Essential Coding Theory Madhu Sudan
6.895
Due: Wednesday September 15, 2004 (Part 1) & Wednesday, September 22, 2004 (Part 2)

Problem Set 1

Instructions

References: In general, try not to run to reference material to answer questions. Try to think about the problem to see if you can solve it without consulting any external sources. If this fails, you may look up any reference material.

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

Writeup: You must write the solutions in latex, by yourselves. Cite all references and collaborators. Explain why you needed to consult any of the references, if you did consult any.

Alternative to writing (on experimental basis): If you prefer to explain your solution(s) in words to me, you may try to find me in my office and do so. (If the class turns to be too big, I might have to withdraw this option.)

Problems

Problems 1-3 form Part 1 of this problem set. Problem 4 is Part 2 of this problem set.

1. (Linear Algebra Review):

   (a) Let

$$H = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}.$$

   Find the largest matrix G of full column rank such that $G \cdot H$ is an all 0 matrix, where all operations are carried out modulo 2.

   (b) What can you say about the minimum distance of the code generated by G, i.e., the code $\{x \cdot G | x \in \{0,1\}^4\}$.

   (c) **(Not to be turned in.)** Give an efficient algorithm for Part (a), i.e., to compute, given an $m \times n$ matrix H, an $n \times k$ matrix G such that $G \cdot H = 0$ (modulo 2).

2. (Probability Review): An instance of the MAX 3SAT problem ϕ consists of m "clauses" $C_1, \ldots, C_m$ on n Boolean variables $x_1, \ldots, x_n$, where a clause is the disjunction of (exactly) 3 distinct literals; and each literal is either a variable x_i or its negation $\neg x_i$. The goal is to find a 0/1 assignment to the n variables that "satisfies" the maximum number of clauses, where a clause is satisfied if at least one of the literals in the clause is set to 1 (and the assignment to a literal x_i is the same as the assignment to the variable x_i, while the assignment to the literal $\neg x_i$ is the complement of the assignment to x_i (i.e., $1 - x_i$).

   EXAMPLE: ϕ may consist of the clauses $C_1 = x_1 \vee x_2 \vee x_3$, $C_2 = x_1 \vee x_2 \vee \neg x_3$, $C_3 = \neg x_2 \vee \neg x_3 \vee x_4$, etc. The assignment $x_1 = x_2 = 1$, and $x_3 = x_4 = 0$. satisfies C_1 and C_2 but not C_3. Setting all variables to 1, satisfies all three clauses.

   PROBLEM: For any MAX 3SAT instance ϕ with m clauses, prove that there exists an assignment satisfying at least $\frac{7}{8} \cdot m$ clauses.

3. (Combinatorics Exercise): Let $E_1 : \{0,1\}^{k_1} \to \{0,1\}^{n_1}$ be an encoding function that maps k_1 bit messages to n_1 bits codewords such that every pair of codewords differ in at least d_1 locations. Similarly let $E_2 : \{0,1\}^{k_2} \to \{0,1\}^{n_2}$ be an encoding function that maps k_2 bit messages to n_2 bits codewords such that every pair of codewords differ in at least d_2 locations. Now consider the map $E_{12} : \{0,1\}^{k_1 \times k_2} \to \{0,1\}^{n_1 \times n_2}$, which views a message M as a $k_1 \times k_2$ matrix and encodes each column first by the map E_1 to get an $n_1 \times k_2$ matrix M_1 and then encodes each row of M_1 by E_2 to get an $n_1 \times n_2$ matrix M_{12} which is the final encoding of M.

   (a) What is the minimum distance of the mapping M_{12}?

   (b) Suppose we reversed the steps above to first encode the rows with E_2 and then encode the columns with E_1. Call this the encoding E_{21}. Give an example of maps E_1 and E_2 for which $E_{12} \neq E_{21}$.

   (c) (Linear algebra workout) Suppose E_1 and E_2 are linear maps; i.e., there exist matrices G_1 and G_2 such that $x \mapsto_{E_i} x \cdot G_i$. Then show that E_{12} is a linear map, and that $E_{12} = E_{21}$.

4. (An Application of Codes): This is a long exercise whose goal is to "derandomize" Problem 2. Specifically the final outcome we seek is a deterministic algorithm to compute, given a MAX 3SAT instance ϕ with m clauses, an assignment that satisfies at least $\frac{7}{8} \cdot m$ clauses of ϕ. We start with some definitions.

   DEFINITION: A probability space on $\{0,1\}^n$ is a function $P : \{0,1\}^n \to [0,1]$ such that $\sum_{\alpha \in \{0,1\}^n} P(\alpha) = 1$. The support of a distribution P is the set of α such that $P(\alpha) > 0$. A probability space is said to be 3-wise independent if for every triple $i, j, k \in \{1, \ldots, n\}$ of distinct indices, the marginal distribution P_{ijk} of P on the (i, j, k)th coordinates is the uniform distribution.[1]

   (a) Let P be a 3-wise independent distribution. Let ϕ be a MAX 3SAT instance with m clauses. Show that there exists an assignment α in the support of P such that α satisfies $\frac{7}{8} \cdot m$ clauses of ϕ.

[1] More elaborately, for $b_1, b_2, b_3 \in \{0,1\}$, let $S_{b_1,b_2,b_3} = \{\alpha \in \{0,1\}^n \mid \alpha_i = b_1, \alpha_j = b_2, \alpha_k = b_3\}$. Now let $P_{ijk}(b_1, b_2, b_3) = \sum_{\alpha \in S_{b_1,b_2,b_3}} P(\alpha)$. This is the marginal distribution of P onto its i, j, kth coordinates. We require this to be uniform, i.e., $P_{ijk}(b_1, b_2, b_3) = \frac{1}{8}$ for every i, j, k, b_1, b_2, b_3.

(b) Given an $m \times n$ matrix H, define an associated probability space P_H, where $P_H(x) = \frac{1}{M}$ if $H \cdot x = \mathbf{0}$ and $P_H(x) = 0$ otherwise.

i. For what value of M does the above satisfy the definition of a probability space. (Note that M is not allowed to depend on x.)

ii. Give a necessary and sufficient condition (using coding theoretic terms) for P_H to be 3-wise independent.

iii. Use the above characterization, to give a 3-wise independent probability space of small support.

(c) Put the above together to describe an efficient deterministic algorithm that computes an assignment satisfying $\frac{7}{8} \cdot m$ clauses given any instance of MAX 3SAT with m clauses.