

# Huffman Coding FAQ

## Why can't TreePrinter access my Node class?

The lab suggests NOT to have the Node class as an inner class (like we usually do). If Node is nested inside HuffmanTree, it only exists inside an *instance* of HuffmanTree. This is not a problem, but inconvenient.

## My friend's encoding tree looks different than mine, which one of us is dumb?

There are multiple different (yet still optimal) Huffman trees possible for a single input text. Differences in how you made the nodes Comparable can change how the PriorityQueue chooses the next node, in cases where the nodes have the same frequency. What matters is that the codes are prefix-free and use the fewest possible bits to encode the given characters. Your tree should be internally consistent (but may look different than a neighbor's).

## Why do I see a '?' character when I print my Huffman tree? There isn't one in the input text!

If you're using Eclipse, this is because Eclipse's console doesn't always know how to display the pseudo-EOF character (integer 256). IntelliJ's console will display the proper Unicode character (Ā). I haven't tested how BlueJ's Terminal (console) handles it.

## Why don't I see anything printed to Eclipse's console when I run my program?

If you print an enormous amount of info (e.g. Hamlet's code tree), the Eclipse console sometimes chokes and doesn't output anything. You'll see the scroll bar move, but you don't see anything printed. You can try to remove Eclipse's console buffer size limit by going to `Window > Preferences`, then the `Run/Debug > Console` section, then you should see an option "Limit console output." Personally, I would just test with smaller files.

## How can I verify that my de-compressed file (\*.new) is exactly the same as the original (\*.txt)?

Eclipse can compare two files, similar to the Unix `diff` command that comes with most Linux distributions. To use, first select the two files (ctrl-click) then right-click and choose `Compare with > Each other`. BlueJ users can use an online diff-checking tool.

## Why does I see extra characters in my \*.new file, e.g. happy hip hophh?

Information is processed in bytes. Suppose your Huffman encoded text used bits 00101; this bit string will be padded with three zeroes until it reaches a full byte (00101000). The pseudo-EOF character is used to indicate where the end of your text should be. See the **Practical Considerations: The Pseudo-EOF** section on page 6 of the background info document for more info.