Dyn Prog 2

Tuesday, March 10, 2020 3:29 PM

Edit Distance

- Distance between two strings
- # number of substitutions, insertions, deletions

3 ⇝ CAT / DOG

CAT
→ DAT
→ DRAT
→ DRT
→ DOT
→ DOG

ALGORITHM
ALTRUISTIC

ALGOR-I-THM
ALT-RUISTIC

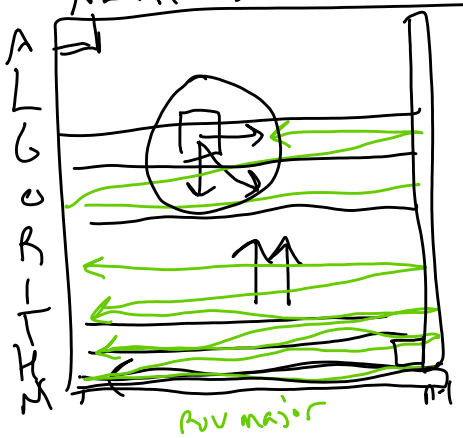
$$Dist(X,Y) = \begin{cases} 0 & \text{if } X=Y \\ |Y| & \text{if } |X|=0 \\ |X| & \text{if } |Y|=0 \end{cases}$$

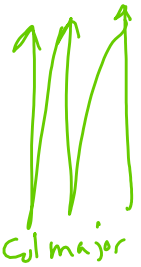
$$\min \begin{cases} 1+Dist(X',Y) & \text{// rem} \quad \text{if } X=aX' \\ 1+Dist(X,Y') & \text{// insert} \quad Y=bY' \\ 1+Dist(X',Y') & \text{// subst } a\to b \quad \text{if } a\neq b \\ 0+Dist(X',Y') & \text{if } a=b \end{cases}$$

X, Y global

```
Dist(i, j):                    // dist before X[i:], Y[j:]
    ~~0 if X=Y~~
    - |Y|-j  if i=|X|
    - |X|-i  if j=|Y|
    otherwise
        min { 1 + Dist(i+1, j)
              1 + Dist(i, j+1)
              (X[i] ≠ Y[j]) + Dist(i+1, j+1)
    j                          // OPT[i,j] = dist(i,j)
```

ALTRUISTIC



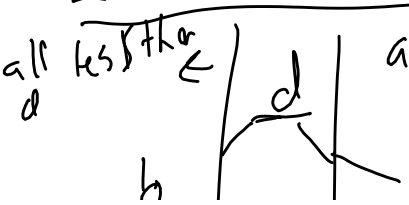



```
Dist(X[0...m-1], Y[0...n-1]): // OPT: (m+1)x(n+1)
    For i in 0...m-1:
        OPT[i, n] := m-i
    For j in 0...n-1:
        OPT[m, j] := n-j
    - For i = m-1 down to 0:
        - For j = n-1 down to 0:
            ins := 1 + OPT[i, j+1]
            del := 1 + OPT[i+1, j]
            sub := (X[i] ≠ Y[j]) + OPT[i+1, j+1]
            OPT[i, j] := min(ins, del, sub)
    return OPT[0, 0]
```

Time: $O(mn)$ Space: $O(mn)$ Could be made $O(m+n)$

Optimal BST

[a, b, d, e, f, h]



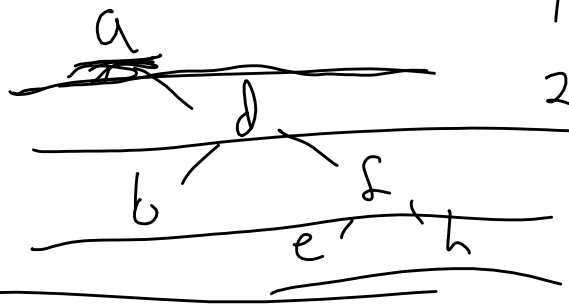
n nodes

search trees $O(\log n)$

a b d e f h

Avg. case a b d e f h = vals

0.5 0.1 0.2 0.05 0.05 0.1 = freq



1. 0.5 a
2. 0.2 d
3. 0.1 b, f
⋮

vals[] is fixed

OptBST(lo, hi): // best avg cost for a subtree containing vals[lo : hi+1]

$$\begin{cases} 0 & lo > hi \\ \min_{k \in lo..hi} \begin{cases} + \text{freq}[i] \\ + \sum_{l \in lo..k-1} \text{freq}[l] \\ + \sum_{r \in k+1..hi} \text{freq}[r] \end{cases} \begin{matrix} + \text{OptBST}(lo, k-1) \\ + \text{OptBST}(k+1, hi) \end{matrix} & lo \le hi \end{cases}$$

$$\left(\sum_{i \in lo..hi} \text{freq}[i]\right) + \text{OptBST}(lo, k-1) + \text{OptBST}(k+1, hi)$$