Advanced Algorithms & Data Structures

AIM OF THE SESSION

To familiarize students with the basic concept of Collision Resolution Techniques

INSTRUCTIONAL OBJECTIVES

This Session is designed to:

1. Demonstrate Collision resolution techniques
2. Describe the types of collision resolution techniques
3. Describe each method

.

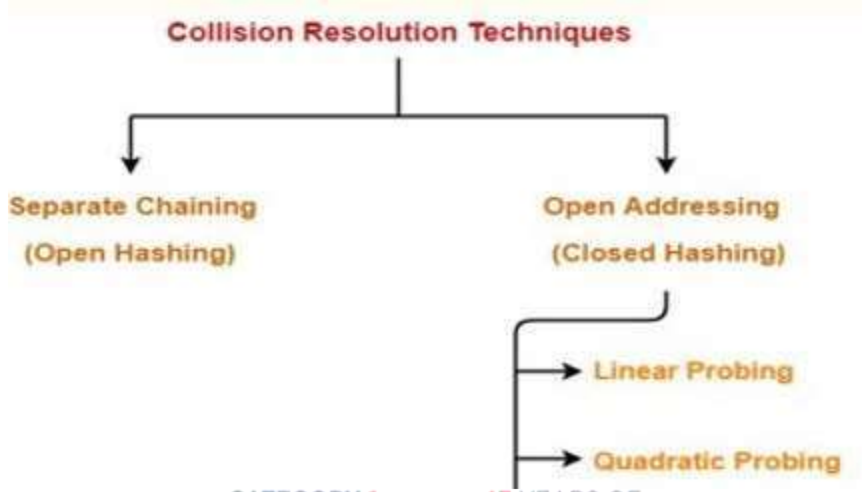
LEARNING OUTCOMES

At the end of this session, you should be able to:

1. Define Collision resolution techniques
2. Describe types of collision resolution techniques.
3. Summarize definition, types of resolution techniques and its applications

Introduction To Collision Resolution Techniques

- Collision occurs when multiple keys are mapped to the same index (slot) in a hash table by the hash function.
- For example, several students having the same locker number in a school, leading to potential conflicts for storage and retrieval.
- The collision creates a problem because each index in a hash table is supposed to store only one value.
- There are several collision resolution techniques to manage the performance of a hash table. The two most popular techniques are:
  1. Open Hashing
     - Separate Chaining
  2. Closed Hashing or Open Addressing
     - Linear Probing
     - Quadratic Probing



Introduction To Collision Resolution Techniques

Open Hashing

- Each Bucket in the Hash table is the head of a Linked List.
- All Elements that hash to a Particular Bucket are Placed on the Buckets Linked List .

Closed Hashing

- Ensures that all elements are stored directly into the Hash Table.

Separate Chaining

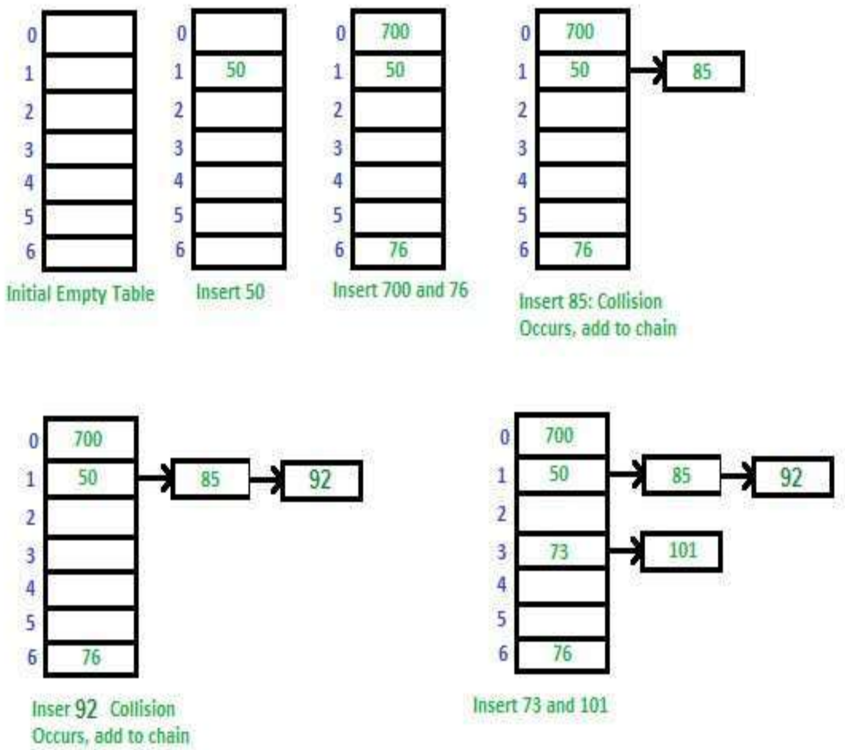
- It is a collision resolution technique that handles collisions in hash tables by storing colliding keys in linked lists.
- Instead of having a single value at each slot, each slot holds a linked list (or chain) of keys that have hashed to the same index.
- The hash function for the Separate Chaining is:

  h(key) = key% table_size.

Example: Consider a sequence of keys as 50, 700, 76, 85, 92, 73, 101 with table_size 7.

h(key) = key% table_size.

h(50) = 50%7 = 1	h(92) = 92%7 = 1
h(700) = 700%7 = 0	h(73) = 73%7 = 3
h(76) = 76%7 = 6	h(101) = 101%7 = 3
h(85) = 85%7 = 1	



Linear Probing

- It's an open addressing collision resolution technique.
- When a collision occurs, it probes sequentially for the next available slot in the hash table.
- The hash function for the Linear Probing to an initial index:

  h(key) = Key% Table_size.
- If a collision occurs at the any index, linearly probe for the next available slot:

  probe_index(key) = (key + i) % Table_size

  Where, i is the probe count, starting from 1 and incrementing for each failed attempt.

Working Mechanism

- **Calculate the hash value:** Use a hash function to determine the initial slot for a key.
- **Check for collision:** If the slot is occupied, start linear probing:
  - Sequentially check the next available slot in the table.
  - Wrap around to the beginning if you reach the end.
- **Insert or search:** Once an empty slot is found, insert the key or search for it.

Linear Probing

Example: Consider a sequence of keys as 71, 64, 56, 36 with table size 10

- The hash function for the Division method is:

  h(key) = Key% table size.

  h(71)=71%10=1

  h(64)=64%10=4

  h(56)=56%10=6

  h(36)=36%10=6

Now, collision occurs at the index 6

probe_index(key) = (key + i) % Table_size

Initially i=1

probe_index(36) =(36+1)%10=7

Now, key 36 is placed at index 7

0	NULL
1	71
2	NULL
3	NULL
4	64
5	NULL
6	56
7	36
8	NULL
9	NULL

Quadratic Probing

- It's a collision resolution technique used in hash tables that employ open addressing.
- When a collision occurs (two keys hash to the same index), quadratic probing probes a sequence of alternate hash table indices using a quadratic function.
- The hash function for the Quadratic Probing to an initial index:

  h(key) = Key% Table_size.

- If a collision occurs at the any index, linearly probe for the next available slot:

  Quaprobe_index(key) = (key + i*i) % Table_size

  Where, i is the probe count, starting from 1 and incrementing for each failed attempt.

Quadratic Probing

Example: Consider a sequence of keys as 67, 90,55,49,17 with table size 10

The hash function for the Division method is:

h(key) = Key% table size.

h(67) = 67%10 = 7

h(90) = 90%10 = 0

h(55) = 55%10 = 5

h(49) = 49%10 = 9

h(17) = 17%10 = 7

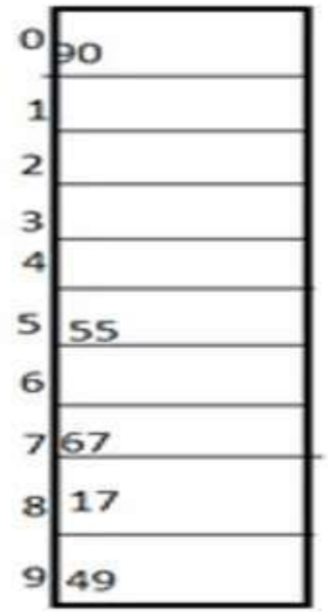
Now, collision occurs at the index 7

Quaprobe_index(key) = (key + i*i) % Table_size

Initially i=1

probe_index(17) =(17+1*1)%10=8

Now, key 17 is placed at index 8



Advantages of a Collision Resolution Techniques

- Enable efficient hash table operations (insertion, deletion, search) in average cases, often achieving O(1) time complexity.
- Allow for flexible hash table usage, even when collisions occur.

Disadvantages

- Can degrade performance if the hash table becomes too full, potentially leading to worst-case O(n) time complexity.
- Require careful selection of a good hash function to minimize collisions and ensure even key distribution.

Applications

- Symbol tables
- Caching

EXAMPLES

1. Let us consider a simple hash function as “key mod 5” and a sequence of keys that are to be inserted are 50, 70, 76, 93. Draw hash table using Linear probing technique.

2. Let us consider table Size = 7, hash function as Hash(x) = x % 7 and collision resolution strategy to be f(i) = i2 . Insert = 22, 30, and 50.

1.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 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

2

2. Which of the following schemes does quadratic probing come under?

a) rehashing
b) extended hashing
c) separate chaining
d) open addressing

SUMMARY

- Collision resolution techniques are crucial in hash tables to effectively deal with collisions occurring when different keys map to the same index.

- Each technique has its own advantages and disadvantages, making the choice depend on specific needs and priorities.

TERMINAL QUESTIONS

1. List out different types of Collision Resolution Techniques?
2. 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 show the resultant hash table?
3. Define collision

Reference Books:

1. Mark Allen Weiss, Data Structures and Algorithm Analysis in C, 2010 , Second Edition, PearsonEducation.
2. 2. Ellis Horowitz, Fundamentals of Data Structures in C: Second Edition, 2015
3. A.V.Aho, J. E. Hopcroft, and J. D. Ullman, “Data Structures And Algorithms”, Pearson Education, First Edition Reprint2003.

Sites and Web links:

1. https://nptel.ac.in/courses/106102064
2. https://in.udacity.com/course/intro-to-algorithms--cs215
3. https://www.coursera.org/learn/data-structures?action=enroll

THANK YOU