Programming Assignment: Huffman Coding

You have not submitted. You must earn 8/10 points to pass.

Deadline Pass this assignment by December 18, 11:59 PM PST

Attention: Once you have submitted your solution, you should see your grade and a feedback about your code on the Coursera website within 10 minutes. If you want to improve your grade, just submit an improved solution.

Download the patmat.zip handout archive file and extract it somewhere on your machine.

Huffman coding is a compression algorithm that can be used to compress lists of characters.

In a normal, uncompressed text, each character is represented by the same number of bits (usually eight). In Huffman coding, each character can have a bit representation of a different length, depending on how common a character is: the characters that appear often in a text are represented by a shorter bit sequence than those being used more rarely. Every huffman code defines the specific bit sequences used to represent each character.

A Huffman code can be represented by a binary tree whose leaves represent the characters that should be encoded. The code tree below can represent the characters A to H.