

CS4335 Design and Analysis of Algorithms
Tutorial 10

Student Name _____ Student Id _____

Question 1. Construct the failure function for the pattern

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
a	b	c	a	B	b	a	b	c	a	b	b	a	b	b	b	b

using the linear time algorithm. How many comparisons are used in order to compute $f(15)$?

Solution:

$q=1, f(1)=0$ (case 1)
 $q=2, P(2)=b, \text{neq } P(f(2-1)+1)=a, f(2)=0$ (case 4)
 $q=3, P(3)=c \text{ neq } P(f(3-1)+1)=P(1)=a, f(3)=0$ (case4)
 $q=4 P(4)=a, \text{eq } P(f(4-1)+1)=P(1)=a, f(4)=f(3)+1=1$ (case 2)
 $q=5 P(5)=b, \text{neq } P(f(5-1)+1)=P(1+1)=b f(5)=2$ (case 2)
 $q=6 P(6)=b \text{ neq } P(f(6-1)+1)=P(2+1)=c$ (case 3)
 $\text{neq } P(q(2)+1)=P(1)=a, f(6)=0$ (case 4)
 $q=7 P(7)=a \text{ eq } P(f(7-1)+1)=P(1)$ (case 2) $f(7)=1$
 $q=8 P(8)=b \text{ eq } P(f(7)+1)=P(2)=b f(8)=2$
 $q=9 P(9)=c \text{ eq } P(f(8)+1)=P(3)=c, f(9)=3$
 $q=10 ? f(10)=4$
 $q=11 ? f(11)=5$
 $q=12 ? f(12)=6$
 $q=13 ? f(13)=7$
 $q=14 ? f(14)=8$
 $q=15 P(15)=b, \text{neq } P(f(14)+1)=P(8+1)=c, (case 3)$
 $\text{neq } P(f(8)+1)=P(2+1)=c (case 3)$
 $\text{neq } P(f(2)+1)=P(0+1)=a, f(15)=0$ (case 4)
 $q=16 P(16)=b \text{ neq } P(f(15)+1)=P(1)=a, q(16)=0$
 $q=17 P(17)=b, \text{neq } P(f(16)+1)=P(1)=a, q(17)=0.$

To compute $f(15)$, three comparison (b, c), (b, c) and (b, a) are required.

Alg:

Case 1: $f(1)$ is always 0.

Case 2: if $P[q] == P[f(q-1)+1]$ then $f(q) = f(q-1) + 1$.

Case 3: if $P[q] \neq P[f(q-1)+1]$ and $f(q-1) \neq 0$ then consider $P[q] ?= P[f(f(q-1))+1]$ (Do it recursively)

Case 4: if $P[q] \neq P[f(q-1)+1]$ and $f(q-1) == 0$ then $f[q] = 0$.

Question 2. Let the text be abababcbabababcc and the pattern be ababc. Find out all the occurrences of the pattern using the KMP scan algorithm?

Answer:

Computing f-value of P=ababc

f(1)=0, f(2)=0, f(3)=1, f(4)=2, f(5)=0

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
T	a	b	a	b	a	b	c	a	b	a	b	a	b	c	c
	a	b	a	b	c										
			a	b	a	b	c								
								a	b	a	b	c			
										a	b	a	b	c	
															a

At i=3 and i=10

Algorithm:

- **i** : indicates that T[i] is the next character in T to be compared (**green arrow**).
 - **q**: indicates that P[q+1] is the next character in P to be compared with T[i] (**red arrow-1**).
1. i=1 and q=0;
 2. **compare** T[i] with P[q+1]
 - case 1:** T[i]==P[q+1]
 - i=i+1;q=q+1;
 - if** q==|P| **then print** "P occurs at i-|P|", **and** q=f(|P|).
 - case 2:** T[i]≠P[q+1] **and** q≠0
 - q=f(q); % the pattern shifts forward
 - case 3:** T[i]≠P[q+1] **and** q==0
 - i=i+1; % the pattern shifts one position forward
 3. **Repeat** step2 until i==|T|.