Universität Bielefeld
Technische Fakultät

AG Genominformatik
Prof. Dr. Jens Stoye
Dr. Lucas Robidou

Sequence Analysis 3

Summer 2024

Exercises

Number 8, Discussion: 2024-June-20

1. (a) Explain the terms *mosaic effect* and *shadow effect*.
   (b) How is a length normalized alignment score defined and what is its meaning?
   (c) Why is the parameter L introduced? What is its effect for the two limit cases $L \to 0$ and $L \to \infty$?
   (d) How can a length normalized alignment be computed efficiently?

2. Given the two strings $x =$ `CAGCTGACAG` and $y =$ `TAGTCTGCAG` and the scores $+2$ for matches, 0 for mismatches and -1 for indels (linear gap costs).
   (a) Compute an optimal local alignment and its score $LAS(x, y)$.
   (b) Compute the value $NAS_L(x, y)$ of the length normalized alignments for $L = 2$ and $L = 3$, using the bisection algorithm presented in class.
   (c) Discuss the effect of the normalization in this example.