Aim 🡪 Longest common subsequence

The Longest Common Subsequence (LCS) problem is a classic problem in computer science and is often encountered in applications such as version control systems, diff tools, and bioinformatics. The goal is to find the longest subsequence that is common to two sequences.

**Problem Definition**

* **Input**: Two sequences (strings) **X** and **Y**.
* **Output**: The length of the longest common subsequence.

**Dynamic Programming Approach**

The dynamic programming approach to solve the LCS problem involves maintaining a 2D array **dp** where **dp[i][j]** represents the length of the longest common subsequence of the first **i** characters of **X** and the first **j** characters of **Y**.

**Steps:**

1. Initialize a 2D array **dp** of size **(m+1) x (n+1)** with all values set to 0, where **m** is the length of **X** and **n** is the length of **Y**.
2. Iterate through each character of both strings:
   * If **X[i-1] == Y[j-1]**, then **dp[i][j] = dp[i-1][j-1] + 1**.
   * Otherwise, **dp[i][j] = max(dp[i-1][j], dp[i][j-1])**.
3. The result will be found in **dp[m][n]**, which gives the length of the longest common subsequence.

**C++ Code Implementation**

Here is the C++ code for finding the length of the Longest Common Subsequence:

cpp

RunCopy code

1#include <iostream>

2#include <vector>

3#include <string>

4#include <algorithm>

5

6using namespace std;

7

8// Function to find the length of the Longest Common Subsequence

9int longestCommonSubsequence(const string& X, const string& Y) {

10 int m = X.length();

11 int n = Y.length();

12

13 // Create a DP table with dimensions (m+1) x (n+1)

14 vector<vector<int>> dp(m + 1, vector<int>(n + 1, 0));

15

16 // Fill the DP table

17 for (int i = 1; i <= m; i++) {

18 for (int j = 1; j <= n; j++) {

19 if (X[i - 1] == Y[j - 1]) {

20 dp[i][j] = dp[i - 1][j - 1] + 1; // Characters match

21 } else {

22 dp[i][j] = max(dp[i - 1][j], dp[i][j - 1]); // Take the maximum

23 }

24 }

25 }

26

27 // The length of the longest common subsequence is in dp[m][n]

28 return dp[m][n];

29}

30

31int main() {

32 string X = "AGGTAB";

33 string Y = "GXTXAYB";

34

35 int length = longestCommonSubsequence(X, Y);

36 cout << "Length of Longest Common Subsequence: " << length << endl;

37

38 return 0;

39}

**Explanation of the Code**

1. **Input Initialization**: We define two strings **X** and **Y** for which we want to find the LCS.
2. **DP Table Creation**: We create a 2D vector **dp** initialized to 0. The size is **(m+1) x (n+1)** to accommodate the base case (0 characters).
3. **Filling the DP Table**:
   * We iterate through each character of both strings starting from index 1.
   * If the characters at the current indices match (**X[i-1] == Y[j-1]**), we set **dp[i][j]** to **dp[i-1][j-1] + 1**.
   * If they do not match, we take the maximum value from either excluding the current character of **X** or **Y**.
4. **Result**: The length of the longest common subsequence is found in **dp[m][n]**.

**Complexity**

* **Time Complexity**: O(m \* n), where **m** is the length of string **X** and **n** is the length of string **Y**.
* **Space Complexity**: O(m \* n) for the DP table.

A close-up of a computer screen

AI-generated content may be incorrect.