Assignment 3

Implement the simple quadratic space **global** alignment algorithm with simple gap costs (so you only need one score array). Use +4 for a character match and -2 for a gap and a mismatch. Have the algorithm printout the score and the alignment.

Now implement the divide and conquer linear space algorithm version of the algorithm above. This algorithm should be like the one we did in class. Check that this linear version produces exactly the same answer. Confirm this with 5 example problem strings. Include the output in your submission.

According to our analysis, the linear space algorithm should have the same complexity, but take about twice as long. Demonstrate that this analysis is correct with some empirical tests. Do an experiment as follows:

For each problem size 256, 512, 1024, 2048, 4096, etc. (the limit will be space of the quadratic algorithm)

Generate a set of random pairs of strings of length problem size from a simple two-character alphabet.

Run both algorithms on this set of problems and record the time it takes to solve the set of problems (turn off the printout). Let linear(n) be the *average* time it takes to solve the problem set of size n and quad(n) the average time it takes to solve the problem set of size n.

Produce a log/log scatter plot comparing the two times that algorithms take. Both axes will be log time. Each point on the graph will be placed at coordinates linear(n),quad(n). You should expect to see a straight line relationship between the points. What is the significance of the slope of the line? What is the significance of the offset of the line? Make sure you label your axis and include a title and key!