CECS 419-519, Writing Assignment 3, Due 8:00 am, February 16th, 2024, Dr. Ebert

Directions

Make sure name is on all pages. Order pages (front and back) so that solutions are presented in their original numerical order. **Please no staples or folding of corners (your papers won't get lost). A paper clip is OK** Show all necessary work and substantiate all claims. Avoid plagiarism.

Problems

1. When an instructor gives a multiple-choice question with four choices, he finds that, on average, 57% of respondents select the best answer, 23% select the second-best answer, 15% select the third-best, and 5% select the worst answer. If the instructor randomly selects someone's answer, on average how much information does the answer convey? Explain. (10 pts)

   Solution. By definition of entropy, the amount of information conveyed is

$$H(0.57, 0.23, 0.15, 0.05) = 0.57\log(\frac{1}{0.57}) + 0.23\log(\frac{1}{0.23}) + 0.15\log(\frac{1}{0.15}) + 0.05\log(\frac{1}{0.05}) = 1.58 \text{ bits.} \quad \square$$

2. Use basic algebra to prove that the entropy function $H(p_1, p_2, p_3)$ satisfies

$$H(p_1, p_2, p_3) = H(p_1, p_2 + p_3) + (p_2 + p_3)H(\frac{p_2}{p_2 + p_3}, \frac{p_3}{p_2 + p_3}).$$

   In words, explain how to interpret this result in terms of a two step information-gathering process. (20 pts)

   Solution. We have

$$H(p_1, p_2 + p_3) + (p_2 + p_3)H(\frac{p_2}{p_2 + p_3}, \frac{p_3}{p_2 + p_3}) =$$

$$p_1 \log\frac{1}{p_1} + (p_2 + p_3)\log\frac{1}{(p_2 + p_3)} + (p_2 + p_3)\left[\frac{p_2}{p_2 + p_3}\log\frac{p_2 + p_3}{p_2} + \frac{p_3}{p_2 + p_3}\log\frac{p_2 + p_3}{p_3}\right] =$$

$$p_1 \log\frac{1}{p_1} + p_2\log\frac{1}{(p_2 + p_3)} + p_3\log\frac{1}{(p_2 + p_3)} + p_2\log\frac{p_2 + p_3}{p_2} + p_3\log\frac{p_2 + p_3}{p_3} =$$

$$p_1 \log\frac{1}{p_1} + p_2\log\frac{1}{(p_2 + p_3)} + p_3\log\frac{1}{(p_2 + p_3)} + p_2\log\frac{1}{p_2} - p_2\log\frac{1}{p_2 + p_3} + p_3\log\frac{1}{p_3} - p_3\log\frac{1}{p_2 + p_3} =$$

$$p_1 \log\frac{1}{p_1} + p_2\log\frac{1}{p_2} + p_3\log\frac{1}{p_3} = H(p_1, p_2, p_3),$$

   where, e.g., in the third-to-last equality we used the fact that $p_2\log(p_2 + p_3) = -p_2\log\frac{1}{p_2 + p_3}$.

In words, to determine the outcome we must first learn if outcome 1 has occurred or if one of outcomes 2 and 3 has occurred. The average information obtained in learning this equals

$$p_1 \log \frac{1}{p_1} + (p_2 + p_3) \log \frac{1}{(p_2 + p_3)}.$$

Moreover, the average amount of additional information that needs to be obtained equals

$$p_1 \cdot 0 + (p_2 + p_3) H(\frac{p_2}{p_2 + p_3}, \frac{p_3}{p_2 + p_3}) = (p_2 + p_3) H(\frac{p_2}{p_2 + p_3}, \frac{p_3}{p_2 + p_3}),$$

since with probability $p_2 + p_3$ we need to determine which of outcomes 2 and 3 occurred. □

3. Is there a binary prefix code of size five and with word lengths 1,2,2,3,3? If yes, then provide such a code. Otherwise, explain why not. Same question for the ternary code of size seven and with word lengths 1,1,2,2,3,3,3. (10 pts each)

   Solution. We have

$$\frac{1}{2} + 2(\frac{1}{4}) + 2(\frac{1}{8}) = 1.25 > 1$$

   which by Kraft's inequality implies no prefix code exists with these lengths. However, since

$$2(\frac{1}{3}) + 2(\frac{1}{9}) + 3(\frac{1}{27}) = 1,$$

   such a ternary prefix code does exist. □

4. Is the following code uniquely decodable: $\{0, 10, 110, 1110, 1011, 1101\}$? Justify your answer. (10 pts)

   Solution. No, since the word $w = 101110$ may be written as either $w = 10 \cdot 1110$ or as $w = 1011 \cdot 10$. □

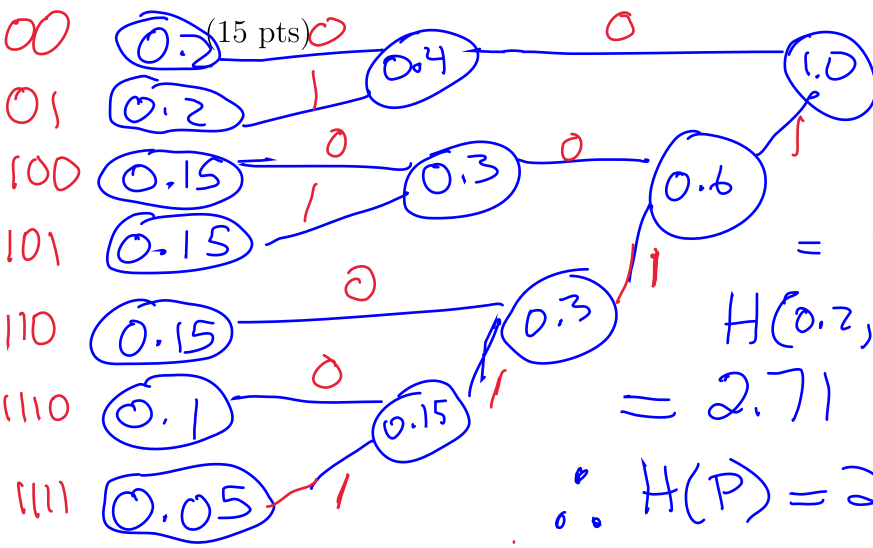
5. Use Huffman's algorithm to provide an optimal average-bit-length code C for the probability distribution

$$P = \{0.2, 0.2, 0.15, 0.15, 0.15, 0.1, 0.05\}.$$

Show the resulting tree and compute both AveLen(C, P) and $H(P)$. Verify that

$$H(P) \leq \text{AveLen}(C, P) \leq H(P) + 1.$$

(15 pts)



$\text{AveLen}(C,P) =$

$(0.4)(2) + (0.45)(3) + (0.15)(4)$

$= 0.8 + 1.35 + 0.6 = 2.75$ bits

$H(0.2, 0.2, 0.15, 0.15, 0.15, 0.1, 0.05)$

$= 2.71$

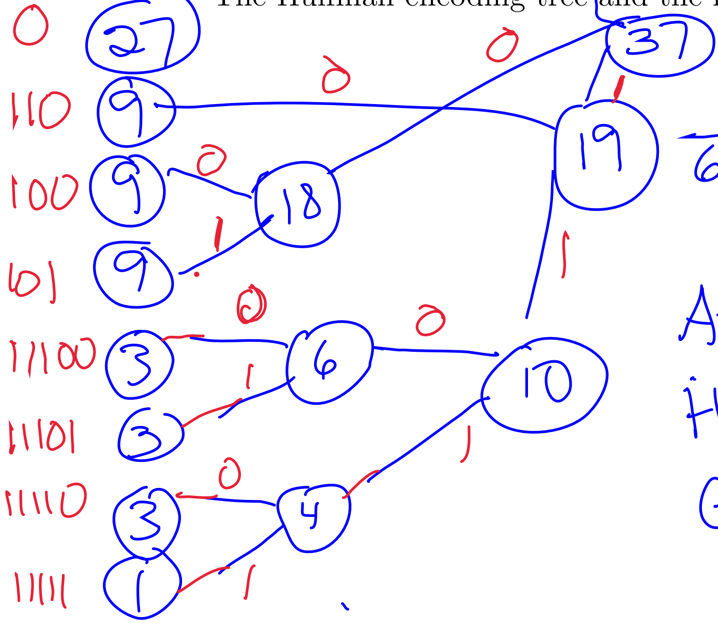
$\therefore H(P) = 2.71 < 2.75 < 2.71 + 1 = 3.71$.

6. Finish Example 4.4 of the Information Theory lecture by applying Huffman's algorithm to the ternary extension of probability distribution $P = \{0.25, 0.75\}$. Compute the average codeword length L for the resulting code and compare $L/3$ with $H(P)$ out to six decimal places. (15 pts)

The probability distribution (shown below without the hidden denominator of 64) that needs encoding is

$$P = \{1, 3, 3, 3, 9, 9, 9, 27\}.$$

The Huffman encoding tree and the resulting code is shown below.



Ave Length $(P, C) = \frac{1}{64}\left[(27)(1) + (27)(3) + (10)(5)\right] = \frac{158}{64} = 2.46875$

AveLen/3 = 0.82292

$H(0.75, 0.25) = 0.81128$

Getting closer?

7. Use the Log-Concavity Lemma to prove that, for any probability distribution $P = \{p_1, \ldots, p_n\}$ and positive real numbers $x_1, \ldots, x_n$,

$$x_1^{p_1} \cdots x_n^{p_n} \le p_1 x_1 + \cdots + p_n x_n.$$

In other words, the geometric mean of n positive real numbers is always less than or equal to their arithmetic mean. Hint: consider the expressions

$$a_j x_j / \sum a_i x_i.$$

(20 pts)

Solution. By the Log-Concavity Lemma, we have

$$\sum_{i=1}^{n} p_i \log \frac{1}{p_i} \le \sum_{i=1}^{n} p_i \log \frac{1}{p_i x_i / (\sum_{j=1}^{n} p_j x_j)} = \sum_{i=1}^{n} p_i \log \frac{\sum_{j=1}^{n} p_j x_j}{p_i x_i}.$$

This is true since, for $i = 1, \ldots, n$,

$$q_i = \frac{p_i x_i}{\sum_{j=1}^{n} p_j x_j}$$

is a probability distribution (verify!). Then, after moving the right sum to the left side and combining logarithms with the same p_i coefficient, we arrive at

$$\sum_{i=1}^{n} p_i \log \frac{x_i}{\sum_{j=1}^{n} p_j x_j} \le 0$$

which in turn can be written as

$$\sum_{i=1}^{n} p_i \log x_i \le \log \sum_{i=1}^{n} p_i x_i$$

which is true iff

$$\log(x_1^{p_1} \cdots x_n^{p_n}) \le \log \sum_{i=1}^{n} p_i x_i \Leftrightarrow$$

$$x_1^{p_1} \cdots x_n^{p_n} \le \sum_{i=1}^{n} p_i x_i. \quad \square$$