

**Programming Assignment 4**

Huffman algorithm using binary tree:

In this program, assume that the sample space consists of no more than 52 samples. Each sample is an English letter and its probability is represented by the frequency that letter appears, which is an integer. You need to build up a binary Huffman tree using Huffman algorithm, and to assign a codeword for each letter. Your program should be capable of translating a string of binary digits  $\{1, 0\}$  into a sequence of letters from the sample space. 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, in this particular case assign a 1 to the left-child (that is the node with the least frequency in the queue), and a 0 to the right-child, if you choose to use a binary tree structure. If two (left and right) have the same number, assign the first read-in 1, and the second read-in 0.
2. Your program should output code-words associated with a letter on the screen alphabetically.
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.***

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

```
9
a 15
b 7
c 8
d 9
e 20
f 6
g 3
h 12
k 4
```

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

```
10111000111011101101011011101111.....
```