1. [40 points] RNA secondary structure prediction. For the following RNA sequence

UCUUAGACCGGG

a. Fill out the DP table in the minimum energy folding algorithm (Zuker) with the following energy function.

\[ E[i, j] = \begin{cases} 
-5, & \text{if } [i, j] \text{ is CG or GC} \\
-4, & \text{if } [i, j] \text{ is AU or UA} \\
-1, & \text{if } [i, j] \text{ is GU or UG} 
\end{cases} \]

b. Show the secondary structure predicted by Zuker algorithm.

2. [60 points] Train a support vector machine to classify these two classes of genes based on microarray gene expression measurements (http://www.cis.udel.edu/~lliao/cis667f05/hw3.data)
You may assume the following format for the input data:

- Each line is tab delimited multicolumn
- The first line is a header line (not data)
- The expression data start at the second line
- The 1st column is gene ID, the last column is class ID: +1, -1, or 0.
  Class ID 0 is for the testing genes.

A support vector machine is defined as

\[ SVM(x) = \text{sign} \left( \sum_{i=0}^{n} \alpha_i y_i K(x_i, x) + b \right), \]

where \( n \) is the number of training cases, and \( K(x, y) \) is a kernel function. You are asked to implement the dual form Rosenblat algorithm to find coefficients \( b \) and \( \alpha_i > 0 \) that correctly classify the training data. Your program should have a command line option –k 0|1 to switch between the following two kernels: a) \( K(x, y) = x \cdot y \); b) \( K(x, y) = (1 + x \cdot y)^2 \).

Your program should output the following.

- The first line is the bias \( b \) of the trained SVM.
- Starting at the second line, each line has two columns delimited by a tab. The first column is gene ID, and the second column is coefficient \( \alpha \) for a training gene and is the predicted class ID (+/-) for a testing gene.

Report if the algorithm terminates during the training with different kernels, and discuss why if it does not.