

# 5 KMP and BM

## Pattern Matching

$S =$  

$P =$  

Find all occurrences of  $p$  in  $S$

### Naive

Try all positions

$O(n \cdot m)$

Can we do better?

### KMP

Use information from pattern



IE we can use the border array to shift faster

How it works



- each match increases  $i$



$j$  decreased using BA

- each mismatch increases  $i$   
- or decreases  $j$



$i$  increased

### Time analysis

- matches  $O(n)$

- mismatches either

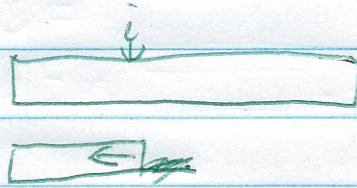
- increase  $i$   $O(n)$

- increase pattern position  $O(n-m)$

Total  $O(n)$

# BM

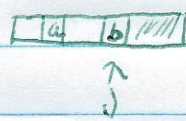
Searches pattern from right-to-left



$O(n \cdot m)$

Use optimizations.

Rightmost array

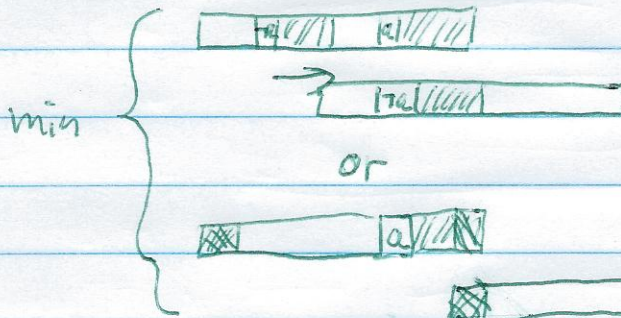


if  $R[a] < j$  we can shift to  $R[a]$



$O(m + |\Sigma|)$

# S



We can use suffix border array

$$SBA = (BA(p^R))^R$$

$\max(R, S)$

still  $O(m \cdot n)$  but fast if few matches  
can run in sublinear time.