

ASSIGNMENT- 03

Advanced Data Structures and Algorithms

1. SEQUENCE ALIGNMENT - Cost

```
1. align_cost(strM, m, strN, n, gapPenalty, mismatchPenalty)
2.     cost[m+1][n+1]
3.     for i 0 to m
4.         cost[i][0] = i * gapPenalty
5.     for i 0 to n
6.         cost[0][i] = i * gapPenalty
7.     for i 1 to m
8.         for j 1 to n
9.             if strM[i-1] == strN[j-1]
10.                if cost[i-1][j-1] < cost[i-1][j] + gapPenalty and cost[i][j-1] <
                    cost[i][j-1] + gapPenalty
11.                    cost[i][j] = cost[i-1][j-1]
12.                else if cost[i-1][j] < cost[i][j-1]
13.                    cost[i][j] = cost[i-1][j] + gapPenalty
14.                else
15.                    cost[i][j] = cost[i][j-1] + gapPenalty
16.            else
17.                if cost[i-1][j-1] + mismatchPenalty < cost[i-1][j] + gapPenalty and
                    cost[i][j-1] + mismatchPenalty < cost[i][j-1] + gapPenalty
18.                    cost[i][j] = cost[i-1][j-1] + mismatchPenalty
19.                else if cost[i-1][j] < cost[i][j-1]
20.                    cost[i][j] = cost[i-1][j] + gapPenalty
21.                else
22.                    cost[i][j] = cost[i][j-1] + gapPenalty
23.     return cost
```

The total cost is given by $\text{cost}[m][n]$

Time Complexity

$$\begin{aligned} T(\text{align_cost}) &= T(\text{loop 0-m}) + T(\text{loop 0-n}) + T(\text{loop 1-m}) * T(\text{loop 1-n}) \\ &= O(m+1) + O(n+1) + O(m)*O(n) \\ &= O(m) + O(n) + O(m*n) \\ &= O(m*n) \end{aligned}$$

Space Complexity

$$\begin{aligned} \text{Only additional space used is by the cost matrix, which is } &(m+1) * (n+1) \\ \text{AdditionalSpace} &= O((m+1)*(n+1)) \\ &= O(m*n) \end{aligned}$$

2. Finding Alignment

```
1. align(cost, strM, m, strN, n, gapPenalty, mismatchPenalty)
2.     align_M = "", align_N = ""
3.     i = m, j = n
4.     while i > 0 and j > 0
5.         if cost[i][j] == cost[i-1][j-1] and strM[i-1] == strN[j-1]
6.             align_M = strM[--i] + align_M
7.             align_N = strN[--j] + align_N
8.         else if cost[i][j] == cost[i][j-1] + gapPenalty
9.             align_M = "_" + align_M
10.            align_N = strN[--j] + align_N
11.        else if cost[i][j] == cost[i-1][j] + gapPenalty
12.            align_M = strM[--i] + align_M
13.            align_N = "_" + align_N
14.        else if cost[i][j] == cost[i-1][j-1] and strM[i-1] == strN[j-1]
15.            align_M = strM[--i] + align_M
16.            align_N = strN[--j] + align_N
17.    while i > 0
18.        align_M = strM[--i] + align_M
19.        align_N = "_" + align_N
20.    while j > 0
21.        align_M = "_" + align_M
22.        align_N = strN[--j] + align_N
23.    return align_M, align_N
```

Time Complexity

$$\begin{aligned} T(\text{align_cost}) &= T(\text{loop } i>0, j>0) + T(\text{loop } i>0) + T(\text{loop } j>0) \\ &= O(m+n) + O(m) + O(n) \\ &= O(m+n) \end{aligned}$$

Space Complexity

Only additional space used is by the align_M and align_N strings, which is m+n for both in worst case.

$$\begin{aligned} \text{AdditionalSpace} &= O(m+n) + O(m+n) \\ &= O(m+n) \end{aligned}$$

For finding alignment with lexicographically least query, as we are making the string in reverse, match is given higher priority and gap in target string is given 2nd priority. This was, in forward order, gap in query string gets higher priority and the given alignment is lexicographically least query string.

PLAGIARISM STATEMENT

I certify that this assignment/report is my own work, based on my personal study and/or research and that I have acknowledged all material and sources used in its preparation, whether they be books, articles, reports, lecture notes, and any other kind of document, electronic or personal communication. I also certify that this assignment/report has not previously been submitted for assessment in any other course, except where specific permission has been granted from all course instructors involved, or at any other time in this course, and that I have not copied in part or whole or otherwise plagiarised the work of other students and/or persons. I pledge to uphold the principles of honesty and responsibility at CSE@IITH. In addition, I understand my responsibility to report honour violations by other students if I become aware of it.

Name: Bhavishya Sharma
Roll No: CS20MTECH12006

Signature:

