**Lindsey Loveland**
11-22-16

Assignment 12

1. Design and conduct an experiment to evaluate the effectiveness of Huffman's algorithm. How is the compression ratio (compressed size / uncompressed size) affected by the number of unique characters in the original file and the frequency of the characters? Carefully describe your experiment, so that anyone reading this document could replicate your results. Submit any code required to conduct your experiment with the rest of your program and make sure that the code is well-commented.   Plot the results of your experiment. Since the organization of your plot(s) is not specified here, the labels and titles of your plots(s), as well as, your interpretation of the plots is critical.
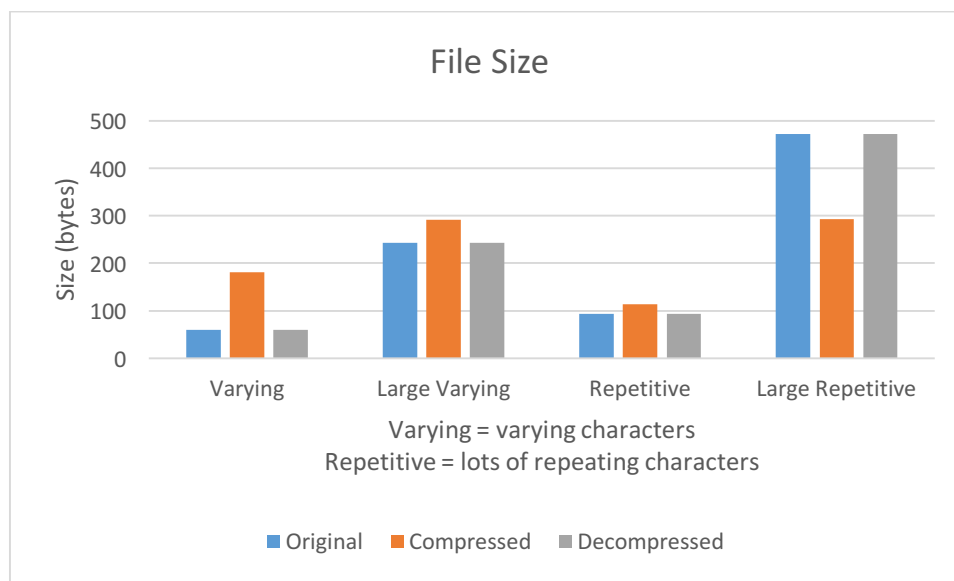
I compared the size of the original, compressed, and decompressed files for files with lots of repeating characters, and files with not a lot of repeating characters.

2. For what input files will using Huffman's algorithm result in a significantly reduced number of bits in the compressed file? For what input files can you expect little or no savings?

Input files with very repetitive content would greatly be reduced, but those files that are either small in size or are hardly repetitive mot likely would not benefit.

3. Why does Huffman's algorithm repeatedly merge the two smallest-weight trees, rather than the two largest-weight trees?

Because it would mess up the encoding if the two largest-weight trees were merged. The point of the Huffman algorithm is to make the characters with the smallest frequency have the longest encoding sequence, by being relatively deep leaf nodes.

4. Does Huffman's algorithm perform lossless or lossy data compression? Explain your answer. (A quick google search can define the difference between lossless and lossy compression).

Huffman's algorithm provides lossless data compression, primarily since we are working with text files. If it was a video, picture, or audio file, it'd be lossy since in those cases, compression does a "summary" of the data which occasionally results in loss of quality.

5. How many hours did you spend on this assignment?

8 hours

Upload your solution (.pdf only) on the assignment page by 11:59pm on Nov 23.