Indraprastha Institute of Information Technology Delhi (IIITD)
Department of Computational Biotechnology

BIO213 − Introduction to Quantitative Biology

ASSIGNMENT-1 (March 13, 2024)

Instructions:

1. The file name for the code should be in the format 2022XXX_Q1_Name.py.
2. The code should be well-commented.
3. No library apart from pandas or numpy is allowed.
4. Keep the sequence and score variables the same as provided in the codes, the rest variables should be renamed accordingly.
5. Submissions after **March 22, 2024** will not be entertained.

The objective of this assignment is to learn the implementation of dynamic programming for finding the best local and global alignments for the given DNA sequences.

DNA sequences to be used: (1) GATGCGCAG, (2) GGCAGTA

Question 1. The code provided is for the global alignment of the DNA sequences using dynamic programming [Scoring scheme used: Match = +2, Mismatch = -3, Gap = -1].   **[30 marks]**

  a) Find the bugs and complete the partial code (comment on your edits).

  b) The code should print the scoring matrix.

  c) Modify the code to get all the optimal alignments with their scores.

  d) Print the best alignment(s) obtained with its corresponding score.

Demo output:

Question 2. Modify the same code for getting the most optimal local alignment (using dynamic programming) for the given DNA sequences and the scoring scheme [Match = +2, Mismatch = -1, Gap = -3].   **[20 marks]**