

Tutorial 07 - Tries & Hashing
CS 240E Winter 2022
University of Waterloo
Monday, February 28th, 2022

1. **Prefix Search:**

Let w_1, \dots, w_k be a set of words, where $n = |w_1| + \dots + |w_k|$. Find, in $O(n)$ time, the longest word w such that w is the prefix of at least two words in $w_1 \dots w_k$.

2. **Universal Hash Functions:**

Recall that a family \mathcal{H} of hash-functions is called *universal* if $P(h(k) = h(k')) \leq \frac{1}{M}$ for all keys $k \neq k'$, and it is said to have *uniform hash-values* if $P(h(k) = i) = \frac{1}{M}$ for all keys k and all slots i . Here ‘probability’ is taken over the random uniform choice of h among \mathcal{H} .

(a) Consider the family \mathcal{H} on slide 4 of module07e:

$$U = \mathbb{Z}_5, M = 2$$

$$h_b(k) = ((k + b) \bmod 5) \bmod 2$$

$$\mathcal{H} = \{h_b : b \in \mathbb{Z}_5\}$$

Choose $b \in \mathbb{Z}_5$ randomly to get hash-function.

Does this have uniform hash-values? Is this universal?

- (b) Assume that \mathcal{H} has uniform hash-values. Prove or disprove: The expected time for an unsuccessful search in hashing with chaining is $O(\alpha)$.
- (c) Assume that \mathcal{H} has uniform hash-values. Prove or disprove: The expected time for a successful search in hashing with chaining is $O(1 + \alpha)$.

3. Ordered Hashing:

For this question we will look at a modified version of open addressing such that, while searching in the hash table, the sequence of keys encountered during the search has some ordering property (though the sequence is not necessarily in sorted order). To achieve this property, we use a modified insert routine: when inserting key k , follow its probe sequence, and if you ever encounter a key k' with $k' > k$, then swap to put k into this position, and proceed to insert k' .

- (a) Carry out this idea for linear probing using $M = 10$, the hash function $h(k, i) = (k + i) \bmod 10$, and the insertion sequence,

31, 26, 16, 23, 11, 30, 20

- .
- (b) Argue that for searching for any key, the keys encountered before it in its probe sequence are all smaller than it. You may assume that no items will ever be deleted from the hash table.
- (c) Give an improved search method for ordered hashing.