

HW 2

Anthony Nardiello, Rob Lewis

September 2019

1 Section 2-7 Question 2.

Can you propose a dynamic programming solution to solve the longest common sub-sequence problem?

There are a multitude of different solutions to solving the longest common sub-sequence problem. However, Dynamic solutions tend to be difficult to find. However, there tends to be a few that could potentially come to mind.

The steps to our Dynamic Programming solution are:

1. We find the lengths, m and n , of two strings, S and T . We also initialize a new array, X , to store all the values we are going to calculate within the solution.
2. This array will be a 2d-array, with $m+1$ making up the rows, and $n+1$ making up the columns. We want to go one above the length of the string to account for the final position.
3. We then iterate, i , over both $m+1$, and then nest an iteration, j , of $n+1$ over it.
4. We then compare i and j , and if they are the same we simply add a Zero into that index position of the array. $X[i][j]=0$
5. If m and n are not the same, we check the letter of the string comparable to the actual value of m and n . If they are the same, the index position of our X array is filled with the sum of the current value at that position and one.
6. If neither if statement is fulfilled, we set the X , $= \max(L[i-1][j], L[i][j-1])$
7. Finally, we return the $L[m][n]$ value, which should be the Longest Common Sub-sequence.

2 Section 2-7. Question 3

Given two sequences S and T, Let G, L, and H be the scores of the optimal global alignment, the optimal local alignment, and an optimal global alignment without counting the beginning space of S and end space of T, respectively.

2.1 Give an example of S and T so that all three scores G, L, and H are different.

One example of two sequences that would have different scores is “GAC” and “GTC”. For this example, the local alignment score is going to be 3, while the Global Alignment is going to be 2. Finally, the H value is going to be -1. Thus, all the values are different for those particular sequences.

2.2 Prove or disprove the statement $L \geq H \geq G$.

By the definition of Local Alignment, we get the following formula, : $2 * (\text{number of matches}) + (-1) * \text{number of insertions, deletions and mismatches}$. For both versions of Global Alignment, the standard evaluations of the Needleman-Wunsch algorithm, which are $(\text{number of matches}) + (-1)(\text{number of mismatches, insertions, or deletions})$. Assume that we have a Sequence S and T with a match at both their first and last characters. As seen in the next, section the Global Alignment with beginning and ending values ends up having a higher value than the Global Alignment without the beginning and ending values. Thus, the statement is false by contradiction.

3 Proof by Contradiction: Number calculation

x = 1
y = -1
z = -1
Strings[0] = GAC
Strings[1] = GTC

- G A C
- 0 -1 -2 -3.
G -1 1 0 -1
T -2 0 0 -1
C -3 -1 -1 1

——-Possible best alignment———
(Mis)match — (Mis)match — (Mis)match —
String S: G A C
String T: G T C
Score: 1
y = -1
z = -1
Strings[0] = AC
Strings[1] = GT

- A C
- 0 -1 -2
G -1 -1 -2
T -2 -2 -2

——-Possible best alignment———
(Mis)match — (Mis)match —
String S: A C
String T: G T
Score: -2

3.1 Longer Nucleotide Sequence Example

```
x = 1
y = -1
z = -1
Strings[0] = cgcgagcattt
Strings[1] = gcctgggctaga
```

```
- c g c g a g c a t t t
- 0 -1 -2 -3 -4 -5 -6 -7 -8 -9 -10 -11
g -1 -1 0 -1 -2 -3 -4 -5 -6 -7 -8 -9
c -2 0 -1 1 0 -1 -2 -3 -4 -5 -6 -7
c -3 -1 -1 0 0 -1 -2 -1 -2 -3 -4 -5
t -4 -2 -2 -1 -1 -1 -2 -2 -2 -1 -2 -3
g -5 -3 -1 -2 0 -1 0 -1 -2 -2 -2 -3
g -6 -4 -2 -2 -1 -1 0 -1 -2 -3 -3 -3
g -7 -5 -3 -3 -1 -2 0 -1 -2 -3 -4 -4
c -8 -6 -4 -2 -2 -2 -1 1 0 -1 -2 -3
t -9 -7 -5 -3 -3 -3 -2 0 0 1 0 -1
a -10 -8 -6 -4 -4 -2 -3 -1 1 0 0 -1
g -11 -9 -7 -5 -3 -3 -1 -2 0 0 -1 -1
a -12 -10 -8 -6 -4 -2 -2 -2 -1 -1 -1 -2
```

——-Possible best alignment———

Deletion — (Mis)match — Insertion — (Mis)match — Insertion — (Mis)match
 — (Mis)match — (Mis)match — (Mis)match — Deletion — (Mis)match — In-
 sersion — (Mis)match — (Mis)match —

String S: c g - c - g a g c a t - t t
 String T: - g c c t g g g c - t a g a
 Score: -2

——-Possible best alignment———

Insertion — (Mis)match — Deletion — (Mis)match — Insertion — (Mis)match
 — (Mis)match — (Mis)match — (Mis)match — Deletion — (Mis)match — In-
 sersion — (Mis)match — (Mis)match —

String S: - c g c - g a g c a t - t t
 String T: g c - c t g g g c - t a g a
 Score: -2

——-Possible best alignment———

Deletion — (Mis)match — (Mis)match — Insertion — Insertion — (Mis)match
 — (Mis)match — (Mis)match — (Mis)match — Deletion — (Mis)match — In-
 sersion — (Mis)match — (Mis)match —

String S: c g c - - g a g c a t - t t
String T: - g c c t g g g c - t a g a
Score: -2

——-Possible best alignment———
Deletion — (Mis)match — Insertion — (Mis)match — Insertion — (Mis)match
— (Mis)match — (Mis)match — (Mis)match — Insertion — (Mis)match —
Deletion — (Mis)match — (Mis)match —
String S: c g - c - g a g c - a t t t
String T: - g c c t g g g c t a - g a
Score: -2

——-Possible best alignment———
Insertion — (Mis)match — Deletion — (Mis)match — Insertion — (Mis)match
— (Mis)match — (Mis)match — (Mis)match — Insertion — (Mis)match —
Deletion — (Mis)match — (Mis)match —
String S: - c g c - g a g c - a t t t
String T: g c - c t g g g c t a - g a
Score: -2

——-Possible best alignment———
Deletion — (Mis)match — (Mis)match — Insertion — Insertion — (Mis)match
— (Mis)match — (Mis)match — (Mis)match — Insertion — (Mis)match —
Deletion — (Mis)match — (Mis)match —
String S: c g c - - g a g c - a t t t
String T: - g c c t g g g c t a - g a
Score: -2

——-Possible best alignment———
Deletion — (Mis)match — Insertion — (Mis)match — Insertion — (Mis)match
— (Mis)match — (Mis)match — (Mis)match — Deletion — (Mis)match —
(Mis)match — Insertion — (Mis)match —
String S: c g - c - g a g c a t t - t
String T: - g c c t g g g c - t a g a
Score: -2

——-Possible best alignment———
Insertion — (Mis)match — Deletion — (Mis)match — Insertion — (Mis)match
— (Mis)match — (Mis)match — (Mis)match — Deletion — (Mis)match —
(Mis)match — Insertion — (Mis)match —
String S: - c g c - g a g c a t t - t

String T: g c - c t g g g c - t a g a
Score: -2

——-Possible best alignment——
Deletion — (Mis)match — (Mis)match — Insertion — Insertion — (Mis)match
— (Mis)match — (Mis)match — (Mis)match — Deletion — (Mis)match —
(Mis)match — Insertion — (Mis)match —
String S: c g c - - g a g c a t t - t
String T: - g c c t g g g c - t a g a
Score: -2

——-Possible best alignment——
Deletion — (Mis)match — Insertion — (Mis)match — Insertion — (Mis)match
— (Mis)match — (Mis)match — (Mis)match — Insertion — (Mis)match —
(Mis)match — Deletion — (Mis)match —
String S: c g - c - g a g c - a t t t
String T: - g c c t g g g c t a g - a
Score: -2

——-Possible best alignment——
Insertion — (Mis)match — Deletion — (Mis)match — Insertion — (Mis)match
— (Mis)match — (Mis)match — (Mis)match — Insertion — (Mis)match —
(Mis)match — Deletion — (Mis)match —
String S: - c g c - g a g c - a t t t
String T: g c - c t g g g c t a g - a
Score: -2

——-Possible best alignment——
Deletion — (Mis)match — (Mis)match — Insertion — Insertion — (Mis)match
— (Mis)match — (Mis)match — (Mis)match — Insertion — (Mis)match —
(Mis)match — Deletion — (Mis)match —
String S: c g c - - g a g c - a t t t
String T: - g c c t g g g c t a g - a
Score: -2

——-Possible best alignment——
Deletion — (Mis)match — Insertion — (Mis)match — Insertion — (Mis)match
— (Mis)match — (Mis)match — (Mis)match — Deletion — (Mis)match —
(Mis)match — (Mis)match — Insertion —
String S: c g - c - g a g c a t t t -
String T: - g c c t g g g c - t a g a

Score: -2

——-Possible best alignment——
Insertion — (Mis)match — Deletion — (Mis)match — Insertion — (Mis)match
— (Mis)match — (Mis)match — (Mis)match — Deletion — (Mis)match —
(Mis)match — (Mis)match — Insertion —
String S: - c g c - g a g c a t t t -
String T: g c - c t g g g c - t a g a
Score: -2

——-Possible best alignment——
Deletion — (Mis)match — (Mis)match — Insertion — Insertion — (Mis)match
— (Mis)match — (Mis)match — (Mis)match — Deletion — (Mis)match —
(Mis)match — (Mis)match — Insertion —
String S: c g c - - g a g c a t t t -
String T: - g c c t g g g c - t a g a
Score: -2

——-Possible best alignment——
Deletion — (Mis)match — Insertion — (Mis)match — Insertion — (Mis)match
— (Mis)match — (Mis)match — (Mis)match — Insertion — (Mis)match —
(Mis)match — (Mis)match — Deletion —
String S: c g - c - g a g c - a t t t
String T: - g c c t g g g c t a g a -
Score: -2

——-Possible best alignment——
Insertion — (Mis)match — Deletion — (Mis)match — Insertion — (Mis)match
— (Mis)match — (Mis)match — (Mis)match — Insertion — (Mis)match —
(Mis)match — (Mis)match — Deletion —
String S: - c g c - g a g c - a t t t
String T: g c - c t g g g c t a g a -
Score: -2

——-Possible best alignment——
Deletion — (Mis)match — (Mis)match — Insertion — Insertion — (Mis)match
— (Mis)match — (Mis)match — (Mis)match — Insertion — (Mis)match —
(Mis)match — (Mis)match — Deletion —
String S: c g c - - g a g c - a t t t
String T: - g c c t g g g c t a g a -
Score: -2