Assignment # 6

Heaps

Submission Dead Line: 24/12/2021

Question 1:

Design a C++ application for compression and decompression of text (.txt) files using Huffman Encoding Algorithm. You will use a Priority queue (**Min Heap created**) and **Binary tree** as Data structures. Text file will contain only characters and numeric data, consider numeric digits as characters.

Compression

You will take name of txt file as input from user, and then follow the following steps for compression.

1. Calculate the frequency (Number of Occurrences) of all unique characters.
2. Create a leaf node for each character and add it to the priority queue (min heap).
3. While there is more than one node in the queue:
    a. Remove two nodes from queue one by one.
    b. Create a new node with these two nodes as children and the frequency of new node is equal to sum of the frequencies of two nodes.
    c. Add the new node to the queue.
4. The remaining node is the root node and the tree is complete.
5. Traverse the constructed binary tree from root to leaves assigning '0' for each left branch and a '1' for each right branch.
6. Generate the binary codes for all characters by traversing all paths till leaf nodes.
7. Encode the txt file according to generated codes and create a Binary file as compressed version.
   Note: You have to store tree of codes along data, which will be required for decompression.

Decompression

You will take name of the binary file as input from user, and then follow the following steps for decompression.

1. Read the compressed Binary File and create the tree for conversion.
2. Decompress the binary file and create a text file as output by matching the encoded binary data.

Validity Check: The decompressed text file should exactly match to the original file, which was compressed by your application.

HAPPY CODING!