

```

1 // Shamim Bin Zahid
2 // Lab 05 | FGP
3 // Roll: 43
4
5 #include <stdio.h>
6 #include <string.h>
7 #include <stdlib.h>
8 #define MAX_CHAR 1000
9
10 char text[100];
11 Node *root = NULL;
12 Node *lastNewNode = NULL;
13 Node *activeNode = NULL;
14 int count=0;
15 int activeEdge = -1;
16 int activeLength = 0;
17 int remainingSuffixCount = 0;
18 int leafEnd = -1;
19 int *rootEnd = NULL;
20 int *splitEnd = NULL;
21 int size = -1;
22
23 struct SuffixTreeNode {
24     struct SuffixTreeNode *children[MAX_CHAR];
25     struct SuffixTreeNode *suffixLink;
26     int start;
27     int *end;
28     int suffixIndex;
29 };
30
31 typedef struct SuffixTreeNode Node;
32
33 Node *newNode(int start, int *end){
34     count++;
35     Node *node =(Node*) malloc(sizeof(Node));
36     int i;
37     for (i = 0; i < MAX_CHAR; i++)
38         node->children[i] = NULL;
39     node->suffixLink = root;
40     node->start = start;
41     node->end = end;
42     node->suffixIndex = -1;
43     return node;
44 }
45
46 int edgeLength(Node *n) {
47     return *(n->end) - (n->start) + 1;
48 }
49
50 int walkDown(Node *currNode){
51     if (activeLength >= edgeLength(currNode)){
52         activeEdge = (int)text[activeEdge+edgeLength(currNode)]-(int)' ';
53         activeLength -= edgeLength(currNode);
54         activeNode = currNode;
55         return 1;
56     }
57     return 0;
58 }
59

```

```

60 void extendSuffixTree(int pos){
61     leafEnd = pos;
62     remainingSuffixCount++;
63     lastNewNode = NULL;
64     while(remainingSuffixCount > 0) {
65         if (activeLength == 0) {
66             activeEdge = (int)text[pos]-(int)' ';
67         }
68         if (activeNode->children[activeEdge] == NULL){
69             activeNode->children[activeEdge] = newNode(pos, &leafEnd);
70             if (lastNewNode != NULL) {
71                 lastNewNode->suffixLink = activeNode;
72                 lastNewNode = NULL;
73             }
74         }
75         else {
76             Node *next = activeNode->children[activeEdge];
77             if (walkDown(next)) {
78                 continue;
79             }
80             if (text[next->start + activeLength] == text[pos]) {
81                 if(lastNewNode != NULL && activeNode != root) {
82                     lastNewNode->suffixLink = activeNode;
83                     lastNewNode = NULL;
84                 }
85                 activeLength++;
86                 break;
87             }
88             splitEnd = (int*) malloc(sizeof(int));
89             *splitEnd = next->start + activeLength - 1;
90
91             Node *split = newNode(next->start, splitEnd);
92             activeNode->children[activeEdge] = split;
93
94             split->children[(int)text[pos]-(int)' '] =
95                 newNode(pos, &leafEnd);
96             next->start += activeLength;
97             split->children[activeEdge] = next;
98
99             if (lastNewNode != NULL) {
100                 lastNewNode->suffixLink = split;
101             }
102             lastNewNode = split;
103         }
104         remainingSuffixCount--;
105         if (activeNode == root && activeLength > 0) {
106             activeLength--;
107             activeEdge = (int)text[pos - remainingSuffixCount + 1]-(int)' ';
108         }
109         else if (activeNode != root) {
110             activeNode = activeNode->suffixLink;
111         }
112     }
113 }
114
115 void printSequence(int i, int j) {
116     for (int k = i; k <= j; k++)
117         printf("%c", text[k]);
118 }
119

```

```

120 void setSuffixIndexByDFS(Node *n, int labelHeight) {
121     if (n == NULL) return;
122     if (n->start != -1) {
123         printSequence(n->start, *(n->end));
124     }
125     int leaf = 1;
126     for (int i = 0; i < MAX_CHAR; i++) {
127         if (n->children[i] != NULL) {
128             if (leaf == 1 && n->start != -1) printf(" [%d]\n", n->suffixIndex);
129             leaf = 0;
130             setSuffixIndexByDFS(n->children[i], labelHeight + edgeLength(n->children[i]));
131         }
132     }
133     if (leaf == 1) {
134         n->suffixIndex = size - labelHeight;
135         printf(" [%d]\n", n->suffixIndex);
136     }
137 }
138
139 void freeSuffixTreeByPostOrder(Node *n) {
140     if (n == NULL) return;
141     for (int i = 0; i < MAX_CHAR; i++)
142         if (n->children[i] != NULL)
143             freeSuffixTreeByPostOrder(n->children[i]);
144     if (n->suffixIndex == -1) free(n->end);
145     free(n);
146 }
147
148 void generateSuffixTree() {
149     size = strlen(text);
150     rootEnd = (int*) malloc(sizeof(int));
151     *rootEnd = - 1;
152     root = newNode(-1, rootEnd);
153     activeNode = root;
154     for (int i=0; i<size; i++)
155         extendSuffixTree(i);
156     int labelHeight = 0;
157     setSuffixIndexByDFS(root, labelHeight);
158     freeSuffixTreeByPostOrder(root);
159 }
160
161 int main(int argc, char *argv[]) {
162     strcpy(text, "ATCATGTCATG");
163     generateSuffixTree();
164     printf("Number of nodes in suffix tree are %d\n",count);
165     return 0;
166 }
167

```