

# Huffman Encoding

Zachary Sparks

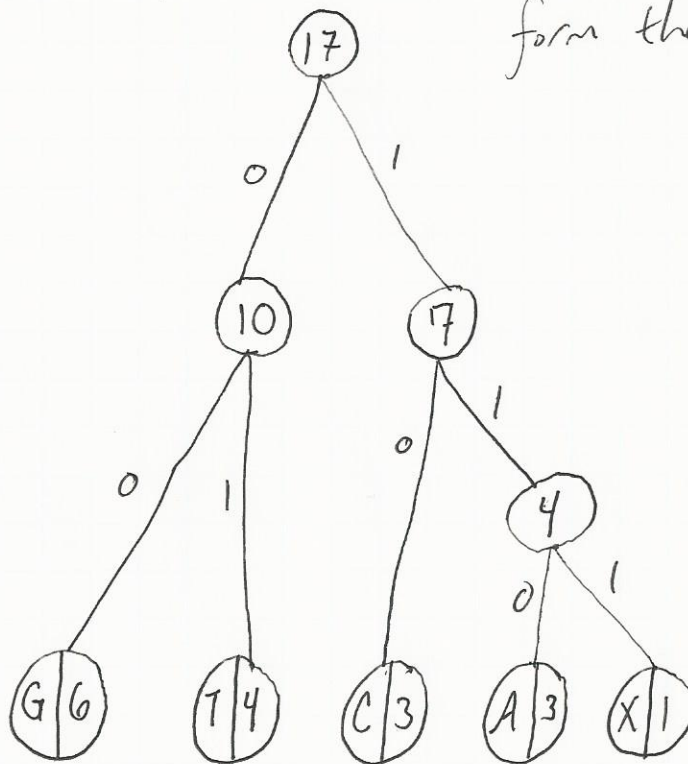
input string: "GGCATTAGGGGCCATX"

frequencies

$\Sigma$	#
G	6
C	3
A	3
T	4
X	1

1. Count frequencies of letters in  $\Sigma$
2. Label leaf nodes of tree with letters + frequencies
3. From bottom up, build tree by creating a parent with two children of the smallest nodes available. Repeat until all nodes have a single root.
4. Follow tree down to leaf nodes concatenating edge values to form the encoding. ✓

Tree



$\Sigma$	Code
G	00
C	10
A	110
T	01
X	111

Discussion a) (con't) Also, if the leaf nodes are arranged from least to greatest, or just shuffled, the output code will be different.

- a) From this example, it is possible to arrive at another encoding. This is because the letters "C" & "A" have the same frequency. As such, a tree could be constructed where their codes were swapped.
- b) In order for a remote machine to be able to decode the coded message, the first encoding table along with the coded message would need to be transmitted.