

Home Work 3

CSE/BENG/BIMM 181, Winter 2012

Given on Mon, January 30, 2012. Due on Mon. February 6, 2012

State your answers legibly and concisely. Your solutions will be graded on correctness, elegance, **clarity** and originality. Your proofs should **avoid getting bogged down in too much detail**. Your algorithms should be as efficient as possible (inefficient algorithms will result in lower grades). Please note that the work handed in must be your own. No late assignments will be accepted without prior arrangement. Every problem is 20 points.

The Following Problems are from textbook:

Problem 7.11 (20 pts)

Design a linear-space algorithm for the Block Alignment problem.

Problem 11.3 & Problem 11.4 (20 pts)

1. Suppose you are given the dinucleotide frequencies in CG-islands and the dinucleotide frequencies outside CG-islands. Design an HMM for finding CG-islands in genomic sequences.
2. Figure 11.7 shows an HMM with two states α and β . When in the α state, it is more likely to emit purines (A and G). When in the β state, it is more likely to emit pyrimidines (C and T). Decode the most likely sequence of states (α/β) for sequence GGCT. Use log-scores, rather than straight probability scores.

Problem 9.11 (20 pts)

Design an efficient algorithm for finding the longest string shared by two given texts.

Problem 9.13 (20 pts)

Design an efficient algorithm that finds a shortest string in text t_1 that does not appear in text t_2 .

Problem I - Local Pairwise Alignment with Scoring Matrix and Affine Gaps (20 pts)

You need to implement local pairwise alignment taking into account a scoring matrix and affine gaps.

Go to this page for full details: http://judge.bioinf.spbau.ru/cgi-bin/new-register?contest_id=5

Register on this site and submit your assignment and it will automatically test your code.