Problem Set 6—Due 9 September 1999

Problem 15 You have a one Megabyte file (8 million bits) with character frequencies as follows:

A	13%
B	6%
C	3%
D	24%
E	8%
F	31%
G	2%
H	13%

Find a Huffman code for these characters, using the algorithm of class. The compute how many bits will be the compressed file if we represent bits using your code.

Problem 16 Use the algorithm shown in class to find the string that results from LZ77-compressing the following string:

$AAAAABCAAABCADDDDBCADDAABCDAA$

If characters, positions and lengths are each encoded in a single byte, what is the length of the compressed string?

Problem 17 Consider the pattern $P = BAABAABCBA$.

(a) Give the finite state machine which allow you to efficiently search a text T for this pattern.

(b) Construct the KMP machine corresponding to the above FSM. (That is, this machine has a single fail pointer from each state.)