

String Matching

Brute Force Algorithm —

$T[1, \dots, n] \rightarrow \text{Text}$

$P[1, \dots, m] \rightarrow \text{Pattern}$

all the character belongs from Σ

$\Sigma = \{0, 1\}$ or $\Sigma = \{a, b, \dots, z\}$

- Pattern occurred with shift s in T
- String matching problem is to find all values of s
i.e. $0 \leq s \leq n-m$

- if $P[1, \dots, m] = T[s+1, \dots, s+m]$

Let $T \Rightarrow$

b	a	c	a	b	c	a	b	c	a	-	-	-	...
---	---	---	---	---	---	---	---	---	---	---	---	---	-----

$P \Rightarrow$

		c	a	b	c				
--	--	---	---	---	---	--	--	--	--

$\downarrow \quad \downarrow$
 $S \Rightarrow 1+1 = 2 \text{ shift}$

after 2 shift we find the match

String matching terminology —

Prefix $\rightarrow x = wy$

Suffix $\rightarrow x = yw$

w is a string

x, y are the any string or number or character

Brute force algorithm -

In Shift $S=0$

$T \Rightarrow$

c	a	c	a	b
---	---	---	---	---

$\downarrow \quad \downarrow \quad \downarrow$
not matched

$P \Rightarrow$

c	a	b
---	---	---

so shift Pattern one character
Position to right

$S=1$

$T \Rightarrow$

c	a	c	a	b
---	---	---	---	---

$\downarrow \quad \downarrow \quad \downarrow$
not matched

$P \Rightarrow$ -

c	a	b
---	---	---

$S=2$

$T \Rightarrow$

c	a	c	a	b
---	---	---	---	---

$\downarrow \quad \downarrow \quad \downarrow$
 $P \Rightarrow$ - -

c	a	b
---	---	---

Pattern matched

so return value of S