CECS 419-519 Midterm, March 8th, 2024, Dr. Ebert

Directions

Choose up to six problems to solve: at most three from each of Parts A and B. Number each page and Make sure your name is on all pages. Show all necessary work and substantiate all claims. Avoid plagiarism.

Part A: Choose at most three of the following problems to solve

1. Provide the instructions of a URM program P that computes the function $f(x) = \lfloor \log_2(x) \rfloor$. For each register used, clearly explain its purpose. (25 pts)

2. Solve each of the following.

   (a) Provide the instructions of the program whose Gödel number is

   $$x = 2^4 + 2^{14} + 2^{301} + 2^{381} - 1.$$

   (12 pts)

   (b) A universal program P_U is simulating a program that has 754 instructions and whose Gödel number is

   $$x = 2^7 + 2^{23} + 2^{63} + 2^{71} + 2^{105} + 2^{141} + \cdots + 2^{c_{754}} - 1.$$

   If the current configuration of the computation of P_x on some input has encoding

   $$\sigma = 2^3 + 2^5 + 2^{10} + 2^{13} + 2^{16} - 1,$$

   then provide the next configuration of the computation *and* its encoding. (13 pts)

3. Using the encoding functions presented in the Computability Basics lecture, provide an arithmetic formula for a function $h(x)$ for which $\phi_{h(x)}(y) = g(x, y) = x$, for all $x \geq 0$. (25 pts)

4. Answer the following.

   (a) Define what it means to be a positive instance of decision problem `Total`. (5 pts)

   (b) The goal is to show that `Total` is undecidable. We assume it is decidable by assuming that its decision function

   $$d_{\text{total}}(x) = \begin{cases} 1 & \text{if } x \text{ is a positive instance of } \texttt{Total} \\ 0 & \text{if } x \text{ is a negative instance of } \texttt{Total} \end{cases}$$

   is total computable. Provide the definition for how to compute the "antagonist" function $g(x)$ based on the value of $f(x)$, where $g(x)$ is used in a diagonalization argument to show that `Total` is undecidable. (8 pts)

(c) By writing the values of $g(0), g(1), \ldots$ in the appropriate cells, verify that function g is different from each computable function ϕ_i, $i = 0, 1, \ldots$, which is a contradiction since g is computable and thus should be equal to at least one of the function. (5 pts)

index\input x	0	1	2	$\cdots$	i	$\cdots$	**total?**
$\phi_0(x)$	2	12	7	$\cdots$	$\uparrow$	$\cdots$	no
$\phi_1(x)$	8	87	36	$\cdots$	96	$\cdots$	yes
$\phi_2(x)$	7	5	0	$\cdots$	$\uparrow$	$\cdots$	no
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\vdots$
$\phi_i(x)$	0	32	65	$\cdots$	5	$\cdots$	yes
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$

(d) How can we be certain that $g(x) \neq \phi_2(x)$? (7 pts)

5. Solve/answer the following.

   (a) Formally state Kleene's 2nd Recursion Theorem. Clearly define each of the symbols used in the statement. (5 pts)

   (b) The proof of Kleene's 2nd Recursion Theorem involves defining a URM program $P = ABC$, where A, B, and C are subprograms that are concatenated together to form P. Explain the role played by each subprogram. Please include information on the effect each has on the program registers. (10 pts)

   (c) Use the `self` programming concept to prove that, for any total computable $f(n)$, there is a number e for which $\phi_e(y) = \phi_{f(e)}(y)$. In other words, the programs that have Gödel numbers e and $f(e)$, respectively, compute the same function. (10 pts)

Part B: Choose at most three of the following problems to solve

6. 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)

7. Consider the following method for constructing a binary encoding of a probability distribution $p_1 \geq p_2 \geq \cdots \geq p_n$. Define q_i by $q_1 = 0$ and, for $i \geq 2$,

$$q_i = p_1 + \cdots + p_{i-1}.$$

Let $m_i = \lceil \log(\frac{1}{p_i}) \rceil$. Finally, let codeword c_i be the binary expansion of q_i truncated at the first m_i bits to the right of the binary point. For example, if $q_i = 37/64$ and $p_i = 1/20$, then $m_i = \lceil \log(20) \rceil = 5$ and c_i is the first 5 bits of q_i, that is $c_i = 10010$, since $q_i = 100101$. Prove the following about the resulting code $\mathcal{C} = \{c_1, \ldots, c_n\}$.

(a) $\mathcal{C}$ is a prefix code. (10 pts)

(b) The average codeword length A satisfies

$$H_2(p_1, \ldots, p_n) \leq A \leq H_2(p_1, \ldots, p_n) + 1.$$

(15 pts)

8. Prove that there is some constant c such that, for every strings x and y,

$$K(xy) \leq 2K(x) + K(y) + c,$$

where xy denotes the concatenation of x with y. Hint: self-terminating code. (25 pts)

9. Recall that a string x is incompressible provided $K(x) \geq |x|$. Prove that there is no program that can generate an infinite list of strings, $s_1, s_2, \ldots$ and for which each string is incompressible. Hint: use the `self` programming concept. (25 pts)

10. Answer the following.

(a) If seven players are playing the Hat Problem Game, using the optimal strategy, and the generated hat-color sequence is 0001011, does the team win? If so, who guesses? If not, explain why. Show work for either case. (10 pts)

(b) If $n = 2^k - 1$ players are playing the Hat Problem Game, using the optimal strategy, prove that the average number of correct guesses made per game equals the average number of incorrect guesses made per game. Hint: use conditional probability. (15 pts)

11. Consider the set $\mathcal{B}$ of all infinite binary sequences s for which every even-length prefix of s has the form $s = b_1 b_1 b_2 b_2 \cdots b_{\frac{n}{2}} b_{\frac{n}{2}}$. In other words, each odd-index bit of the prefix equals the previous even-index bit. Prove that $\mathcal{B}$ is a constructive null set by defining an appropriate constructive null cover for $\mathcal{B}$. (25 pts)