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)

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)

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)

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

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 $\text{AveLen}(C, P)$ and $H(P)$. Verify that

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

   (15 pts)

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)

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

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

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)