2004 Paper 10 Question 4

Data Structures and Algorithms

(*a*) Describe how the Lempel Ziv text compression algorithm works, illustrating your answer by deriving the sequence of numbers and corresponding bit patterns it would generate when applied to a string starting with the following 24 characters:

ABCDABCDABCDABCDABCDABCD ...

You may assume that the initial table is of size 256 (containing bytes 0 to 255) and that the codes for "A", "B", "C" and "D" are 65, 66, 67 and 68, respectively.

[12 marks]

(*b*) Estimate how many bits the algorithm would use to encode a string consisting of 1000 repetitions of the character "A". [8 marks]