week6_lec1

Main Ideas

- Edit Distance

Edit Distance

Problem : Two strings A of length m and B of length n. Transform A into B with minimum number of operations. The operations are:

1. Delete a character from A.

2. Insert a character into A.

3. Change some character in A into a new character.

The Edit Distance between two strings is the minimal number of operations required to convert one string to the other.

Dynamic Programming Solution

We can use dynamic programming to solve this.

Input: Two text strings A of length m and B of length n.

Before going to a solution, let us consider the possible operations for converting A into B.

1. If $m > n$, we need to remove some characters of A.

2. If $m == n$, we need to convert some characters of$A$.

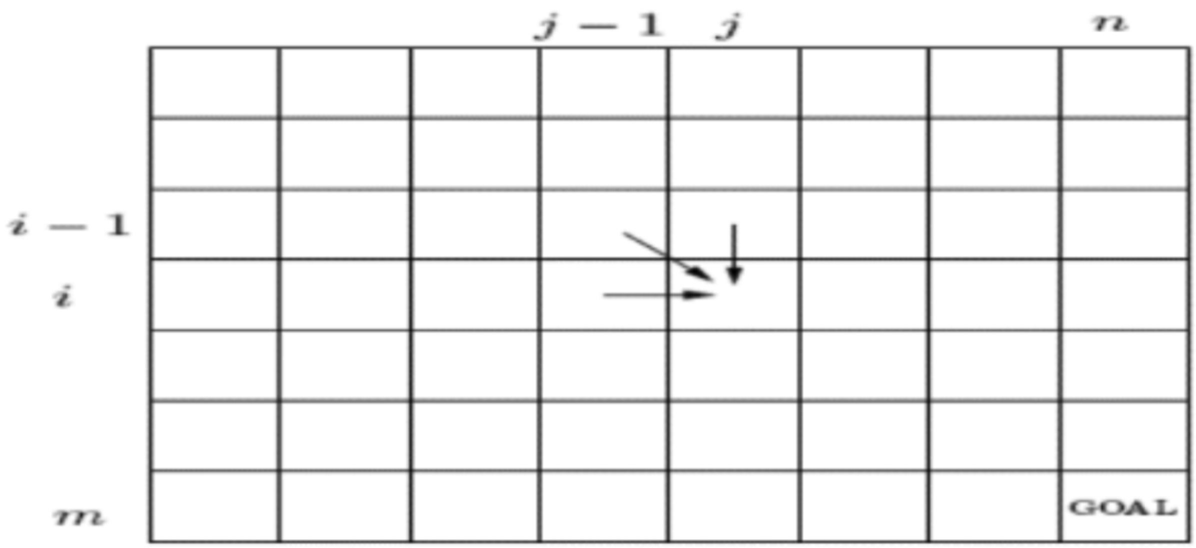
3. If $m < n$, we need to insert some characters in A.

We generate the recursive formulation of the problem.

Let, $T(i, j)$represents the minimum cost required to transform first i characters of A to first j characters of$B$. That means, $A[1...i] to B[1...j]$ $T(i, j) = min(1 + T(i-1, j); 1 + T(i, j-1); T(i-1, j-1) + diff(i, j))$

$diff(i, j)$ is 0 if the characters are equal and 1 otherwise

DP table is



We have the following cases

- If we insert i^{th} character in A, then convert these i characters of A to $j$characters of B.
- If we delete i^{th} character from A, then we have to convert remaining $i-1$ characters of A to$j$ characters of B.
- If $A[i] == B[j]$, then we have to convert the remaining $i-1$ characters of A to $j-1$ characters of B.
- If $A[i] \neq B[j]$, then we have to replace i^{th} character of A to j^{th}character of B and convert remaining $i-1$ characters of A to $j-1$characters of B.

Subproblems :

Range of i : 1 ≤ i ≤ m

Range of j : 1 ≤ i ≤ n

No. of subproblems : m*n

Time Complexity for each subproblem : O(1)

Total Time Complexity : $O(mn)$

Space Complexity : $O(mn)$ where m is number of rows and n is number of columns in the given matrix.