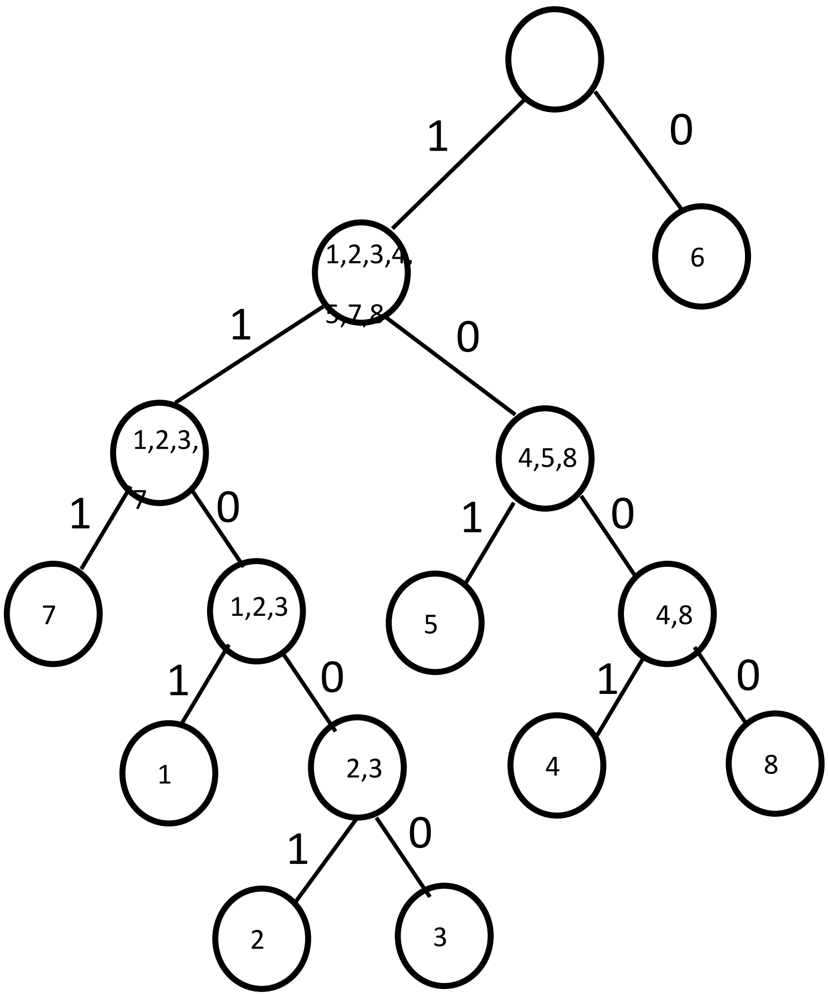
Quiz 2

Solution:

1) below is the designed binary tree for problem 1:

Therefore, the code of different characters are:

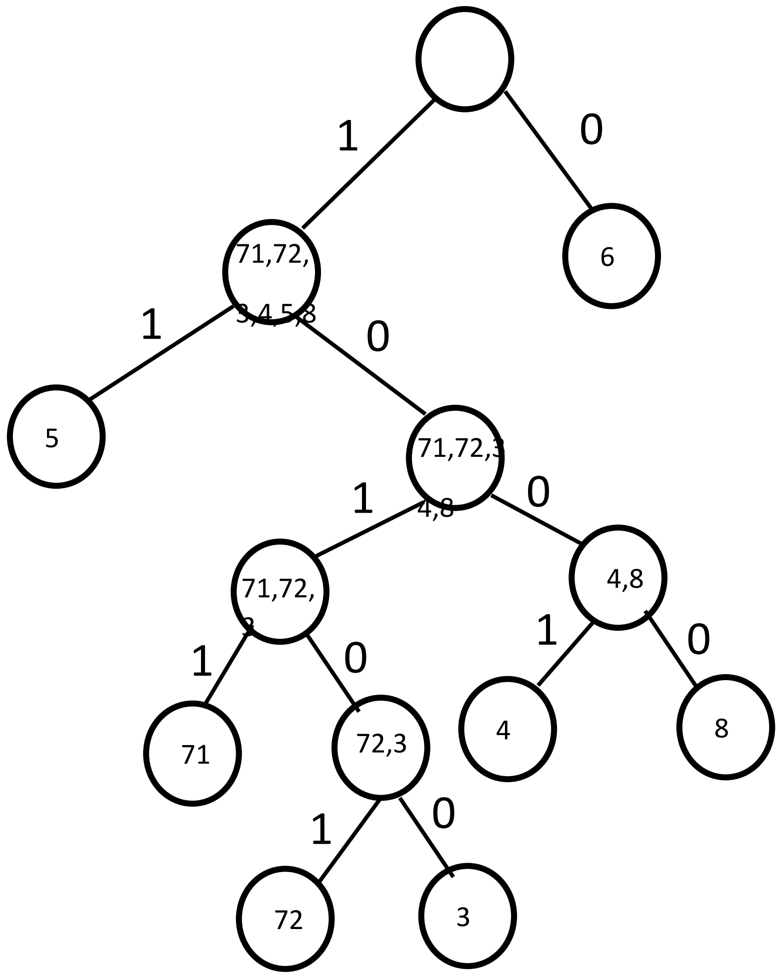
c_1: 1101, c_2:11001, c_3:11000, c_4:1001, c_5:101, c_6:0, c_7:111, c_8:1000



2) due to these constraints, we treat c_7c_1 and c_7c_2 as two new characters c_1' and c_2', and their corresponding probability will be $f_1' = 0.06$ and $f_2' = 0.04$. Now we have 7 instead of 8 characters: $c_1', c_2', c_3, c_4, c_5, c_6, c_8$

Below is the designed binary tree. Therefore, the code of different characters are:

c_1': 1011, c_2':10101, c_3: 10100, c_4: 1001, c_5:11, c_6:0, c_8:1000



In 2), we simply ignore c_7 and only encode the remaining 7 symbols. In another word, this is equivalent to assigning a codeword of length 0 to symbol c_7. Therefore, the frequency of c7 is still 0.1, but its corresponding codeword length is 0. The remaining 7 symbols keep the same frequencies.

Now we compute the number of bits of N number of symbols.

1) we have 0.06N c_1, 0.04N c_2, 0.03N c_3, 0.08N c_4, 0.2N c_5, 0.4N c_6, 0.1N c_7 and 0.09N c_8. Based on the bit number of different symbols, the total bit number M_1 is calculated as:

$$M_1 = 4 * 0.06N + 5 * 0.04N + 5 * 0.03N + 4 * 0.08N + 3 * 0.2N + 1 * 0.4N + 3 * 0.1N + 4 * 0.09N = 2.57N$$

2) Similarly, we compute the bit number M_2 of 2).

$$M_2 = 4 * 0.06N + 5 * 0.04N + 5 * 0.03N + 4 * 0.08N + 2 * 0.2N + 1 * 0.4N + 4 * 0.09N = 2.07N$$

Obviously, M_2 is smaller than M_1