Lab 5

**1. One method of reducing bandwidth use is to compress the data being transmitted. Let A = {a/20, b/15, c/5, d/15, e/45} be the alphabet and its frequency distribution. Use Huffman coding algorithm to compute the optimal coding for each character. What is the average number of bits/symbol of the codes?**

1a.

Sort: {c – 5, b – 15, d – 15, a – 20, e - 45}

Build Tree:

c(5) + b(15) -> N(20)

d(15) + N(20) -> N(35)

a(20) + N(35) -> N(55)

e(45) + N(55) -> N(100)

N(100) = Root(Symbol frequency)

Using the tree, make table:

e = 0

a = 10

b = 110

c = 1110

d = 1111

Plug into equation:

L = (45/100 × 1) + (20/100 × 2) + (15/100 × 3) + (5/100 × 4) + (15/100 × 4)

L= (0.45 × 1) + (0.2 × 2) + (0.15 × 3) + (0.05 × 4) + (0.15 x 4)

L = 0.45 + 0.4 +0.45 +0.2 +0.6 = **2.1 bits/symbol**

**2. One method of reducing bandwidth use is to compress the data being transmitted. Use the LZW algorithm to compress the string: BABAABAAA. Note that the dictionary is initialized with 256 entries, each consisting of a single character string and its corresponding ASCII code. The uppercase A has ASCII value 65 in decimal, and upper case B has ASCII value 66 in decimal, respectively. Draw diagrams to aid your explanation if appropriate.**

2a.

A -> 65 B-> 66

BABAABAAA

Go throught the word, left to right.

BA -> 256

AB -> 257

BAA -> 258

AA -> 259

Compressed output: {256, 258, 258, 65}

**3. Briefly explain how delta compression works and give an application as example where delta compression is used.**

3a.

Delta compression works by only encoding the changes between successive file or data streams. It compares two files, old and new and makes a delta file. The delta file is all the changes made to the original and if you want to get the new version you add the delta to the original.

Git uses delta compression. Git stores the changes between pushes and applies the delta to the original, rather than storing the whole file.