Write a program to read a multiple sequence alignment and output mutual information.

1) Read a gapless DNA multiple sequence alignment in FASTA format.
2) For each pair of positions, calculate MI as the sum over all possible base-pairs (including non-canonical) of the frequency of the base-pair times the log-base-2 of the frequency of the base-pair divided by the expected frequency of the base-pair. The expected frequency of the base-pair is the product of the frequencies of the bases. (Be sure you don't divide by zero.) The result is a matrix of MI values. (See Lecture 13)
3) Have your program write out the MI result matrix in HTML format, as follows. (This example is in standard HTML. If you want to "get fancy" and use HTML5, be my guest, but you don't get extra credit.) Set height and width (in pixels) to a number between 1 and 10 depending on the size of the screen and the length of the input sequences. Set the cell color to a shade of blue if the basepair is GC, a shade of green if it is AT, and a shade of red if it is non-Watson-Crick. Set color to black for the diagonal. Colors are expressed in 2-digit hex, in red-green-blue (RGB) order, as shown in the example below. Set negative MI value to zero. Scale positive MI to the range 0-255 by multiplying by 255/MAXMI, where MAXMI is maximum MI value. Convert scaled MI to 2-digit hex using a built-in function (if one exists) or write your own. For example, if myhex(mi) is the function to convert scaled MI to 2-digit hex, then shades of blue are expressed in HTML as "#%s%sFF"%(myhex(mi),myhex(mi)) where myhex(0.0) returns "FF" and myhex(1.0) returns "00". Be sure you understand the HTML table format below. Verify that your file is valid HTML by opening it in a browser.
4) Upload your program (as "program") and your HTML result files (as "result1" and "result2") for the input data linked to "HW4 small msa" and "HW4 large msa" on the course website, using the homework server.