

<b>Name</b>	Anand Tiwari
<b>UID no.</b>	2021700068
<b>Experiment No.</b>	4

<b>AIM:</b>	Implement Dynamic Programming -Longest Common Subsequence.
<b>Program</b>	
<b>PROBLEM STATEMENT:</b>	Implement Longest Common Subsequence code using Dynamic Programming.
<b>ALGORITHM/ THEORY:</b>	<p>The longest common subsequence (LCS) is defined as the longest subsequence that is common to all the given sequences, provided that the elements of the subsequence are not required to occupy consecutive positions within the original sequences.</p> <p>If X and Y are the two given sequences then, Z is the common subsequence of X and Y if Z is a subsequence of both X and Y. Furthermore, Z must be a <b>strictly increasing sequence</b> of the indices of both X and Y.</p> <p>Example 1: if</p> <p><math>X = \{B, C, D, A, A, C, D\}</math></p> <p>Then, <math>\{A, D, B\}</math> cannot be a subsequence of X as the order of the elements is not the same (i.e. not strictly increasing sequence).</p> <p>Example 2: if</p> <p><math>X = \{B, C, D, A, A, C, D\}</math></p> <p><math>Y = \{A, C, D, B, A, C\}</math></p> <p>Then, the common subsequences are <math>\{B, C\}</math>, <math>\{C, D, A, C\}</math>, <math>\{D, A, C\}</math>, <math>\{A, A, C\}</math>, <math>\{A, C\}</math>, <math>\{C, D\}</math>, ...</p> <p>Among these subsequences, <math>\{C, D, A, C\}</math> is the longest common subsequence. We are going to find this longest common subsequence using dynamic programming.</p> <p>Dynamic Programming Steps:</p> <p>1] Optimal substructure of an LCS:</p>

Let  $X = \langle x_1, x_2, \dots, x_m \rangle$  and  $Y = \langle y_1, y_2, \dots, y_n \rangle$  be sequences, and let  $Z = \langle z_1, z_2, \dots, z_k \rangle$  be any LCS of  $X$  and  $Y$ .

1. If  $x_m = y_n$ , then  $z_k = x_m = y_n$  and  $Z_{k-1}$  is an LCS of  $X_{m-1}$  and  $Y_{n-1}$ .
2. If  $x_m \neq y_n$ , then  $z_k \neq x_m$  implies that  $Z$  is an LCS of  $X_{m-1}$  and  $Y$  or  $z_k \neq y_n$  implies that  $Z$  is an LCS of  $X$  and  $Y_{n-1}$ .

Thus, the longest common subsequence tells us that an LCS of two sequences contains within it an LCS of prefixes of the two sequences. Thus, the LCS problem has an optimal-substructure property.

2] A recursive solution:

$$\begin{aligned} c[i,j] &= 0 && \text{if } i = 0 \text{ or } j = 0 \\ &= c[i-1,j-1] + 1 && \text{if } i, j > 0 \text{ and } x_i = y_j \\ &= \max(c[i,j-1], c[i-1,j]) && \text{if } i, j > 0 \text{ and } x_i \neq y_j \end{aligned}$$

3] Computing the length of an LCS: Memorize the solutions in a matrix/table.

4] Constructing an LCS: Backtrace the table to find the LCS.

The method of dynamic programming reduces the number of function calls. It stores the result of each function call so that it can be used in future calls without the need for redundant calls.

In the above dynamic algorithm, the results obtained from each comparison between elements of  $X$  and the elements of  $Y$  are stored in a table so that they can be used in future computations.

So, the time taken by a dynamic approach is the time taken to fill the table (i.e.  $O(m*n)$ ). Whereas, the recursion algorithm has the complexity of  $2^{\max(m,n)}$ .

## PROGRAM:

```
#include <bits/stdc++.h>
using namespace std;

/* Returns length of LCS for X[0..m-1], Y[0..n-1] */
int lcs(char* X, char* Y, int m, int n,
        vector<vector<int>> & dp)
{
    if (m == 0 || n == 0)
        return 0;

    if (X[m - 1] == Y[n - 1])
        return dp[m][n] = 1 + lcs(X, Y, m - 1, n - 1, dp);
```

```

        if (dp[m][n] != -1) {
            return dp[m][n];
        }
        return dp[m][n] = max(lcs(X, Y, m, n - 1, dp),
                               lcs(X, Y, m - 1, n, dp));
    }

    /* Driver code */
    int main()
    {
        char X[] = "AGGTAB";
        char Y[] = "GXTXAYB";

        int m = strlen(X);
        int n = strlen(Y);
        vector<vector<int> > dp(m + 1, vector<int>(n + 1, -1));
        cout << "Length of LCS is " << lcs(X, Y, m, n, dp);

        return 0;
    }

```

## RESULT:

```

Enter string 1: abcdef
Enter string 2: acbcf
The Longest Common Subsequence is abcf

```

## WORKING:

$S_1 = [a b c d e f]$ ,  $S_2 = [a c b e f]$

	$j_2$	a	b	c	d	e	f
$i_1 =$	0	0	0	0	0	0	0
a	0	1	1	1	1	1	1
c	0	1	1	2	2	2	2
b	0	1	2	2	2	2	2
c	0	1	2	3	3	3	3
f	0	1	2	3	3	3	4

abc f

#### CONCLUSION:

We used dynamic programming approach to find the Longest Common Subsequence between two strings. We also analyzed what are the steps of dynamic programming. We saw how Dynamic programming method to solve LCS is better than Recursive method to solve LCS.