CSCI1820 - Algorithmic Foundations of Computational Biology
Spring 2017
Midterm Exam Class Sumary

March 18, 2017

Lecture Topics

• Chapter 1: Sequence Alignment Algorithms and the BLAST Algorithm

Pairwise Sequence Alignment Algorithms

– Alignment algorithms structure: Edit graphs, the principle of optimality, dynamic programming, heuristic interpretation of the alignment score, the Waterman notation for alignments
– Needleman-Wunsch Pairwise Sequence Global Alignment Algorithm (global alignment)
– Smith-Waterman Pairwise Sequence Local Alignment Algorithm (local alignment)
– Pairwise Sequence Global Affine-Gap Alignment

BLAST Algorithm

– The BLAST Algorithm phases: Seeding, Extension, Evaluation
– The random walk interpretation of scored pairwise alignment

Statistical Theory

– The Karlin-Altschul statistical theory of local alignment (un-gapped local alignment), Maximum-Segment Pairs statistics
– Substitution Matrices for local alignment and Information Theory - The Margaret Dayhoff PAM Matrices algorithm (Margaret Dayhoff is known as the "mother and father of Bioinformatics")
- Simple random walks theory - the analytical formula for the walk to end up at one end of the interval

- Chapter 2: Genome Assembly Algorithms and Haplotype Assembly Algorithms

**Sequence Assembly Algorithms**

- de Bruijn graphs and Eulerian paths and cycles
- Genome Assembly algorithms structure: reads/fragments, genome coverage, sequencing errors, genomic repeats, genomic contamination, Poisson statistics for depth of coverage
- Idury-Waterman Assembly Algorithm and Poisson statistics of DNA k-mers (only what we covered so far in class so far)