

CSCI3160: Special Exercise Set 5

Prepared by Yufei Tao

Problem 1. Consider the alphabet Σ with letters a, b, c, d, e, f, g, h whose frequencies are 3%, 5%, 10%, 12%, 14%, 16%, 18%, and 22%, respectively. Use Huffman's algorithm to find a prefix code on Σ that has the smallest average length.

Problem 2. Consider an alphabet Σ that contains n letters with their frequencies given, where n is a power of 2. Prove: the prefix code constructed using Huffman's algorithm has an average length of *at most* $\log_2 n$.

Problem 3. Describe how to implement Huffman's algorithm to ensure a worst-case time complexity of $O(n \log n)$, where n is the size of the alphabet Σ .

Problem 4* (Textbook Exercise 16.3-7). Consider an alphabet Σ of n letters with their frequencies given. The prefix code we construct using Huffman's algorithm is *binary* because each letter $\sigma \in \Sigma$ is mapped to a string that consists of only 0's and 1's. Now, we want the code to be *ternary*, namely, each letter $\sigma \in \Sigma$ is mapped to a string that consists of three possible characters: 0, 1, or 2. As before, the code must be prefix-free. Give an algorithm to find an encoding with the smallest average length.