1.(25pts) Write a PERL program that takes in a seriesof words from the command line and outputs the number of letters and the average length of a word.

>> perl Letters.pl . # %

Letters: 0

Average Length: 1.00

>> perl Letters.pl There once was a man from Nantucket

Letters: 29

Average Length: 4.14

2.(25pt) Given a FASTA file,write a PERL program that outputs the reverse compliment of the string.

>> perl ReverseComplement.pl example1.fa

CGATAATTTCCTCACTATTATCGGAATAAAGCGGATAAGGGCACAGGTGACAGCGAACACTAATATATTAGCACGGACCTCATCTTTGGAGCGCTCCTCAATTCTAATCGGCTCAATGGACAGTGTGCGAAACAT

3.(30 pts) You are given the following DNA sequence, which is believed to contain a small protein-coding gene.

GGAGGCGTAA AATGCGTACT  GGTAATGCAA  ACTAATGG

* If this sequence is fully transcribed (used as a codingstrand), what is the corresponding mRNA sequence?
* Which region of the mRNA do you think can be translatedinto a protein (hint: Can you identify the start codon and stop codon from themRNA sequence?)
* What is the protein sequence encoded by the gene?
* If the reverse-complementary strand of the DNA sequenceis also transcribed, what will be the mRNA sequence?
* Do you think the reverse-complementary strand can encodea protein?

4.(20 pts) Considerthe sequence v = TACGGGTAT and w = GGACGTACG. Assume that the match score is+1, and the mismatch and gap penalties are -1.

* Fill out the dynamic programming table for a global alignment between v and w. Draw arrowsin the cells to store traceback information. What is the score of the optimal global alignment and what alignment(s) achieves this score?