

Data Compression using Huffman's algorithm.

1 Project description

See the description of this problem in the book for further details.

Write a program that implements Huffman's algorithm for Data Compression. Please refer to the slides for a description of the algorithm.

Your program should read a (large) text from a file, ideally a text with at least two pages. In the first pass you should build a frequency table for each letter/character, including blank space and punctuation signs.

You should then implement a coding/decoding function, based on the Huffman coding method. The output for the coding phase should be in a file, consisting of a preamble and the encoded version of the text.

The preamble should consist of a sequence of lines of type
<character/symbol><codeword> one for each character/symbol (appearing in the text), followed by a blank line, then the encoded text.

The decoded file should just print the result of the decoding process.

2 What to turn in

You will turn in

1. *a short written report* containing: A description of the significant choices/issues in the design of your code.
2. The sourcecode of your program.

You may turn in the document via elearning.e-uvt.ro (Class moodle).

3 Coding standards

A percentage of your grade will be based on the quality of your code, so pay attention to it. Discuss changes (if any) you made to programs presented in class. Take extra care in documenting the code you are implementing on your own. Properly modularize the code (for instance implement separate functions for significant parts of the program).

4 Deadline

June 3rd, 2020.