COSC494/594 study guide

Multiple choice questions

1. Given the sequence: CATAT. What is the reverse compliment? (n = 5)
   a. GTATA
   b. GGATA
   c. ATATG
   d. None of the above

2. If doing end-gap free alignment, where would you start the traceback?
   a. Last column
   b. Last row
   c. M,n cell
   d. Last row and last column

3. Given the sequence ATGA, what is the probability of A under multinomial model?
   a. 0.5
   b. 0.66
   c. 0.25
   d. 0.33

4. Use the following scoring scheme to evaluate the alignment below. Match: +1 Mismatch: –1 Gap: –2 Sequences: Sequence A: A C G T A Sequence B: A C T A Proposed Alignment: A C G T A/A C - T A What is the alignment score for the sequences above? Choices:
   a. +1
   b. +2
   c. +3
   d. +4

Short response:

A) What are the dinucleotide frequencies of a given sequence?

B) Given the criteria, score the following alignment of these two sequences.

C) Given the sequence: CATAT. What is the probability of $P(T|A)$?

D) Given a sequence, compute its multinomial probabilities

E) In your own words, explain the difference between local and global alignment?

F) Given the sequence ATGACATTTG, give the reverse compliment. (n = 4)

G) Draw the forward algorithm chart and fill in the first 3 probabilities

H) Why do we use log probabilities for Markov Models?

I) How do you determine the best possible global alignment sequence given a table filled with scores?

J) What are the differences between global, local, and end gap free alignment of a sequence? Compute the score of the end gap free sequence alignment for the following sequences.

K) Why is order important when creating a Markov chain model?

Longer response questions:

I. Given the sequences: AGGTC, TGCCC. Calculate the best local alignment score. Matches are +5, mismatches are -2, and gaps are -3.

II. Create a HMM with your own transition probabilities and initial pi probabilities / Draw a Hidden Markov Model (like the dog, barbeque house question)

III. Given a sequence of coin flips/rolls and a model, fill out a table using forward algorithm.

IV. Draw a Markov model given. 1 > 1 = 0.1 1 > 2 = 0.8 1 > 3 = 0.1 2 > 1 = 0.5 2 > 2 = 0.25 2 > 3 = 0.35 3 > 1 = 0.3 3 > 3 = 0.7 (Attached image of the answer I couldn't attach D:)

V. Perform a global alignment between the following two sequences using the Needleman–Wunsch algorithm. Fill out the dynamic programming alignment table completely and then provide the traceback path that yields the optimal alignment. Scoring scheme: Match = +1 Mismatch = –1 Gap = –2 Sequences: Sequence A: A C G T Sequence B: A G T (n = 4)

VI. Draw out a Hidden Markov Model (HMM) using the provided description: • There are two 6-sided dice. One is fair and the other is loaded where it lands on 5 half the time. The chances of it going from fair to loaded is 0.1 and it going from loaded to fair is 0.4.

VII. Use the Viterbi algorithm to find the most likely hidden state sequence (e.g., decoding) under one of the HMM questions above.