

ESERCIZIO 4

$S = ababca$

→ Costruire il suffix array

ababca\$

babca\$

abca\$

bca\$

ca\$

a\$

\$

→

	SA	LCP
\$	7	0
a\$	6	1
ababca\$	1	2
abca\$	3	0
babca\$	2	1
bca\$	4	0
ca\$	5	/

→ Ricerca pattern $P = ab$

• Se usiamo la ricerca stile binary search facciamo $O(\log n)$ ricerche e ogni volta confrontiamo al più P caratteri → $O(p \log n) + O(n)$ per prendere le occorrenze.

Facciamo la ricerca di P e di $P\#$ per trovare la posizione i e j .

Binary Search:

①

L → \$
a \$

$$\text{strcmp}(P, SA[1], p) = 0$$

quindi $R = 1$

②

L → \$
M → a \$
R → ababca \$
abca \$

$$\text{strcmp}(ab, SA[2], 2) > 0$$

M
ababca \$
abca \$
babca \$
bca \$
L → ca \$

③ $\text{strcmp}(ab, SA[3], 2) = 0$

L → ababca \$
R → abca \$

→ L → ababca \$
R → abca \$

Restituisci la posizione
 $L = 3 = i$

Allo stesso modo cerca $P \#$ e
trova la posizione $j = 4$

Numero occorrenze = $4 - 3 + 1 = 2$

Se voglio anche le stringhe le prendo
prendo tempo $O(n)$

→ Alternativa: Algoritmo de evita confronti
 e mi costo time: $O(p + \log n)$
 + $O(occ)$ per prendere le occorrenze.

$$P = ab$$

L → \$

a \$

ababca\$

M abca\$

babca\$

bca\$

R → ca\$

$$L_{lcp} = lcp(L, \pi) = \emptyset$$

$$R_{lcp} = lcp(R, \pi) = \emptyset$$

$$l = lcp(P, L) = \emptyset$$

$$r = lcp(P, R) = \emptyset$$

$$LCP(L, R) = \emptyset$$

Dato che $l = L_{lcp} = \emptyset$ allora partiamo
 a confronto del carattere \emptyset

Quindi confrontiamo $P[0:]$ e $Sup_{\pi} SA[\pi][0:]$
 e poi $R = \pi$

L → \$

M → a \$

ababca\$

R → abca\$

$$LCP(L, R) = \emptyset$$

$$L_{lcp}[\pi] = \emptyset$$

$$R_{lcp}[\pi] = 2$$

$$l = \emptyset$$

$$M = 2$$

$$M \geq l$$

Confrontiamo M con $R_{\text{lep}}[M]$ hanno lo stesso valore e allora il confronto tra P e $\text{SuffSA}[M]$ parte dal 3° carattere. In questo caso il confronto dice che $b > \$$ quindi $L = M + 1$

$L, M \rightarrow ababca\$$
 $R \rightarrow abca\$$

$$L_{\text{lep}}[M] = 7$$

$$R_{\text{lep}}[M] = 2$$

$$l = 2$$

$$M = 2$$

$$M = l$$

$$M = R_{\text{lep}}[M]$$

quindi ci sono 2 caratteri in comune,
 $\text{strcmp}(P, \text{SA}[M], p) = \emptyset \rightarrow R = M$
 e trovo la posizione $L = 3$.