Lecture 18 Sept 25 CSL 630

Viterbi's expectation maximization

Given an edge labelled directed graph $G=(V,E)$ with weights, and a starting vertex v_0 and a string $\sigma_1 \sigma_2 \dots \sigma_n$ over the labels, we want to find the most profitable path starting from v_0

$$v_0 \xrightarrow{\sigma_1} v_{i_1} \xrightarrow{\sigma_2} v_{i_2} \cdots \xrightarrow{\sigma_n} v_{i_n}$$

length 1 path is easy

For length > 1, we can write an inductive defn

—— end of DP ——

Searching

Binary search : static
(set of elements is fixed)

Balanced search trees : dynamic dictionary

n elements $O(\log n)$ search
update (insert / delete)

Comparison search requires total ordering

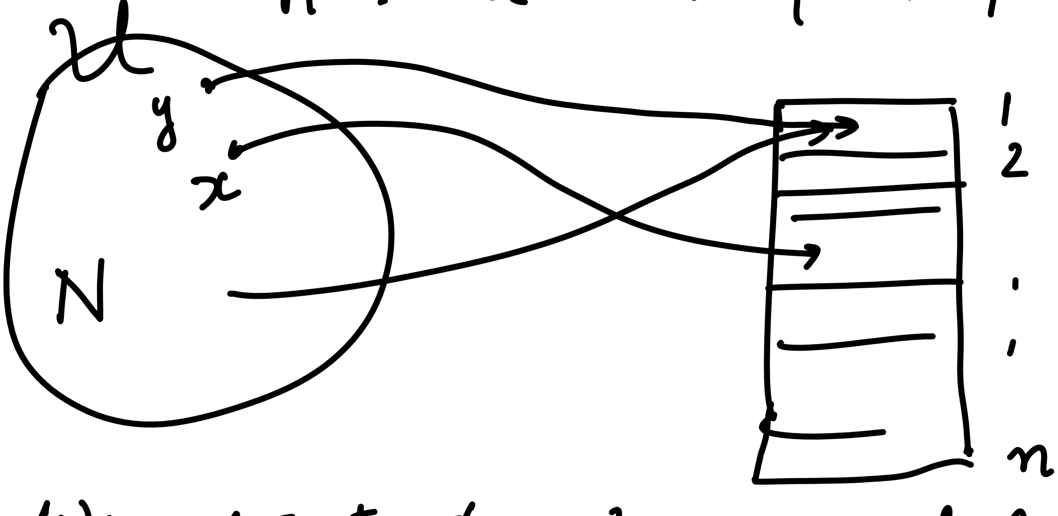
$O(n)$ space

$\mathcal{U}$: set of all possible elements

Hashing $x \in \mathcal{U}$ Hash table T

hash function $h(x)$

$h : x \to \{1, 2, \ldots n\}$



set S

$|S| \leq n$

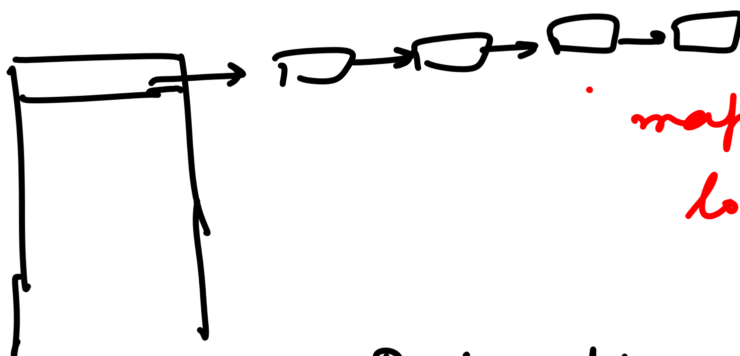
$S \subset \mathcal{U}$

We want to store a set of n elements in T $n \ll N$

Conflict resolution techniques

1. Open addressing: find the "next" available location

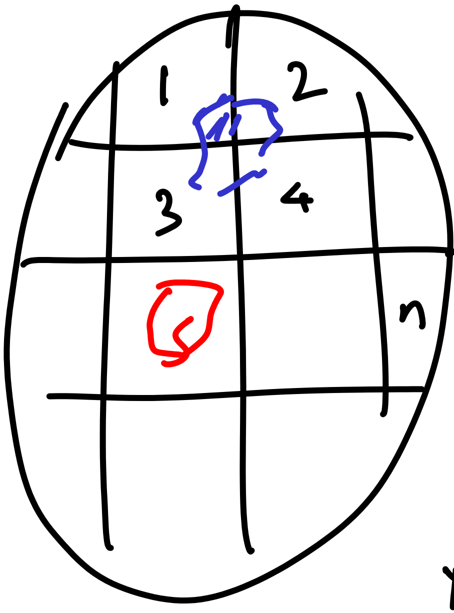
2. Chaining: maintain a linked list with the location

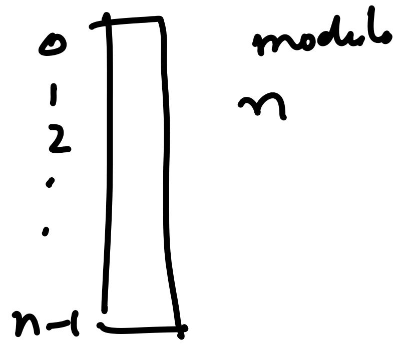


different elements mapped to the same location

Both the conflict resolution schemes could degenerate to linear search.

Some location of the table will be the range of at least N/m elements





Most popular hash function including modulo n equipartitions

$\mathcal{U}$ $\quad |h^{-1}(1)| = |h^{-1}(2)| = \cdots |h^{-1}(n)|$

Suppose S is chosen randomly (uniformly) from $\mathcal{U}$

What is the expected length of chain in location 1 $= \frac{1}{n} \times n = 1$

(for equipartition)

hw Show that all chains are $\leq O(\log n)$ with prob $1 - \frac{1}{n}$