Imperial College London – Department of Computing

MSc in Computing Science

580: Algorithms
Tutorial: Hash Tables

1. An open address hash table T has $m = 12$ slots and uses the hash function $h(k) = k \bmod m$. Assuming collisions are resolved using linear probing, draw the table after inserting the following keys, in this order: 82, 7, 47, 17, 49, 150, 34, 61, 107, 6.

Answer:

	0	1	2	3	4	5	6	7	8	9	10	11
T	34	49	61	107		17	150	7	6		82	47

2. A hash table T has a constant load factor, uses a hash function h and the chaining method of collision resolution. Assume the following non-uniform hashing: the probability of a key k hashing to $h(k) = 1$ is $1/2$; the probabilities of k hashing to any other slot are all equal. What is the expected time complexity for an unsuccessful search if T contains N objects?

Answer: If T currently has m slots, the probability that an object x is in the chain $T[i]$ is

$$P\{i\} = \begin{cases} 0.5 & \text{when } i = 1 \\ 0.5/(m-1) & \text{otherwise} \end{cases}$$

So, the expected length of the chain at $T[i]$ is

$$l[i] = \begin{cases} N/2 & \text{when } i = 1 \\ N/2(m-1) & \text{otherwise} \end{cases}$$

The probability that the search key k will hash to i is the same as the probability that x will be found in $T[i]$. So, the expected number of keys that k will be compared to is:

$$\frac{N}{4} + \sum_{i=2}^{m} \frac{N}{4(m-1)^2} = \frac{N}{4} + \frac{N}{4(m-1)}$$

Since T has a constant load factor, $N/4(m-1)$ is also constant. So, the time complexity of an unsuccessful search is

$$\frac{N}{4} + \Theta(1) = \Theta(N).$$