Algorithms

Chapter 8: Hashing

GATE CS PYQ

Solved by Monalisa

Section 5: Algorithms

- Searching, sorting, hashing. Asymptotic worst case time and space complexity. Algorithm design techniques : greedy, dynamic programming and divide-and-conquer . Graph traversals, minimum spanning trees, shortest paths
- Chapter 1:Algorithim Analysis:-Algorithm intro , Order of growth ,Asymptotic notation, Time complexity, space complexity, Analysis of Recursive & non recursive program, Master theorem]
- Chapter 2:Brute Force:-Sequential search, Selection Sort and Bubble Sort , Radix sort, Depth first Search and Breadth First Search.
- Chapter 3: Decrease and Conquer :- Insertion Sort, Topological sort, Binary Search .
- Chapter 4: Divide and conquer:-Min max problem , matrix multiplication ,Merge sort ,Quick Sort , Binary Tree Traversals and Related Properties .
- Chapter 5: Transform and conquer:- Heaps and Heap sort, Balanced Search Trees.
- Chapter 6: Greedy Method:-knapsack problem , Job Assignment problem, Optimal merge, Hoffman Coding, minimum spanning trees, Dijkstra's Algorithm.
- Chapter 7: Dynamic Programming:-The Bellman-Ford algorithm ,Warshall's and Floyd's Algorithm ,Rod cutting, Matrix-chain multiplication ,Longest common subsequence ,Optimal binary search trees
- Chapter 8: Hashing.
- Reference : Introduction to Algorithms by Thomas H. Cormen
- Introduction to the Design and Analysis of Algorithms, by Anany Levitin
- My Note

- uniform hash functions
- *dynamic* hashing
- Chaining
- linear probing

Index	Key
0	87
1	11
2	13
3	36
4	92
5	4
6	71
7	14
8	
9	
10	43

- **GATE IT 2008 | Question: 48**
- Consider a hash table of size 11 that uses open addressing with linear probing. Let h(k)=k mod 11 be the hash function used. A sequence of records with keys 43, 36, 92, 87, 11, 4, 71, 13, 14 is inserted into an initially empty hash table, the bins of which are indexed from zero to ten. What is the index of the bin into which the last record is inserted?
- A.3 B.4 C.6 D.7
- h(43)=43 mod 11=10
- h(36)=36 mod 11=3
- h(92)=92 mod 11=4
- h(87)=87 mod 11=10
- h(11)=11 mod 11=0
- h(4)=4 mod 11=4
- h(71)=71 mod 11=5
- h(13)=13 mod 11=2
- h(14)=14 mod 11=3

- Ans : D.7

GATE CS 2009 | Question: 36

- The keys 12,18,13,2,3,23,5 and 15 are inserted into an initially empty hash table of length 10 using open addressing with hash function h(k)=k mod 10 and linear probing. What is the resultant hash table?

A.

0	
1	
2	2
3	23
4	
5	15
6	
7	
8	18
9	

B.

0	
1	
2	12
3	13
4	
5	5
6	
7	
8	18
9	

C.

0	
1	
2	12
3	13
4	2
5	3
6	23
7	5
8	18
9	15

D.

0	
1	
2	2, 12
3	13, 3, 23
4	
5	5, 15
6	
7	
8	18
9	

- h(12)=12 mod 10=2
- h(18)=18 mod 10=8
- h(13)=13 mod 10=3
- h(2)=2 mod 10=2
- h(3)=3 mod 10=3
- h(23)=23 mod 10=3
- h(5)=5 mod 10=5
- h(15)=15 mod 10=5
- Ans : C

- **GATE CS 2010 | Question: 52**
- A hash table of length 10 uses open addressing with hash function h(k)=k mod 10, and linear probing. After inserting 6 values into an empty hash table, the table is shown as 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
- h(42)=42 mod 10=2 ,h(23)=23 mod 10=3,h(34)=34 mod 10=4, h(46)=46 mod 10=6
- Above are in place
- h(52)=52 mod 10=2 but present at 5 means after 42,23 & 34.
- 46 hash value doesn't affect 52.
- h(33)=33 mod 10=3 but present at 7 so after filled up 2..6 hash value.
- Ans : C.46,34,42,23,52,33

GATE CS 2010 | Question: 53

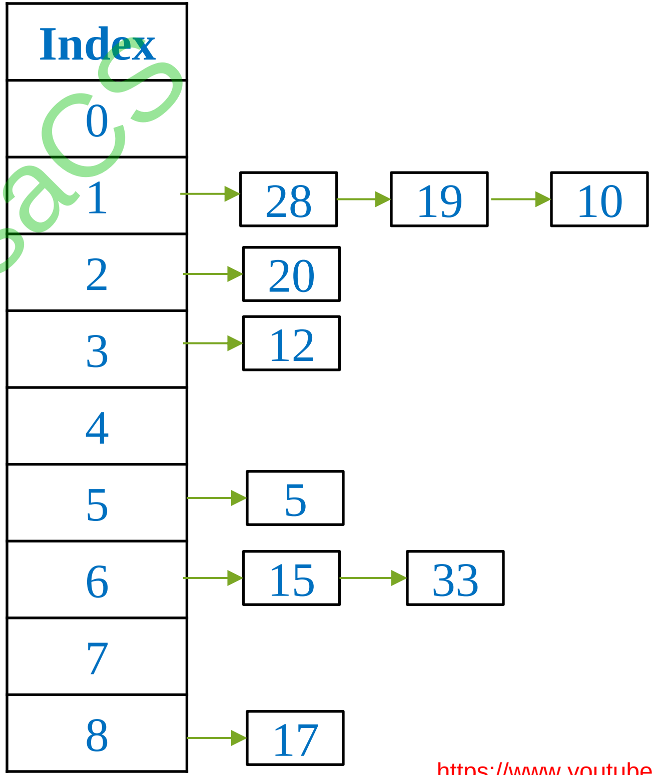
- A hash table of length 10 uses open addressing with hash function h(k)=kmod 10, and linear probing. After inserting 6 values into an empty hash table, the table is shown as below

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

- 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?
- A.10 B.20 C.30 D.40
- 33 must be after all ,so position fixed.
- For 52 two possibilities .Can come after 42,23,34,46 or just after 42 ,23,34.
- If 52 comes after 42,23,34,46
- These 4 can insert in any order 4!=24 ways.
- If 52 comes after 42,23,34
- These 3 can insert in any order 3!=6 ways.
- So total 24+6=30 insertion sequences.
- Ans : C.30

GATE CS 2014 Set 1 | Question: 40

- Consider a hash table with 9 slots. The hash function is h(k)=k mod 9. The collisions are resolved by chaining. 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
- h(5)=5 mod 9=5 h(28)=28 mod 9=1
- h(19)=19 mod 9=1 h(15)=15 mod 9=6
- h(20)=20 mod 9=2 h(33)=33 mod 9=6
- h(12)=12 mod 9=3 h(17)=17 mod 9=8
- h(10)=10 mod 9=1
- Maximum chain length=3
- Minimum chain length=0
- Average chain length=9/9=1
- Ans: (A)3,0, and 1



GATE CS 2014 Set 3 | Question: 40

- 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 3 insertions?
- A.$(97\times97\times97)/100^3$ B.$(99\times98\times97)/100^3$
- C.$(97\times96\times95)/100^3$ D.$(97\times96\times95)/(3!\times100^3)$
- Uniform hashing function is a hypothetical hashing function that evenly distributes items into the slots of a hash table. Moreover, each item to be hashed has an equal probability of being placed into a slot, regardless of the other elements already placed.
- We have 100 slots each of which are picked with equal probability by the hash function.
- So, to avoid first 3 slots, the hash function has to pick from the remaining 100 slots.
- And repetition is allowed, since chaining is used: meaning a list of elements are stored in a slot and not a single element.
- So ,required probability $=\frac{97}{100}\times\frac{97}{100}\times\frac{97}{100}$
- Ans : A.$(97\times97\times97)/100^3$

- **GATE CS 2015 Set 2 | Question: 33**
- 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. $h(i)=(11*i^2) \bmod 10$ D. $h(i)=(12*i^2) \bmod 10$

i	$h(i)=i^2 \bmod 10$	$h(i)=i^3 \bmod 10$	$h(i)=(11*i^2) \bmod 10$	$h(i)=(12*i^2) \bmod 10$
1	$1^2 \bmod 10=1$	$1^3 \bmod 10=1$	$(11*1^2) \bmod 10=1$	$(12*1^2) \bmod 10=2$
2	$2^2 \bmod 10=4$	$2^3 \bmod 10=8$	$(11*2^2) \bmod 10=4$	$(12*2^2) \bmod 10=8$
3	$3^2 \bmod 10=9$	$3^3 \bmod 10=7$	$(11*3^2) \bmod 10=9$	$(12*3^2) \bmod 10=8$
4	$4^2 \bmod 10=6$	$4^3 \bmod 10=4$	$(11*4^2) \bmod 10=6$	
5	$5^2 \bmod 10=5$	$5^3 \bmod 10=5$	$(11*5^2) \bmod 10=5$	
6	$6^2 \bmod 10=6$	$6^3 \bmod 10=6$	$(11*6^2) \bmod 10=6$	
7		$7^3 \bmod 10=3$		
8		$8^3 \bmod 10=2$		
9		$9^3 \bmod 10=9$		

- Ans: B. $h(i)=i^3 \bmod 10$

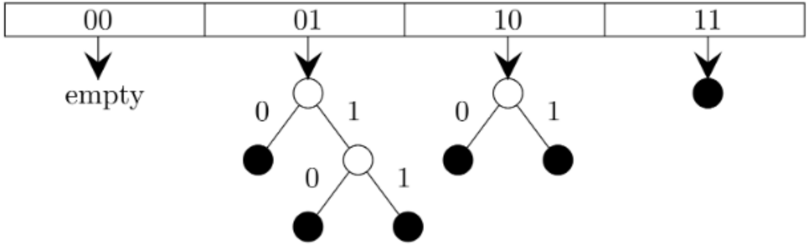
- **GATE CS 2015 Set 3 | Question: 17**
- Given that hash table *T* with 25 slots that stores 2000 elements, the load factor α for T is _______.
- load factor α =n/m
- n=2000,m=25
- α = 2000/25=80
- Ans :80

- **GATE CS 2020 | Question: 23**
- Consider a double hashing scheme in which the primary hash function is $h_1(k)=k \bmod 23$, and the secondary hash function is $h_2(k)=1+(k \bmod 19)$. Assume that the table size is 23. Then the address returned by probe 1 in the probe sequence (assume that the probe sequence begins at probe 0) for key value k=90 is_______________.
- $h(k,i)=(h_1(k)+ih_2(k)) \bmod m$
- $h_1(90)=90 \bmod 23=21$
- $h_2(90)=1+(90 \bmod 19)=15$
- $h(90,1)=(h_1(90)+1*h_2(90)) \bmod 23$
- $=(21+15) \bmod 23$
- $= 36 \bmod 23$
- 13
- Ans :13

- **GATE CS 2021 Set 1 | Question: 47**
- Consider a *dynamic* hashing approach for 4 bit integer keys:

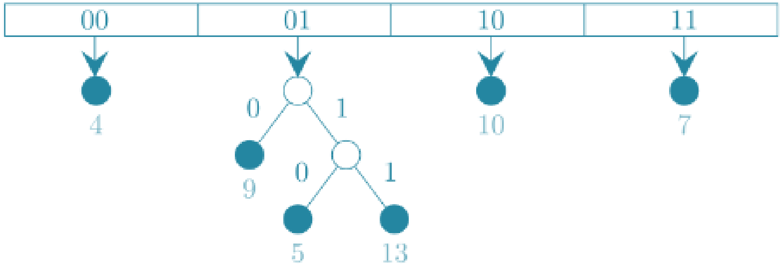
1. There is a main hash table of size 4.
2. The 2 least significant bits of a key is used to index into the main hash table.
3. Initially, the main hash table entries are empty.
4. Thereafter, when more keys are hashed into it, to resolve collisions, the set of all keys corresponding to a main hash table entry is organized as a binary tree that grows on demand.
5. First, the 3rd least significant bit is used to divide the keys into left and right subtrees.
6. To resolve more collisions, each node of the binary tree is further sub:divided into left and right subtrees based on the 4th least significant bit.
7. A split is done only if it is needed, i.e., only when there is a collision.

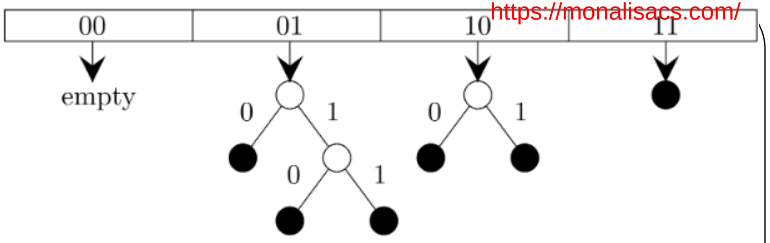
- Consider the following state of the hash table.



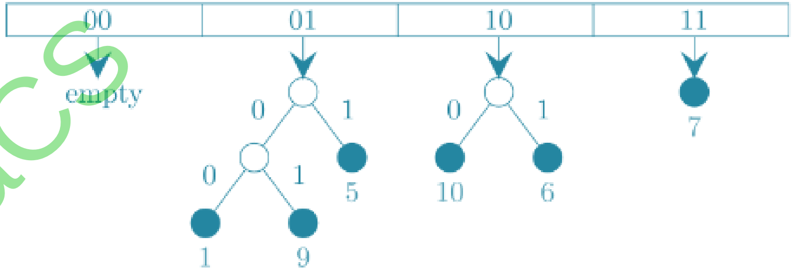
- Which of the following sequences of key insertions can cause the above state of the hash table (assume the keys are in decimal notation)?
- A.5,9,4,13,10,7 B.9,5,10,6,7,1
- C.10,9,6,7,5,13 D.9,5,13,6,10,14

- A)5:0101,9:1001,4:0100,13:1101,10:1010,7:0111

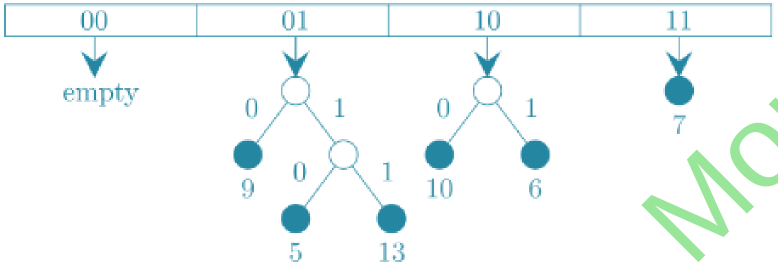




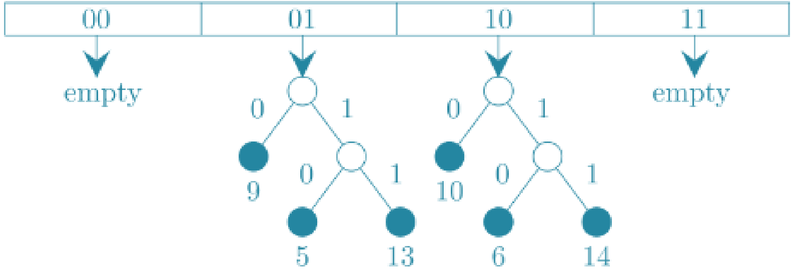
- B)9:1001,5:0101,10:1010,6:0110,7:0111,1:0001



- C)10:1010,9:1001,6:0110,7:0111,5:0101,13:1101



- D)9:1001,5:0101,13:1101,6:0110,10:1010,14:1110



- Ans : C.10,9,6,7,5,13

- **GATE CS 2022 | Question: 6**
- Suppose we are given *n* keys, *m* hash table slots, and two simple uniform hash functions h_1 and h_2. Further suppose our hashing scheme uses h_1 for the odd keys and h_2 for the even keys. What is the expected number of keys in a slot?
- A. $\frac{m}{n}$ B. $\frac{n}{m}$ C. $\frac{2n}{m}$ D. $\frac{n}{2m}$
- Let x odd keys and y even keys. $x+y=n$
- h_1 and h_2 hash values are in m.
- Both h_1,h_2 are Simple Uniform Hash Functions, So, using any of them, for any key k, each slot is equally likely.
- Hence, the expected number of keys in a slot is n/m.
- Ans : B. $\frac{n}{m}$

- **GATE CS 2023 | Question: 10**
- An algorithm has to store several keys generated by an adversary in a hash table. The adversary is malicious who tries to maximize the number of collisions. Let k be the number of keys, m be the number of slots in the hash table, and k >m.
- Which one of the following is the best hashing strategy to counteract the adversary?
- (a) Division method, i.e., use the hash function h(k) = k mod m.
- (b) Multiplication method, i.e., use the hash function h(k) =$\lfloor$m (kA – $\lfloor kA \rfloor$)$\rfloor$, where A is a carefully chosen constant.
- (c) Universal hashing method.
- (d) If k is a prime number, use Division method . Otherwise, use Multiplication method.
- There is a chance that all the keys are hashed in same slot, yielding an average retrieval time θ(n) .This is the worst case behavior.
- The only effective way to improve the situation is to choose the hash function randomly in such a way that is independent of the key.
- This approach is called universal hashing.
- Ans :(C)

- **GATE DA 2024 | Question: 6**
- Match the items in **Column 1** with the items in **Column 2** in the following table:

Column 1	**Column 2**
(p) First In First Out (q) Lookup Operation (r) Last In First Out	(i) Stacks (ii) Queues (iii) Hash Tables

- (A) (p) - (ii), (q) - (iii), (r) - (i)
- (B) (p) - (ii), (q) - (i), (r) - (iii)
- (C) (p) - (i), (q) - (ii), (r) - (iii)
- (D) (p) - (i), (q) - (iii), (r) - (ii)
- Ans: (A) (p) - (ii), (q) - (iii), (r) - (i)

- **GATE DA 2024 | Question: 11**
- Consider performing uniform hashing on an open address hash table with load factor $\alpha = \frac{n}{m} < 1$, where n elements are stored in the table with m slots. The expected number of probes in an unsuccessful search is at most $\frac{1}{1-\alpha}$.Inserting an element in this hash table requires at most _______ probes, on average.
- (A) $\ln(\frac{1}{1-\alpha})$ (B) $\frac{1}{1-\alpha}$ (C) $1+\frac{\alpha}{2}$ (D) $\frac{1}{1+\alpha}$
- Inserting an element in this hash table requires at most $\frac{1}{1-\alpha}$ probes, on average.
- Its till unsuccessful
- Ans: (B) $\frac{1}{1-\alpha}$