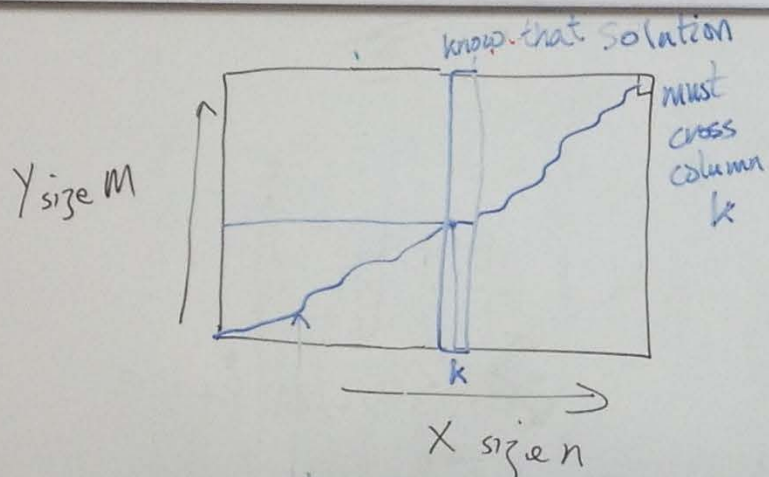


Cs5050 notes 01 28 14



alignment

Problem

algorithm is Space bound!

time  $\approx nm$

space  $\approx nm$

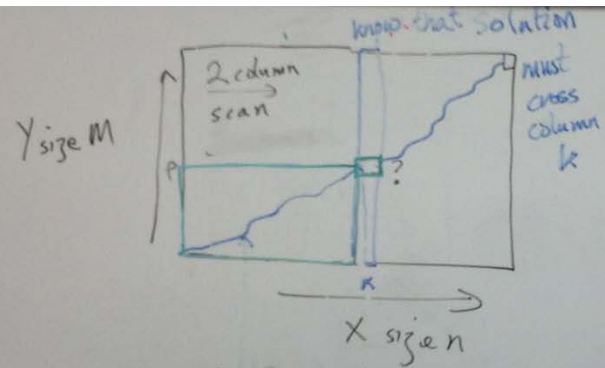
$n = m = 10^9$

$nm = 10^{18}$  characters

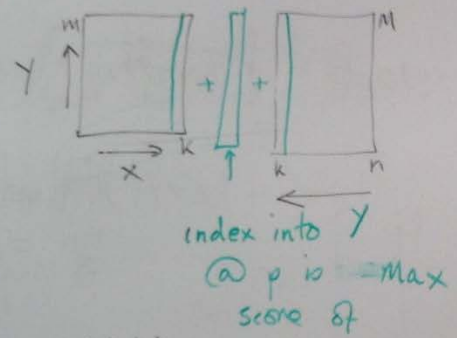
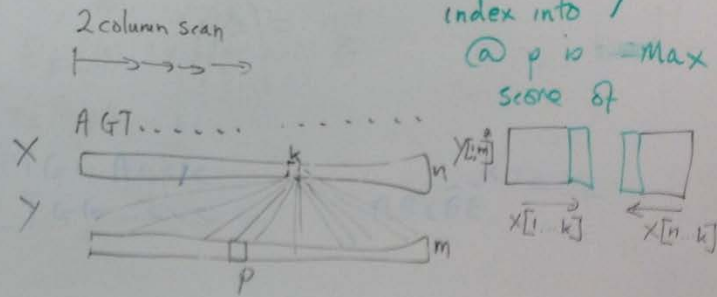
trace back.  
solution size  
 $\approx n + m$

A G T G G A c c c  
A \_ T G G \_ c c c

\_ \_ \_ \_ X Y Z Z Z  
A B C D E \_ \_ \_ \_

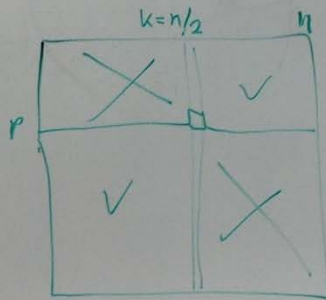


+ 2 column scan  
all prefix of  $Y$   
matched to  $1 \dots k$  of  $X$



### Problem

- + trace back  
need for the alignment
- + need all the scores  
at the same time!



### Ideas

- + only need  $2n$  space  
for the score
- + trade time for space
- + Solution is small  $n+m$
- + Try one char alignment @ a time

+ to find one character in  $m$  requires  $n \times m$  steps

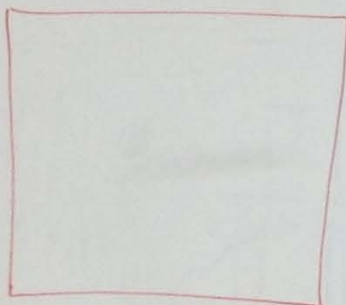
+ run  $n$  times to get all alignments

$$n \times n \times m = n^2 m$$

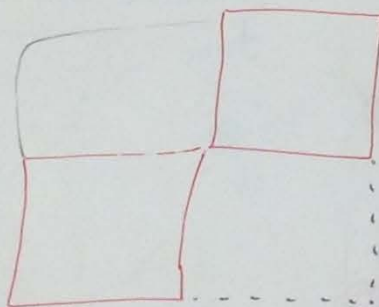
$$n = m = 10^9$$

$$10^{27}$$

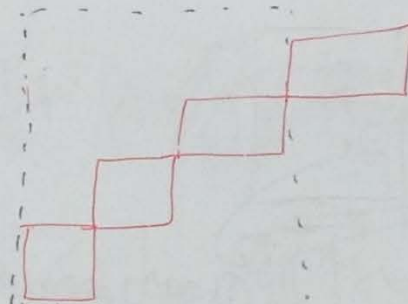
time increased by  $\times n$



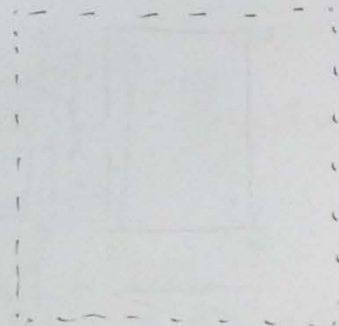
$n \times m$



$\frac{n \times m}{2}$



$\frac{n \times m}{4}$



$\frac{n \times m}{8}$

$$\text{time} = n \times m + \frac{n \times m}{2} + \frac{n \times m}{4} + \frac{n \times m}{8} + \dots$$

$$= n \times m \left( 1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{8} \right)$$

$$= n \times m \left( 1 + \sum_{i=1}^{\log_2 n} \frac{1}{2^i} \right)$$

$$\text{time} < n \times m \left( 1 + \sum_{i=1}^{\infty} \frac{1}{2^i} \right)$$

$$f = \sum_{i=1}^{\infty} \frac{1}{2^i}$$