

Hash Function

Hash function is a function that maps any big number or string to a small integer value.

- Hash function takes the data item as an input and returns a small integer value as an output.
- The small integer value is called as a hash value.
- Hash value of the data item is then used as an index for storing it into the hash table.

Properties of Hash Function-

The properties of a good hash function are-

- It is efficiently computable.
- It minimizes the number of collisions.
- It distributes the keys uniformly over the table.

Types of Hash Function

1. Division Method:

Choose a number m smaller than the number of n of keys in k (The number m is usually chosen to be a prime number or a number without small divisors, since this frequently a minimum number of collisions).

The hash function is:

$$h(k) = k \bmod m$$

$$h(k) = k \bmod m + 1$$

For Example: if the hash table has size m = 12 and the key is k = 100, then $h(k) = 4$. Since it requires only a single division operation, hashing by division is quite fast.

2. Mid Square Method:

The key k is squared. Then function H is defined by

1. $H(k) = L$

Where L is obtained by deleting digits from both ends of k^2 . We emphasize that the same position of k^2 must be used for all of the keys.

3. Folding Method:

The key k is partitioned into a number of parts k_1, k_2, \dots, k_n where each part except possibly the last, has the same number of digits as the required address.

Then the parts are added together, ignoring the last carry.

$$H(k) = k^1 + k^2 + \dots + k^n$$

Example: Company has 68 employees, and each is assigned a unique four-digit employee number. Suppose L consist of 2-digit addresses: 00, 01, and 02....99. We apply the above hash functions to each of the following employee numbers:

1. 3205, 7148, 2345

(a) Division Method: Choose a Prime number m close to 99, such as $m = 97$, Then

1. $H(3205) = 4, H(7148) = 67, H(2345) = 17.$

That is dividing 3205 by 17 gives a remainder of 4, dividing 7148 by 97 gives a remainder of 67, dividing 2345 by 97 gives a remainder of 17.

(b) Mid-Square Method:

$$k = 3205 \quad 7148 \quad 2345$$

$$k^2 = 10272025 \quad 51093904 \quad 5499025$$

$$h(k) = 72 \quad 93 \quad 99$$

Observe that fourth & fifth digits, counting from right are chosen for hash address.

(c) Folding Method: Divide the key k into 2 parts and adding yields the following hash address:

1. $H(3205) = 32 + 50 = 82$ $H(7148) = 71 + 84 = 55$
2. $H(2345) = 23 + 45 = 68$