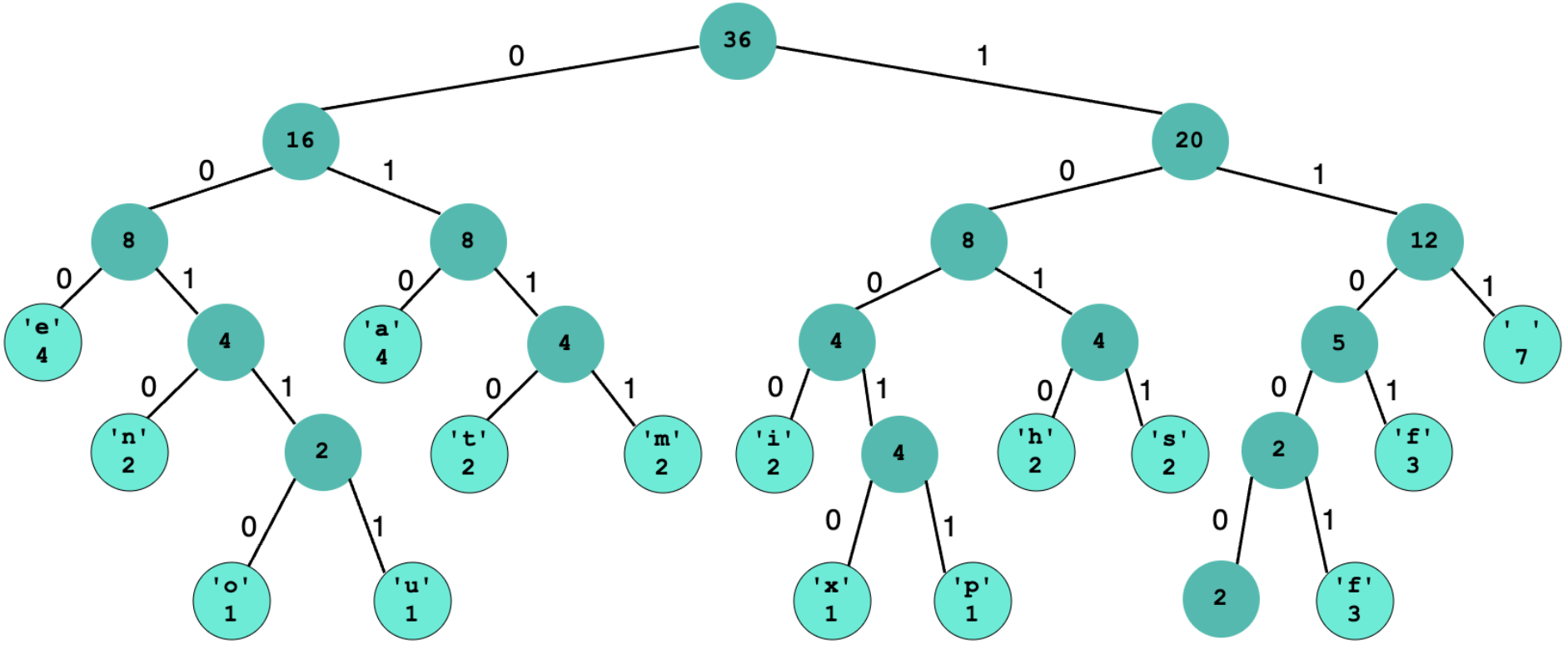
Week 25 Extension Question: Huffman Trees

A Huffman code is a type of lossless compression. It encodes characters as binary sequences. Each character sequence represents a path on the coded binary tree which leads to a leaf node character: 0 represents following the left child and 1 following the right child.
E.g. the character "u" is encoded as `00111` in the Huffman Tree below:



(a) Why do you think characters like ‘e’, ‘a’ and ‘[space]’ are at a lesser depth than other characters?

(b) Why do we not have to put any special character or break between the sequences for each character?

(c) Decode the following sequence of bits: `01101010100010111110110110000000001111000101111100100000100110`