Bioinformatics Computing
CSE40532/60532
Homework #2

Reading:
1. Read section 3.8 (week of 9/4)
2. Read Chapter 4 and handouts on Viterbi algorithm (week of 9/11)
3. Read Chapter 2 and Glimmer paper (week of 9/18)

Problems: (due 9/14)

1. Download the two genes mentioned in Problem #1 in Section 3.9 (X79493 and AY707088) from fruit fly and human, respectively.

2. Write a program called "globalign" to compute the global alignment score of the two sequences above using the following parameters: +2 for a match, -1 for a mismatch, and -2 for a gap. Perform the traceback and display the alignment in a user-friendly format. Please place the source code in your dropbox, and tell the instructor how to compile and run it in an accompanying write up. (10 points)

3. Compute the local alignment between X79493 and AY707088 using a modified program "localign". Perform the traceback and display the alignment in a user-friendly format. Submit your code with instructions on how to run and compile. Discuss in your write up how the local alignment compares to the global alignment from #2 (10 points).

4. Compute just the score of the optimal end-gap free alignment of the two mitochondrial sequences from HW #1: *NC_012920* and AF254446 in a third program called "egfalign." Traceback and display will not be required.

5. BLAST the two sequences from #1 using the bl2seq utility at NCBI BLAST. Summarize the BLAST alignment in your writeup (alignment length, identity, e-value, etc.) and discuss how it confirms or refutes your result from #3 (3 points).

6. Download the "mystery" sequence from the course website. Report its BLAST hits using blastn and blastx at the NCBI website in your writeup (1 point each).

Required for CSE60532:

7. Compute the end-gap free alignment of the two mitochondrial sequences from HW #1: *NC_012920* and AF254446 using only linear space. Report the score using the scheme used for problem #2 above (+2 match, -1 mismatch, -1 for a gap) in your writeup (6 points)