Computer Science & IT

Data Structure & Programming

Graph & Hashing

Lecture No. 02

Recap of Previous Lecture

Topic Graph Representation

Topic Graph Traversal (BFS)

Topic Direct Addressing

Topic

Topic

Slide

Topics to be Covered

- **Topic** Hashing.
- **Topic** Hash table
- **Topic** Hash-function
- **Topic** Collision
- **Topic** Chaining, Linear probing

Slide

Topic : Hashing

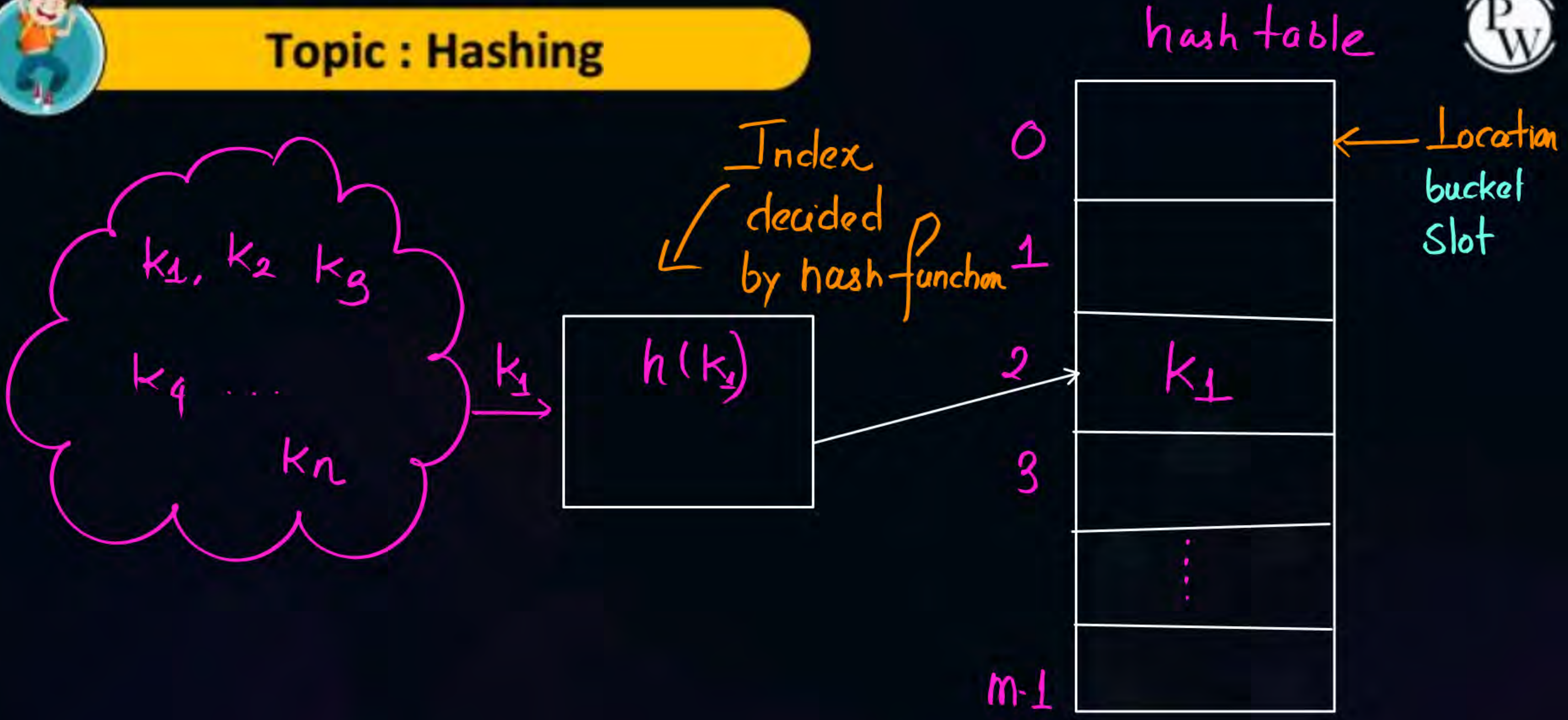
Searching Best Method Direct Addressing . Constant time

Rather than using a big Size table lets use a Small Size table called Hash table.

* rather than mapping the key k at k^{th} Index we store (map) h(k) (Hash function)

Slide

Topic : Hashing



Slide

Topic : Hashing

* Each faster access location of Hash table is called bucket/slot
* Hash function should uniformally distribute the keys.
* The table size m, if m is prime No then it leads to better distribution of keys.
* No. of keys to be mapped is n

Slide

Topic : Hashing

$\boxed{\text{fundamental}}$

hash function $h(k) = k \bmod 7$

41. $h(41) = 41 \bmod 7 = 6$

108 $h(108) = 108 \bmod 7 = 3$

	0
	1
	2
108	3
	4
	5
41	6

Slide

Topic : Hashing

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

(A) $h(i) = i^2 \bmod 10$

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

(C) $h(i) = (11*i^2) \bmod 10$

(D) $h(i) = (12*i^2) \bmod 10$

Home work

0
1
2
3
4
5
6
7
8
9

Slide

Topic : Hashing

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

(A) $h(i) = i^2 \bmod 10$

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

(C) $h(i) = (11*i^2) \bmod 10$

(D) $h(i) = (12*i^2) \bmod 10$

Home work

0	$0^2 \bmod 10 = 0$
1	$1^2 \bmod 10 = 1$
2	$2^2 \bmod 10 = 4$
3	$3^2 \bmod 10 = 9$
4	$4^2 \bmod 10 = 6$
5	$5^2 \bmod 10 = 5$
6	$6^2 \bmod 10 = 6$
7	$7^2 \bmod 10 = 9$
8	$8^2 \bmod 10 = 4$
9	$9^2 \bmod 10 = 1$

$11^2 \bmod 10$

$= \underline{1}$

Slide

Topic : Hashing

$i^2 \bmod 10$

0	0,10
1	1,9
2	
3	
4	2,8
5	5
6	4,6
7	
8	
9	3,7

$i^3 \bmod 10$

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

uniform

0	
1	
2	
3	
4	
5	
6	
7	
8	
9	

0	
1	
2	
3	
4	
5	
6	
7	
8	
9	

Slide

Load factor Load factor is ratio between No. of keys to Size of table. (α)

$$\boxed{\alpha = \frac{n}{m}}$$

#Q Given a has table T with 25 slots that stores 2000 elements, the load factor α for T is 80.

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

Two keys mapped to Same Location leads to collision.

$k^2 \bmod 10$

$h(2) = 2^2 \bmod 10 = 4$

$h(8) = 8^2 \bmod 10 = 4$

is called collision

Collision Resolution Technique

1 open Hashing

* chaining or seperate chaining

2 Closed Hashing or open Addressing

* Linear probing
* quadratic probing
* Double Hashing

Slide

open hashing collided key stores outside of table.

Seperate chaining is an example of open Hashing where each time collision occurs we create a New Node and connect to chain.

Slide

Example

Let us see the following example to get better idea. If we have some elements like $\{15, 47, 23, 34, 85, 97, 65, 89, 70\}$. And our hash function is $h(x) = x \bmod 7$.

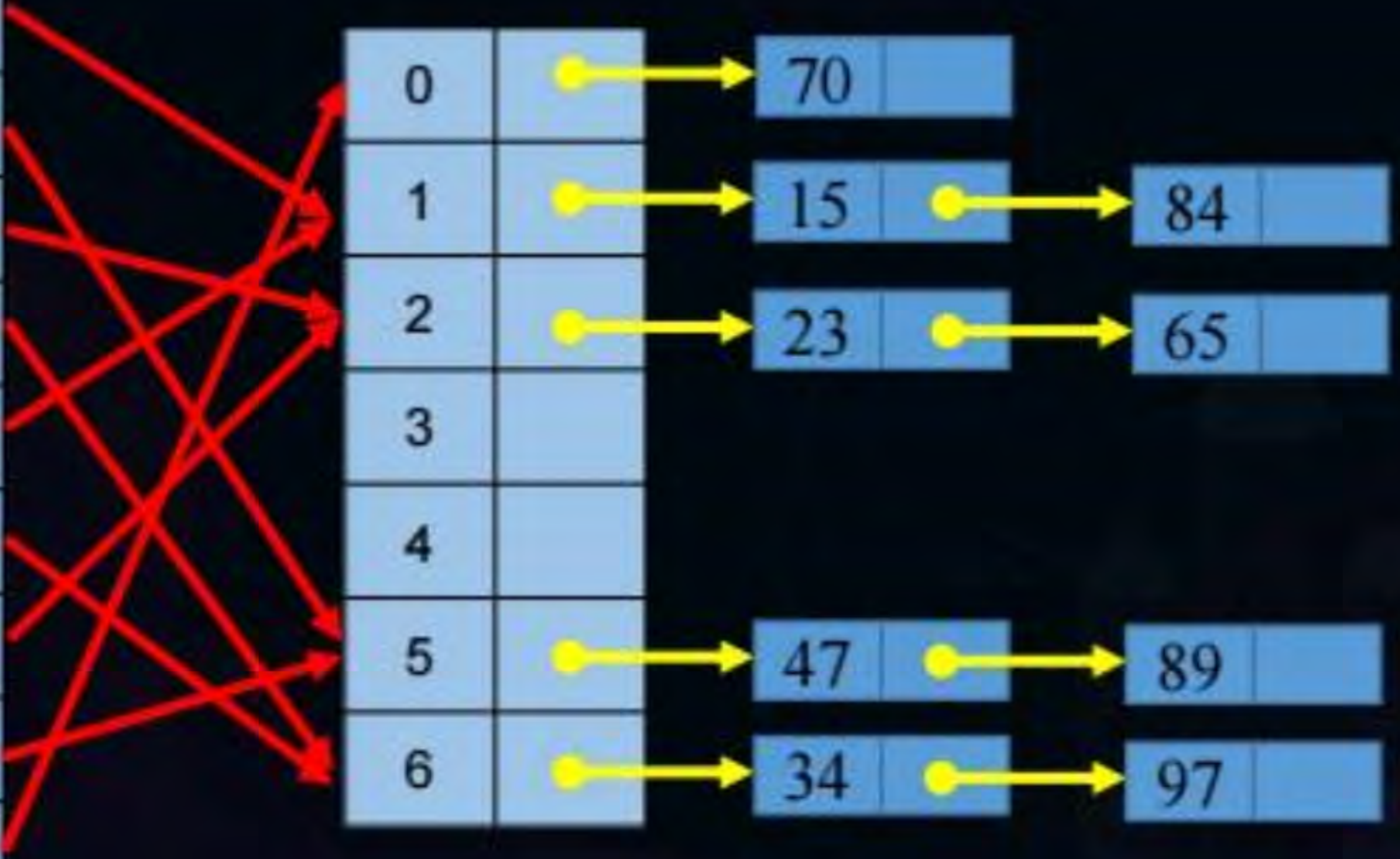
Array of linked List

Insert at begin

Slide

Topic : Hashing

$Key - x$	$h(x) = x \bmod 7$
15	1
47	5
23	2
34	6
85	1
97	6
65	2
89	5
70	0



Slide

Topic : Hashing

#Q. Consider a hash table with 9 slots. The hash function is The collisions are resolved by chaining.

$\boxed{h(k) = k \bmod 9}$

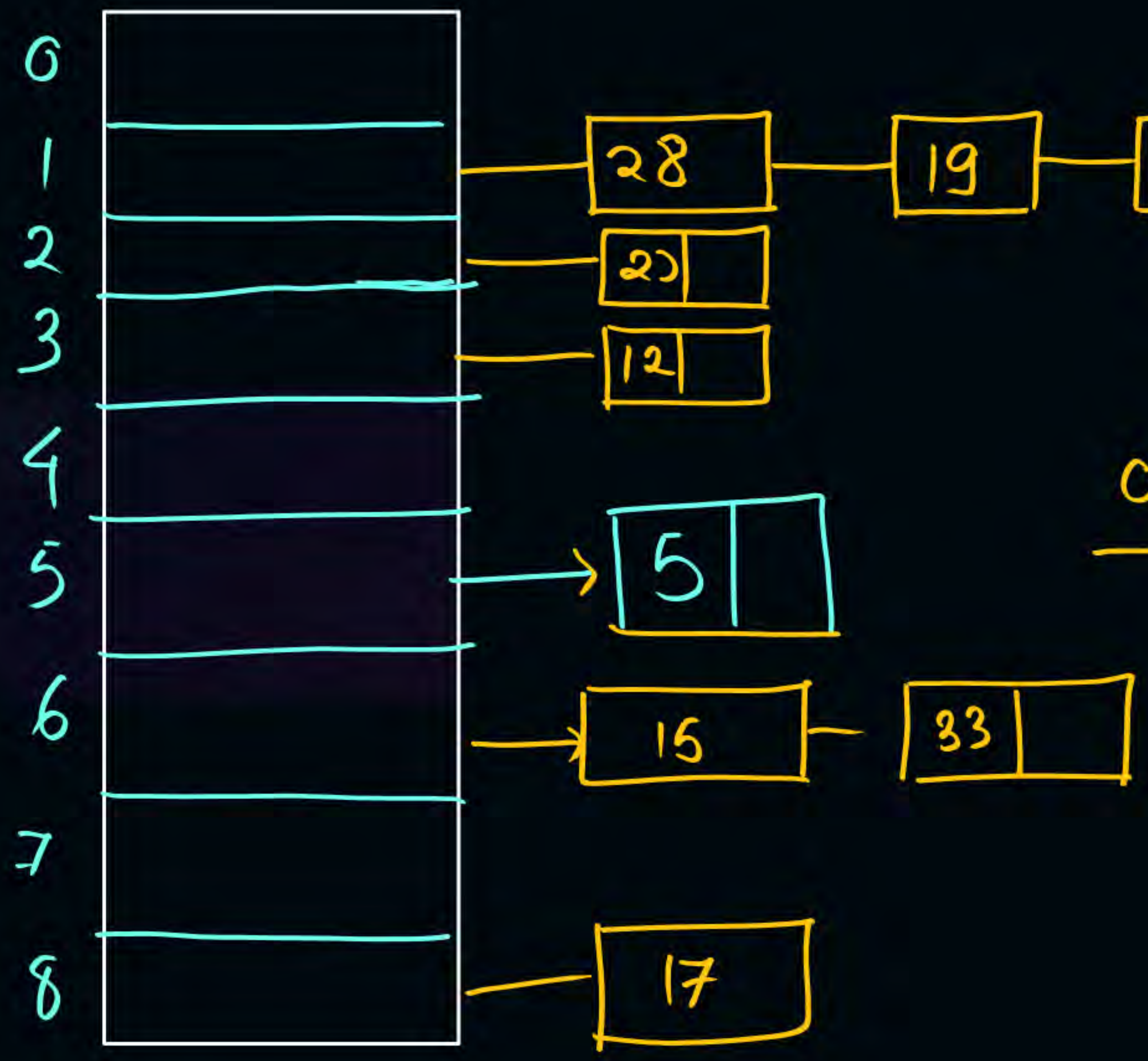
The following 9 keys are inserted in the order: 5, 28, 19, 15, 20, 33, 12, 17, 10. The maximum, minimum, and average chain lengths in the hash table, respectively, are

(A) 3, 0, and 1   (B) 3, 3, and 3

(C) 4, 0, and 1   (D) 3, 0, and 2

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

k mod 9



3, 0, 1

$$\frac{0+3+1+1+1+2+1}{9}$$

$9/9 = 1$

Topic : Hashing

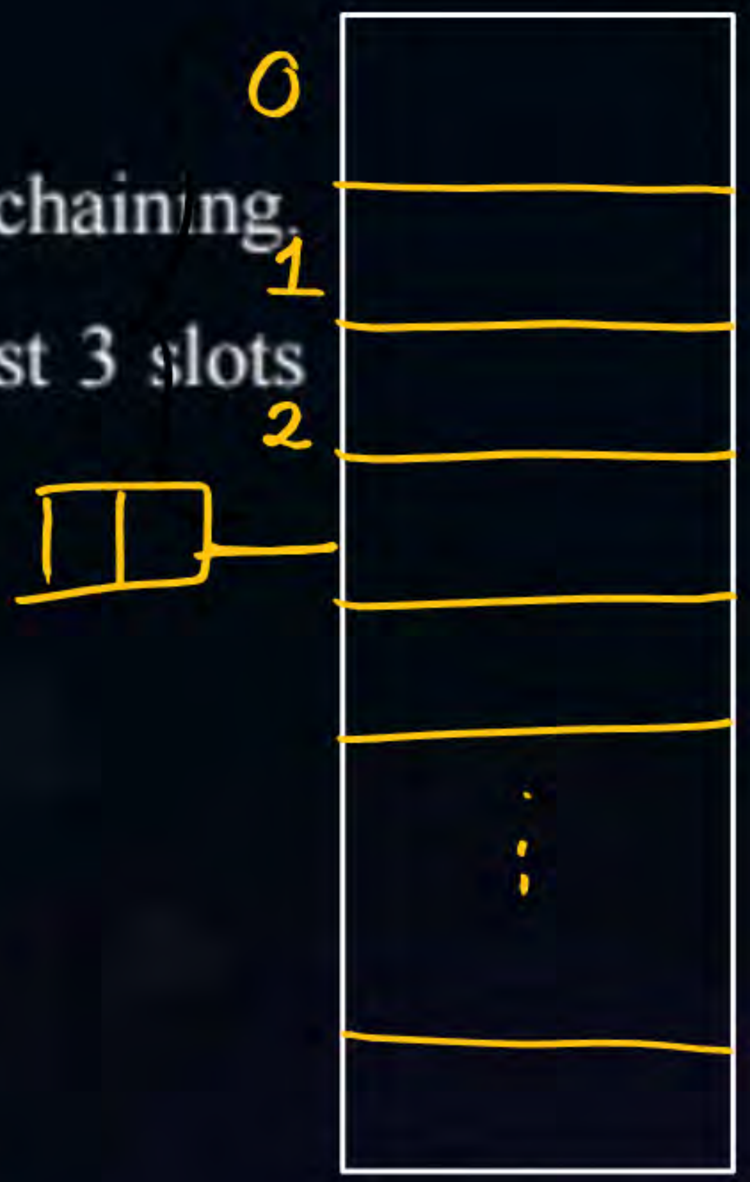
Q. Consider a hash table with 100 slots. Collisions are resolved using chaining. Assuming simple uniform hashing, what is the probability that the first 3 slots are unfilled after the first three insertions?

(A) (97 x 97 x 97)/100^3

(B) (99 x 98 x 97)/100^3

(C) (97 x 96 x 95)/100^3

(D) (97 x 96 x 95)/(3! x 100^3)



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

chaining does not block the slot

Slide

Topic : Hashing

Generalized result

$$\frac{m-1}{m} \times \frac{m-1}{m} \times \cdots \times \frac{m-1}{m}$$

1st 2nd K

$$\left(\frac{m-1}{m}\right)^K$$

Suppose there are m buckets
uniform distribution is used
What is the probability that slot
1 is empty after K^{th} Insertion

Slide

H.W

Suppose there are m buckets
uniform distribution is used
what is the probability No collision occurs in K^{th} Insertion

Topic : Hashing

Closed Hashing or open Addressing

Linear probing : In closed Hashing collided key store within the table at different Location.

In Linear probing if collision occurs then we linearly Search for empty position one after another.

Slide

Linear Probing Example

Consider the the hash table of size 13, the hash table is initially empty and hash $h(k) = k \ mod \ 13$ is used. Collision is resolved by *linear probing*.

Insert keys:

18, 41, 22, 44, 59, 32, 31, 73

in to hash table and show the resultant hash table.

0	70
1	
2	41
3	
4	
5	18
6	44
7	59
8	32
9	22
10	31
11	73
12	25

59 mod 13 = 7

44 mod 13 = 5

31 mod 13 = 5

32 mod 13 = 6

73 mod 13 = 8

25 mod 13 = 12

70 mod 13 = 5

Slide

Linear Probing Example

18, 41, 22, 44, 59, 32, 31, 73

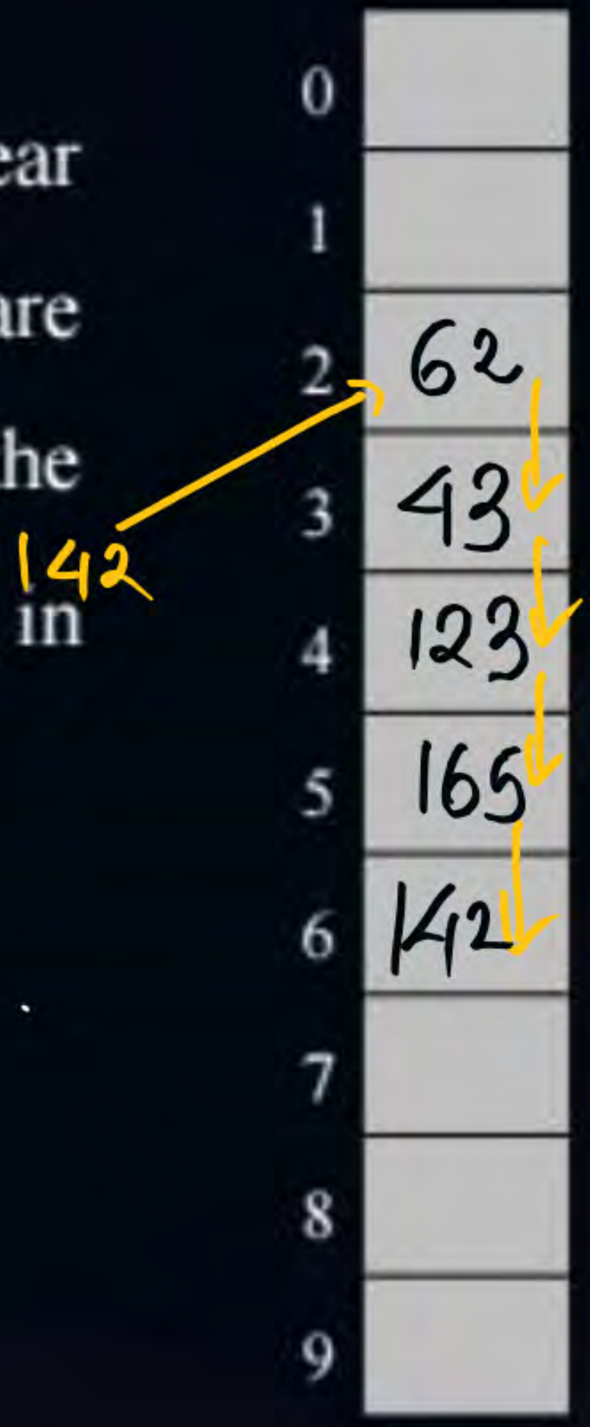
Key	$h(x) = x \bmod 13$
18	5
41	2
22	9
44	5
59	7
32	6
31	5
73	8

	0
	1
41	2
	3
	4
18	5
44	6
59	7
32	8
22	9
31	10
73	11
	12

Slide

Topic : Hashing

A hash table contains 10 buckets and uses linear probing to resolve collisions. The key values are integers and the hash function used is key % 10. if the values 43, 165, 62, 123, 142 are inserted in the table, in what location would the key value 142 be inserted?



(A) 2

(B) 3

(C) 4

(D) 6

Slide

Topic : Hashing

A hash table of length 10 uses open addressing with hash function h(k) = kmod10, and linear probing. After inserting 6 values into an empty hash table, the table is as 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 the 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

Slide

Topic : Hashing

42, 23, 34, 52, 46, 33

	A	B	C	D
	42	42	42	42
	52	23	23	33
	34	34	34	
		52	52	
	46	33	46	46
			33	

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

A) 46, 42, 34, 52, 23, 33 ✗

B) 34, 42, 23, 52, 33, 46 ✗

C) $\boxed{46, 34, 42, 23, 52, 33}$ ✓

D) 42, 46, 33, 23, 34, 52 ✗

Slide

Topic : Hashing

42, 23, 34, 52, 46, 33

A	B	C	D
42	42	42	42
23	23	23	33
34	34	34	
52	52	52	
46	33	46	46
33		33	

α

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

A) 46, 42, 34, 52, 23, 33 α

B) 34, 42, 23, 52, 33, 46 α

C) $\boxed{46, 34, 42, 23, 52, 33}$ ✓

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

Slide

Topic : Hashing

How many different insertion sequences of the key values using the same hash function and linear probing will result in the hash table shown above?

HW

(A) 10

(B) 20

(C) 30

(D) 40

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

Slide

2 mins Summary

Topic — Hashing

Topic — Collision

Topic — Chaining

Topic — Linear probing

Topic

Slide

THANK - YOU

Slide