

check1.txt:

line 98 10

line 99 10

line 100 10

line 101 10

Encoded string:

```
00011011000110110001101100011011000110110001101100011011000110110
001101100011011
```

check2.txt:

line 98 4

line 99 8

line 100 16

line 101 32

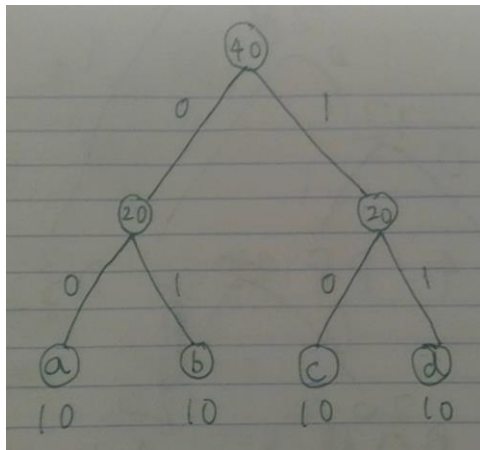
Encoded string:

[illegible]

## Tree building strategy:

Sort the leaves from smallest to largest frequencies. Connect two smallest nodes into one parent node with their sum of frequencies and no symbol on it. Continue this step until there is only one available node. Label left child as 0 and right child as 1. Encode the specific symbol with the tracing paths from the parent to the node.

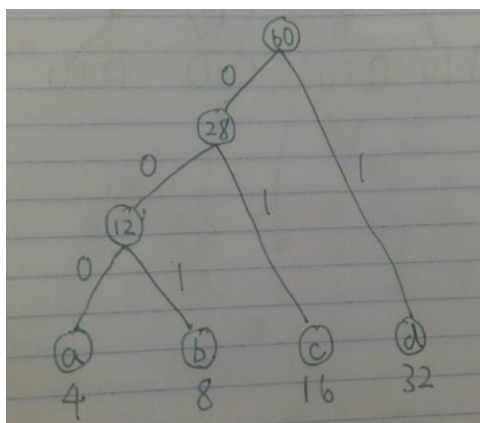
check1.txt



Encoded string:

```
00011011000110110001101100011011000110110001101100011011000110110
001101100011011
```

check2.txt



Encoded string:

[illegible]

Efficient header design:

We decide to use a header that skips all the characters with 0 frequencies. We will make a pair of elements consisting of first with its indexes relative to its ascii values and second with its frequencies. The file will just go through the “unique” symbols and skip all the 0 ones. This will let us compress a rather smaller one and the header is more efficient.