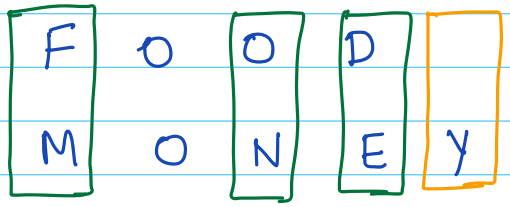
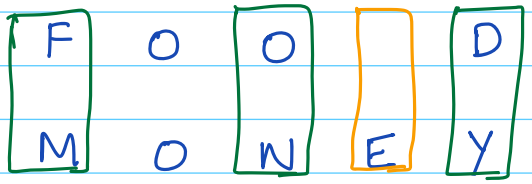
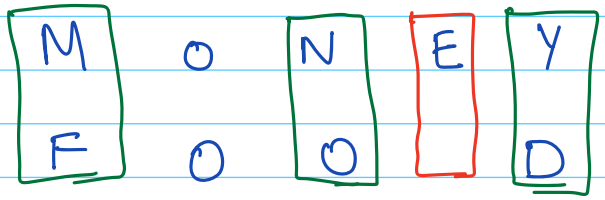
* Edit distance

Given two strings $A[1,..,n]$ $B[1,2,..,m]$
what is the min # of
- additions,
- deletions
- substitutions

necessary to convert A to B







Recurrence

Edit(i, j) = edit distance of $A[1,..,i]$ and $B[1,2,..,j]$

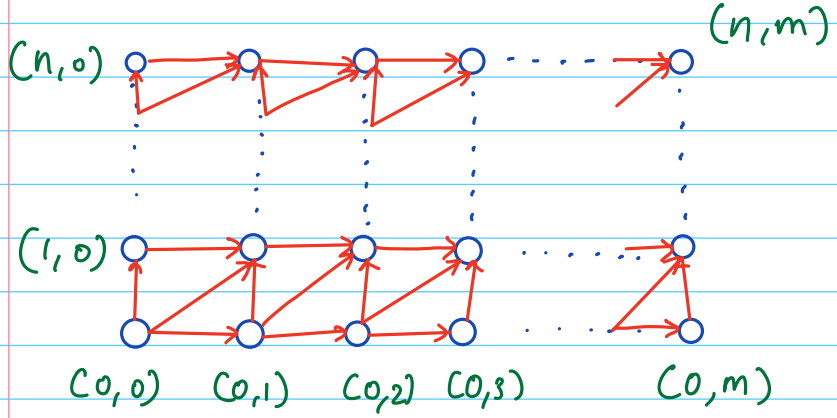
* delete $A[i]$
* insert $B[j]$
* substitute depending on whether $A[i] = B[j]$

- delete A[i]: Edit(i-1, j) + 1
- insert B[j] Edit(i, j-1) + 1
- substitution $\begin{cases} \text{Edit}(i-1, j-1) + 1 & \text{if } A[i] \neq B[j] \\ \text{Edit}(i-1, j-1) & \text{if } A[i] = B[j] \end{cases}$

↳ Edit(i-1, j-1) + [[A[i] ≠ B[j]]]

$$\text{Edit}(i,j) = \begin{cases} i & \text{if } j = 0 \\ j & \text{if } i = 0 \\ \min \begin{cases} \text{Edit}(i-1, j) + 1 \\ \text{Edit}(i, j-1) + 1 \\ \text{Edit}(i-1, j-1) + [[A[i] \neq B[j]]] \end{cases} \end{cases}$$

What is the dependency digraph?



Edge weighted graph with edge weights 1/0

Shortest path from (0,0) to (m,n)

Improving space using divide-&-conquer

- Compute the sequence with min edit distance in $O(mn)$ time and $O(m+n)$ space.

Shortest path from $(0,0)$ to (n,m)

$$= \min_i \{ \text{shortest path from } (0,0) \text{ to } (i, m/2) + \text{Shortest path from } (i, m/2) \text{ to } (n,m) \}$$

Shortest path from $(0,0)$ to $(i, m/2)$

$$= \text{Edit}(i, m/2)$$

Shortest path from (i,j) to (n,m)

$$= \tilde{\text{Edit}}(i,j)$$

$$\tilde{\text{Edit}}(i,j) = \begin{cases} n-i & \text{if } j=m \\ m-j & \text{if } i=n \\ \min \begin{cases} \tilde{\text{Edit}}(i+1, j) \\ \tilde{\text{Edit}}(i, j+1) \\ \tilde{\text{Edit}}(i+1, j+1) + [A(i) \neq B(i)] \end{cases} \end{cases}$$