Dynamic Programming for DNA Sequence Alignment

Xinyu Cheng, Zhihao Wang, Xueni Tang, Jieni Tang

August 2025

Contents

1 Introduction

1.1 DNA Sequence Alignment

DNA sequence alignment is a fundamental task to compare and arrange DNA sequences, in order to uncover regions of similarities in these sequences. These similarities often allows researchers to infer homologies between genes and proteins, that is, these regions may share a evolutionary history. These homologies can provide insights into their functions and evolutionary relationships. In the context of bioinformatics, DNA sequence alignment is crucial for tasks such as gene prediction, functional annotation, and phylogenetic analysis.

Traditionally, DNA sequence alignment was achieved using brute-force methods that compare all possible alignments between sequences. However, these brute-force methods pose significant challenges, in terms of time and computational power, as the number of possible alignments grows exponentially with the length of the sequences. This makes it almost impractical for real-world applications. However, one way to approach DNA sequence alignment is by finding the longest common subsequence (LCS) between two sequences. Here, a subsequence is defined as a sequence that can be derived from another sequence by deleting some elements without changing the order of the remaining elements.

1.2 Dynamic Programming

1.3 Project Structure Overview

2 Introduction

2.1 DNA Sequence Alignment

abc

2.2 The Role of Dynamic Programming

abc

2.3 Project Structure Overview

References

[1]

[2]