CS 240E: Structures and Data Management **Winter 2021**

Tutorial 6: Tries, Hashing

Warmup. Draw the compressed trie containing the following keys: 10, 101, 1001, 10010, 10011, 1110, 1111, 11100, 111110, 111101.

1. Consider sorting the following base-4-numbers: 300, 211, 112, 230, 1, 0, 12, 101, 233, 110.

a) Illustrate how you would sort them with MSD-radix sort, by drawing the recursion tree and the subarray in each recursion.

b) Show the corresponding 4-way pruned trie.

c) Show that the expected time to insert a base-4-number into a 4-way pruned trie is less than $\log_4 n + O(1)$, assuming all numbers have been uniformly chosen. You may assume the numbers have been padded with 0s so that all numbers begin with the same place value.

2. Consider a hash table of size 7. For each of the scenarios below, insert the keys $14, 10, 20, 13, 7, 17$, then delete 14 and search for 13.

a) Linear Probing with $h(k) = k \bmod 7$.

b) Double Hashing with $h_1(k) = k \bmod 7$ and $h_2(k) = (k \bmod 5) + 1$.

3. Suppose that we use double hashing to resolve collisions, i.e., we use the hash function $h(k, i) = (h_1(k) + ih_2(k)) \bmod m$. Show that if m and $h_2(k)$ have greatest common divisor $d \geq 1$ for some key k, then an unsuccessful insertion for key k examines $\frac{1}{d}^{th}$ of the hash table before returning to slot $h_1(k)$. Thus, when $d = 1$, i.e., m and $h_2(k)$ are relatively prime, then the insertion of k can only fail if every entry of the hash table is occupied.