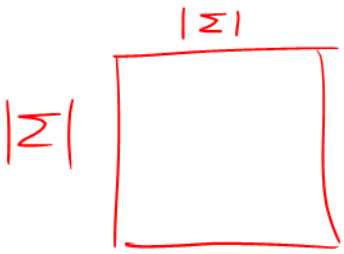
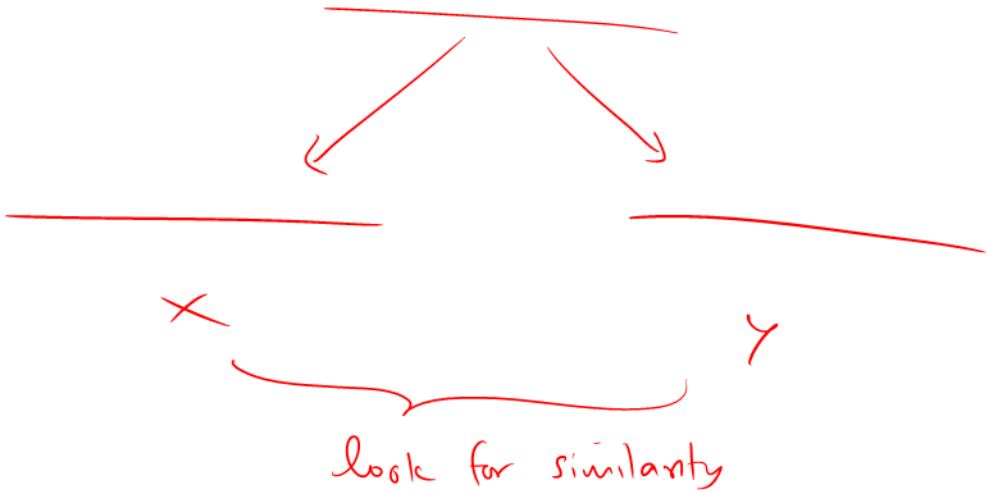


Scoring matrix

Σ = alphabet
= {A, C, G, T}

gap penalties

linear model : gap = d



① global alignment

X: A C T G A
Y: A T A

→

A C T G A
A - T - A

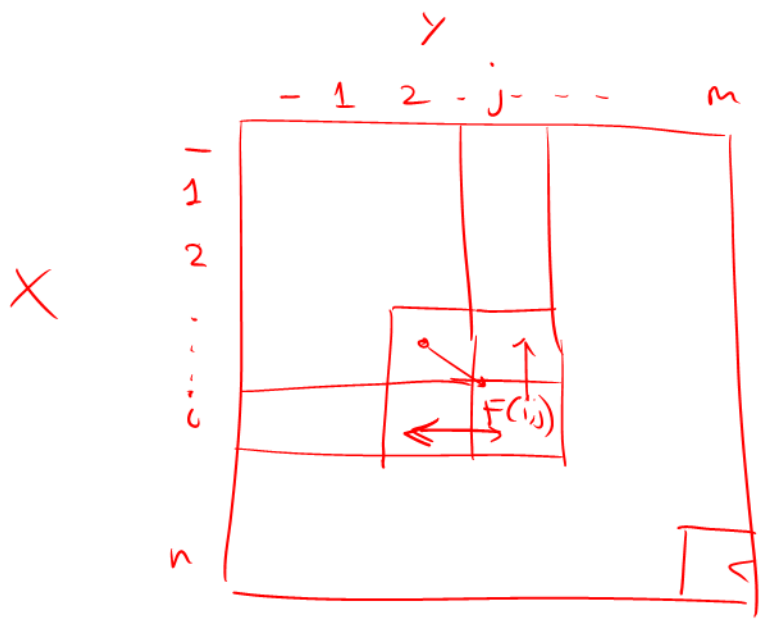
Score

② local alignment

X: A C T A A T A G C
Y: C A T A G T

Task:

find the optimal
highest scoring
sub sequence of X with
subsequence of Y



local case:

$$F(i,j) = \max \begin{cases} 0 \\ F(i-1, j-1) + s(x_i, y_j) & \nwarrow \\ F(i, j-1) + d & \leftarrow \text{ gap in } x \\ F(i-1, j) + d & \uparrow \text{ gap in } y \end{cases}$$

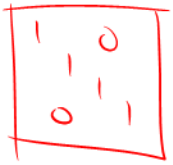
To trace back and recover the best local alignment

Start from the cell with the highest value

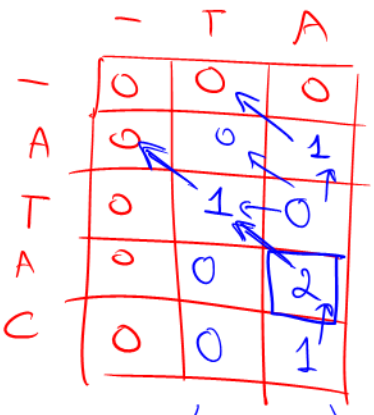
match = 1

mismatch = 0

gap = -1



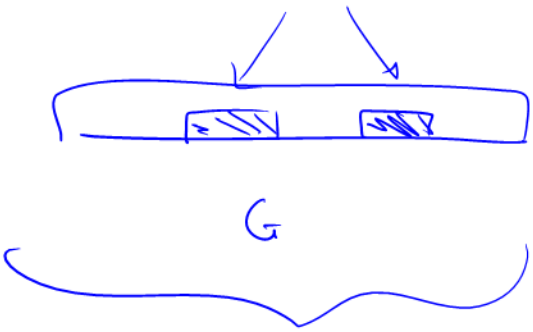
X



X A T A C

Y ← T A →

2



Affine gap model

d : gap opening

e : gap extension $\quad e < d$

$O(n \cdot m)$ time method

Space: 3 time as much

keep extension matrices for x and y

$M \equiv F$

I_x = I_y = } keep track of extensions

$$M(i,j) = \max \begin{cases} M(i-1,j-1) + s(x_i, y_j) \\ I_x(i-1,j-1) + s(x_i, y_j) \\ I_y(i-1,j-1) + s(x_i, y_j) \end{cases}$$

diagonal case

$$I_x(i,j) = \max \begin{cases} M(i,j-1) + d \\ I_x(i,j-1) + e \end{cases}$$

$M(i,j-1) + d$ ⟵ gap opening

$I_x(i,j-1)$: already opened in the past

e: extending