

## Iqb assignment report

1) a) The generated bidimensional array obtained for computing the global alignment is as follows->

```
PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL SQL CONSOLE
Python + - [ ] [ ] ... ^ x

PS C:\Users\Shyama Goel\OneDrive\Desktop\iqb_final> & "C:/Users/Shyama Goel/AppData/Local/Programs/Python/Python310/python.exe" "c:/Users/Shyama Goel/OneDrive/Desktop/iqb_final/1.py"
Resulting DP Matrix:
=====
[0, -1, -2, -3, -4, -5, -6, -7]
[-1, 2, 1, 0, -1, -2, -3, -4]
[-2, 1, 0, -1, 2, 1, 0, -1]
[-3, 0, -1, -2, 1, 0, 3, 2]
[-4, -1, 2, 1, 0, 3, 2, 1]
[-5, -2, 1, 4, 3, 2, 1, 0]
[-6, -3, 0, 3, 2, 5, 4, 3]
[-7, -4, -1, 2, 1, 4, 3, 2]
[-8, -5, -2, 1, 4, 3, 2, 5]
[-9, -6, -3, 0, 3, 6, 5, 4]
=====
Highest score of the alignment is: 4
PS C:\Users\Shyama Goel\OneDrive\Desktop\iqb_final>
```

b) Yes, there can be more than one possibility of aligning the given sequence GATGCGCAG, GGCAGTA in order to make the highest score possible as there are many repeating, overlapping terms. Moreover, alignment score of 4 can be obtained from more than one path.

c) All the possible alignments are as follows->

There are total 6 possible alignments that are generated.

```
PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL SQL CONSOLE Python
PS C:\Users\Shyama Goel\OneDrive\Desktop\iqb_final> & "C:/Users/Shyama Goel/AppData/Local/Programs/Python/Python
sers/Shyama Goel/OneDrive/Desktop/iqb_final/lc.py"
=====
GATGC-G-CAG
|  |  |  |
G--G-CAGT-A-
=====
GATGC-GC-AG
|  |  |  |
G--GCAG-TA-
=====
GATGCGCAG--
|  |  |  |
---G-GCAGTA
=====
GATGCGCAG--
|  |  |  |
G---GCAGTA
=====
GATGCGCAG--
|  |  |  |
G--G--CAGTA
=====
GATGCGCAG--
|  |  |  |
G--GC--AGTA
=====
The alignment score for the above sequences with given penalties is: 4
PS C:\Users\Shyama Goel\OneDrive\Desktop\iqb_final>
```

2) Yes, changing the scoring scheme will definitely modify the results obtained. here, we are changing our penalties

[Match = +2, Mismatch = -1, Gap = -3]. Due to the swapping of mismatch and gap penalties, the matrix obtained will also change and the highest score obtained would change accordingly. As we do our calculations and compute the matrix,

upper =  $dp[i-1][j] + \text{gap\_penalty}$

diag\_aln =  $dp[i-1][j-1] + \text{mismatch\_penalty}$ , therefore changes are going to be implemented in the program. This is because the mismatch penalties have been reduced and the gap has been increased. As a result, different optimal alignments between the two DNA sequences will result from changing the gap and mismatch penalties in the scoring system.

Our matrix generated now would be as follows->

```
PS C:\Users\Shyama Goel\OneDrive\Desktop\iqb_final> & "C:/Users/Shyama Goel/AppData/Local/Programs/Python/Python39-64/Scripts/python.exe" C:/Users/Shyama Goel/OneDrive/Desktop/iqb_final/2.py
Resulting DP Matrix:
=====
[0, -3, -6, -9, -12, -15, -18, -21]
[-3, 2, -1, -4, -7, -10, -13, -16]
[-6, -1, 1, -2, -2, -5, -8, -11]
[-9, -4, -2, 0, -3, -3, -3, -6]
[-12, -7, -2, -3, -1, -1, -4, -4]
[-15, -10, -5, 0, -3, -2, -2, -5]
[-18, -13, -8, -3, -1, -1, -3, -3]
[-21, -16, -11, -6, -4, -2, -2, -4]
[-24, -19, -14, -9, -4, -5, -3, 0]
[-27, -22, -17, -12, -7, -2, -5, -3]
=====
Highest score of the alignment is: -3
=====
PS C:\Users\Shyama Goel\OneDrive\Desktop\iqb_final>
```

The sequences obtained would be->

-GAT

GGCA

GATG

G-GC

GATGC-

G--GC-

GATGCGCA

G—GCAGT

GATGC-GCAG

G—GCAGTA-

These are all the possible sequences that would be obtained on changing the scoring scheme.

4)The changes that are required in the program for performing local rather than global pairwise sequence alignment are as follows->

1)In the local alignment, the elements of the 0<sup>th</sup> row and column were initialised by 0, whereas for the global alignment, they were initialised with 0, -1, -2, and so on. All the -ve values are initialised to zero in local alignment.

2)For the other entries, -ve values are available for global alignment but not for local alignment.

3) In the local alignment, the maximum of the neighbouring cells with added penalties and 0 and no negative is present. In the global alignment, only the maximum value of the adjacent cells with added penalties are taken into account when populating the table.

4)When supporting in a global alignment, the maximum value is always in the last cell, whereas in a local alignment, the maximum value can be located anywhere in the matrix and is not limited to the last cell.

While the maximum value can only be accommodated in one cell in global alignment, it can be found in numerous cells in local alignment. When you are at (0,0) in the global alignment, the base case is achieved.

- Anywhere in local alignment is feasible.

3 1) a) The generated bidimensional array obtained for computing the local alignment is as follows->

```
PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL SQL CONSOLE
PS C:\Users\Shyama Goel\OneDrive\Desktop\iqb_final> & "C:/Users/Shyama Goel/AppData/Local/Programs/Python/Python38-64/Python.exe" & "C:/Users/Shyama Goel/OneDrive/Desktop/iqb_final/3.py"
Resulting DP Matrix:
=====
[0, 0, 0, 0, 0, 0, 0, 0]
[0, 2, 2, 0, 0, 2, 0, 0]
[0, 0, 1, 1, 2, 0, 1, 2]
[0, 0, 0, 0, 0, 1, 2, 0]
[0, 2, 2, 0, 0, 2, 0, 1]
[0, 0, 1, 4, 1, 0, 1, 0]
[0, 2, 2, 1, 3, 3, 0, 0]
[0, 0, 1, 4, 1, 2, 2, 0]
[0, 0, 0, 1, 6, 3, 1, 4]
[0, 2, 2, 0, 3, 8, 5, 2]
=====
=====
```

b) All the optimal alignments with their score are->

```
PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL SQL CONSOLE
PS C:\Users\Shyama Goel\OneDrive\Desktop\iqb_final> & "C:/Users/Shyama Goel/AppData/Local/Programs/Python/Python38-64/Python.exe" & "C:/Users/Shyama Goel/OneDrive/Desktop/iqb_final/3c.py"
=====
The score for sequences is: 8
GCAG
|||
GCAG
=====
PS C:\Users\Shyama Goel\OneDrive\Desktop\iqb_final>
```