FIT2004 S2_2016 Tute Week 9 Solutions

Question 1:

First, we have to list down all possible cyclic rotations of the given text "woolloomooloo$"

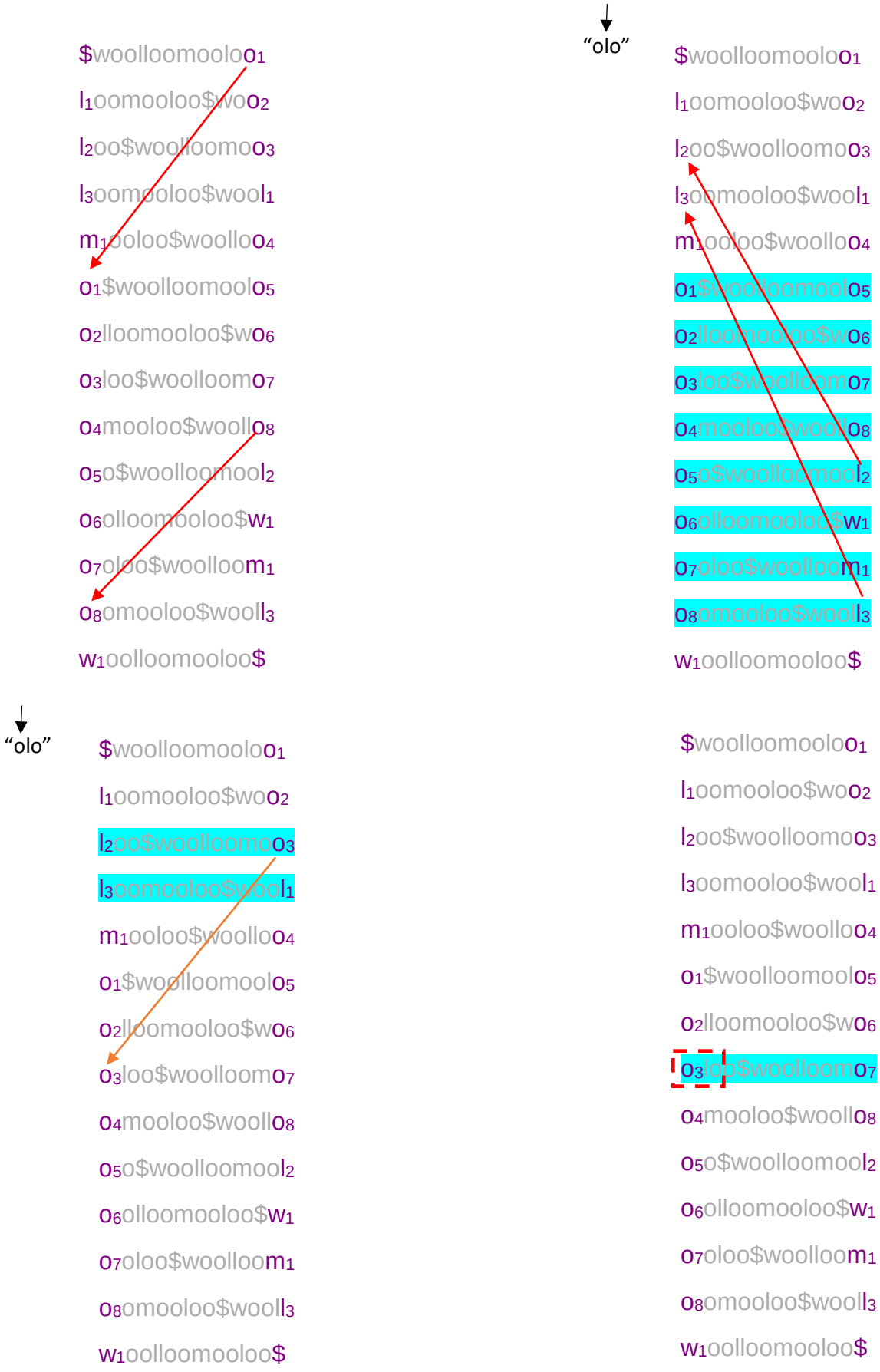
woolloomooloo$	After Sorting we get,	$woolloomooloo
$woolloomooloo		lloomooloo$woo
o$woolloomoolo		loo$woolloomoo
oo$woolloomool		loomooloo$wool
loo$woolloomoo		mooloo$woolloo
oloo$woolloomo		o$woolloomoolo
ooloo$woolloom		olloomooloo$wo
mooloo$woolloo		oloo$woolloomo
omooloo$woollo		omooloo$woollo
oomooloo$wooll		oo$woolloomool
loomooloo$wool		oolloomooloo$w
lloomooloo$woo		ooloo$woolloom
olloomooloo$wo		oomooloo$wooll
oolloomooloo$w		woolloomooloo$

The last column in red of the sorted matrix is Burrows-Wheeler Transform

Question 2:

a) Using BWT, find substring "olo"

- Initially the range contains all rows of BWT
- Starting from right most character "olo"
- Find first occurrence of character in the range and the last occurrence of character in the range in the Last column
- Find the corresponding **characters** in the first column and update the range


$woolloomoolo o1
l1oomooloo$woo o2
l2oo$woolloomo o3
l3oomooloo$wool l1
m1ooloo$woollo o4
o1$woolloomool o5
o2lloomooloo$w o6
o3loo$woolloom o7
o4mooloo$woll o8
o5o$woolloomool2
o6olloomooloo$w1
o7oloo$woolloom m1
o8omooloo$wool l3
w1oolloomooloo$
"olo"
$woolloomoolo o1
l1oomooloo$woo o2
l2oo$woolloomo o3
l3oomooloo$wool l1
m1ooloo$woollo o4
o1$woolloomool o5
o2lloomooloo$w o6
o3loo$woolloom o7
o4mooloo$woll o8
o5o$woolloomool2
o6olloomooloo$w1
o7oloo$woolloom m1
o8omooloo$wool l3
w1oolloomooloo$
"olo"
$woolloomoolo o1
l1oomooloo$woo o2
l2oo$woolloomo o3
l3oomooloo$wool l1
m1ooloo$woollo o4
o1$woolloomool o5
o2lloomooloo$w o6
o3loo$woolloom o7
o4mooloo$woll o8
o5o$woolloomool2
o6olloomooloo$w1
o7oloo$woolloom m1
o8omooloo$wool l3
w1oolloomooloo$

b) Using BWT, find substring “oll”
Starting from right most character “oll”

$\$\text{woolloomoolo}\,o_1$
$l_1\,\text{oomooloo\$wo}\,o_2$
$l_2\,\text{oo\$woolloomo}\,o_3$
$l_3\,\text{oomooloo\$woo}\,l_1$
$m_1\,\text{ooloo\$woollo}\,o_4$
$o_1\,\text{\$woolloomool}\,o_5$
$o_2\,\text{lloomooloo\$w}\,o_6$
$o_3\,\text{loo\$woolloom}\,o_7$
$o_4\,\text{mooloo\$wooll}\,o_8$
$o_5\,\text{o\$woolloomoo}\,l_2$
$o_6\,\text{olloomooloo\$}\,w_1$
$o_7\,\text{oloo\$woolloo}\,m_1$
$o_8\,\text{omooloo\$wool}\,l_3$
$w_1\,\text{oolloomooloo}\,\$$

“oll”

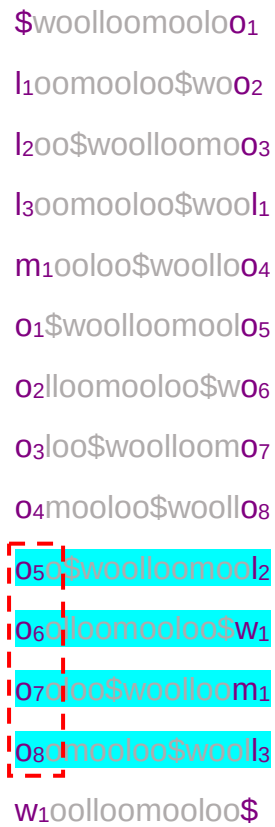
$\$\text{woolloomoolo}\,o_1$
$l_1\,\text{oomooloo\$wo}\,o_2$
$l_2\,\text{oo\$woolloomo}\,o_3$
$l_3\,\text{oomooloo\$woo}\,l_1$
$m_1\,\text{ooloo\$woollo}\,o_4$
$o_1\,\text{\$woolloomool}\,o_5$
$o_2\,\text{lloomooloo\$w}\,o_6$
$o_3\,\text{loo\$woolloom}\,o_7$
$o_4\,\text{mooloo\$wooll}\,o_8$
$o_5\,\text{o\$woolloomoo}\,l_2$
$o_6\,\text{olloomooloo\$}\,w_1$
$o_7\,\text{oloo\$woolloo}\,m_1$
$o_8\,\text{omooloo\$wool}\,l_3$
$w_1\,\text{oolloomooloo}\,\$$

“oll”

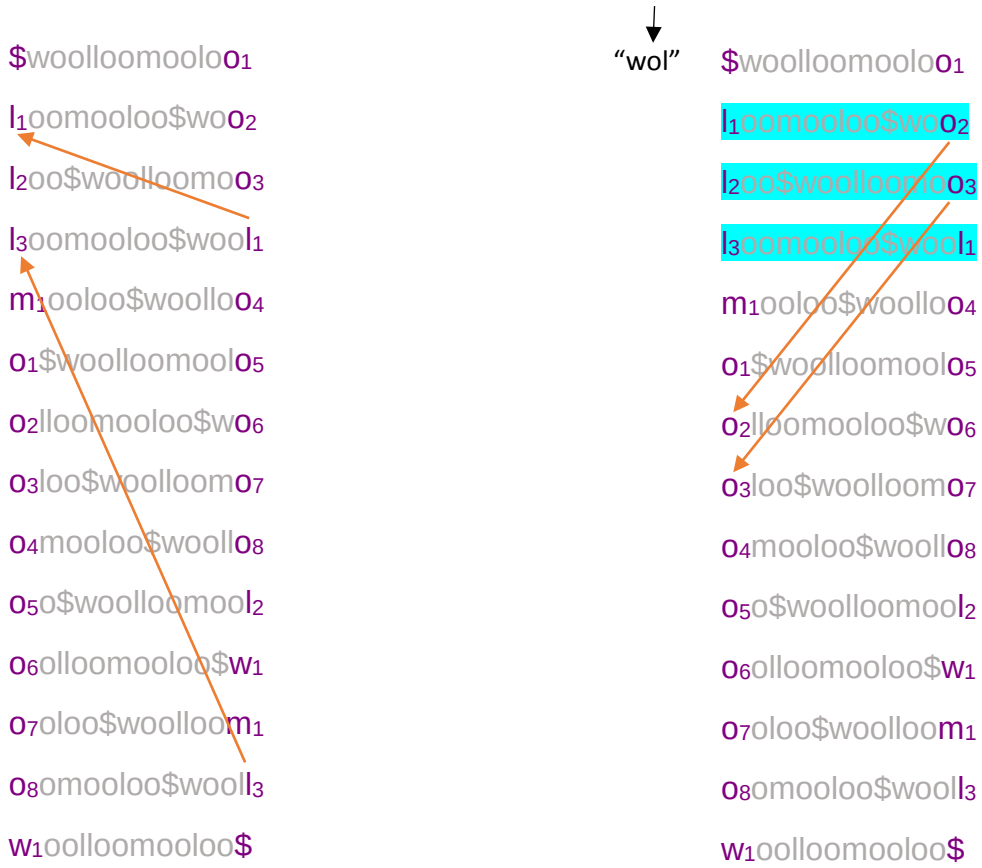
$\$\text{woolloomoolo}\,o_1$
$l_1\,\text{oomooloo\$wo}\,o_2$
$l_2\,\text{oo\$woolloomo}\,o_3$
$l_3\,\text{oomooloo\$woo}\,l_1$
$m_1\,\text{ooloo\$woollo}\,o_4$
$o_1\,\text{\$woolloomool}\,o_5$
$o_2\,\text{lloomooloo\$w}\,o_6$
$o_3\,\text{loo\$woolloom}\,o_7$
$o_4\,\text{mooloo\$wooll}\,o_8$
$o_5\,\text{o\$woolloomoo}\,l_2$
$o_6\,\text{olloomooloo\$}\,w_1$
$o_7\,\text{oloo\$woolloo}\,m_1$
$o_8\,\text{omooloo\$wool}\,l_3$
$w_1\,\text{oolloomooloo}\,\$$

$\$\text{woolloomoolo}\,o_1$
$l_1\,\text{oomooloo\$wo}\,o_2$
$l_2\,\text{oo\$woolloomo}\,o_3$
$l_3\,\text{oomooloo\$woo}\,l_1$
$m_1\,\text{ooloo\$woollo}\,o_4$
$o_1\,\text{\$woolloomool}\,o_5$
$o_2\,\text{lloomooloo\$w}\,o_6$
$o_3\,\text{loo\$woolloom}\,o_7$
$o_4\,\text{mooloo\$wooll}\,o_8$
$o_5\,\text{o\$woolloomoo}\,l_2$
$o_6\,\text{olloomooloo\$}\,w_1$
$o_7\,\text{oloo\$woolloo}\,m_1$
$o_8\,\text{omooloo\$wool}\,l_3$
$w_1\,\text{oolloomooloo}\,\$$

c) Using BWT, find substring “oo”
Starting from right most character “oo”



d) Using BWT, find substring “wol”
Starting from right most character “wol”



“wol”

$woolloomoolo$1
l1oomooloo$woo2
l2oo$woolloomoo3
l3oomooloo$wool1
m1ooloo$woollo4
o1$woolloomool5
o2lloomooloo$wo6
o3loo$woolloomo7
o4mooloo$woollo8
o5o$woolloomool2
o6olloomooloo$w1
o7oloo$woolloom1
o8omooloo$wooll3
w1oolloomooloo$

There is no ‘w’ in the last column of the range. So there is no ‘wol’ substring.

Question 3:

Source vertex: Z

Discovered
Z,0
W, ~~18~~ ~~17~~ 16
Y, 3
X, 8
V, ~~18~~ 13
U, ~~19~~ 18
T, 19
S, ~~23~~ 22
R, ~~29~~ 24
Q, ~~31~~ 23
P, ~~29~~ 28

Finalized
Z, 0
Y, 3
X, 8
V, 13
W, 16
U, 18
T, 19
S, 22
Q, 23
R, 24
P, 28

Question 4:

Below is a sample linked list representation:

A4	L2
A6	L4 L5
T3	L1 L2
T5	L2 L3
T7	L4 L5
L1	L2 L3 L4 L5 L6 L7 L8
L2	L2 L3 L4 L5 L6 L7 L8
L3	L4 L5 L6 L8
L4	L4 L6 L7 L8
L5	L6 L7 L8
L6	L6 L7
L7	L7
L8	L8
L9	L9