ECS 452: In-Class Exercise # 3

Instructions

Name		ID (last 3 digits)	

1. Separate into groups of no more than three students each. **The group cannot be the same as any of your former groups.** Only one submission is needed for each group.
2. **Write down all the steps** that you have done to obtain your answers. You may not get full credit even when your answer is correct without showing how you get your answer.
3. **Do not panic.**

1. Consider a random variable X which has five possible values. Their probabilities are shown in the table below.

x	$p_X(x)$		$c(x)$	$\ell(x)$	$\log_2 \dfrac{1}{p_X(x)}$
a	0.2		10	2	2.3219
e	0.1		111	3	3.3219
ℓ	0.1		110	3	3.3219
n	0.3		00	2	1.7370
r	0.3		01	2	1.7370

(Diagram: Step 1, Step 2, Step 3, Step 4 combining probabilities: 0.2 (0), 0.1(1)/0.1(0) → 0.2(1) → 0.4(0/1), 0.3(0)/0.3(1) → 0.6 → 1; "Can be swapped" between 00 and 01.)

a. Find a binary Huffman code (without extension) for this random variable.
    Put the values of the codewords and the codeword lengths in the table above.
    Note that the codeword for the source symbol "ℓ" is required to be 110. (So, you do not have complete freedom on assigning 0 and 1 to the edges.)

    Step 1: We start with {0.2,0.1,0.1,0.3,0.3}.
    We first combine the two smallest probabilities which are 0.1 and 0.1.
    Now, we have {0.2,0.2,0.3,0.3}.

    Step 2: We combine the two smallest probabilities which are 0.2 and 0.2. This gives {0.4,0.3,0.3}

    Step 3: We combine the two smallest probabilities which are 0.3 and 0.3. This gives {0.4,0.6}

    Step 4: There are only two probabilities left. So, we combine them. It's reassuring to have the final sum being 1.

    Step 5: The codeword for ℓ must be 110. So, we assign the 0 and 1 to the edges such that this is the corresponding codeword. This also gives the values on the paring edges.

    Step 6: The remaining edges have some flexibility. Here we choose to assign 0 to the upper edge and 1 to the lower edge.

b. Find the expected codeword length (per source symbol) when your Huffman coding in part (a) is used (without extension).

$$\mathbb{E}[\ell(X)] = \sum_x \ell(x)p(x) = \left(2 \times (0.2 + 0.3 + 0.3)\right) + \left(3 \times (0.1 + 0.1)\right)$$
$$= (2 \times 0.8) + (3 \times 0.2) = 1.6 + 0.6 = 2.2 \text{ [bits per source symbol]}$$

c. Calculate $\sum_x \left(p_X(x) \left(\log_2 \dfrac{1}{p_X(x)} \right) \right)$

(Note that $\log_2 c = \dfrac{\log_{10} c}{\log_{10} 2} = \dfrac{\ln c}{\ln 2}$.)

$$\approx (0.2 \times 2.3219) + (2 \times 0.1 \times 3.3219) + (2 \times 0.3 \times 1.7370)$$
$$= 0.4644 + 0.6644 + 1.0422 = 2.1710 \text{ [bits per source symbol]}$$