Objectives for 11/15/2017 (Wed)

- Huffman trees
  - Exercise
char vs unsigned char

char: -128 to 127
   okay for vanilla English
unsigned char: 0 to 255
   needed for ñ è ñ etc.
   or binary files (e.g. images)
   cannot store EOF (-1).

This is why fgetc returns int
not char or unsigned char.
It needs all of those and EOF
Why do we need priority queue?

Hypothetically, try reverse:

Goal: More frequent letters higher on tree ⇒ shorter codes for more frequent letters (characters).
typedef struct TreeNode {
    unsigned char ch;
    int frequency;
    struct TreeNode* left;
    struct TreeNode* right;
} TreeNode;

typedef struct ListNode {
    TreeNode* root;
    int* next;
} ListNode;

Code book

| f | 00 |
| m | 01 |
| u | 1  |

root is like a PQ from backwards on prev page

You know it's a leaf if left and right are NULL. Leaf's have a character.
Code book

<table>
<thead>
<tr>
<th>f</th>
<th>m</th>
<th>u</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>00</td>
<td>01</td>
</tr>
</tbody>
</table>

Huffman code from your Huffman tree

Many of these, one for each character.