COP5536 Advanced Data Structures

Assignment 2

Prof. Sartaj Sahni

Due Date: 12/02/2019

Answers will be graded on correctness, elegance, efficiency, and other quality measures.

Last name: CHAUBEY

First name: VIKAS

UFID: 3511-5826

Problem-1 =

Given range in [0 ----- 63]
which could be written as [0 --- 64)

insert (12,49) ⟶ (12,49) [0 --- 64)

insert (30,12) ⟶

(30,12) [0,64)

(12,49) [0..32)

as we know that
12 < 49
Hence (30,12)
replaces (12,49)

insert (20,1) ⟶

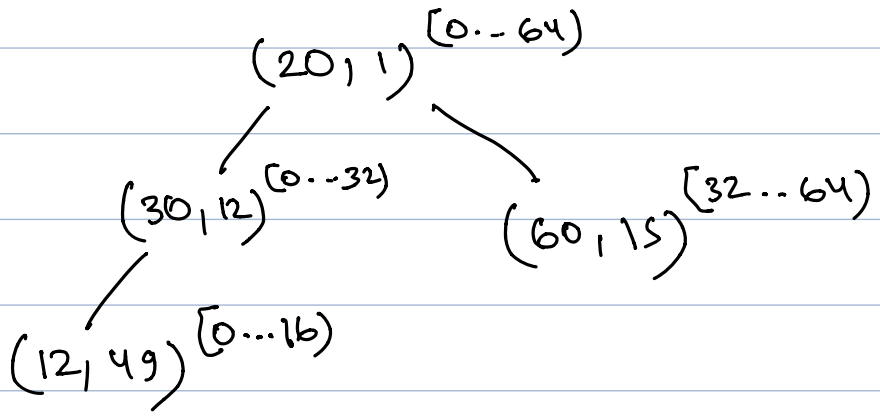
(20,1) [0..64)

(30,12) [0..32)

(12,49) [0..16)

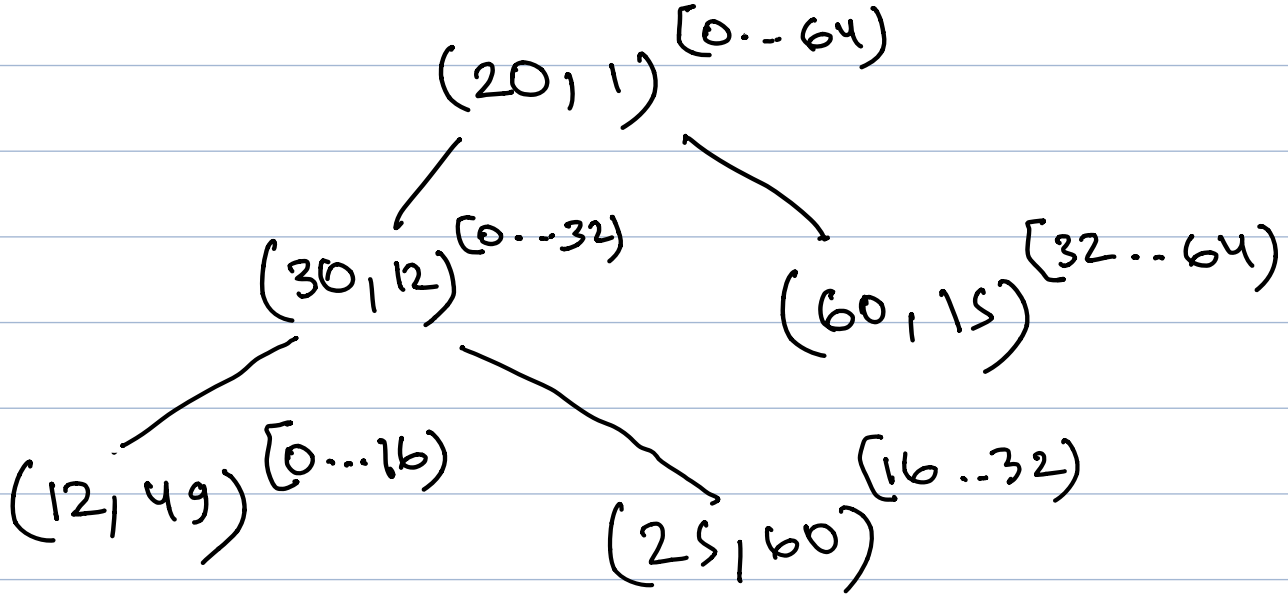
as (1 < 12)
(20,1) replaces
(30,12)
as (12 < 49)
(12,49) becomes
the left child of
(30,12)

insert (60, 15) →



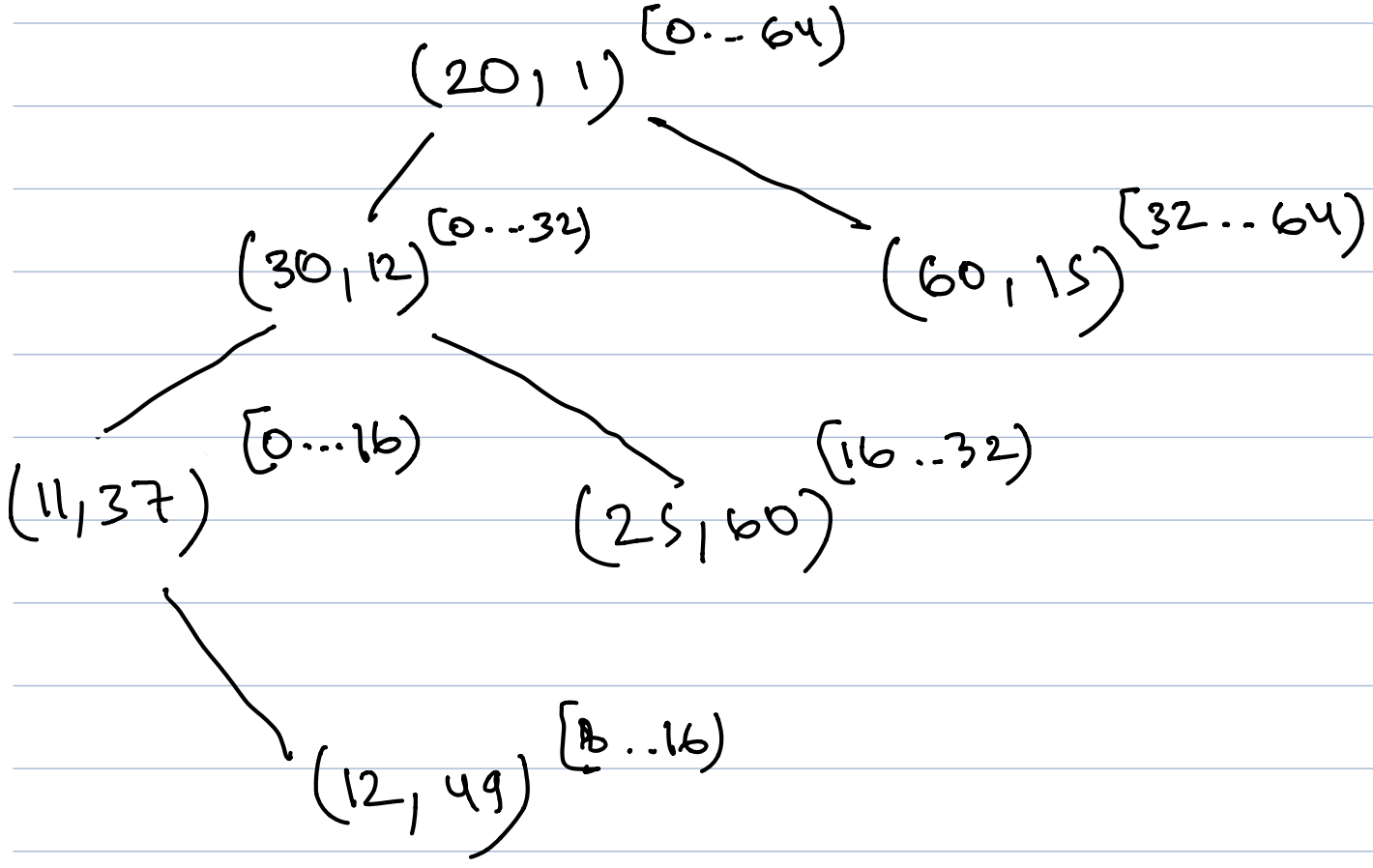
Here (32 < 60 < 63) Hence (60, 15) is right child.

insert (25, 60) →



(25, 60) is right child of (30, 12) as 16 < 25 < 32

insert (11,37) →

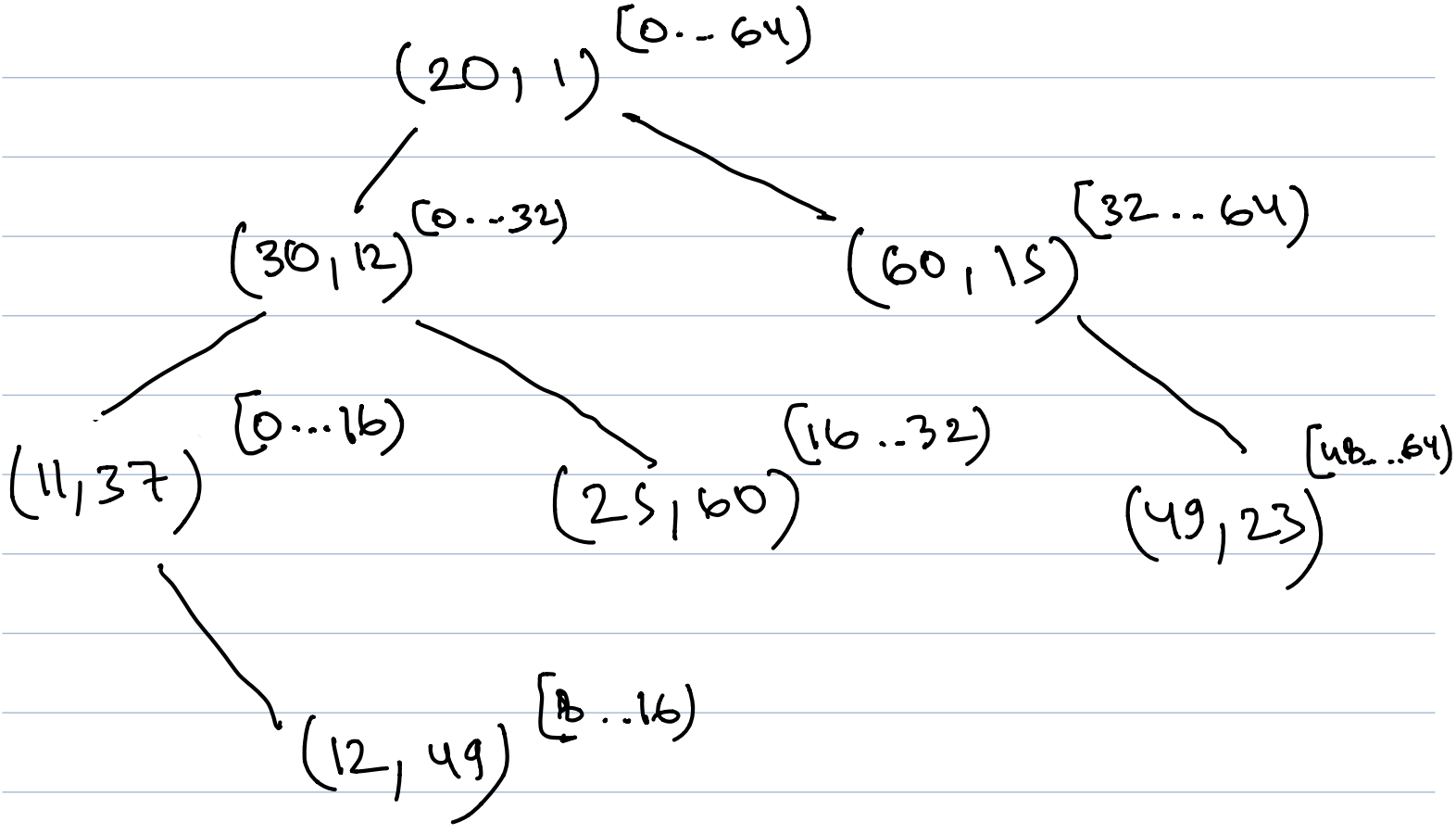


(11,37) replaces (12,49) as 37 < 49

and (12,49) becomes right child of

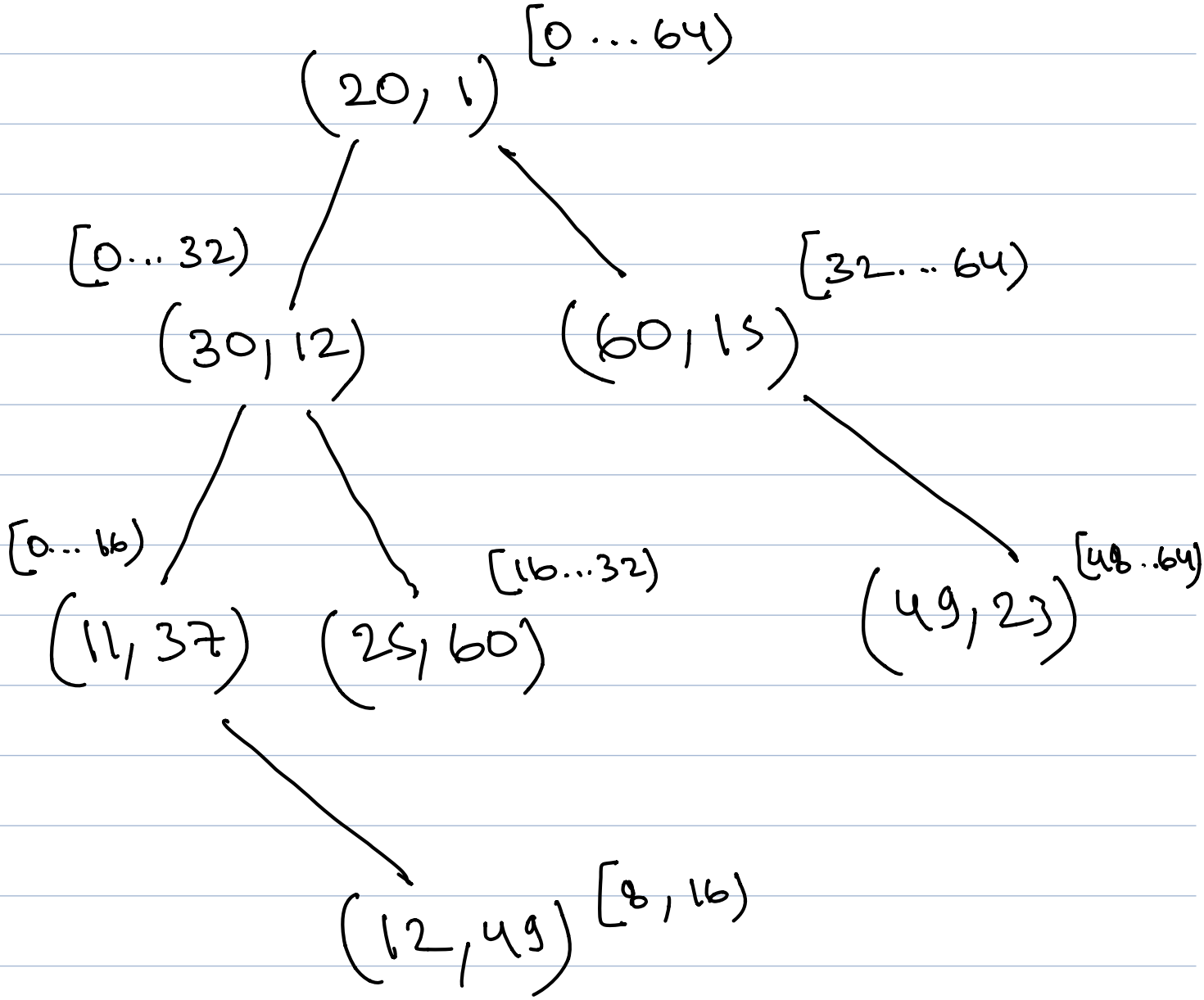
8 < 12 < 16.

insert (49, 23) =



(49, 23) becomes the right child of (60, 15)

final Min-RPST is ⇒



PROBLEM-2 -

(a) = Given that

$$n = 100,000$$

$$M = 5,000$$

$$u = 1000$$

To minimize P(u) we know that

$$h = 0.693 \times \frac{M}{u}$$

$$= 0.693 \times \frac{5000}{1000}$$

$$= 0.693 \times 5$$

$$= 3.465$$

Hence h could be either 3 or 4.

(b) = $P(u) = A * B$

where $A = \left(1 - \frac{1}{n}\right)^u$

$$B = \left(1 - \left(1 - \frac{1}{M}\right)^{uh}\right)^{h}$$

for $(h=3)$ we have,

$$P(u) = \left(1 - \frac{1}{100{,}000}\right)^{1000} \times \left(1 - \left(1 - \frac{1}{5000}\right)^{1000\times 3}\right)^{3}$$

$$= \left(\frac{99999}{100{,}000}\right)^{1000} \times \left(1 - \left(\frac{4999}{5000}\right)^{3000}\right)^{3}$$

$$= 0.99 \times (1 - 0.5487)^{3}$$

$$= 0.99 \times (0.4512)^{3}$$

$$= 0.99 \times 0.0918$$

$$= 0.0909 \quad (\text{when } h=3)$$

for $(h=4)$ we have,

$$P(4) = \left(1 - \frac{1}{100,000}\right)^{1000} \times \left(1 - \left(1 - \frac{1}{5000}\right)^{1000\times 4}\right)^4$$

$$= \left(\frac{99999}{100,000}\right)^{1000} \times \left(1 - \left(\frac{4999}{5000}\right)^{4000}\right)^4$$

$$\approx 0.99 \times (1 - 0.4492)^4$$

$$= 0.99 \times (0.5508)^4$$

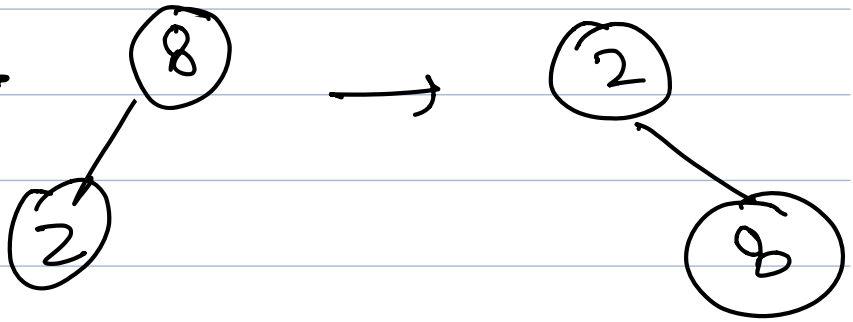
$$= 0.99 \times 0.09203$$

$$= 0.09111 \quad (\text{when } h=4)$$

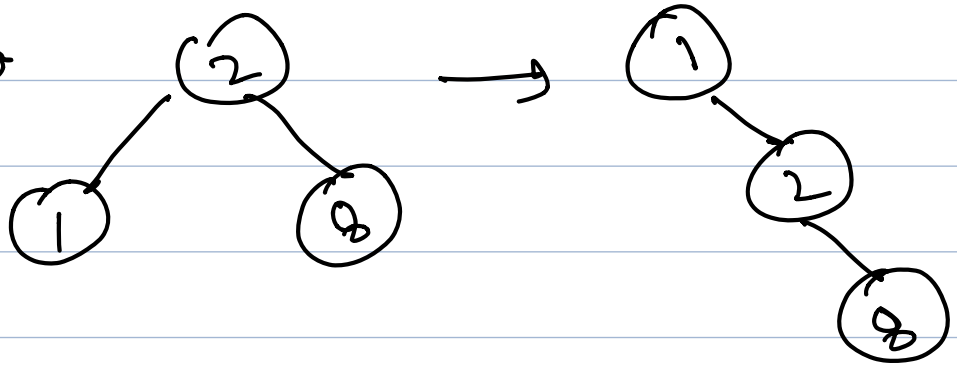
PROBLEM-3 =

insert 8 ⟶

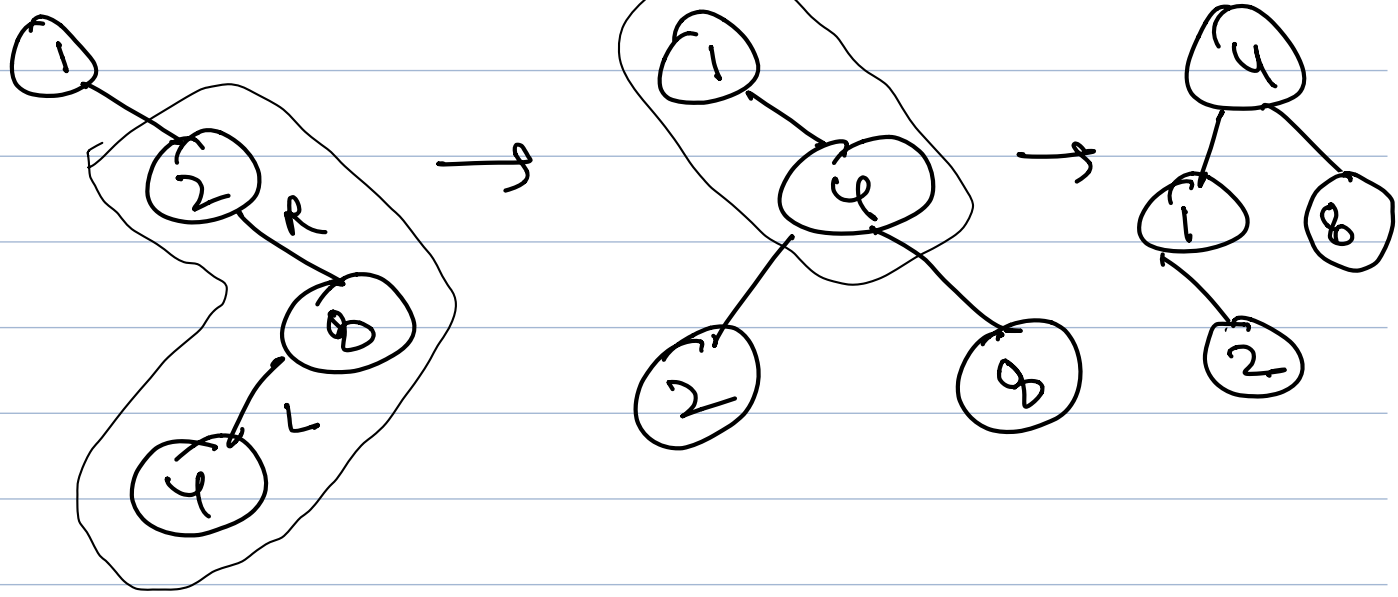
insert 2 ⟶



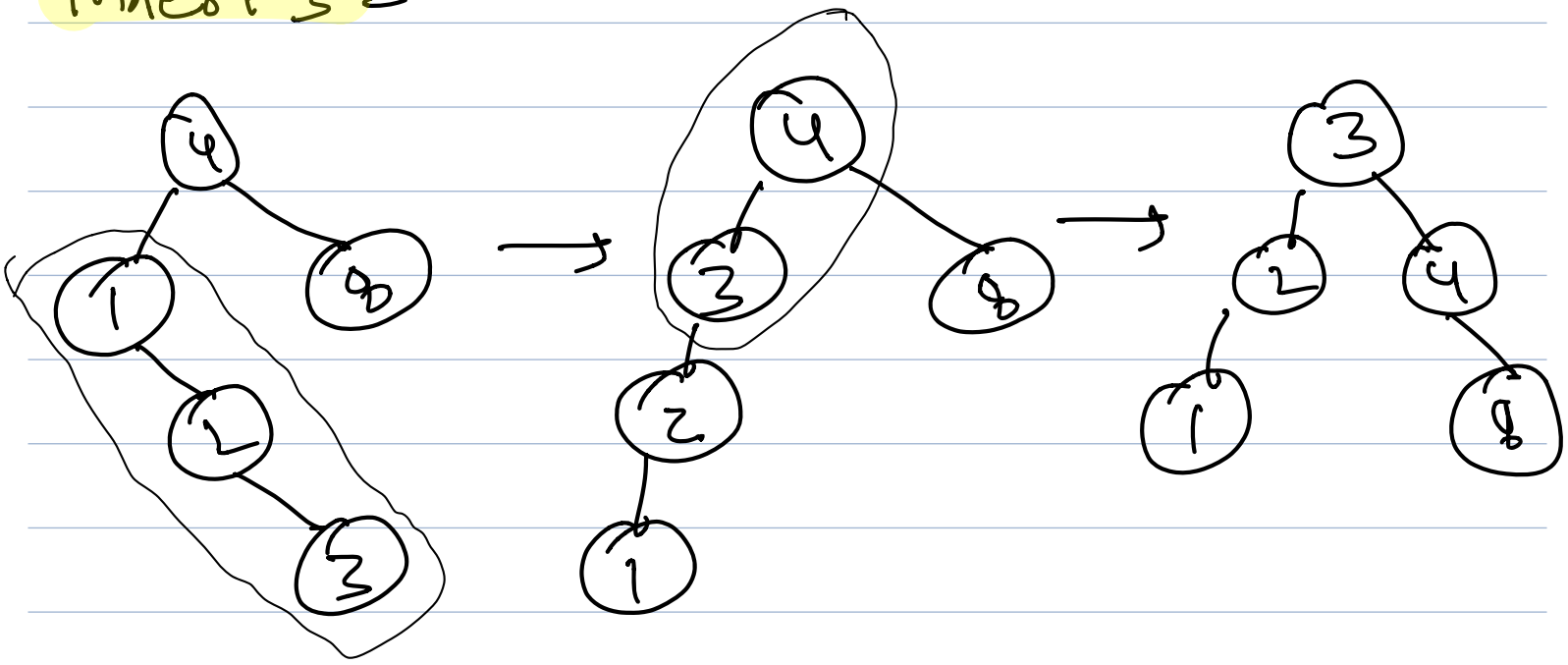
insert 1 ⟶



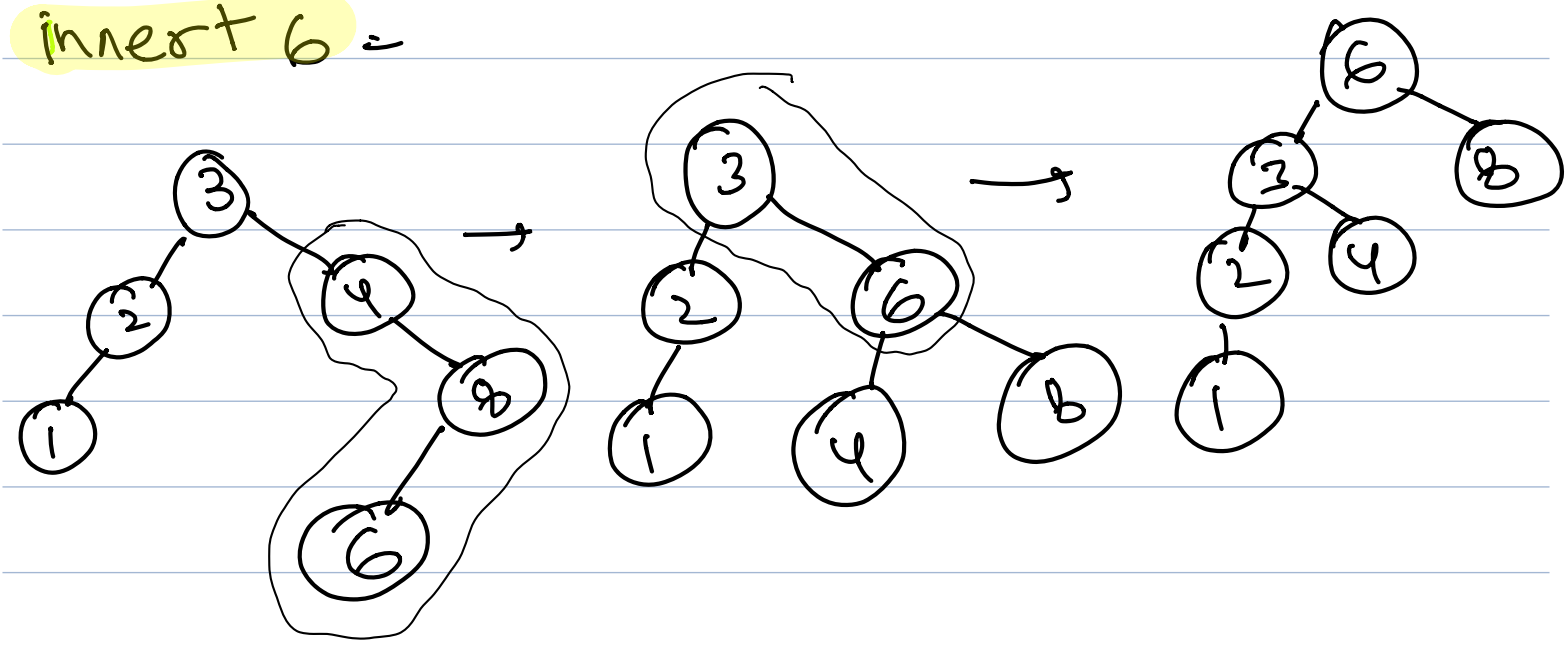
insert 4 ⟶



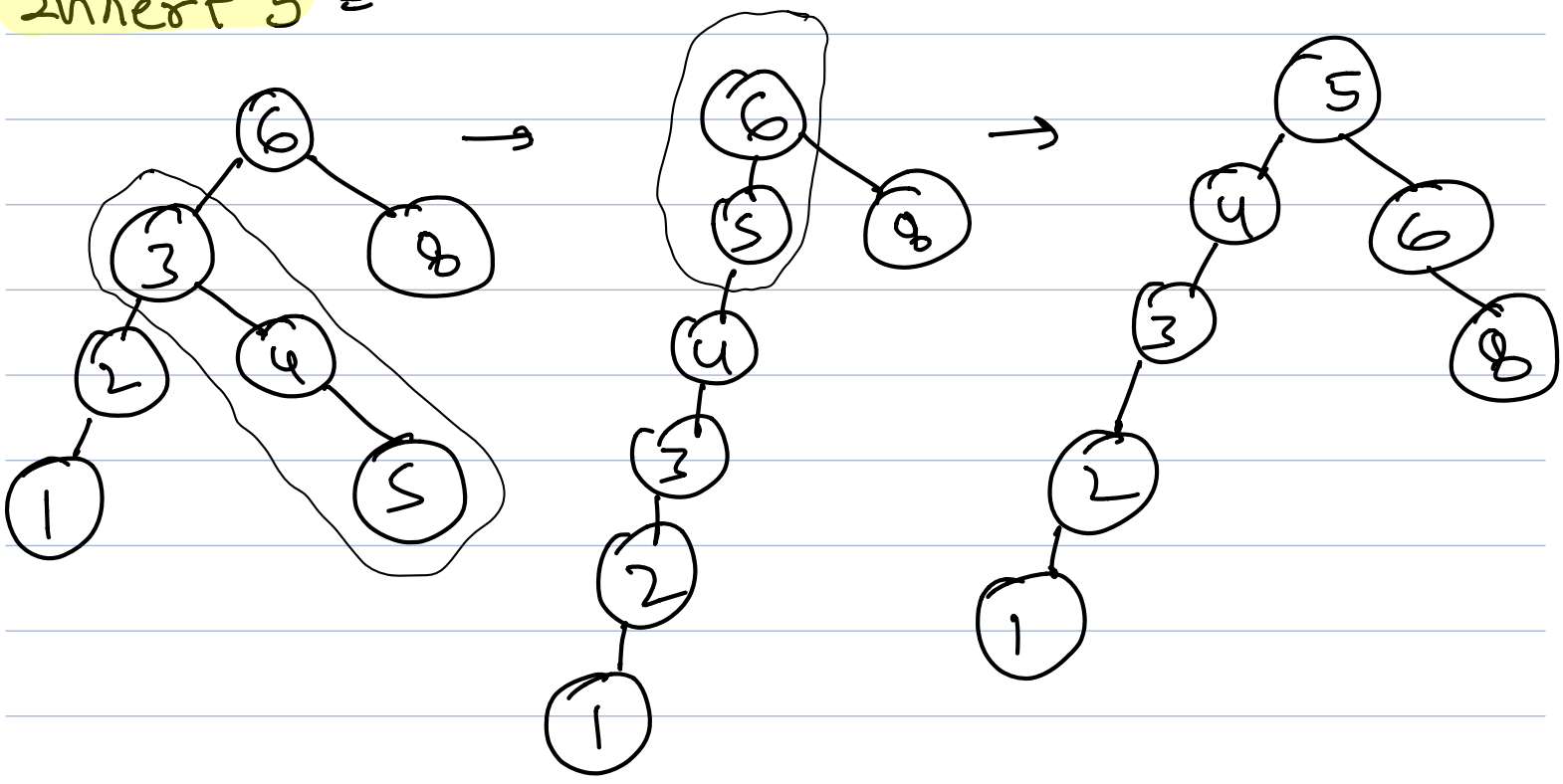
insert 3 =



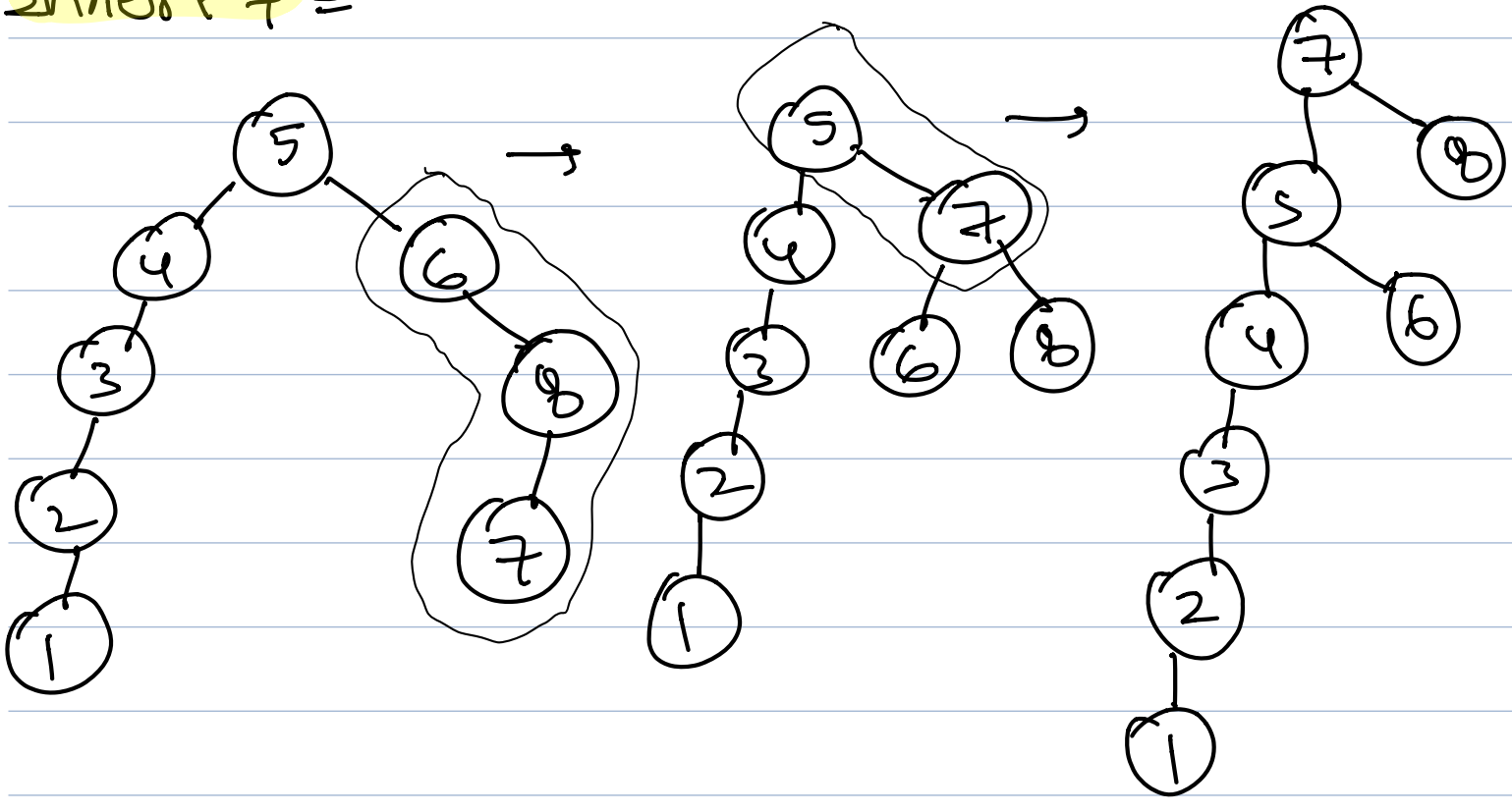
insert 6 =



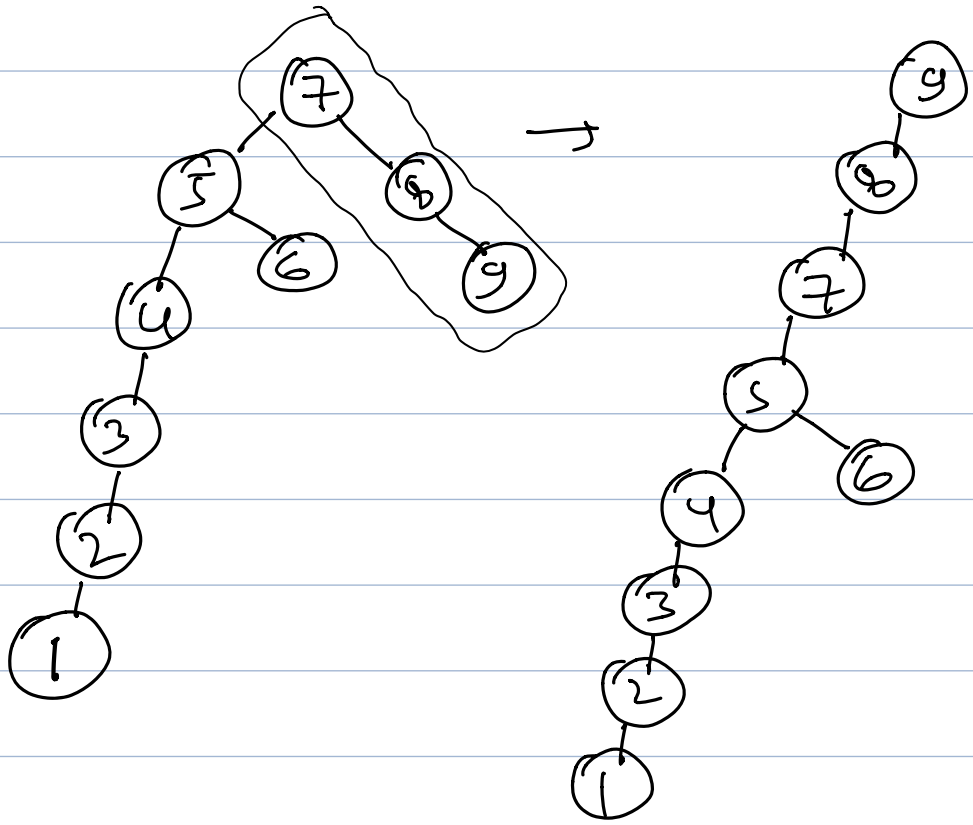
Innert 5 =



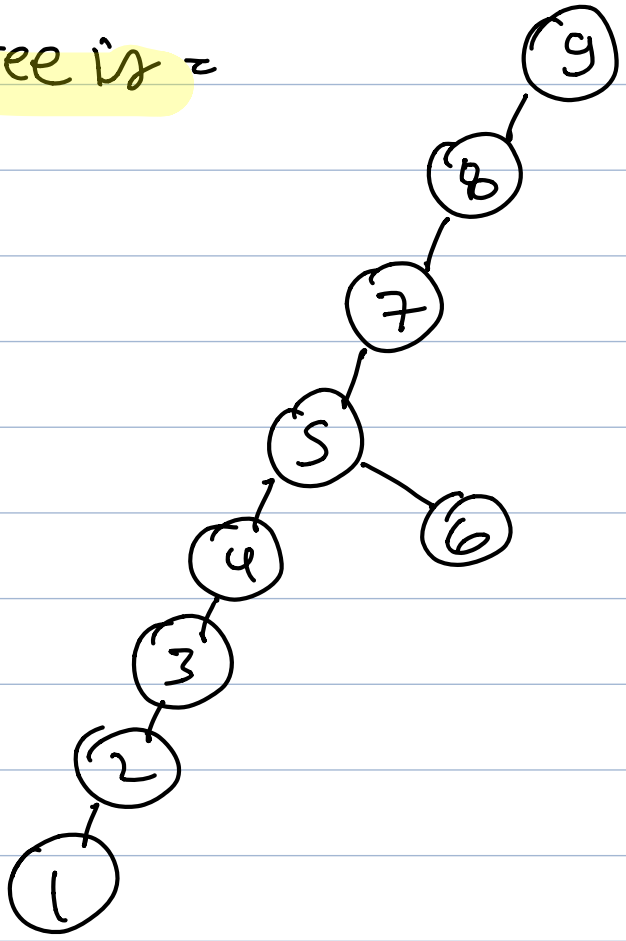
Innert 7 =



Insert 9 =



Resultant Splay tree is =



Part (b) = Multiple String Suffix tree:

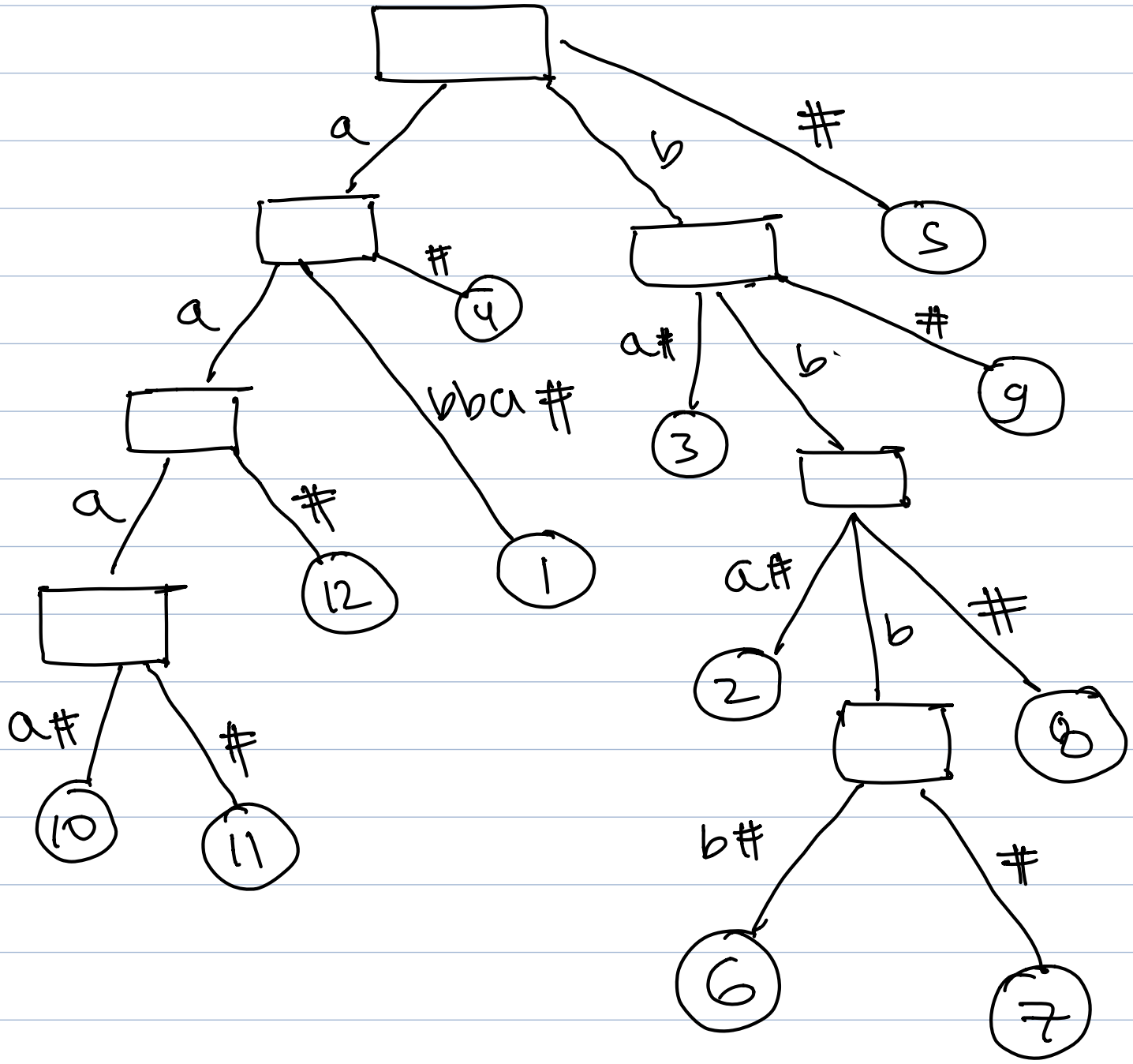
Given Strings abba, bbbb, aaaa

following are possible Substrings:

(1) = abba #	(6) = bbbb #	(10) = aaaa #
(2) = bba #	(7) = bbb #	(11) = aaa #
(3) = ba #	(8) = bb #	(12) = aa #
(4) = a #	(9) = b #	(13) = a #
(5) = #	(5) = #	(5) = #

(Please see below)

Multiple String Suffix tree:



PROBLEM 4 =

(a) = Deletion from the segment tree can be done using following pseudocode.

```
delete (s, e, v)
{
    if (s <= s(v) && e(v) <= e)
        delete (s, e, v) from v;

    else{
        if (s < (s(v) + e(v))/2)
            delete (s, e, v.leftchild)
        if (e > (s(v) + e(v))/2)
            delete (s, e, v.rightchild)
    }
}
```

Based on above code, check if the given node is contained in node, if no delete it in that node. if not, descend further and check whether it is contained or not.

Similarly check and find nodes where the given line segment is contained and delete it in that node.

Given a line segment interval we traverse through left and right boundary. left boundary is the number of nodes traversed in the path to reach left most leaf node. and right boundary is the number of nodes traversed in the path to reach right most leaf node. In worst case L, R all are ancestor of L & R & possibly the other child of each of three ancestor are reached.

So total number of nodes visited are (4logn) it is even less than 4logn as root is counted twice.

Hence the complexity to delete a line segment is O(logn) when delete from each node list can be done in O(1) time.

Part(b)=

let N be the set of nodes of line segment in interval [i, j], N doesnot contain 3 node that are at same level.

In order to prove that lets assume that line segment is stored in three nodes say x, y, z. in the same level. lets take the middle node y.

The sibling of y is also in the range between x and z or it can be either x or z. nodes are stored in x, y, z. or all of its leaf nodes are highlighted. In such case the sibling of y's leaf nodes should also be highlighted. In such case the sibling of y's leaf nodes should also be highlighted as it is also part of the line segment. so node should be stored in sibling of y also. But property of segment tree states that we can not store in a node and its sibling tree.

So this argument proves that the N can not contain 3 nodes that are at same level.