Alignment Scoring Algorithm Report

In this endeavor, our objective was to design a dynamic programming algorithm for the Sequence Alignment Problem, where the closeness of two gene sequences is measured through optimal alignment. The algorithm takes two input strings, `x` and `y`, and employs a scoring matrix to evaluate their alignment, considering the possibility of gaps denoted by "-".

Approach:

We implemented a dynamic programming approach, leveraging a matrix to store intermediate scores. The algorithm iterates through each position in the matrix, calculating scores based on matches, gaps in sequence `x`, and gaps in sequence `y`. The maximum score at each position is then chosen, ensuring that the resulting matrix captures the optimal alignment scores.

Subsequently, a backtracking mechanism traces the path with the highest scores, constructing the optimal alignment of the input sequences.

Time Complexity:

The time complexity of our algorithm is O(mn), where `m` and `n` represent the lengths of the input sequences `x` and `y`, respectively. This efficiency is achieved by exploring and filling in the dynamic programming matrix in a systematic manner, ensuring that each entry is computed only once.

Example:

For instance, when applied to the sequences "ATGCC" and "TACGCA," and a scoring matrix `δ` provided, the algorithm yields the highest-scoring alignment.

Conclusion:

Our dynamic programming approach efficiently addresses the Sequence Alignment Problem, providing an optimal alignment with a time complexity proportional to the product of the input sequence lengths.