ECAP770

ADVANCE DATA STRUCTURES

Ashwani Kumar

Assistant Professor

Learning Outcomes

After this lecture, you will be able to

- understand basics of hashing,

- hash table.

Hashing

- It is the process of converting an input of any length into a fixed size string or a number using an algorithm.

- In hashing, a hash function is used that converts a given key to a smaller number and uses the small number as an index in a table called a hash table.

Hashing

- Hashing data structure is used to store and find data efficiently using array.

- It is a technique or process of mapping keys, values into the hash table by using a hash function. It is done for faster access to elements.

- The efficiency of mapping depends on the efficiency of the hash function used.

Hashing

- Hashing is a technique that is used to uniquely identify a specific object from a group of similar objects

- In colleges, each student is assigned a unique registration number that can be used to retrieve information about student.

- In libraries, each book is assigned a unique number that can be used to determine information about the book, such as its exact position in the library or the users it has been issued to etc.

Time complexity

- Time complexity in linear search is O(n)

40	10	30	60	70	20	90	80	50

- Time complexity in binary search is O(log n)

10	20	30	40	50	60	70	80	90

- Time complexity in hashing is O(1)

Hashing

- It is used to better searching method as compared to the linear or binary search.

- Hashing allows to update and retrieve any data entry in a constant time $O(1)$.

- Constant time $O(1)$ means the operation does not depend on the size of the data.

Steps to implement hashing

- An element is converted into an integer by using a hash function. This element can be used as an index to store the original element, which falls into the hash table.

- The element is stored in the hash table where it can be quickly retrieved using hashed key.

Steps to implement hashing

hash = hash func(key)

index = hash % array_size

- The hash is independent of the array size and it is then reduced to an index (a number between 0 and array_size − 1) by using the modulo operator (%).
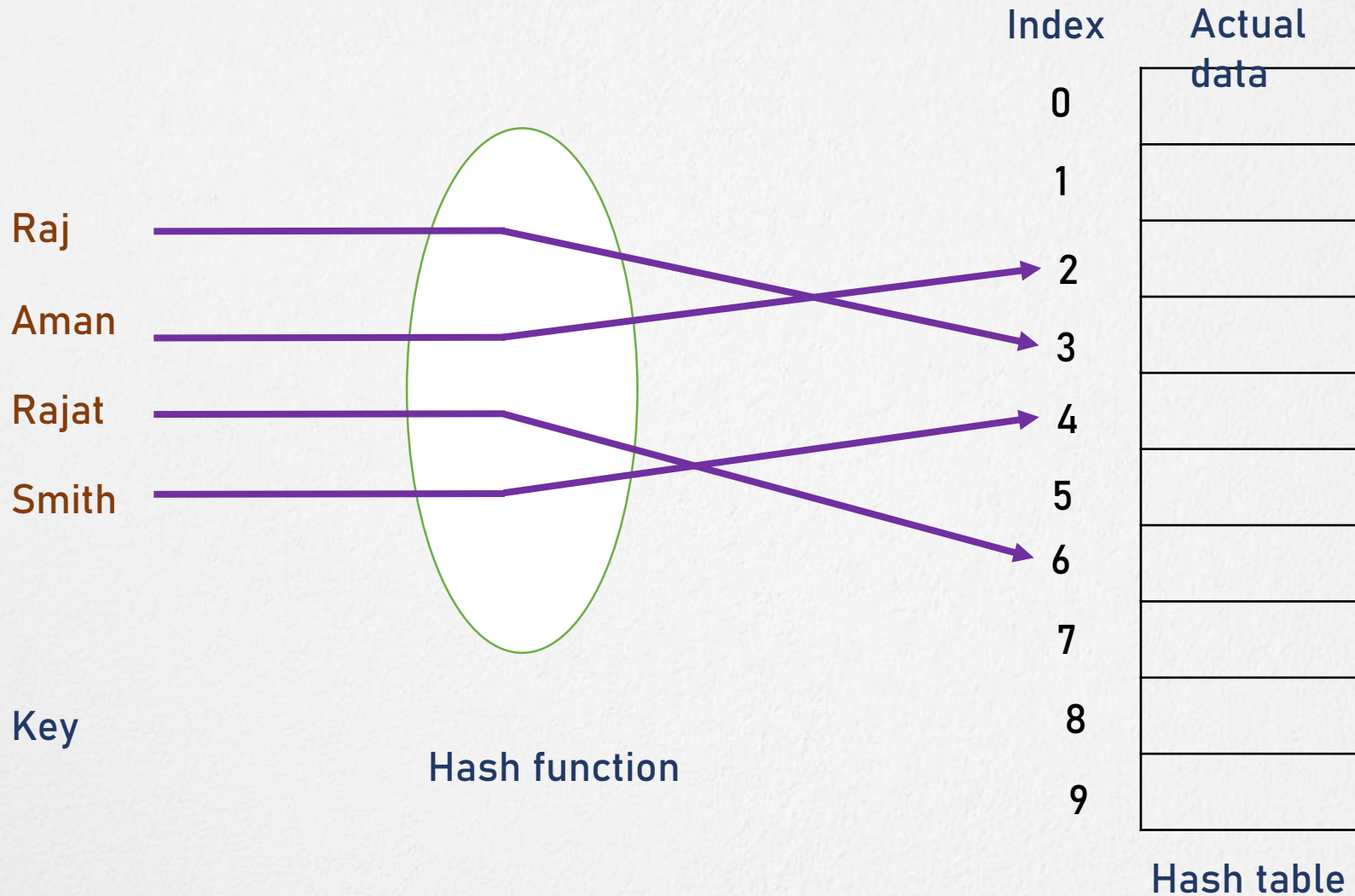
Hash table

- A Hash table is a data structure that stores information, and the information has basically two components.

  – key and value.

- The hash table can be implemented with the help of an associative array

Hash table

- It uses a hash function to compute an index into an array of buckets or slots from which the desired value can be found.

- It is an array of list where each list is known as bucket.

- It contains value based on the key.

- Hash table is synchronized and contains only unique elements.

Hash table

Index | Actual data

0

1

Raj

Aman

Rajat

Smith

2

3

4

5

6

7

8

Key

9

Hash function

Hash table

Hash function

- A hash function is any function that can be used to map a data set of an arbitrary size to a data set of a fixed size, which falls into the hash table.

- The values returned by a hash function are called hash values, hash codes, hash sums, or hashes

Terminology: Hash table

- Data bucket – Data buckets are memory locations where the records are stored. It is also known as unit of storage.

- Hash index – It is an address of the data block. A hash function could be a simple mathematical function to even a complex mathematical function.

Terminology: Hash table

- Linear Probing – Linear probing is a fixed interval between probes. In this method, the next available data block is used to enter the new record, instead of overwriting on the older record.

- Double Hashing –Double hashing is a computer programming method used in hash tables to resolve the issues of has a collision.

Terminology: Hash table

- Quadratic probing– It helps you to determine the new bucket address. It helps you to add Interval between probes by adding the consecutive output of quadratic polynomial to starting value given by the original computation.

Hash function

- Methods for calculating the hash function

  - Division method

  - Folding method

  - Mid square method

  - Multiplication Method

Array

- The array holds all the key-value entries in the table.

- The size of the array should be set according to the amount of data expected.

Operations of a hash table

- Search – Searches an element in a hash table.

- Insert – inserts an element in a hash table.

- delete – Deletes an element from a hash table.

That's all for now…