

Chapter 16

16.1 Consider a string 'ABCX'. List out the number of sub-sequences and substrings.

Solution:

Some of the possible substrings are

'A'

'AB'

'ABC'

'ABCX'

16.2 Consider the string 'ABCD'. List out the prefix and suffix of the string.

Solution:

A prefix is formed by deleting zero or more characters from the end of the string .

ABCD	BCD
ABC	CD
AB	D
A	E
E	
(Prefix)	(Suffix)

16.3 Consider the string

$s = \text{'ABCD'}$

$t = \text{'RMST'}$

Find $s + t$ and $t + s$ length of s and t and substrings of s and t .

Solution:

Length of $(s) = 4$

Length of $(t) = 4$

$s + t = \text{'ABCDRMST'}$

$t + s = \text{'RMSTABCD'}$

Substrings of $s = A, AB, ABC, ABCD$

Substrings of $t = R, RM, RMS, RMST$

16.4 Consider the following text T and pattern P

Text : THIS IS AN EXAMPLE

Pattern: AN

Apply a naïve matching algorithm and illustrate the intermediate steps.

Solution:

THIS IS AN EXAMPLE
AN

THIS IS AN EXAMPLE
AN

THIS IS AN EXAMPLE
AN

THIS IS AN EXAMPLE
AN

THIS IS AN EXAMPLE
AN

THIS IS AN EXAMPLE
AN

THIS IS AN EXAMPLE
AN

Report match.

16.5 Find LCS between strings A and B .

a) EXAMPLE APE

		E	X	A	M	P	L	E
		+0	+0	+0	+0	+0	+0	+0
A		+0	∧0	\1	<1	<1	<1	<1
P		+0	∧0	∧1	∧1	\2	\2	<2
E		+0	\1	<1	∧1	∧1	∧2	\3

∴ The LCS of two strings is 3.

b) ACGTCCAT CCCT

		A	C	G	T	C	C	A	T
	+0	+0	+0	+0	+0	+0	+0	+0	+0
C	+0	$\wedge 0$	$\backslash 1$	< 1	< 1	$\backslash 1$	$\backslash 1$	< 1	< 1
C	+0	$\wedge 0$	$\backslash 1$	$\wedge 1$	< 1	$\backslash 2$	$\backslash 2$	< 2	< 2
C	+0	$\wedge 0$	$\backslash 1$	$\wedge 1$	$\wedge 1$	$\backslash 2$	$\backslash 3$	< 3	< 3
T	+0	$\wedge 0$	$\wedge 1$	$\wedge 1$	$\backslash 2$	$\wedge 2$	$\wedge 3$	$\wedge 3$	$\backslash 4$

The LCS of two strings is 4

16.6 Consider the following sequences.

A : acgabdc

B : bdc

Solution:

hash(bdc)=691

hash(y[0..2])=689

attempt 1:

acgabdc

...

hash(y[1..3])=699

attempt 2:

acgabdc

...

hash(y[2..4])=704

attempt 3:

acgabdc

...

hash(y[3..5])=684

attempt 4:

acgabdc

...

hash(y[4..6])=691

BDC

attempt 5:

acgabdc

...

acgaBDC

String length: 7

Pattern length: 3

Attempts: 5

Character comparisons: 3

16.7 Find the prefix function for the following strings using the KMP algorithm.

a. ABCTA

The prefix function value is

0 0 0 0 1

b. ABCTABC

THE prefix value is

0 0 0 0 1 2 3

16.8 Apply the KMP algorithm for matching the following strings and illustrate the intermediate steps:

a) A : EXAMPLE

B : AMP

EXAMPLE
AMP

EXAMPLE Comparison 1
AMP

EXAMPLE Comparison 2
AMP

EXAMPLE Comparison 5 found. Report success.
AMP Comparison 6

EXAMPLE Comparison 7
AMP

EXAMPLE Comparison 7 , Stop.
AMP

b) A : ACGTCCAT
B : TCCA

ACGTCCAT
TCCA Comparison = 0

ACGTCCAT
TCCA Comparison = 1

ACGTCCAT
TCCA Comparison = 2

ACGTCCAT
TCCA Comparison = 7 Found, report success.

ACGTCCAT
TCCA Comparison = 8 Stop.

16.9 Apply Boyer-Moore algorithm for the following strings and illustrate the intermediate steps.

a) EXAMPLE and AMP

EXAMPLE

AMP

Comparison = 0

Compare A and P Comparison = 1

EXAMPLE

AMP

Comparison = 4

Found and report Success.

b) A : A C G T C C A T

B : T C C A

A C G T C C A T

T C C A

Comparison = 0

A C G T C C A T

T C C A

Compare T and A

Comparison = 1

Comparison = 5 , Stop.

16.10 Find the edit distance between the following strings directly.

a) FOOD and FODE

Distance is 2.

		-1	0	1	2	3
			F	O	D	E
-1		✓ 0	1	2	3	4
0	F	1	✓ 0	1	2	3
1	O	2	2	✓ 0	1	2
2	O	3	2	1	✓ 1	2
3	D	4	3	2	1	✓ 2

Food –

Fo–de

Two operations are required to find distance.

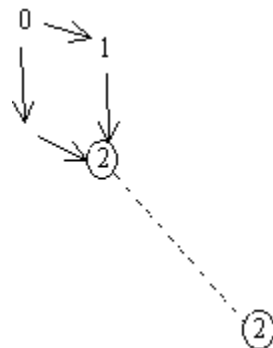
b) SAMPLE and EXAMPLE

Solution:

		-1	0	1	2	3	4	5
			S	A	M	P	L	E
-1		①	1	2	3	4	5	6
0	E	①	①	2	3	4	5	5
1	X	2	②	2	3	4	5	6
2	A	3	3	②	3	4	5	6
3	M	4	4	3	②	3	4	5
4	P	5	5	4	3	②	3	4
5	L	6	6	5	4	3	②	3
6	E	7	7	6	5	4	3	②

E X A M P L E

S – A M P L E



∴ Two operations are required to make it same.