Project Abstraction: Huffman Encoding and Decoding

By Wenhao Tan, Chenghui Xue, and Chuefeng Vang

Abstract

The purpose of this project is to build an encoding and decoding application using Huffman coding. Huffman coding is an encoding technique which builds a tree based on the frequencies of all the characters in the input text. Left child in a path will be considered as a “0”, while right child will be a “1”. The higher frequency one character has, the closer it will be to the root. It means characters with higher frequency will be represented with less bits, which will require less storage space for the entire text than representing each character using the same number of bits. Every node which represents a character will not have any child nodes. It means every path to a character is unique.

For encoding, the application will take a text file as input. The output will be two text files. One contains the result of encoding, while the other one will contain the key to rebuild the tree for decoding. For decoding, the application will take a encoded file with the key file and output the original text.

The data structures we plan to use for the application are tree and stack/queue. For the sorting part of the application, we plan to use quick sort.

Tentative Schedule (9 weeks)

Week 1: Write the abstract of project

Week 2: Assign parts to members

Week 3 - 4: Build codes of each part, then collect them together

Week 5- 6: Debug code and test, start designing Graphic User Interface(GUI)

Week 7: Wrap up the whole project

Week 8: Final report of project

Week 9: Finish report