

Programming Assignment 4

Huffman Coding:

In your program, assume that the sample space consists of no more than 26 samples. Each sample is an English letter and its probability is represented by the frequency that letter appears, that is an integer. You need to build a binary Huffman tree using Huffman algorithm, and to assign one codeword for each letter. Your program should also translate a string of binary digits {1,0} into a sequence of letters from the sample space, based on the binary Huffman tree. There are two input files. The first file **a4-1.txt** consists of:

1. an integer at the first line indicating the number of samples within sample space.
2. starting from 2nd line, each line contains an English letter and its frequency (i.e. its associated probability).

It is a general algorithm so your program can handle all possible cases as long as the sample space consists of English letters (no matter how many are there in **a4-1.txt**).

The second input file **a4-2.txt** consists of a string of binary digits {1,0}.

Note:

1. When a codeword is assigned to a letter, always assign a **0** to the left_child (that is the node with the least frequency in the queue), and a **1** to the right_child. If two (left and right) have the same number, assign the first read-in 0, and the second read-in 1.
2. Your program should print codewords associated with a letter on the screen alphabetically, with the letter at the left, followed by a space and then the codeword. Each letter from the sample space should start with a new line.
3. Your program should output a sequence of letters on the screen after the program translates the string of {1,0} in **a4-2.txt** into a sequence of letters.
4. ***This is a single program. You may have many functions within it. All are within one cpp program.***

Here is an example of input file **a4-1.txt**:

```
9
a 26
b 139
c 42
d 88
e 98
f 121
g 35
m 999
y 1999
```

An example of input string in **a4-2.txt**

01000111000100010010100100010000.....