Algorithms Design and Analysis [ETCS-301]

Dr. A K Yadav
Amity School of Engineering and Technology
(affiliated to GGSIPU, Delhi)
akyadav1@amity.edu
akyadav@akyadav.in
www.akyadav.in
+91 9911375598

November 9, 2019

Knuth-Morris-Pratt algorithm

- This algorithm avoids computing the transition function δ altogether
- Its matching time is $\Theta(n)$
- Its preprocessing time is $\Theta(m)$
- The array π allows us to compute the transition function δ efficiently.
- π is prefix function for a pattern $P[1..m]$ defined as $\pi : \{1, 2, \ldots, m\} \rightarrow \{0, 1, \ldots, m-1\}$ such that

  $$\pi[q] = \max\{k : k < q, P_k \sqsupset P_q\}$$

- $\pi[q]$ is the length of the longest prefix of P that is a proper suffix of P_q.
- $P_q \sqsupset T_{s+q}$
- $k = \pi[q]$
- $s = s + (q - k)$

Algorithm I

COMPUTE-PREFIX-FUNCTION(P)

1. $m = length(P)$
2. let $\pi[1..m]$ a new array
3. $\pi[1] = 0$
4. $k = 0$
5. for $q = 2$ to m
6.   while $k > 0$ and $P[k+1] \neq P[q]$
7. $k = \pi[k]$
8.   if $P[k+1] == P[q]$
9. $k = k + 1$
10. $\pi[q] = k$
11. return π

Algorithm II

KMP-MATCHER(T, P)

```
n = length(T)
m = length(P)
π = COMPUTE-PREFIX-FUNCTION(P)
q = 0
for i = 1 to n
    while q > 0 and P[q + 1] ≠ T[i]
        q = π[q]
    if P[q + 1] == T[i]
        q = q + 1
    if q == m
        print Pattern occurs with shift i − m
        q = π[q]
```

Thank you

Please send your feedback or any queries to akyadav1@amity.edu, akyadav@akyadav.in or contact me on +91 9911375598