

Suffix_Array_and_LCP_Array.cpp

```

1  /**
2   *   author:  devesh95
3   *
4   **/
5  #include <bits/stdc++.h>
6  using namespace std;
7
8  #define int          long long int
9  #define double       long double
10 #define F            first
11 #define S            second
12 #define pb           push_back
13 #define lb           lower_bound
14 #define ub           upper_bound
15 #define si           set <int>
16 #define vi           vector <int>
17 #define vvi          vector <vi>
18 #define pii          pair <int, int>
19 #define vpi          vector <pii>
20 #define mii          map <int, int>
21 #define sz(v)        ((int) v.size())
22 #define form(i, a, b) for (int i=a; i<(b); i++)
23 #define forn(i, a)    for (int i=0; i<(a); i++)
24
25 ///////////////////////////////////////////////////////////////////
26 // Suffix Array and LCP Array - Detailed Explanation
27 //
28 // Purpose: Efficiently handle string processing tasks such as
29 //          substring search, pattern matching, and finding
30 //          the longest repeated substring.
31 //
32 // -----
33 // Problem Statement:
34 // Given a string `s`, construct the suffix array and the LCP
35 // (Longest Common Prefix) array.
36 //
37 // -----
38 // Steps:
39 // 1. Build the Suffix Array:
40 //    - A suffix array is an array of integers giving the starting positions
41 //      of suffixes of a string in lexicographical order.
42 //    - For example, for the string "banana", the suffix array is [5, 3, 1, 0, 4, 2].
43 //      This means the suffixes starting at indices 5, 3, 1, 0, 4, and 2 are in
44 //      lexicographical order.
45 //
46 //    - To build the suffix array:
47 //      a) Initialize the suffix array with the indices of the string.
48 //      b) Sort the suffixes based on the first character.
49 //      c) Iteratively sort the suffixes based on the first 2^k characters.
50 //         This is done by comparing pairs of ranks and updating the ranks
51 //         after each iteration.

```

```

52 //
53 // 2. Build the LCP Array using Kasai's Algorithm:
54 //   - The LCP array stores the lengths of the longest common prefixes
55 //     between consecutive suffixes in the suffix array.
56 //   - For example, for the string "banana" with suffix array [5, 3, 1, 0, 4, 2],
57 //     the LCP array is [0, 1, 3, 0, 0, 2].
58 //
59 //   - To build the LCP array:
60 //     a) Initialize the rank array, which stores the rank of each suffix.
61 //     b) Iterate through the string and compute the LCP values by comparing
62 //        characters of the suffixes and using previously computed LCP values
63 //        to optimize the process.
64 //
65 // -----
66 // Applications:
67 //   - Pattern matching in strings: Quickly find occurrences of a pattern.
68 //   Example: Using binary search on the suffix array to find the pattern.
69 //
70 //   - Finding the longest repeated substring: Identify repeated sequences.
71 //   Example: The maximum value in the LCP array gives the length of the longest
72 //   repeated substring.
73 //
74 //   - Solving various string-related problems efficiently:
75 //   Example: Finding the number of distinct substrings, finding the lexicographical
76 //   order of substrings, etc.
77 //
78 // Tips:
79 // i) Suffix array construction can be done in  $O(n \log n)$  time using a combination
80 //    of sorting and rank updating techniques.
81 //
82 // ii) LCP array construction using Kasai's algorithm is  $O(n)$  and leverages the
83 //     previously computed suffix array and rank array to efficiently compute
84 //     the LCP values.
85 ///////////////////////////////////////////////////////////////////
86
87 // Function to build the suffix array
88 vector<int> buildSuffixArray(const string &s) {
89     int n = s.size();
90     vector<int> suffixArray(n), rank(n), temp(n);
91
92     // Initialize suffix array and rank array
93     for (int i = 0; i < n; ++i) {
94         suffixArray[i] = i;
95         rank[i] = s[i];
96     }
97
98     // Sort suffixes based on first  $2^k$  characters
99     for (int k = 1; k < n; k *= 2) {
100         auto cmp = [&](int a, int b) {
101             if (rank[a] != rank[b])
102                 return rank[a] < rank[b];
103             int ra = (a + k < n) ? rank[a + k] : -1;
104             int rb = (b + k < n) ? rank[b + k] : -1;
105             return ra < rb;

```

```
106     };
107     sort(suffixArray.begin(), suffixArray.end(), cmp);
108
109     // Update rank array based on sorted suffixes
110     temp[suffixArray[0]] = 0;
111     for (int i = 1; i < n; ++i) {
112         temp[suffixArray[i]] = temp[suffixArray[i - 1]] + cmp(suffixArray[i - 1],
suffixArray[i]);
113     }
114     rank = temp;
115 }
116
117 return suffixArray;
118 }
119
120 // Function to build the LCP array using Kasai's algorithm
121 vector<int> buildLCPArray(const string &s, const vector<int> &suffixArray) {
122     int n = s.size();
123     vector<int> rank(n), lcp(n);
124
125     // Build rank array from suffix array
126     for (int i = 0; i < n; ++i) {
127         rank[suffixArray[i]] = i;
128     }
129
130     int h = 0;
131     // Build LCP array
132     for (int i = 0; i < n; ++i) {
133         if (rank[i] > 0) {
134             int j = suffixArray[rank[i] - 1];
135             while (i + h < n && j + h < n && s[i + h] == s[j + h]) {
136                 ++h;
137             }
138             lcp[rank[i]] = h;
139             if (h > 0) --h;
140         }
141     }
142
143     return lcp;
144 }
145
146 void solve() {
147     string s;
148     cin >> s;
149     int n = s.size();
150
151     // Build suffix array and LCP array
152     vector<int> suffixArray = buildSuffixArray(s);
153     vector<int> lcpArray = buildLCPArray(s, suffixArray);
154
155     cout << n << ' ';
156     // Output the suffix array
157     for (int i = 0; i < n; ++i) {
158         cout << suffixArray[i] << " ";
```

```
159     }
160     cout << endl;
161
162     // Output the LCP array
163     for (int i = 0; i < n; ++i) {
164         cout << lcpArray[i] << " ";
165     }
166     cout << endl;
167 }
168
169 int32_t main() {
170     ios_base::sync_with_stdio(0); cin.tie(0); cout.tie(0);
171
172     #ifndef ONLINE_JUDGE
173         freopen("input.txt", "r", stdin);
174         freopen("output.txt", "w", stdout);
175     #endif
176     clock_t z = clock();
177     int t = 1;
178     //cin >> t;
179     while (t--) {
180         solve();
181     }
182     cerr << "Run Time : " << ((double)(clock() - z) / CLOCKS_PER_SEC);
183     return 0;
184 }
```