**Decoding Huffman codes**

In 1952 David Huffman discovered an optimal compression algorithm for data. The algorithm assigns binary codes to items of data and generates a unique coding scheme based on that data. After transmittal, the binary code and its coding scheme are decoded to get back the original message. Today's .zip files and .jpegs use variants of the Huffman compression scheme.

In this lab you will only decode a Huffman message. As an example, suppose the message is

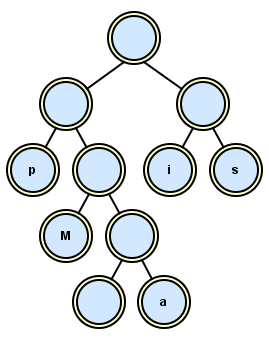
01010111110111110000010011001001110011

|  |  |
| --- | --- |
| <space> | 0110 |
| M | 010 |
| a | 0111 |
| i | 10 |
| p | 00 |
| s | 11 |

and its Huffman coding scheme is

Then the decoded message is:

\_\_M\_ \_i\_ \_s\_ \_s\_ \_i\_ s\_\_ s\_\_ \_\_i\_ \_\_p \_p\_\_ \_\_i\_ \_\_\_ \_\_M\_ \_a\_\_ \_\_p\_ \_\_s\_



In this lab the scheme is a text file which you use to construct a binary tree. The zeroes and ones describe the path to get to the leaf, where a "0" means a left child and a "1" means a right child. To be consistent with the test in CodePost, all the non-leaf nodes must store empty strings ("") and each leaf node must store the letter as a string ("M"). The scheme above turns into a tree looking like this:

After the scheme builds the tree, the tree decodes the binary message. You follow the zeroes and ones until you reach a node, a leaf, which stores the letter as a string. As you can see for yourself, each path is unique, i.e., each letter’s left-right code arrives at exactly one letter.

**Assignment**

The shell deHuffman prompts the user to read two files, the binary message and the Huffman coding scheme. Write one method that builds the tree and another method that decodes the binary message.

At the beginning, decode the two files message.maips.txt and scheme.maips.txt and see if your tree looks like the tree above. Then decode the other text files in the folder. See the next page.

What binary message (middle part)? fault  
ALWAYS ACKNOWLEDGE A FAULT. THIS WILL THROW THOSE IN AUTHORITY OFF THEIR GUARD AND GIVE YOU AN OPPORTUNITY TO COMMIT MORE.

What binary message (middle part)? laughter  
[Humanity] has unquestionably one really effective weapon--laughter. Power, money, persuasion, supplication, persecution--these can lift at a colossal humbug--push it a little--weaken it a little, century by century; but only laughter can blow it to rags and atoms at a blast. Against the assault of laughter nothing can stand.

What binary message (middle part)? money  
The holy passion of Friendship is of so sweet and steady and loyal and enduring a nature that it will last through a whole lifetime, if not asked to lend money.

What binary message (middle part)? richest  
The richest person is not the one who has the most, but the one who needs the least.

What binary message (middle part)? ships  
There are big ships and there are little ships, but there is no ship like friendship.

What binary message (middle part)? tj  
TJHSSTS

What binary message (middle part)? today  
Today is the first day of the rest of your life.

What binary message (middle part)? vegetables  
I'm gonna be round my vegetables. I'm gonna chow down my vegetables. I love you most of all, My favorite vege-table.