 300, 400, 500, 600, 700, 800

$(m+n)$ pointers

Space complexity.

Linear Probing

$m = 12$

Keys : 31, 26, 43, 27, 34, 12, 46, 14, 58

7 2 ~~7~~ 8 3 10 0 ~~10~~ 11 ~~2~~ ~~3~~ 4 ~~10~~ ~~11~~ ~~0~~ 1

Insert, Search (i)

(ii) Deletion 26

→ Search 14

$h(14) = 14 \bmod 12 = 2$

Re-hash

0	12
1	58
2	~~26~~
3	27
4	14
5	
6	
7	31
8	43
9	
10	34
11	46

Deletion is very difficult in LP, QP & DH

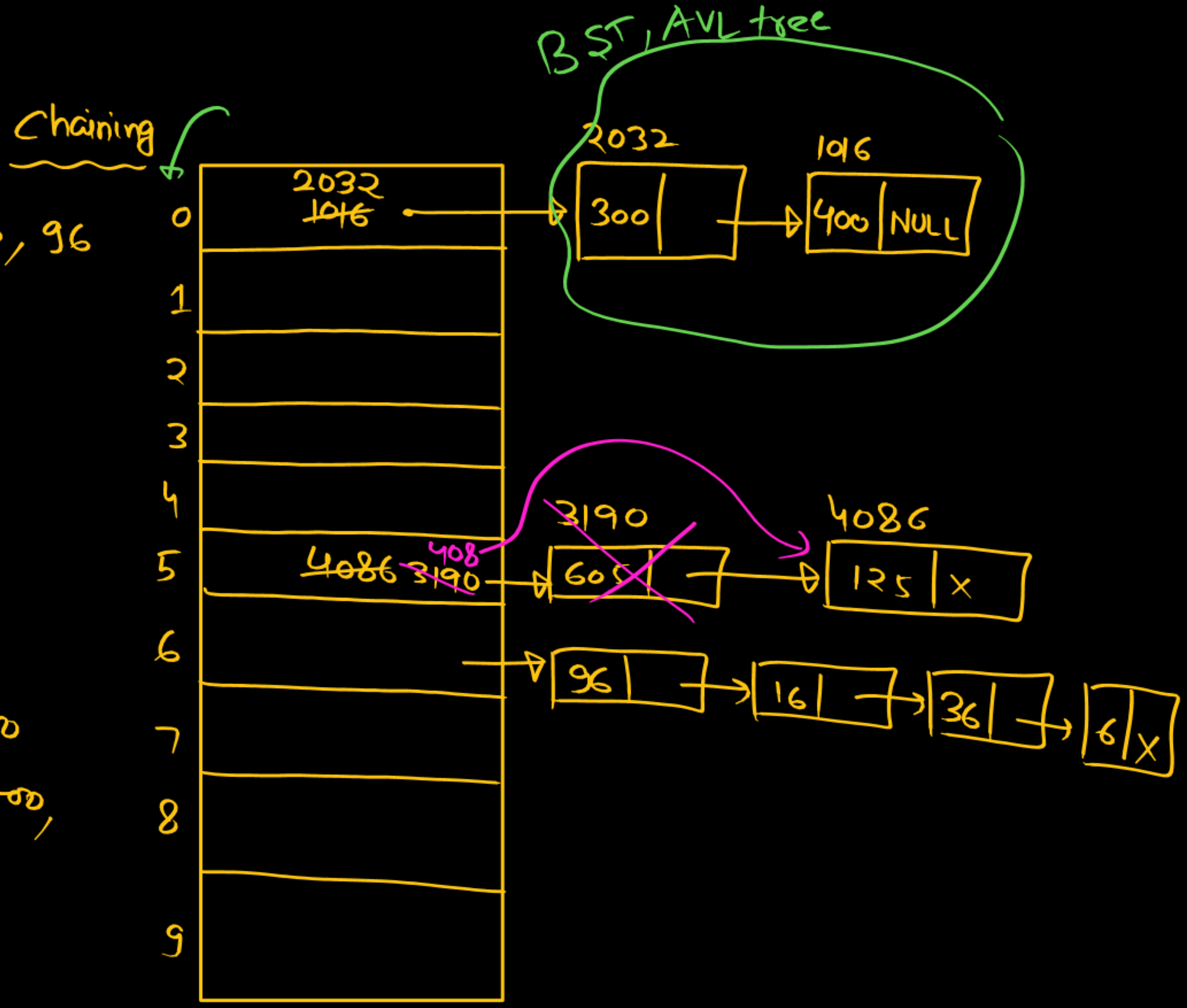
deletion is very easy in Chaining.

$m = 10$ $h(k) = k \bmod 10$

Keys: 400, 300, 125, 605, 6, 36, 16, 96

$h(400) = 0$

$h(300) = 0$] Collission



Worst case

100, 200, 300, 400

1500, 600, 700,

800

$(m+n)$ pointers

Space complexity.

Gate 2020

Double hashing

$h_1(k) = k \bmod 23$

$h_2(k) = 1 + (k \bmod 19)$

$m = 23$

Address returned by probe 1 in the probe seq (assume probe seq. begins at probe 0) for key $k = 90$ is ________

Collission no $\Rightarrow$ 1

$k = 90$

$h_1(90) = 90 \bmod 23$

$= 21$

$h_2(90) = 1 + 90 \bmod 19$

$= 15$

$H(90, 1) = ?$

$H(90, 1) = (h_1(90) + 1 \cdot h_2(90)) \bmod 23$

$= (21 + 15) \bmod 23$

$= 36 \bmod 23$

$= 13$

Gate-2015

25 slots

2000 elements

the load factor α is ______

$$\alpha = \frac{2000}{25} = 80$$

Gate-2015

Which one of the following hash functions on integers will distribute keys most uniformly over 10 buckets numbered 0 to 9 for i ranging from 0 to 2020?

~~a)~~ $h(i) = i^2 \bmod 10$

b) $h(i) = i^3 \bmod 10$

~~c)~~ $h(i) = (11 \times i^2) \bmod 10$

d) $h(i) = (12 \times i) \bmod 10$

$11 \times i^2 \bmod 10$

$= i^2 \bmod 10$

unit digit

	unit digit
$(\_\_1)^2 =$	1
$(\_\_2)^2 =$	4
$(\_\_3)^2 =$	9
$(\_\_4)^2 =$	6
$(\_\_5)^2 =$	5
$(\_\_6)^2 =$	6
$(\_\_7)^2 =$	9
$(\_\_8)^2 =$	4
$(\_\_9)^2 =$	1
$(\_\_0)^2 =$	0

bucket $\rightarrow$ 0, 1, 4, 5, 6, 9

Empty $\rightarrow$ 2, 3, 7, 8

Gate-2015

Which one of the following hash functions on integers will distribute keys most uniformly over 10 buckets numbered 0 to 9 for i ranging from 0 to 2020?

~~a)~~ $h(i) = i^2 \bmod 10$

b) $h(i) = i^3 \bmod 10$

~~c)~~ $h(i) = (11 \times i^2) \bmod 10$

~~d)~~ $h(i) = (12 \times i) \bmod 10$ → all odd numbered bucket will remain empty

$(12 \times i)$
↓
Even no. mod 10
↓
Even no

i	$i^3 \bmod 10$
$(\_\_1)^3$	1
$(\_\_2)^3$	8
$(\_\_3)^3$	7
$(\_\_4)^3$	4
$(\_\_5)^3$	5
$(\_\_6)^3$	6
$(\_\_7)^3$	3
$(\_\_8)^3$	2
$(\_\_9)^3$	9
$(\_\_0)^3$	0

Gate - 2014

Hash table with 100 slots. Collisions are resolved using $\boxed{\text{chaining.}}$

Assume simple uniform hashing, what is the probability that the first 3 slots are unfilled after first 3 insertions?

a) $(97 \times 97 \times 97)/100^3$ ✓

b) $(99 \times 98 \times 97)/100^3$

c) $(97 \times 96 \times 95)/100^3$

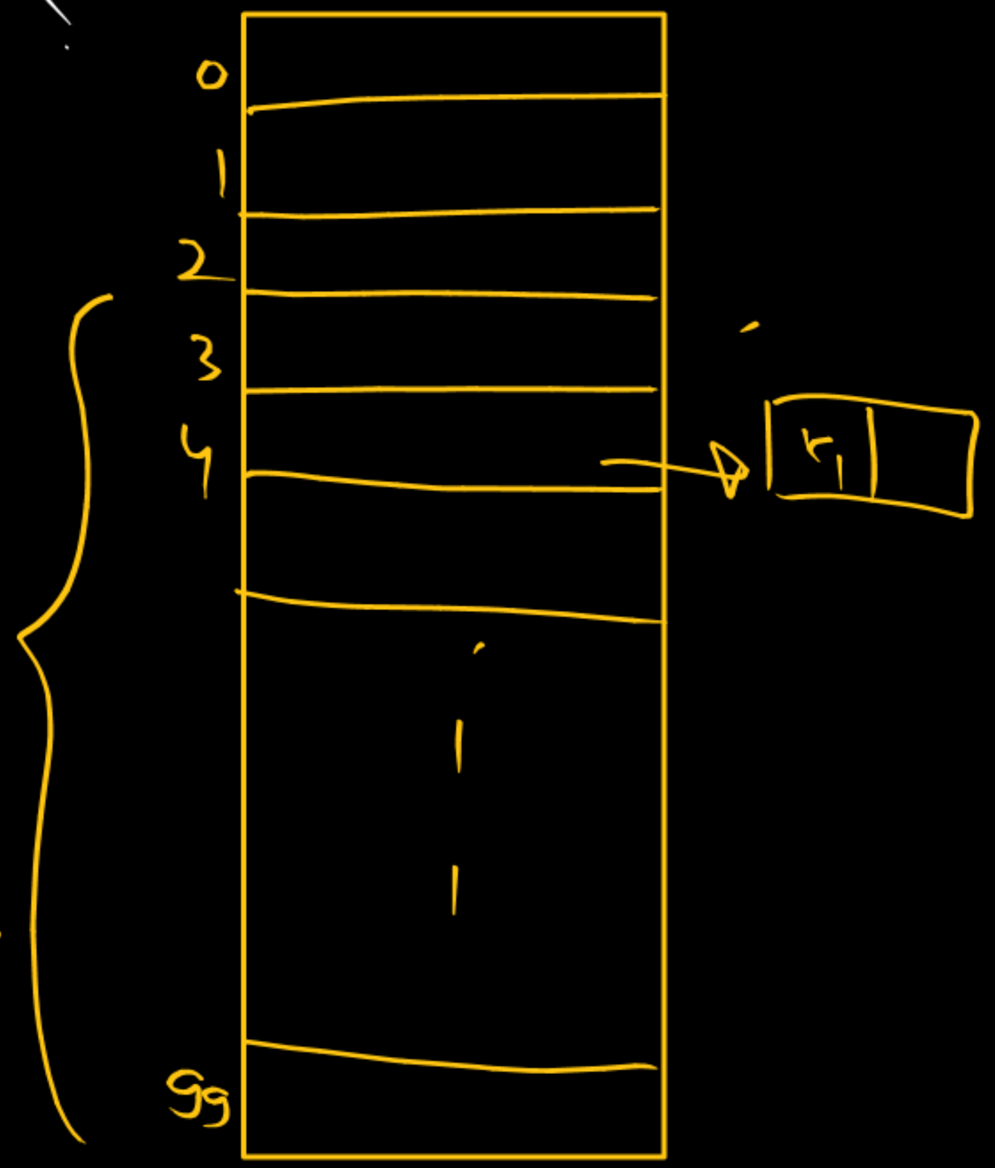
d) $(97 \times 96 \times 95)/3! \times 100^3$

1st $\rightarrow \frac{97}{100}$

2nd $\rightarrow \frac{97}{100}$

3rd $\rightarrow \frac{97}{100}$

Prob $= \frac{97}{100} \times \frac{97}{100} \times \frac{97}{100}$



Gate-2014

m = 9 slots. $h(k) = k \bmod 9$

CR → chaining.

Keys: 5, 28, 19, 15, 20, 33, 12, 17, 10

(hash values above keys: 5, 1, 1, 6, 2, 6, 3, 8, 1)

The max, min & avg. chain length in the hash table are

a) 3, 0, 1 ✓

b) 3, 3, 3

c) 4, 0, 1

d) 3, 0, 2

max → 3

min → 0

0	x
1	3
2	1
3	1
4	x
5	1
6	1
7	x
8	1
9	x

$$avg = \frac{0+3+1+1+0+1+1+0+1+0}{9} = \frac{9}{9} = 1$$

3, 0 and 1

Gate 2010

hash table of size 10 uses open addressing with hash func. $h(K) = K \bmod 10$ and linear probing. After inserting 6 values into an Empty hash table, the table is shown below:

0	
1	
2	42
3	23
4	34
5	52
6	46
7	33
8	
9	

Which one of the following choices gives a possible order in which the key values could have been inserted in table?

a) 46, 42, 34, 52, 23, 33

b) 34, 42, 23, 52, 33, 46

c) 46, 34, 42, 23, 52, 33

d) 42, 46, 33, 23, 34, 52

THANK - YOU