



Sunbeam Institute of Information Technology
Pune and Karad
PG - DESD

Module – Data Structures

Trainer - Devendra Dhande

Email – devendra.dhande@sunbeaminfo.com



Sunbeam Infotech

www.sunbeaminfo.com

Hash Table

- The implementation of hash tables is frequently called as **Hashing**.
- Hashing is a technique used for performing insertions, deletions and finds in constant average time.
- The ideal hash table is an **array of some fixed size**, containing the **keys**, where each key is a string with an associated value.
- Each Key is mapped into some number in the range 0 to TableSize-1 and placed in the appropriate cell (slot).
- The mapping of keys with its corresponding cell is called a **hash function**.

0	
1	
2	
3	John 25000
4	Phil 31500
5	
6	Dave 27500
7	Mary 28000
8	
9	

In this example, john hashes to 3,
 phil hashed to 4, dave hashes to 6
 and marry hashes to 7.



Sunbeam Infotech

www.sunbeaminfo.com

Hash Table

- **Hash function**

- The mapping is called a hash function.
- It is mathematical function of the key that yields slot of the hash table where key-value is stored.
- Ideally it should be simple to compute and should ensure that any two distinct keys get different cells.
- Since there are a finite number of cells and infinite supply of keys, this is clearly impossible and thus we seek a hash function that distributes the keys evenly among the cells.
- Simplest example is: $f(k) = k \% \text{size}$.

- **Collision:**

- There is possibility that two keys hash to the same value(cell). This is called collision.
- Must be handled using one of the **collision handling technique**.
 - **Open Addressing**
 1. Linear Probing
 2. Quadratic Probing
 3. Double Hashing
 - **Closed Addressing**
 1. Chaining / Separate chaining



Sunbeam Infotech

www.sunbeaminfo.com

Hash Table

- **Open Addressing:**

- All key-value pairs are stored in the hash table itself.
- If key (to find) is not matching with the key in the slot calculated by hash function, it is probed in next possible slot using one of the following.
- **Linear Probing:** In linear probing, if collision occurs next free slot will be searched/probed linearly.
- **Quadratic Probing:** In quadratic probing, if collision occurs next free slot will be searched/probed quadratically.
- **Double Hashing:** In double hashing, if collision occurs next free slot will be searched/probed by using another hash function, so two hash functions can be used to find next/probe next free slot.



Sunbeam Infotech

www.sunbeaminfo.com

Hash Table

- **Load Factor = n / m**
 - n = Number of key-value pairs to be inserted in the hash table
 - m = Number of slots in the hash table
 - If $n < m$, then load factor < 1
 - If $n = m$, then load factor $= 1$
 - If $n > m$, then load factor > 1
- **Limitations of Open Addressing**
 - Open addressing requires more computation.
 - Cannot be used if load factor is greater than 1 (i.e. number of pairs are more than number of slots in the table).
- **Chaining/Separate Chaining:**
 - Another collision handling technique.
 - Each slot of hash table holds a collection of key-values for which hash value of keys are same.
 - This collection in each slot is also referred as bucket.
 - Chaining is simple to implement, but requires additional memory outside the table.



Sunbeam Infotech

www.sunbeaminfo.com

Thank you!

Devendra Dhande

<devendra.dhande@sunbeaminfo.com>



Sunbeam Infotech

www.sunbeaminfo.com