Computer Science 341
Discrete Mathematics

Final Exam
Outline of Solutions

Problem 1 [15 points]

A Huffman code for 3 symbols consists of the codewords 0, 10 and 11. Find the number of $\{0, 1\}$ strings of length n that are valid encodings with these codewords. In other words, find the number of distinct strings of length n that can be obtained by concatenating (copies of) the strings 0, 10 and 11. For example, 0101110 is a valid string (0 10 11 10) but 1100101 is not.

Solution:

Denote the number of valid strings of length n by a_n. We will find a recurrence formula for a_n. Divide the set of all valid strings of length n into 3 classes:

(a) valid strings that start with 0;
(b) valid strings that start with 10;
(c) valid strings that start with 11.

Clearly every valid string belongs to exactly one class. Let us count number of strings in the class (a). Observe that a string $0d_2 \ldots d_n$ is valid if and only if the string $d_2 \ldots d_n$ is valid. Therefore the number of strings in the class (a) equals the number of valid strings of length $n - 1$, which equals a_{n-1}. Similarly, number of strings in class (b) equals a_{n-2}, and number of strings in class (c) also equals a_{n-2}. We get the folowing recurrence relation:

$$a_n = a_{n-1} + 2a_{n-2}.$$

The roots of the characteristic equation $\lambda^2 - \lambda - 2 = 0$ are -1 and 2. So $a_n = A\,(-1)^n + B\,2^n$ for some constants A and B.

Now $a_1 = 1$ (the only valid string is 0), $a_2 = 3$ (the valid strings are 00, 01, and 10). Solving for A and B, we get $a_n = (-1)^n/3 + 2/3 \cdot 2^n$.

Answer: $a_n = (-1)^n/3 + 2/3 \cdot 2^n$.

Problem 2 [20 points]

Consider a cycle of length n with vertices $v_1, \ldots, v_n$. v_i is adjacent to v_{i+1} for $1 \le i \le n-1$ and v_n is adjacent to v_1. We would like to assign one of k colors to each vertex in the cycle such that adjacent vertices get different colors. Find the number of ways of doing this (as a function of n and k).

Solution:

Let T_n be the number of colorings. Note that

$$\begin{aligned} T_n =& \text{number of proper colorings of a path on } n \text{ vertices} \\ & - \text{number of proper colorings of a path on } n \text{ vertices, s.t. colors of } v_1 \text{ and } v_n \text{ are the same.} \end{aligned}$$

The first term equals $k(k-1)^{n-1}$ (there are k ways to color the first vertex, then there are $k-1$ ways to color the second vertex, there are $k-1$ ways to color each consecutive vertex). To compute the second term note that we can replace v_1 and v_n with a single vertex and get a proper coloring of a cycle on $n-1$ vertices. In fact, this replacement defines a bijection. Therefore, the second term equals T_{n-1}. We have

$$T_n = k(k-1)^{n-1} - T_{n-1} = k(k-1)^{n-1} - k(k-1)^{n-2} + T_{n-2} = \ldots$$

$$T_3 = k(k-1)(k-2).$$

Note that values of T_1 and T_2 are not well-defined since there are no cycles on 1 and 2 vertices. However, it is consistent to put $T_2 = k(k-1)$, and $T_1 = 0$ — these values give the correct value for T_3.

Expand the formula for T_n:

$$T_n = k\left((k-1)^{n-1} + \cdots + (-1)^n(k-1) + T_1\right).$$

Add up the geometric progression

$$T_n = k(k-1)\frac{(k-1)^{n-1} + (-1)^n}{(k-1)-(-1)} = (k-1)^n + (-1)^n(k-1).$$

Answer: $(k-1)^n + (-1)^n(k-1)$.

Problem 3 [20 points]

A game is played on an $n \times n$ board, as follows: Here (i, j) denotes the square in row i and column j. The goal is to move a special piece (Tiger) from the bottom left corner square $(1, 1)$ to the upper right corner square (n, n). In each move, Tiger can move one square up or one square to the right, i.e. from (i, j), he can move to $(i + 1, j)$ or $(i, j + 1)$. To complicate matters, some squares on the board are marked as forbidden and Tiger is not allowed to land on them. At the beginning of the game, each square on the board (other than $(1, 1)$ and (n, n)) is marked as forbidden at random with probability $1/2$. These choices are made independently for every square. Let p_n be the probability that Tiger can move from $(1, 1)$ to (n, n) avoiding any forbidden squares. Find $\lim_{n\to\infty} p_n$.

Hint: Think of the paths from $(1, 1)$ to (n, n).

Solution:

There are $\binom{2(n-1)}{n-1} = O(\frac{2^{2n}}{\sqrt{n}})$ paths from the field $(1, 1)$ to the field (n, n) (see problem 1a, precept 2, `http://www.cs.princeton.edu/courses/archive/fall05/cos341/Precepts/precept2sol.pdf`). Each path consists of $2n-3$ squares (not counting the corner fields). Each of these squares is forbidden with probability $\frac{1}{2}$. Therefore, the probability that a path survives (e.g. none of its fields is forbidden) is $\frac{1}{2^{2n-3}}$. By the union bound, the probability that at least one path survives is at most $O(\frac{2^{2n}}{\sqrt{n}} \cdot \frac{1}{2^{2n-3}}) = O(\frac{1}{\sqrt{n}})$. Therefore, $p_n = O(\frac{1}{\sqrt{n}})$ tends to 0 as $n \to \infty$.

Answer: $\lim_{n\to\infty} p_n = 0$.

We used an approximation formula for binomial coefficients:

$$\binom{2n}{n} = O\left(\frac{2^{2n}}{\sqrt{n}}\right).$$

This formula can be derived from the Stirling Approximation for $n!$ as follows:

$$\binom{2n}{n} = \frac{(2n)!}{n!n!} = \frac{\sqrt{2\pi(2n)}\left(\frac{2n}{e}\right)^{2n}(1+o(1))}{\left(\sqrt{2\pi n}\left(\frac{n}{e}\right)^{n}(1+o(1))\right)\left(\sqrt{2\pi n}\left(\frac{n}{e}\right)^{n}(1+o(1))\right)}$$
$$= \frac{2^{2n}}{\sqrt{\pi n}}(1+o(1)) = O\left(\frac{2^{2n}}{\sqrt{n}}\right).$$

Problem 4 [25 points]

A random tree on n vertices is formed in the following way: The vertices are added to the tree in the sequence $v_1, \dots, v_n$. The first vertex v_1 is considered the root. The ith vertex v_i is added to the tree by connecting it to one of the previous $i-1$ vertices chosen uniformly and at random from amongst $v_1, \dots, v_{i-1}$. Let P_i be the length of the path (i.e. number of edges on the path) from v_i to the root.

(a) [5 points] Let $R_n = \mathbf{E}[P_n]$. Write a recurrence relation for R_n.

(b) [7 points] Use the recurrence for R_n to compute $\mathbf{E}[P_n]$.
Write down an exact expression (not necessarily in closed form) and also find a closed form function $f(n)$ such that $\mathbf{E}[P_n] = \Theta(f(n))$.

(c) [5 points] Let $S_n = \mathbf{E}[(P_n)^2]$. Write a recurrence relation for S_n.

(d) [8 points] Use the recurrence for S_n to compute $E[(P_n)^2]$.
Write down an exact expression (not necessarily in closed form) and also find a closed form function $g(n)$ such that $\mathbf{E}[(P_n)^2] = \Theta(g(n))$

Solution:
a) We have

$$\begin{aligned} R_{n+1} = \mathbf{E}[P_{n+1}] &= \sum_{i=1}^{n} \mathbf{E}[P_{n+1} | v_{n+1} \text{ is connected to } v_i] \Pr[v_{n+1} \text{ is connected to } v_i] \\ &= \sum_{i=1}^{n} \mathbf{E}[P_i + 1 | v_{n+1} \text{ is connected to } v_i] \cdot \frac{1}{n} \\ &= \frac{1}{n} \sum_{i=1}^{n} (R_i + 1) = 1 + \frac{\sum_{i=1}^{n} R_i}{n} \end{aligned}$$

Answer: $R_{n+1} = 1 + \dfrac{\sum_{i=1}^{n} R_i}{n}$.

b) We have

$$\begin{aligned} nR_{n+1} &= n + (R_1 + \cdots + R_n); \\ (n-1)R_n &= (n-1) + (R_1 + \cdots + R_{n-1}). \end{aligned}$$

Subtracting the second identity from the first, we get

$$nR_{n+1} = nR_n + 1.$$

So

$$R_{n+1} = nR_n + 1/n.$$

Note that $R_1 = 0$. So
$$R_{n+1} = 1 + \frac{1}{2} + \cdots + \frac{1}{n} = H_n.$$
Recall that $H_n = \ln n + O(1)$.

Answer: $R_n = H_{n-1} = \Theta(\ln n)$.

c) We have

$$\begin{aligned}
S_{n+1} = \mathbf{E}[P_{n+1}^2] &= \sum_{i=1}^{n} \mathbf{E}[P_{n+1}^2 | v_{n+1} \text{ is connected to } v_i] \Pr[v_{n+1} \text{ is connected to } v_i] \\
&= \sum_{i=1}^{n} \mathbf{E}[(P_i+1)^2 | v_{n+1} \text{ is connected to } v_i] \cdot \frac{1}{n} \\
&= \frac{1}{n} \sum_{i=1}^{n} \left(\mathbf{E}[P_i^2] + \mathbf{E}[2P_i] + 1 \right) \\
&= \frac{1}{n} \sum_{i=1}^{n} (S_i + 2R_i + 1)
\end{aligned}$$

Now, by part a, and then by part b,
$$\frac{2}{n} \sum_{i=1}^{n} R_i = 2(R_{n+1} - 1) = 2(H_n - 1).$$

We conclude
$$S_{n+1} = 2H_n - 1 + \frac{1}{n} \sum_{i=1}^{n} S_i.$$

Answer: $S_{n+1} = 2H_n - 1 + \frac{1}{n} \sum_{i=1}^{n} S_i$.

d) Write recurrence relation for S_{n+1} and S_n:
$$S_{n+1} = 2H_n - 1 + \frac{1}{n} \sum_{i=1}^{n} S_i;$$
$$S_n = 2H_{n-1} - 1 + \frac{1}{n-1} \sum_{i=1}^{n-1} S_i.$$
Multiply the second formula by $\frac{n-1}{n}$ and subtract from the first formula:
$$S_{n+1} - \frac{n-1}{n} S_n = 2H_n - 2\frac{n-1}{n} H_{n-1} - 1 + \frac{n-1}{n} + \frac{S_n}{n}.$$
We get
$$S_{n+1} = S_n + 2H_n - 2\frac{n-1}{n} H_{n-1} - \frac{1}{n}.$$

Since $S_1 = \mathbf{E}[P_1^2] = 0$, we have

$$\begin{aligned} S_{n+1} &= \sum_{i=1}^{n}(S_{i+1} - S_i) = \sum_{i=1}^{n}\left(2H_i - 2\frac{i-1}{i}H_{i-1} - \frac{1}{i}\right) \\ &= 2\left(\sum_{i=1}^{n} H_i - \sum_{i=1}^{n-1}\frac{i}{i+1}H_i\right) - H_n = H_n + 2\sum_{i=1}^{n-1}\frac{H_i}{i+1}. \end{aligned}$$

Now,

$$2\sum_{i=1}^{n-1}\frac{H_i}{i+1} = 2\sum_{i=1}^{n-1}\sum_{j=1}^{i}\frac{1}{i+1}\cdot\frac{1}{j} = 2\sum_{i=2}^{n}\sum_{j=1}^{i-1}\frac{1}{i}\cdot\frac{1}{j} = \left(\sum_{i=1}^{n}\frac{1}{i}\right)^2 - \sum_{i=1}^{n}\frac{1}{i^2} = H_n^2 - \sum_{i=1}^{n}\frac{1}{i^2}.$$

Therefore,

$$S_{n+1} = H_n^2 + H_n - \sum_{i=1}^{n}\frac{1}{i^2}.$$

Since $H_n = \ln n + o(1)$, and $\sum_{i=1}^{n}\frac{1}{i^2}$ is bounded by $\sum_{i=1}^{\infty}\frac{1}{i^2} = \frac{\pi^2}{6}$, we get

$$S_{n+1} = \Theta((\log n)^2).$$

Answer:

$$S_{n+1} = H_n^2 + H_n - \sum_{i=1}^{n}\frac{1}{i^2}.$$

$$S_{n+1} = \Theta((\log n)^2).$$

Problem 5 [20 points]

Consider a sequence of graphs constructed as follows. G_1 consists of a single vertex and G_2 consists of two vertices connected by an edge. Graph G_{i+1} is constructed from G_i by adding some vertices and edges as follows. Suppose G_i has vertex set $V = \{v_1, \ldots, v_n\}$. Then G_{i+1} is constructed by adding vertices $U \cup \{x\}$ where $U = \{u_1, \ldots, u_n\}$. In addition to the edges already present in G_i, the following edges are added to G_{i+1}: For every i, u_i is connected to x and u_i is connected to every neighbor of v_i. Note that G_3 obtained from this process is the 5-cycle. The goal of this problem is to analyze the number of colors in a proper coloring of G_k. Recall that a proper coloring of a graph is an assignment of colors to vertices so that adjacent vertices get different colors.

(a) [5 points] Prove that G_k has a proper coloring with at most k colors.

(b) [15 points] Prove that any proper coloring of G_k must use at least k colors.

Hint: If G_{k+1} can be colored with k colors, show that G_k can be colored with $k-1$ colors.

Solution:

a) We prove by induction on k that G_k has a proper coloring with at most k colors.

Base case $k = 1$*:* G_1 consists of a single vertex, which can be colored with one color.

Inductive step: Suppose G_k has a proper coloring with k colors. We color G_{k+1} with $k+1$ colors as follows:

1) We color G_k (e.g. vertices in V) with k colors (we can do this by the inductive hypothesis).

2) Color each vertex u_i with the same color as v_i.

3) Finally, color v_0 with a new color $k + 1$.

We need to show that this coloring is proper, e.g. there are no adjacent vertices of the same color.

1) Clearly, any adjacent vertices v_i and v_j are colored with different colors.

2) v_0 is colored with the color which differs from all other colors.

3) Any two vertices in U are not adjacent.

4) If vertices u_i and v_j are adjacent, then v_i and v_j are also adjacent, and thus they are colored with different colors. Since u_i and v_i are colored with the same color, u_i and v_j are colored with different colors.

We showed that the coloring is proper. This concludes the inductive step.

b) We prove by induction on k that G_{k+1} cannot be colored with k colors.

Base case $k = 1$*:* G_1 clearly cannot be colored with 0 colors.

Inductive step: Suppose that G_k cannot be colored with $k - 1$ colors. We will prove that G_{k+1} cannot be colored with k colors. Assume the contrary.

Then consider a proper coloring of G_{k+1} with k colors. Denote the color of the vertex v_0 by c. If none of the vertices in G_k are colored with the color c we are done: G_k is colored with at most $k-1$ colors. Otherwise, let us recolor all the vertices colored with the color c. Namely, if v_i is of color c, recolor it with the color of u_i. Note that u_i is not colored with c, since u_i is adjacent to v_0. Thus the new coloring of G_k uses at most $k-1$ colors.

We claim that the new coloring is a proper coloring of G_k (but not necessarily of G_{k+1}). Indeed, let v_i and v_j be two adjacent vertices. If none of them have been recolored, they are colored in different colors. Note that it is not possible that the both of them have been recolored, since these vertices are adjacent, and thus both of them could not be colored with the color c. Suppose that v_i has been recolored. Since v_i and v_j are adjacent, u_i and v_j are also adjacent and are colored with different colors. But u_i is colored with the same color as v_i, so v_i and v_j are of different colors.

We get a proper $k-1$-coloring of G_k, which contradicts to the inductive hypothesis.