Circle

Description automatically generated with medium confidence

**MARMARA UNIVERSITY**

**FACULTY of ENGINEERING**

**COMPUTER ENGINEERING DEPARTMENT**

**CSE 4065 – Computational Genomics**

**Programming Assignment # 2**

**Group Members**

150116012 Rahim Gün

150116056 Emre YİĞİT

150116016 Berk Köylü

1 - Implementation Details

In this assignment, we implemented pairwise sequence alignment by using Needleman–Wunsch algorithm.

In out project, we created SequenceAlignment class and inside the class we have alignSequences method that makes the sequence alignment. Method takes the sequences, match score, mismatch score and indel score as arguments. Inside the method, firstly, we created two-dimensional matrix with length of the strings. And then, we filled the first row and first column with the indel score by summing cumulatively. And then, we filled the other cells of the matrix by selecting the highest cumulative score that comes from match/mismatch, insert or delete one by one. Score of each cell is calculated depending on the match, mismatch and indel score given by the user. Finally, the lower right cell of the alignment matrix gives the optimum alignment between the given sequences. Code snipped below shows the implementation of the algorithm inside the project.

Text

Description automatically generated

Figure 1 - Implementation of Needleman Wunsch Algorithm

After calculation of the sequence matrix, we need to backtrack from sink to source to find the optimum alignment sequence. Inside the SequenceAlignment class, we created backtrackMatrix method that traverse the matrix backward from sink to source by comparing the cell value with three possible sources match/mismatch, insert, delete to see where it comes from. If it is match/mismatch, then nucleotides are aligned. If it is insert, first sequence is aligned with gap. If it is delete, second sequence is aligned with gap. Code below shows the implementation of the backtracking algorithm and its arguments. And then, method prints the aligned sequences on the terminal.

Text

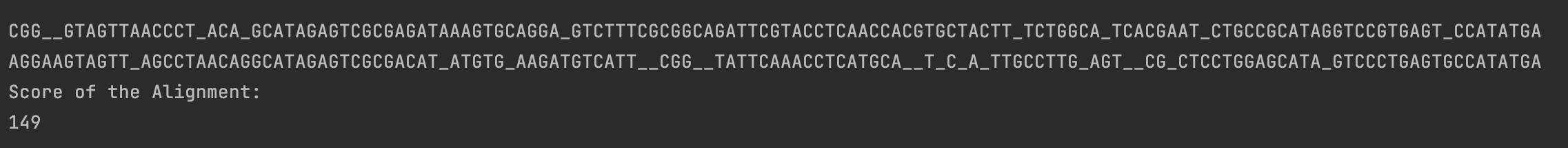
Description automatically generated

Figure 2 – Implementation of the Backtracking Algorithm

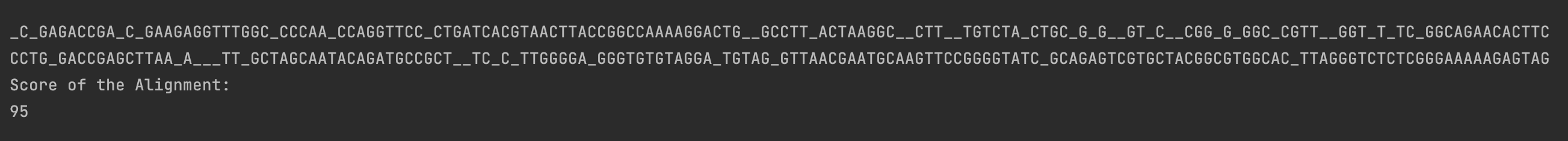
2 – Program Outputs

Sequence-alignment-result-output.txt file also shared with submitted assignment folder.

Output of test1.seq



Output of test2.seq



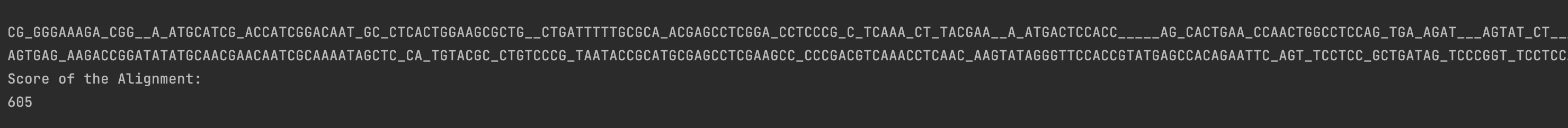
Output of test3.seq

Text

Description automatically generated

Output of test4.seq

Sequences do not fit the screen shot but we provide the outputs of the program as sequence-alignment-result-output.txt file inside the submitted folder.



Output of test5.seq

