

Assignment 9: Symbol Tables
OCS25 Data Structures in Java Dr. Made
Vishak Srikanth
November 15 2021

Question 1 Symbol Tables	1
Trace of inserting the letters of the word EASYQUESTION into an initially empty table using SequentialSearchST	1
Trace of inserting the letters of the word EASYQUESTION into an initially empty table using BinarySearchST	2
Question 2: Binary Search Tree	2
Trace when we insert the keys M I D T E R M Q U E S T I O N, in that order (associating the value i with the ith key, as per the convention discussed in class) into an initially empty binary tree.	2
Sequence of BSTs that result when we delete the keys from the tree in question 2a, one by one, in the order they were inserted	8
Sequence of BSTs that result when we delete the keys from the tree in question 2.a, one by one, in alphabetical order.	14
Sequence of BSTs when deleting the keys from the tree in question 2a, one by one, successively deleting the root key.	20
Give the sequences of nodes examined when the methods in BST are used to compute each of the following quantities for the tree in question 2a	24
a. floor("P")	25
b. select(5)	25
c. ceiling("V")	26
d. rank("S")	26

Question 1 Symbol Tables

a) Trace of inserting the letters of the word EASYQUESTION into an initially empty table using SequentialSearchST

		Symbol Table											
key	value	First											
E	0	E 0											
A	1	A 1	->	E 0									
S	2	S 2	->	A 1	->	E 0							
Y	3	Y 3	->	S 2	->	A 1	->	E 0					
Q	4	Q 4	->	Y 3	->	S 2	->	A 1	->	E 0			
U	5	U 5	->	Q 4	->	Y 3	->	S 2	->	A 1	->	E 0	
E	6	U 5	->	Q 4	->	Y 3	->	S 2	->	A 1	->	E 6	
S	7	U 5	->	Q 4	->	Y 3	->	S 7	->	A 1	->	E 6	
T	8	T 8	->	U 5	->	Q 4	->	Y 3	->	S 7	->	A 1	->
I	9	I 9	->	T 8	->	U 5	->	Q 4	->	Y 3	->	S 7	->
O	10	O 10	->	I 9	->	T 8	->	U 5	->	Q 4	->	Y 3	->
N	11	N 11	->	O 10	->	I 9	->	T 8	->	U 5	->	Q 4	->

Legend:

Red Nodes are New

Black Nodes are Accessed during Sequential Search

Grey Nodes are untouched

The values in orange are changed values when the same key is found

Number of compares = # of black nodes = $0+1+2+3+4+5+6+4+6+7+8+9 = 55$

(counting for each step as each key is inserted into Symbol table above)

b) Trace of inserting the letters of the word EASYQUESTION into an initially empty table using BinarySearchST

		key[]										value[]										Number of Compares		
key	value	0	1	2	3	4	5	6	7	8	9	N	0	1	2	3	4	5	6	7	8	9	in rank() calls	additional in put() calls
E	0	E										1	0										0	0
A	1	A	E									2	1	0									1	0
S	2	A	E	S								3	1	0	2								2	0
Y	3	A	E	S	Y							4	1	0	2	3							2	0
Q	4	A	E	Q	S	Y						5	1	0	4	2	3						2	0
U	5	A	E	Q	S	U	Y					6	1	0	4	2	5	3					3	0
E	6	A	E	Q	S	U	Y					6	1	6	4	2	5	3					3	1
S	7	A	E	Q	S	U	Y					6	1	6	4	7	5	3					3	1
T	8	A	E	Q	S	T	U	Y				7	1	6	4	7	8	5	3				3	0
I	9	A	E	I	Q	S	T	U	Y			8	1	6	9	4	7	8	5	3			3	0
O	10	A	E	I	O	Q	S	T	U	Y		9	1	6	9	10	4	7	8	5	3		3	0
N	11	A	E	I	N	O	Q	S	T	U	Y	10	1	6	9	11	10	4	7	8	5	3	4	0
																						29	2	
Total Number of compares = 31																								

Total Number of compares = 31

Legend:

Entries in Red are new items inserted

Entries in Black are moved to the right

Entries in Grey did not move

Entries in orange are changed values

N= number of keys in binarySearchST

Note: Number of compares includes both compares in rank() and put()

(the compares in put() occur when we have same key appear multiple times)

Question 2: Binary Search Tree

a) Trace when we insert the keys M I D T E R M Q U E S T I O N, in that order (associating the value i with the i^{th} key, as per the convention discussed in class) into an initially empty binary tree.

The trace is color coded as follows:

- At each step the key and value being inserted into tree using put method are shown along with number of compares (cumulative)

2. The key inserted during the current step is shown in **RED** while the keys that were compared to place the key in its current position are shown in **bolded black**. The nodes that are not explored (i.e that are untouched) to place the current element are shown in dark grey
3. When the same key is encountered, the value for its node is updated to reflect the latest position where the letter was encountered and such changed values are marked in **orange**

BST after adding letter key: M with value: 0 Total Compares: 0

M:0

BST after adding letter key: I with value: 1 Total Compares: 1

M:0

/

I:1

BST after adding letter key: D with value: 2 Total Compares: 3

M:0

/

/

I:1

/

D:2

BST after adding letter key: T with value: 3 Total Compares: 4

M:0

/ \

/ \

I:1

T:3

/

D:2

BST after adding letter key: E with value: 4 Total Compares: 7

M:0

/ \

/ \

/ \

I:1

T:3

/

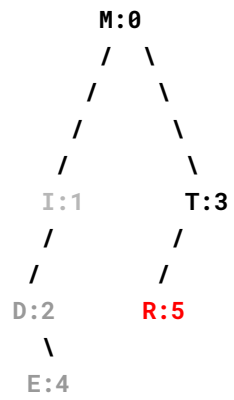
/

D:2

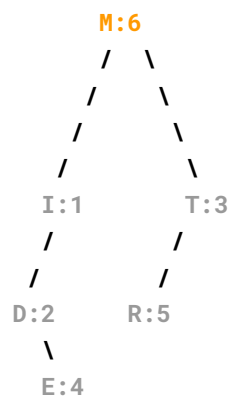
\

E:4

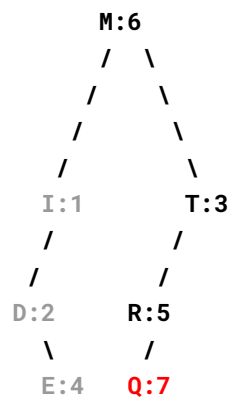
BST after adding letter key: R with value: 5 Total Compares: 9



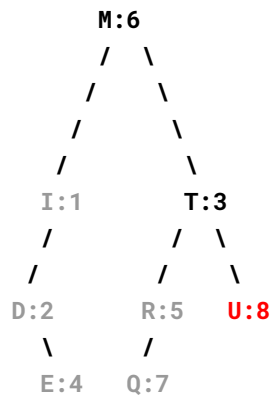
BST after adding letter key: M with value: 6 Total Compares: 10



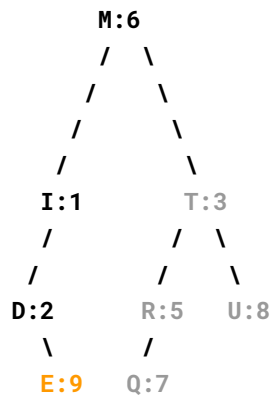
BST after adding letter key: Q with value: 7 Total Compares: 13



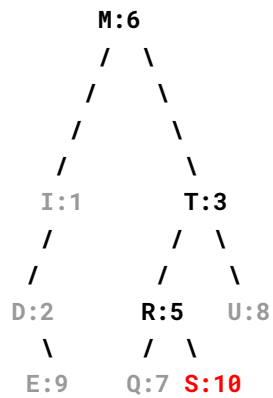
BST after adding letter key: U with value: 8 Total Compares: 15



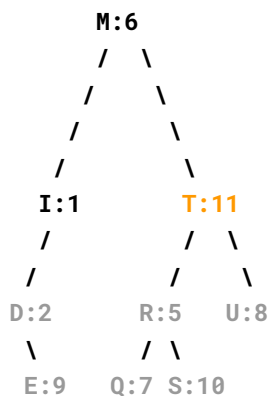
BST after adding letter key: E with value: 9 Total Compares: 19



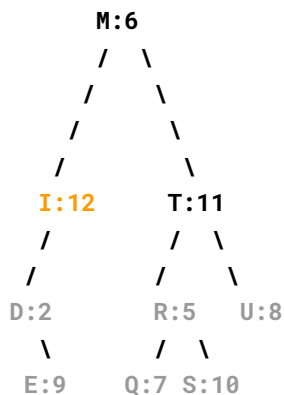
BST after adding letter key: S with value: 10 Total Compares: 22



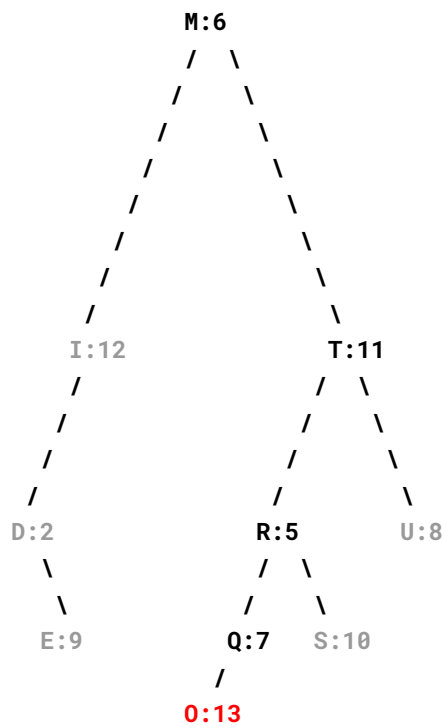
BST after adding letter key: T with value: 11 Total Compares: 24



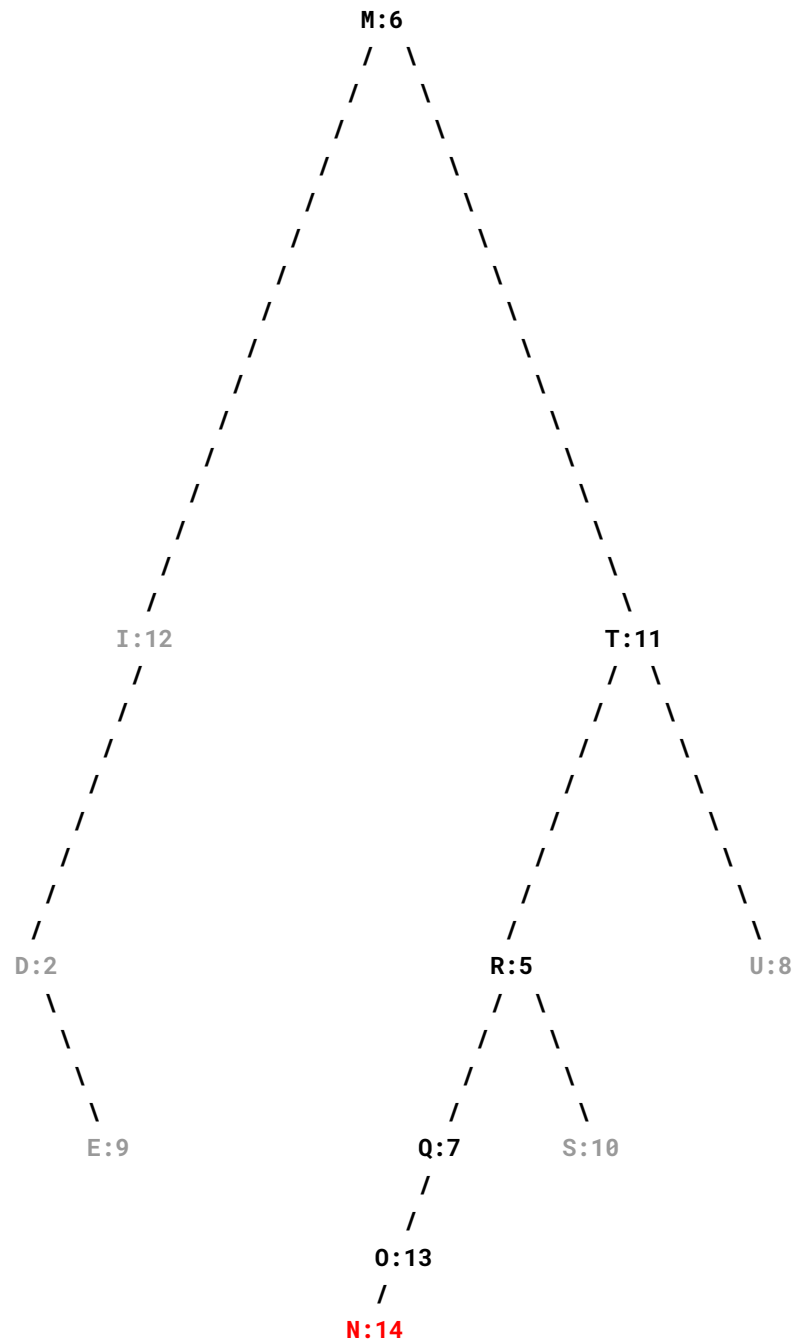
BST after adding letter key: I with value: 12 Total Compares: 26



BST after adding letter key: O with value: 13 Total Compares: 30



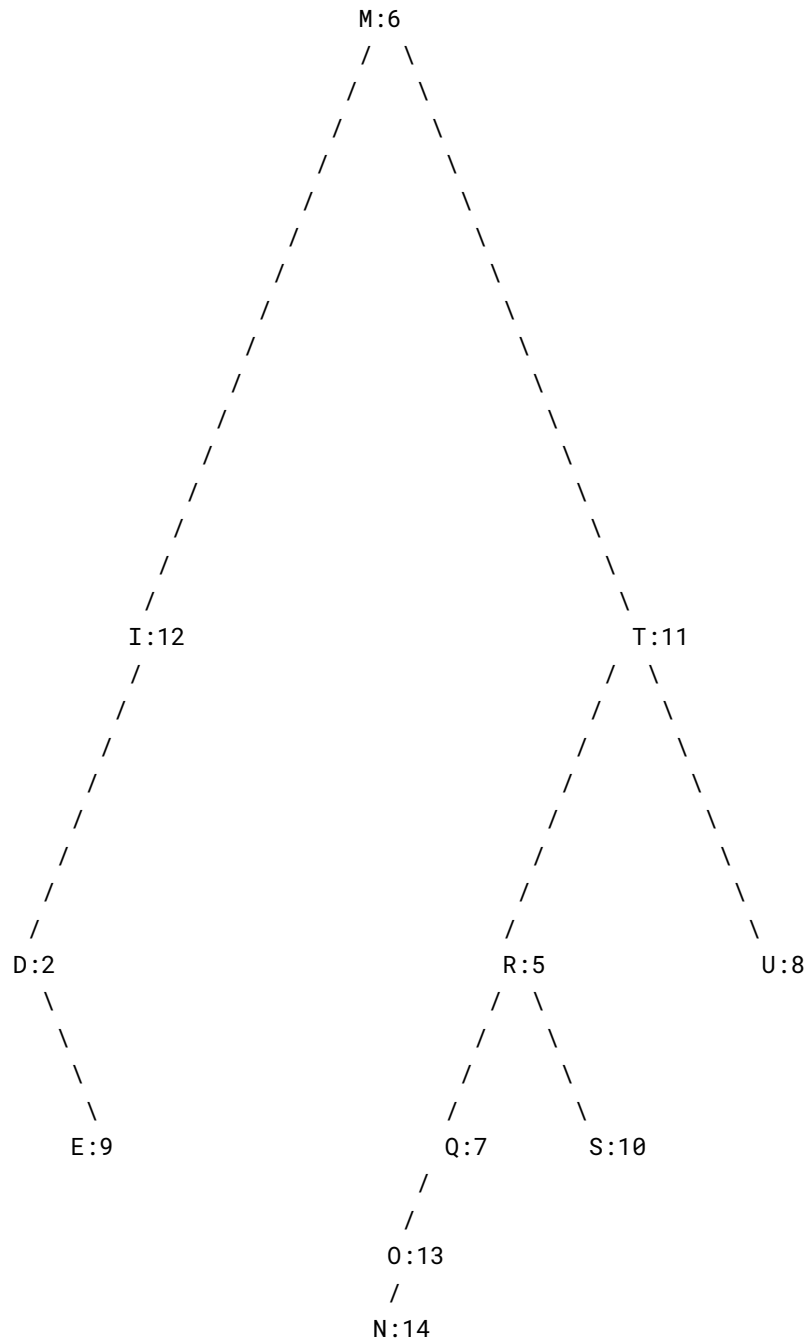
BST after adding letter key: N with value: 14 Total Compares: 35



As we can observe from the final resulting BST above there were **35 comparisons** done to place the letters of the word M I D T E R M Q U E S T I O N into the BST in the order in which they are present in the word.

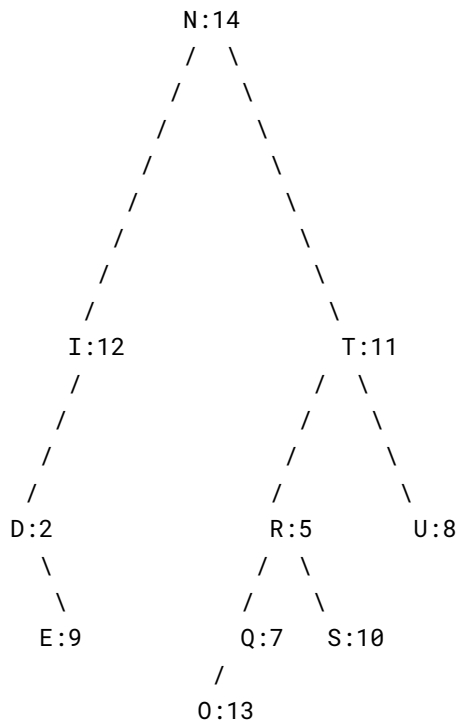
b) Sequence of BSTs that result when we delete the keys from the tree in question 2a, one by one, in the order they were inserted

Final BST after adding all letters in order they appear in: MIDTERMQUESTION

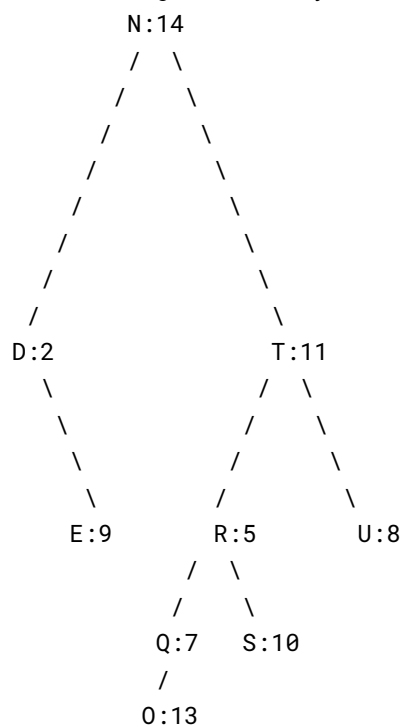


 BST sequence when deleting letters one by one in the order they were first inserted
 from MIDTERMQUESTION

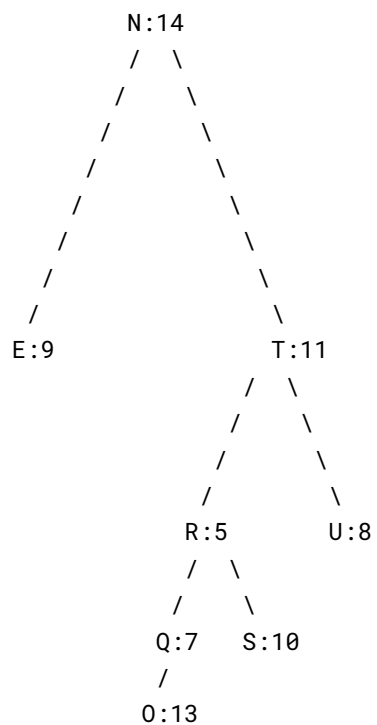
 BST after deleting letter key: M



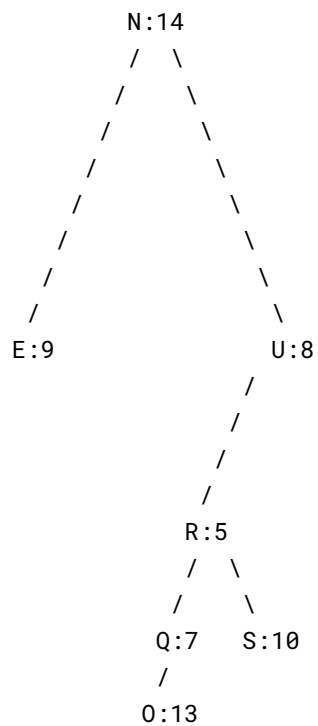
 BST after deleting letter key: I



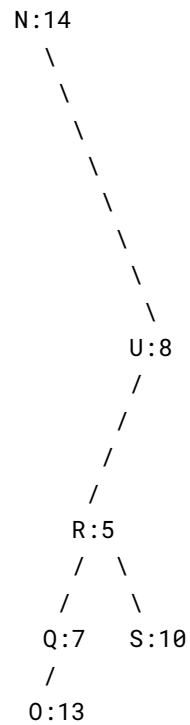
BST after deleting letter key: D



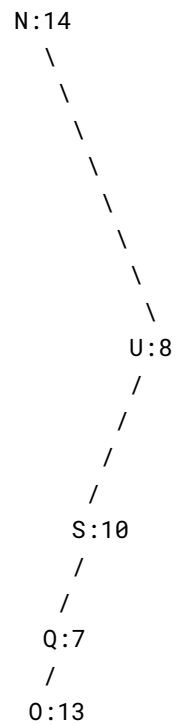
BST after deleting letter key: T



BST after deleting letter key: E



BST after deleting letter key: R



A vertical zigzag line with labels at the ends and midpoints. The labels are: N:14 at the top, U:8 on the right side, S:10 on the left side, Q:7 on the left side, and O:13 at the bottom.

N:14
U:8
S:10
O:13

$$\begin{array}{c} \text{N:14} \\ \backslash \\ \backslash \\ \text{S:10} \\ / \\ \text{O:13} \end{array}$$

BST after deleting letter key: E

```

N:14
 \
  \
   S:10
  /
 O:13

```

BST after deleting letter key: S

```

N:14
 \
  O:13

```

BST after deleting letter key: T

```

N:14
 \
  O:13

```

BST after deleting letter key: I

```

N:14
 \
  O:13

```

BST after deleting letter key: O

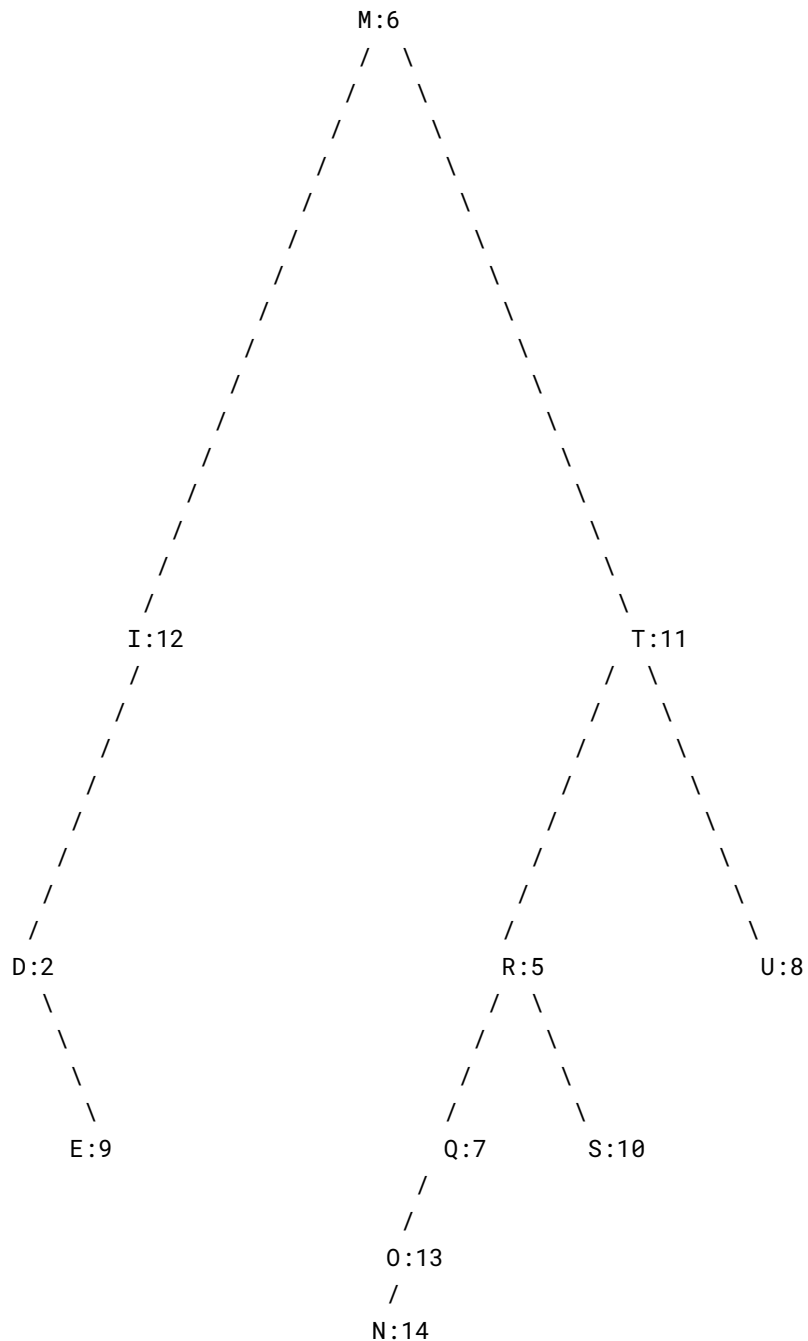
N:14

BST after deleting letter key: N

Tree is Empty!

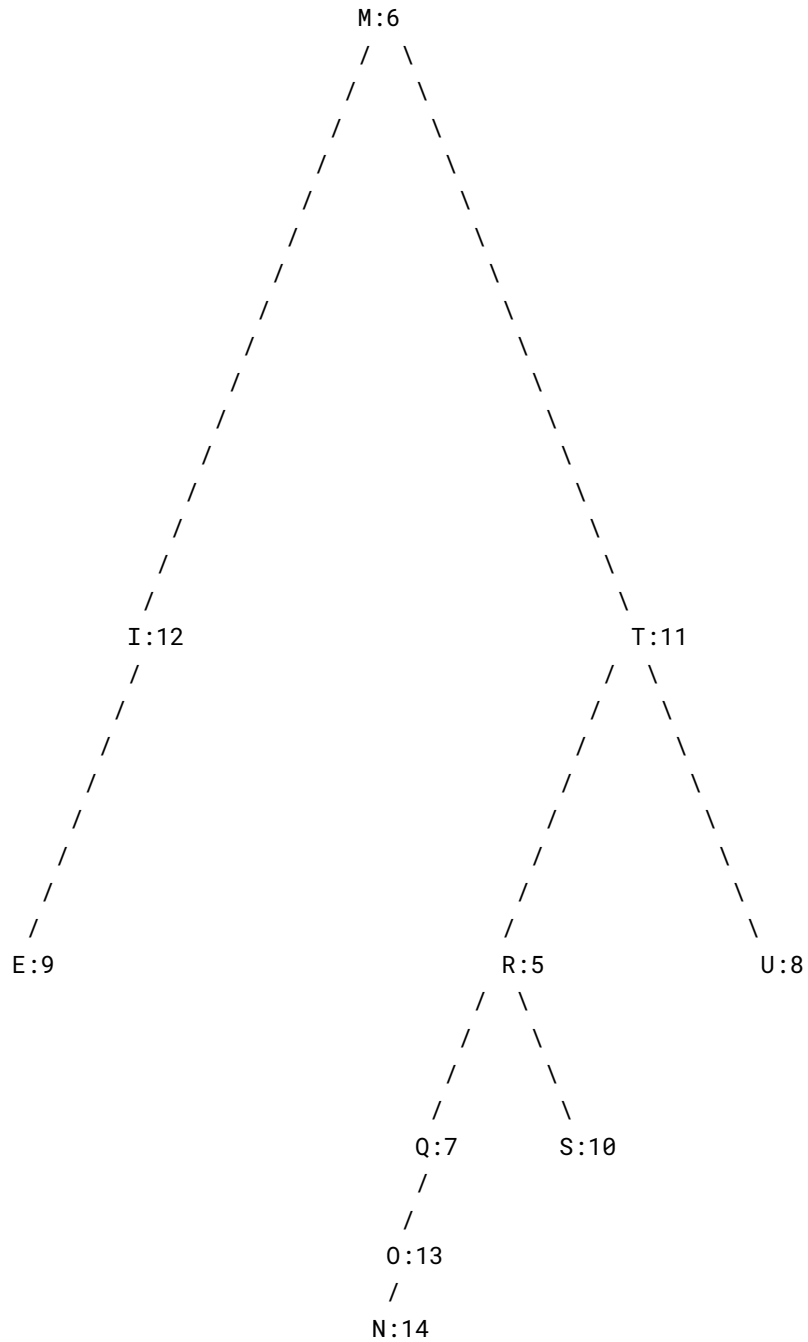
c) Sequence of BSTs that result when we delete the keys from the tree in question 2.a, one by one, in alphabetical order.

Starting BST from 2(a) after adding all letters in order they appear in MIDTERMQUESTION

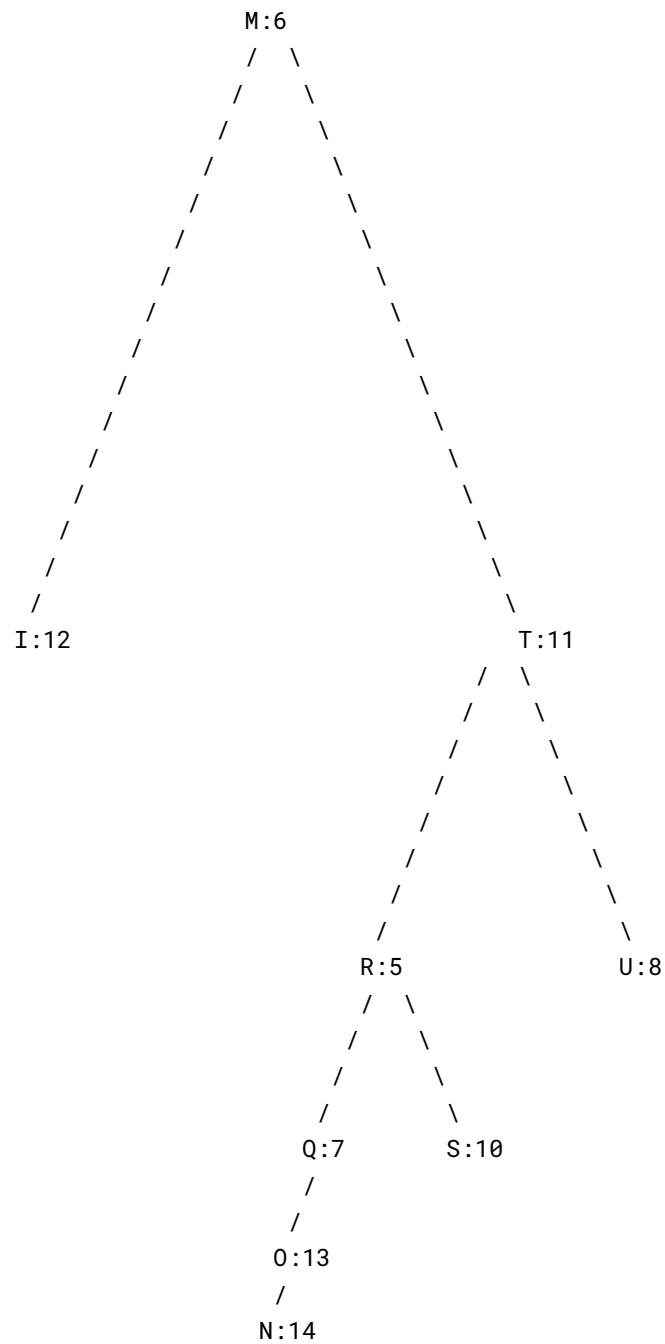


 BST sequence when deleting letters one by one in alphabetical order : [D, E, I, M, N,
 O, Q, R, S, T, U]

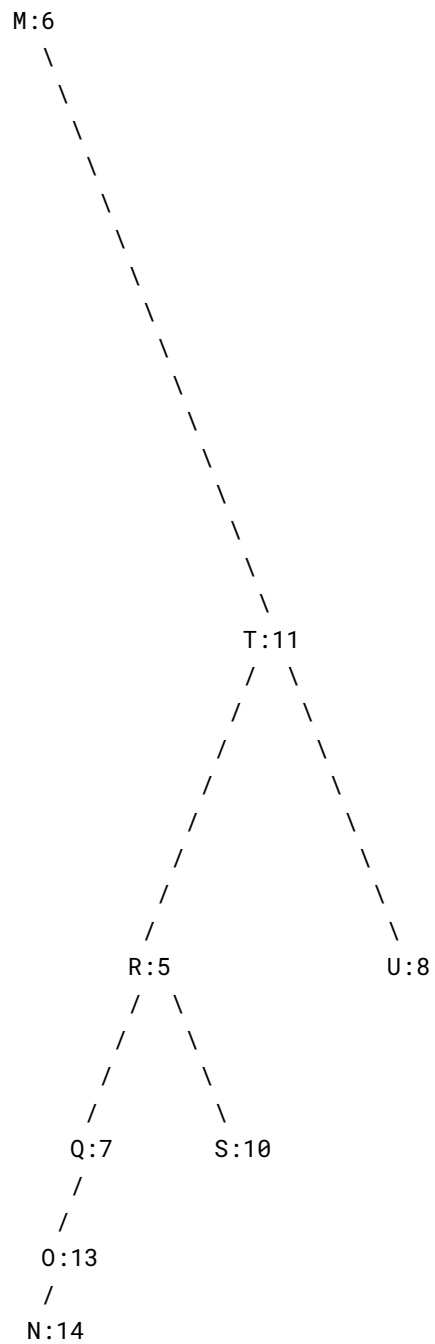
BST after deleting letter key: D



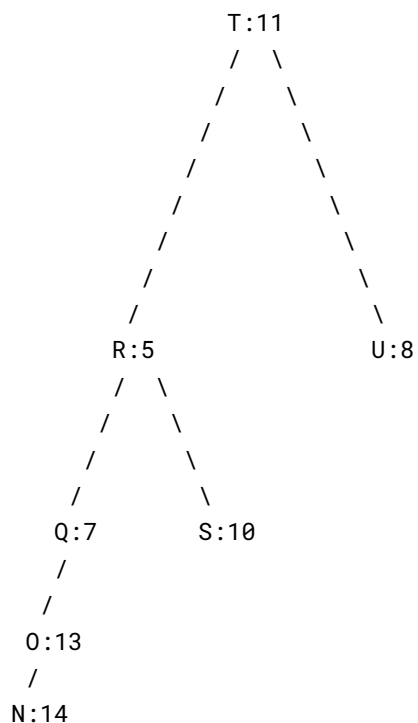
BST after deleting letter key: E



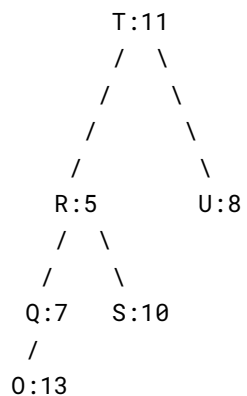
BST after deleting letter key: I



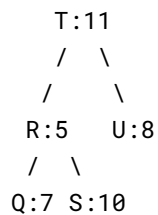
BST after deleting letter key: M



BST after deleting letter key: N



BST after deleting letter key: O



BST after deleting letter key: Q

```

  T:11
 /  \
/    \
R:5   U:8
 \
  S:10

```

BST after deleting letter key: R

```

  T:11
 /  \
S:10 U:8

```

BST after deleting letter key: S

```

  T:11
   \
    U:8

```

BST after deleting letter key: T

U:8

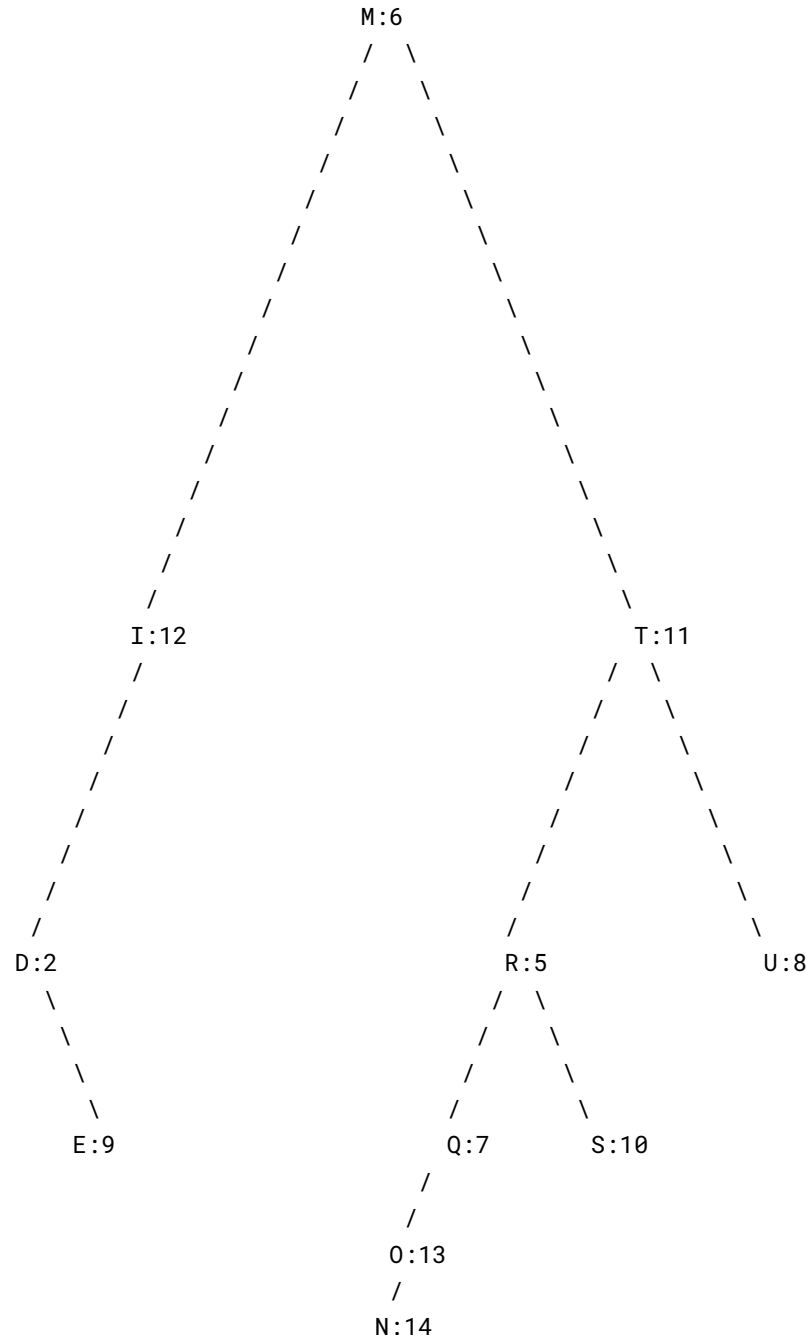
BST after deleting letter key: U

Tree is Empty!

d) Sequence of BSTs when deleting the keys from the tree in question 2a, one by one, successively deleting the root key.

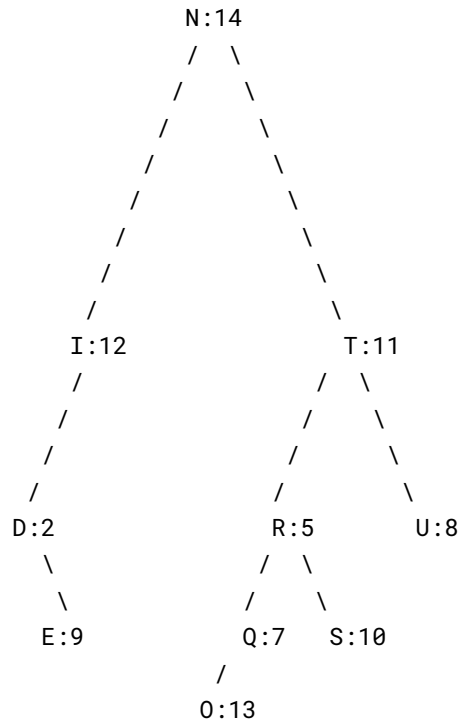
The logic for the deletion is to find the successor element and replace the deleted root with its successor. The successor is the **smallest element in the right subtree** and if there is no right subtree we find the **largest element from the left subtree** and replace the root element being deleted with this successor element. We show the tree that results from deleting the root element successively at each step

Starting BST from 2(a) after adding all letters in order they appear in MIDTERMQUESTION

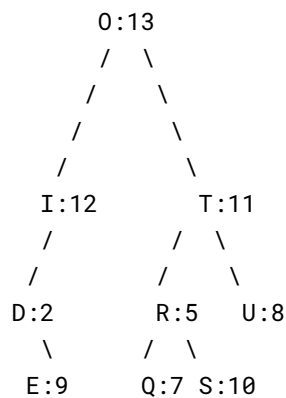


BST sequence when deleting letters by successively deleting the key at the root

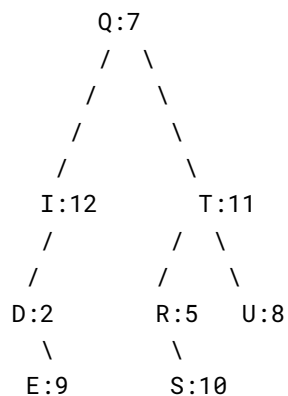
BST after deleting letter at the root: M



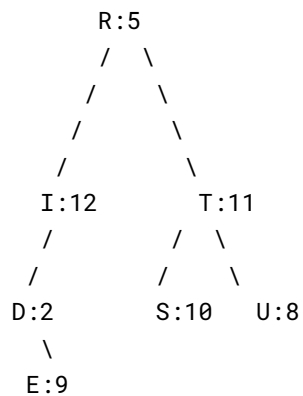
BST after deleting letter at the root: N



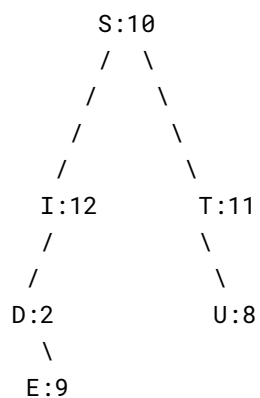
BST after deleting letter at the root: 0



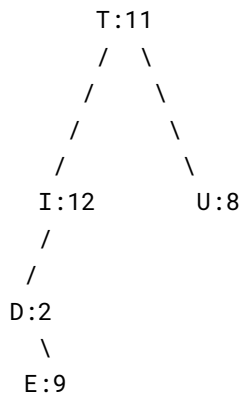
BST after deleting letter at the root: Q



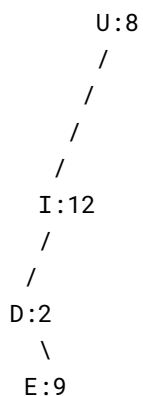
BST after deleting letter at the root: R



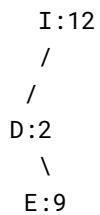
BST after deleting letter at the root: S



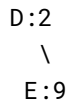
BST after deleting letter at the root: T



BST after deleting letter at the root: U



BST after deleting letter at the root: I



BST after deleting letter at the root: D

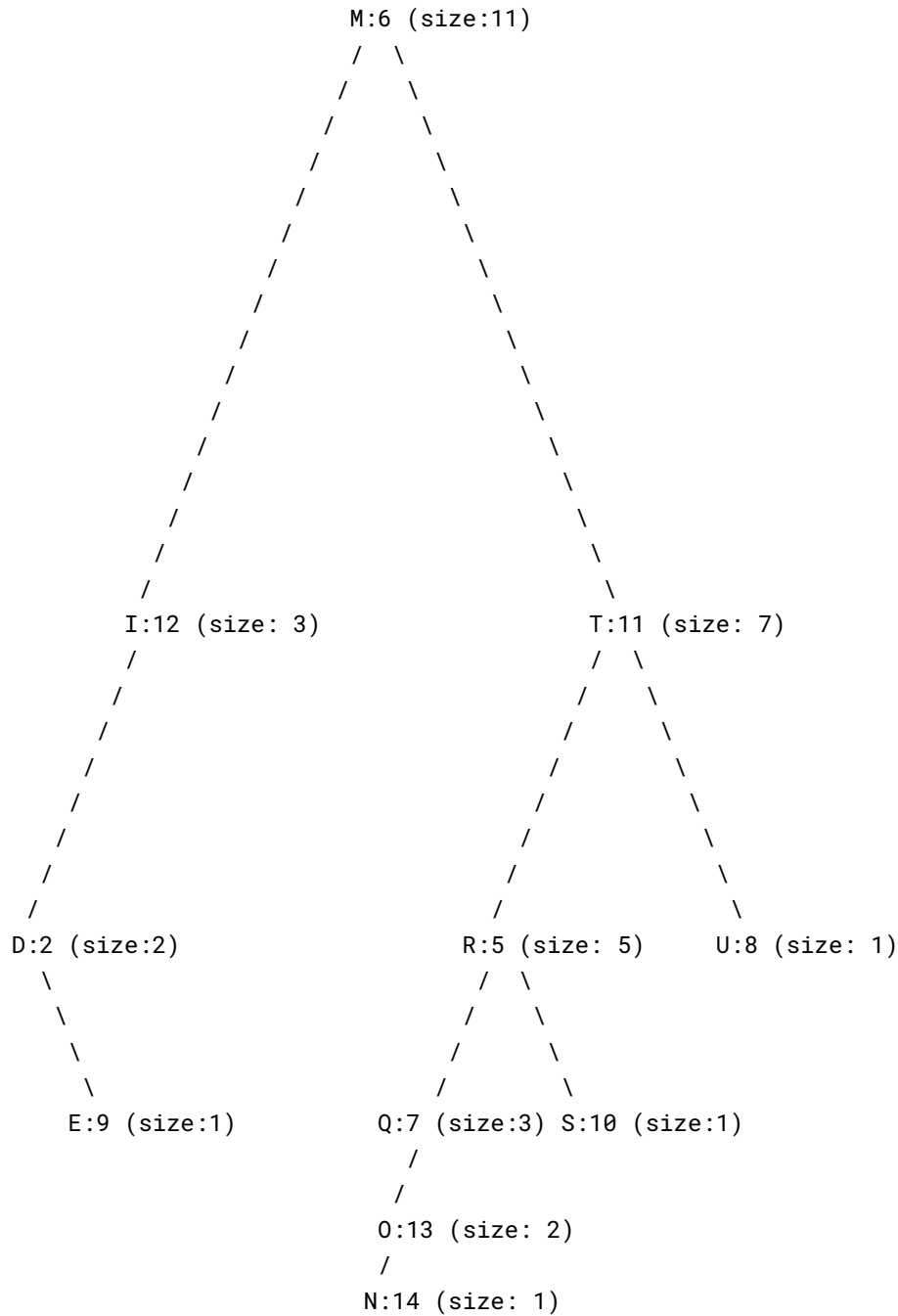
E:9

BST after deleting letter at the root: E

Tree is Empty!

e) Give the sequences of nodes examined when the methods in BST are used to compute each of the following quantities for the tree in question 2a

Starting BST from 2(a) after adding all letters in order they appear in MIDTERMQUESTION



a. floor("P")

When we call floor("P"), we start with root and compare "P" with the key of root node "M". If the element we are searching for "P" is less than "M" (this means the compare function on key yields < 0), we examine the left subtree. If the comparison yields > 0 we examine the right subtree. We continue to recursively examine each subtree until we find a match or we are at just last element just below "P" which has no right subtree (which is "O"). With the BST from 2(a), we will walk down from M \rightarrow T \rightarrow R \rightarrow Q \rightarrow O and yield "O" as the result.

```
*****
BST sequence of nodes examined when calling: floor("P")
[M, T, R, Q, O]
The result from calling: floor(P):
0
*****
```

b. select(5)

When we call select(5), we want to get the key in the symbol table of a given rank which has the property that there are 5 keys in the symbol table that are smaller. In other words, the key we are looking for is the 6th smallest key in the symbol table. We start with root and look first at size of the left subtree. If the left subtree's size > 5 we recursively look at the size of its left subtree until we find a node whose rank matches. If the size of left subtree < 5 we look at the right subtree for remaining size (i.e.) rank - leftSize - 1. We continue this recursive process at each node until we find the desired node with rank of 5. With the BST from 2(a), we will walk down from M \rightarrow T \rightarrow R \rightarrow Q \rightarrow O and yield "O" as the result since it is the 6th smallest element (Elements D, E, I, M, and N are smaller than O in this tree).

```
*****
BST sequence of nodes examined when calling: select(5)
[M, T, R, Q, O]
The result from calling: select(5):
0
*****
```

c. ceiling("V")

When we call ceiling("P"), we start with root and compare "V" with the key of root node "M". If the element we are searching for "V" is less than "M" (this means the compare function on key yields < 0), we examine the left subtree. If the comparison yields > 0 we examine the right subtree. We continue to recursively examine each subtree until we find a match or we are at the highest element in the BST which has no right subtree (which is "U"). With the BST from 2(a), we will walk down from M- \rightarrow T- \rightarrow U and we cannot find any element that is larger than the one we are searching for so it will result in an error since we have overshot the largest element. In this case we need to print an error message that the element whose ceiling we are looking for, "V" in this case, is larger than the largest element in the BST, "U" in this case, so no ceiling can be found as a result.

```
*****
BST sequence of nodes examined when calling: ceiling("V")
[M, T, U]
V is larger than the largest element in BST: U
No ceiling can be found!
*****
```

d. rank("S")

When we call rank("S"), we want to get the count of the number of keys in the symbol table which are strictly smaller than "S". In other words, the key we are looking for is the number of the elements which have a smaller key in the symbol table than "S". We start with root and compare the root with the key we are searching for. If the key is already at the root element we simply return the size of the left subtree. If the element whose rank we are searching for, "S" in this case, is less than root (this means the compare function on key yields < 0), we examine the left subtree. If the comparison yields > 0 we examine the right subtree. When we examine the right subtree we just add the size of left subtree + 1 to the rank recursive call to account for the exact rank of the final element in the tree. We continue to recursively examine each subtree until we find a match. With the BST from 2(a), we will walk down from M- \rightarrow T- \rightarrow R- \rightarrow S. S has a rank of 8 since there are 8 elements smaller than it in the BST (D, E, I, M, N, O, Q, R)

We will get:

```
rank(S)    = size(M.leftsubtree) + 1 + rank(T)
            = 3 + 1 + rank(T.leftsubtree)
            = 3 + 1 + size(R.leftsubtree) + 1 (since right subtree has S has only element)
            = 3 + 1 + 3 + 1
            = 8
```

BST sequence of nodes examined when calling: rank("S")

[M, T, R, S]

The result from calling: rank("S"):

8
