CECS 419-519, Writing Assignment 6, Due 8:00 am, March 22nd, 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. Solve the following.

   (a) Given a probability distribution with n probabilities, Huffman's algorithm may be applied to obtain an optimal average-length ternary code over the alphabet $\{0, 1, 2\}$, quaternary code over $\{0, 1, 2, 3\}$, and, in general, an optimal b-ary code over the alphabet

   $$\{0, 1, \ldots, b - 1\}.$$

   When going from base 2 to base b, the algorithm undergoes two changes: i) in Steps 2 and beyond, the b least probabilities are merged to form a new probability that is equal to their sum, and ii) in Step 1, the k least probabilities are merged, $2 \leq k \leq b$, where k is chosen so that, in all subsequent steps, exactly b least probabilities are merged until there is only a single probability (namely 1.0) remaining. Provide a formula for computing k. Hint: express your answer as a congruence between two quantities modulo some third quantity, i.e. of the form $X \equiv Y \bmod Z$, for some values of X, Y, and Z that depend on n, b, and k. (10 pts)

   (b) Apply your formula from part a to obtain an optimal average-length ternary code for the probability distribution $P = \{0.3, 0.3, 0.15, 0.1, 0.05, 0.05, 0.03, 0.02\}$. Verify that your initial reduction size k satisfies the equation from part a. (10 pts)

   (c) Compute the average codeword length for the code you obtained in part b and show that it is within 1 of the $H_3(P)$, the base-3 entropy of P. (5 pts)

2. An instance of the `Max Cut` decision problem is a simple graph $G = (V, E)$ and an integer $k \geq 0$. The problem is to decide if there is a way to color the vertices of G using the colors red and blue and results in there being at least k edges $e = (u, v)$ for which u and v are assigned different colors.

   (a) For a given instance (G, k) of `Max Cut`, describe a certificate in relation to (G, k). (5 pts)

   (b) Provide a semi-formal verifier algorithm that takes as input i) an instance (G, k) of `Max Cut`, ii) a certificate for (G, k) as defined in part a, and decides if the certificate is valid for (G, k). (10 pts)

(c) Use the size parameters $m = |E|$ and $n = |V|$ to describe the running time of your verifier from part b. Defend your answer in relation to the algorithm you provided for the verifier. (5 pts)

3. Answer the following.

   (a) For the mapping reduction f : `Subset Sum` $\rightarrow$ `Set Partition` defined in lecture, determine $f(S, t)$ for `Subset Sum` instance $(S = \{2, 6, 8, 9, 12, 14, 18, 26, 30, 33\}, t = 64)$. Show work. (5 pts)

   (b) Verify that both (S, t) and $f(S, t)$ are both positive instances of their respective decision problems. Provide solutions for both instances. (10 pts)

4. Given `3SAT` instance

$$\mathcal{C} = \{(x_1, \overline{x_3}, \overline{x_5}), (x_1, x_3, \overline{x_4}), (x_1, x_2, \overline{x_3}), (x_1, x_3, x_5), (\overline{x_1}, x_2, x_3), (\overline{x_1}, x_2, \overline{x_4}),$$

$$(\overline{x_1}, \overline{x_2}, x_4), (\overline{x_1}, \overline{x_2}, \overline{x_5}), (\overline{x_1}, x_3, x_5), (\overline{x_1}, \overline{x_3}, x_4), (x_2, x_4, \overline{x_5}),$$

$$(\overline{x_2}, \overline{x_3}, x_5), (\overline{x_2}, x_3, \overline{x_4})(x_2, x_3, \overline{x}_4), (x_1, x_2, x_4), (\overline{x}_1, \overline{x}_3, \overline{x}_4)\},$$

   answer the following questions.

   (a) For the mapping reduction f from `3SAT` to `Clique` presented in lecture, if $f(\mathcal{C}) = (G, k)$, then how many vertices does G have? How many edges does it have? Explain and show work. Hint: there is a clever way for calculating the number of edges without having to enumerate them. What is the value of k? (10 pts)

   (b) Does G have a k-clique? If yes, provide the vertices of the clique (for clarity make sure to indicate the vertex group for each vertex). In any case, defend your answer. (10 pts)

5. Suppose you have acccess to a polynomial-time algorithm $\mathcal{A}$ that decides the `Set Partition` problem. Use $\mathcal{A}$ to devise a polynomial-time algorithm that computes the partitioning set $A \subset S$ for any positive instance S of `SP`. Please do your best to carefully explain the key idea(s) behind your algorithm before reverting to (sigh!) pseudocode. Every problem has at least one "key idea" in its solution. Start with the key idea and *not* the boring details. (20 pts)