4.5.2017

$OPT(i,j)$

$$\begin{array}{lll} X: & x_1 \; x_2 \; \ldots & \boxed{\begin{array}{c} x_i \\ | \\ y_j \end{array}} \\ Y: & y_1 \; y_2 \; \ldots & \end{array} \checkmark$$

$$\begin{array}{ll} X: & x_1 \; x_2 \ldots x_{i-1} \; \boxed{\begin{array}{c} x_i \\ - \end{array}} \\ Y: & y_1 \; y_2 \ldots y_j \end{array} \checkmark$$

$$\begin{array}{c} x_1 \; x_2 \ldots x_i \\ \times \\ y_1 \; y_2 \ldots y_j \end{array}$$

- x_i and y_j cannot be matched with othe characters

$$\begin{array}{ll} X: & x_1 \; x_2 \ldots x_i \\ Y: & y_1 \ldots y_{j-1} \; \boxed{\begin{array}{c} - \\ y_j \end{array}} \end{array} \checkmark$$

$$\boxed{OPT(i,j) = \min\{\alpha_{x_i y_j} + OPT(i-1, j-1),\ \delta + OPT(i-1, j),\ \delta + OPT(i, j-1)\}}$$

<u>Example</u> Sequence Alignment

$X =$ mean

$Y =$ name

$\delta = 2$

matching cost
- same symbols, cost $= 0$
- vowel & different vowel, cost $= 1$
- consonant & different consonant, cost $= 1$
- vowel & consonant, cost $= 3$

Find an optimal alignment of X and Y.

Solution

A

X	i		—	n	a	m	e
	4	n	8	6	5	4	6
	3	a	6	5	3	5	5
	2	e	4	3	2	4	4
	1	m	2	1	3	4	6
	0	—	0	2	4	6	8
			0	1	2	3	4

Y, j

$\delta = 2$

$A[i,0] = i\delta$ for each i

$A[0,j] = j\delta$ for each j

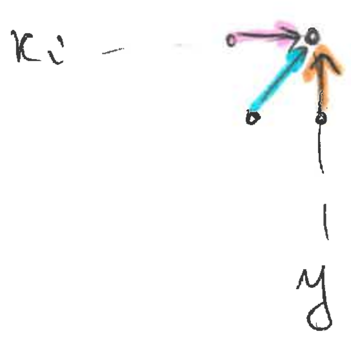
$$A[i,j] = \min\{\underbrace{\alpha_{x_i y_j} + A[i-1,j-1]}, \underbrace{\delta + A[i-1,j]}, \underbrace{\delta + A[i,j-1]}\}$$

$A[1,1] = \min\{1+0, 2+2, 2+2\} = 1$

$A[2,1] = \min\{3+2, 2+1, 2+4\} = 3$

Cost of an optimal alignment is $A[m,n] = A[4,4] = 6$

Find the alignment?



X	x_i	x_i	—
Y	y_j	—	y_j

optimal alignment:

X: m e a n -

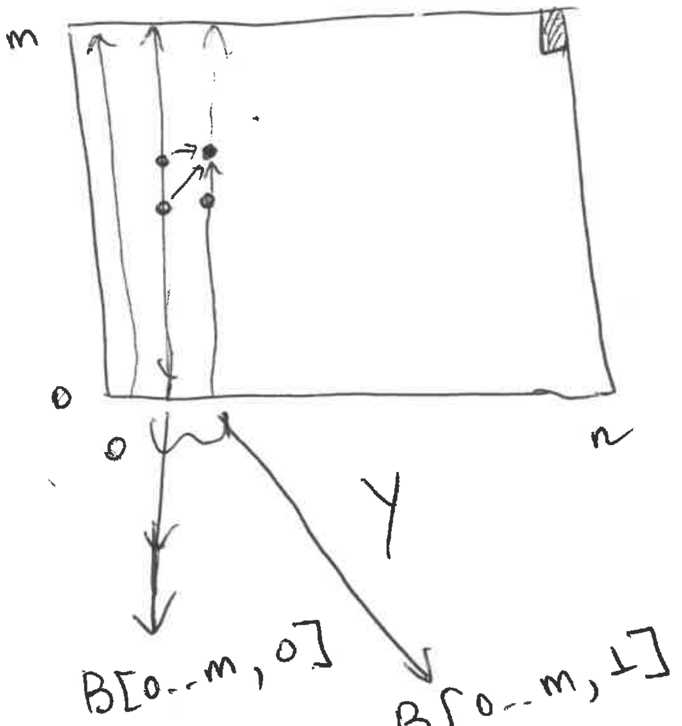
Y: n - a m e

(1,1), (3,2), (4,3)

cost = 1+2+0+1+2 = **6**

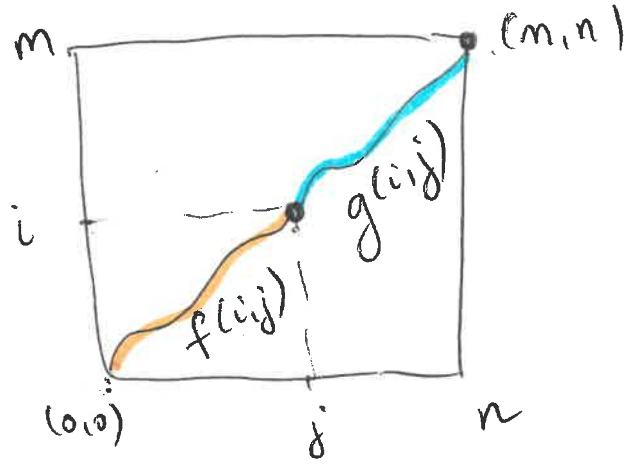
Space-efficient alignment

table A



array B [0..m, 0..1] $\Rightarrow$ size $(m+1) \times 2$

$\Theta(m)$

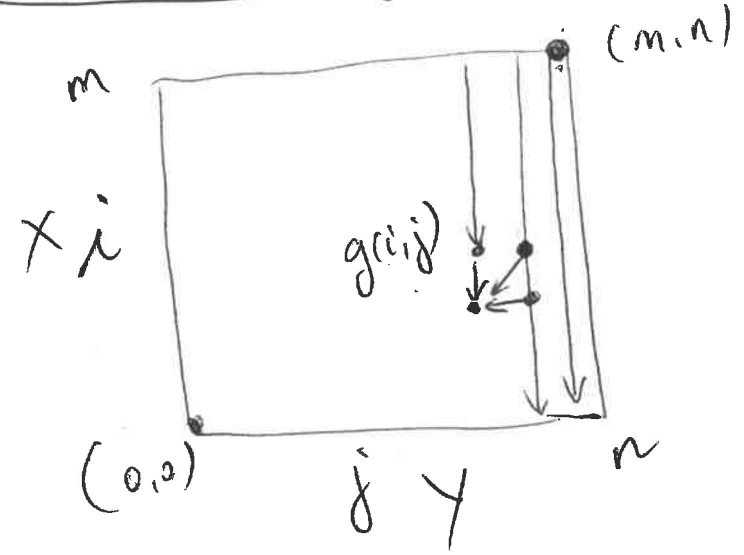


$f(i,j) = OPT(i,j)$

$g(i,j)$ - cost of shortest-path between (i,j) and (m,n)

$f(i,j) + g(i,j)$ → cost of a shortest path from $(0,0)$ to (m,n) which goes through (i,j)

Backward Alignment



$RT = O(m \cdot n)$

$Space = O(m \cdot n)$

$$g(i,j) = \min\{\alpha_{x_{i+1} y_{j+1}} + g(i+1, j+1),\ \delta + g(i, j+1),\ \delta + g(i+1, j)\}$$

- can have Backward-Space-Efficient-Alignment (uses only 2 columns) => space $O(m)$