Homework 3
Due 2006/11/23

Oded Regev
Department of Computer Science
Tel Aviv University

Instructions

Collaboration: Collaboration is allowed, but limit yourselves to groups of size at most three.

References: Try not to run to reference material to answer questions (this also includes the web!). Try to think about the problem to see if you can solve it without consulting any external sources. If this fails, you may ask me for a hint, or look up any reference material.

Writeup: You must write the solutions by yourselves. For each question, cite all references used (or write 'none') and collaborators (or write 'alone'). Explain why you needed to consult any of the references.

Deadline: The deadline is strict.

Problems[1]

1. **Hamming codes:**
   (a) Describe explicitly the parity check matrix and the generator matrix in systematic form of a Hamming code of block length $2^\ell - 1$. What is its message length, rate, distance, relative distance?
   (b) Show that if C is a t-error-correcting code (not necessarily linear) in $\{0,1\}^n$, then $|C| \leq 2^n/\mathrm{vol}(n,t)$, where $\mathrm{vol}(n,t) = \sum_{i=0}^{t} \binom{n}{i}$ is the volume of the Hamming ball of radius t.
   (c) Conclude that the Hamming codes of Part (a) achieve the highest possible rate for their distance.

2. **Improving the Hamming code:** Let C be an $(n,k,d)_2$ code for some odd d. Construct the code C' by adding a 'parity bit' to each codewords: i.e., a bit whose value is set to the XOR of all n other bits.
   (a) Find the parameters of C' (message length, block length, distance). What are its error detection/correction abilities?
   (b) What happens if we add another parity bit? What happens over non-binary alphabets?
   (c) Show that for any $l \geq 1$ there exists a $[2^l, 2^l - l - 1, 4]_2$ code. Briefly describe the generating/parity check matrix of this code.

3. **Extending the Hamming code:** For any prime power q, find a family of perfect q-ary codes of minimum distance 3.

4. **Pairwise independent spaces:**
   (a) Let H be the $\ell \times (2^\ell - 1)$ parity check matrix of a binary Hamming code. Show that the set of vectors $\{\mathbf{x}H \mid \mathbf{x} \in \{0,1\}^\ell\}$ forms a pairwise independent space (i.e., that for any $i, j \in [2^\ell - 1]$, $i \neq j$, the pair $(\mathbf{y}_i, \mathbf{y}_j)$ where $\mathbf{y}$ is chosen uniformly from that set, is distributed uniformly on $\{0,1\}^2$).
   (b) Extra credit question: Show that any pairwise independent space on n bits must contain at least $n+1$ points (and hence Part (a) is tight).

5. **Derandomization:** A directed cut in a directed graph $G = (V,E)$ is an ordered partition $(S, \overline{S})$ of V. The size of the directed cut is the number of edges $(u,v) \in E$ with $u \in S$ and $v \in \overline{S}$.
   (a) Show that every graph has a directed cut of size at least $|E|/4$.
   (b) Give a deterministic polynomial time algorithm to find such a directed cut in a given graph. (There are two natural solutions to this problem – one that involves pairwise independence and one that doesn't. Guess which one I want!)

[1] Questions 1-6 are based on Madhu Sudan's problems

6. **The hat problem:** n players enter a room and a red or yellow hat is placed on each player's head. The color of each hat is determined by an independent coin toss. Each person can see the other players' hats but not his own. No communication of any sort is allowed, except for an initial strategy session before the game begins. Once they have had a chance to look at the other hats, the players must simultaneously guess the color of their own hats or pass. The group shares a $10 million prize if at least one player guesses correctly and no player guesses incorrectly. Your goal is to find a strategy for the group that maximizes their chances of winning the prize.

   Before you go on, try to obtain probability $3/4$ for $n = 3$.

   (a) Let $\mathcal{G}$ be the family of all directed graphs G on vertex set $\{0,1\}^n$ satisfying that for any edge (u, v) in G, u and v differ in at most one coordinate. For $G \in \mathcal{G}$, let $K(G)$ be the number of vertices of G with in-degree at least one, and out-degree zero. Show that the maximum probability of winning the hat problem is given by $\max_{G\in\mathcal{G}} K(G)/2^n$.

   (b) Show that $K(G)/2^n$ is at most $\frac{n}{n+1}$ for any $G \in \mathcal{G}$.

   (c) Show that if $n = 2^\ell - 1$, then there exists $G \in \mathcal{G}$ with $K(G)/2^n = \frac{n}{n+1}$.

7. **One-way communication complexity of comparison:** Alice has a number x, and Bob has a number y, both taken from the set $\{0, 1, \ldots, 2^n - 1\}$. A protocol for comparison using one-way communication is a protocol in which Alice sends Bob an m-bit message (based on her input x), and then Bob outputs a yes/no answer based on his input and on the received message. The protocol should be such that if $y > x$ then Bob outputs yes with probability at least $2/3$, and that if $y \leq x$, Bob outputs no with probability at least $2/3$. The parties are allowed to use randomization (Alice in her message, and Bob in his decision). Show that in any such protocol we must have $m \geq (1 - H(2/3))n$. Hint: use the lower bound on weak RAC from last week's homework.

8. **Constructing random access codes:** Recall that an (n, m, p) random access code is a randomized function f from $\{0,1\}^n$ to $\{0,1\}^m$ and randomized functions $g_1, \ldots, g_m : \{0,1\}^m \to \{0,1\}$ such that

$$\forall x \in \{0,1\}^n, i \in [n], \quad \Pr_{y\sim f(x)}[g_i(y) = x_i] \geq p.$$

   (a) Show that for any $\varepsilon > 0$, $p < 1/2$, and large enough n, there exists a subset A of $\{0,1\}^n$ of size at most $2^{(1-H(p)+\varepsilon)n}$ with the property that for any $x \in \{0,1\}^n$ there exists a $y \in A$ within Hamming distance pn. In other words, $\Delta(x, A) \leq pn$ for all x. (Such sets are known as covering codes.)

   (b) For any $\varepsilon > 0$, $p > 1/2$, and large enough n, find an (n, m, p) random access code with $m \leq (1 - H(p) + \varepsilon)n$ satisfying the weaker property

$$\forall x \in \{0,1\}^n, \quad \Pr_{y\sim f(x), i\sim[n]}[g_i(y) = x_i] \geq p,$$

   i.e., decoding only has to be correct on an average bit.

   (c) Extra credit question: improve your construction to an actual random access code.

9. **A tough one:** (bonus; no deadline) Alice and Bob play the following game. In each round, one fair coin is tossed. Just before the coin is tossed, Alice and Bob simultaneously declare their guess for the result of the coin toss. They win the round if both guessed correctly. The goal is to maximize the expected fraction of rounds won, as the number of rounds goes to infinity.

   So far, the answer is obviously 50%: for instance, Alice and Bob can always declare 'heads' and clearly they can't do anything better. However, the story is that Alice has amazing precognitive abilities so she knows in advance the results of all the coin tosses. Can you come up with a good strategy? (Notice that each player knows the other player's votes in all previous rounds.) The best strategy succeeds on roughly 81% of the rounds.