Course - CSE 208

Data Structure and Algorithms II - Sessional

Name - Anik Saha
Student ID - 2005001
Report Topic: Hashing Offline

Date: 18/08/23

Hash Function 1 : Polynomial Rolling Hash

Parameters: p = 47

```
int get_polynomial_hash(string s, int p, int m)
{
    int n = s.length();
    int hash_value = 0;
    for(int i=0;i<n;i++)
    {
        char ch = s[i];
        hash_value = (hash_value*p + (ch - 'a'))%m;
    }
    return hash_value;
}
```

Hash Function 2 : Bitwise Interleaved Hash

Parameters: p = 1000000007

```
int get_bitwise_interleaved_hash(string s, int p, int m)
{
    int n = s.length();
    int hash_value = p;
    for(int i=0;i<n;i++)
    {
        char ch = s[i];
        hash_value = (hash_value<<5)^(hash_value>>27)^ch;
    }
    while(hash_value<0)  // make negative values positive
    {
        hash_value = (hash_value + m)%m;
    }
    return hash_value%m;
}
```

Auxiliary Hash Function :

Parameters: prime_2 = 71, prime_3 = 1000000007

```
int get_aux_hash(string s) // O(1) // Can never be zero // prime_1 must be less than table_size
{
    int n = s.length();
    int salt = prime_3*(s[0] - 'a') + s[n-1] - 'a';// getting a big number from first and last two
    return prime_2 - ( salt % prime_2 );   // ensuring that aux_hash < m
}
```

Constants for Custom Probing:

c1 = 37
c2 = 43

Hash Function Performance Report:

Count of Unique Words = 10000

Load Factor: 0.50

Uniqueness Percentage for Hash Function 1: **78.47 %**
Uniqueness Percentage for Hash Function 2: **78.36 %**

Load Factor: 0.70

Uniqueness Percentage for Hash Function 1: **71.87 %**
Uniqueness Percentage for Hash Function 2: **71.44 %**

Load Factor: 0.90

Uniqueness Percentage for Hash Function 1: **65.81 %**
Uniqueness Percentage for Hash Function 2: **66.28 %**

N’ = 5000

HashTable Performance Report:

Hash Table Size	Collision Resolution Method	Hash1		Hash2	
		# of Collisions	Average Probes	# of Collisions	Average Probes
5000	Chaining	9915	2.4810	10003	2.4780
	Double Hashing	40886	7.8340	35009	7.9230
	Custom Probing	45770	8.0100	55570	10.8490

N’ = 10000

HashTable Performance Report:

Hash Table Size	Collision Resolution Method	Hash1		Hash2	
		# of Collisions	Average Probes	# of Collisions	Average Probes
10000	Chaining	4974	1.4840	5039	1.4840
	Double Hashing	64500	7.3070	62337	7.9220
	Custom Probing	61573	7.3930	57393	7.2590

N’ = 20000

HashTable Performance Report:

Hash Table Size	Collision Resolution Method	Hash1		Hash2	
		# of Collisions	Average Probes	# of Collisions	Average Probes
20000	Chaining	2484	1.2030	2490	1.2440
	Double Hashing	3911	1.3930	3924	1.4180
	Custom Probing	3843	1.3730	3807	1.4050