

CSIS 1119B/C Introduction to Data Structures and Algorithms  
Programming Assignment Two  
Due Date: 18 Apr 2019 (Thu) 23:59

You are required to write a Huffman code encoder.

Details:

- Name your program as “hmcoder”.
- Usage of hmcoder:  
    hmcoder [input\_file]  
    where input\_file is the file containing the input message.
- The input message contains normal English text. You need to encode all English letters together with all printable characters such as “.”, “;”, “!”, space etc. The capital letter and the corresponding small letter are treated as two symbols. For example, “A” and “a” are regarded as two symbols.
- The output of hmcoder should have two output files, “code.txt” and “encodmsg.txt”.
- The output file “code.txt” stores the codewords for each symbol and the average number of bits used for each symbol. The format of this output file is as follows:

Each row contains the codeword for a separate symbol, the symbols are listed according to the increasing order of the corresponding ASCII code value. The last row shows the average number of bits used for each symbol.

E.g.

Space: 000

!: 001

.....  
A: 0100

.....  
z: 1000

Ave = 3.43 bits per symbol

- The output file “encodmsg.txt” stores the encoded message with 80 0/1 characters on each row (except the last row).
- \*Important\*: (i) We use the letter with the smallest ASCII code as the representative of a subtree; and (ii) in case of ambiguity, select the two trees with the smallest representatives and when combining the two trees, the one with the smaller representative will be on the left.
- A sample input file and the corresponding output files will be given in the course moodle page soon.

Hand in the following (via Moodle):

- The source code and the executable of your program (also specify clearly the environment for which your program can be executed at the beginning of your source code as comments. Failing to do so may receive 0 marks even your program is correct.)