Essential Coding Theory Madhu Sudan
6.895
Due: Wednesday October 13, 2004

Problem Set 2

Instructions

- For reference, collaboration, writeup etc. follow instructions for PS 1.

Problems

1. Shannon Capacity: As mentioned in class, Shannon's coding theorem covers a broad class of channels. Here we explore a slight variant of the channel described in lectures and analyze its capacity.

   Consider a channel that takes as inputs bits, and outputs bits as follows: On input $b \in \{0, 1\}$, the channel outputs $b' = b$ with probability $\frac{1}{2}$ and outputs 0 with probability $\frac{1}{3}$ and 1 with probability $\frac{1}{6}$. (And this process is independent for each transmitted bit.)

   (a) Lower bound the capacity of the channel. I.e., give an encoding function $E : \{0, 1\}^{Rn} \rightarrow \{0, 1\}^n$ and a decoding function $D : \{0, 1\}^n \rightarrow \{0, 1\}^{Rn}$ such that $\Pr_{m,\text{Channel}}[D(\text{Channel}(E(x)) = x]$ with probability close to 1. Your aim should be to get as high a value of R as possible.

   (b) Upper bound the capacity of the channel. I.e., show that for any encoding, decoding functions as above, the probability that $\Pr_{m,\text{Channel}}[D(\text{Channel}(E(x)) = x]$ goes to 0.

2. Shannon vs. Hamming:

   (a) Prove Shannon's theorem for the binary symmetric channel with flip probability p, holds even the encoding function E is restricted to being linear.

   (b) Prove that if E is linear and the probability of decoding error is exponentially small ($2^{-\epsilon n}$), then the image of E is an asymptotically good code.

3. Many proofs of the Gilbert-Varshamov Bound: First, some "terminology" we did not introduce in lecture. The theorem: "There exists a family of codes of rate R and relative distance δ, satisfying $R \geq 1 - H_2(\delta)$" is often referred to as the Gilbert-Varshamov (or Varshamov-Gilbert) Bound. Gilbert first discovered this for general codes; and then Varshamov showed the bound held even for linear codes. By now we know how to prove this bound by looking at smaller and smaller space of codes. The proofs below enumerate many possible proofs of this bound. You only need to turn in Part (a) below.

   (a) Maximal codes: An $(n, k, d)_q$ code C is maximal if for every word $w \in \Sigma^n$, there exists a codeword $c \in C$ such that $\Delta(w, c) < d$ (where $|\Sigma| = q$). Prove that for every maximal $(n, k, d)_q$ code, $q^n \leq q^k \cdot \text{Vol}_q(d-1, n)$. Conclude that for every q, there exists a family of

q-ary codes of rate R and relative distance δ satisfying $R \geq 1 - H_q(\delta)$. Conclude further that such a code can be found in $q^{O(n)}$ time.

(b) Linear codes (Varshamov bound - 1): Show that there exists an $[n, k, d]_q$ linear code C achieving the bound $q^n \leq q^k \cdot \mathrm{Vol}_q(d-1, n)$.

(c) Varshamov bound - 2: (Due to Sergey Yekhanin.) For q, n, m, d satisfying $q^m > \mathrm{Vol}_q(d-2, n-1)$, show that there exists a matrix $H \in \mathbb{F}_q^{n \times m}$ such that no subset of $d-1$ or fewer rows of H are linearly dependent. Conclude that there exist $[n, k, d]_q$ codes satisfying $q^n \leq q^k \cdot \mathrm{Vol}_q(d-2, n)$. (This construction is thus the right extension of the Hamming construction when $d > 3$. Hamming's code was *perfect* because he also showed that for every n, k, d code $q^n \leq q^k \cdot \mathrm{Vol}_q((d-1)/2, n)$. Notice the tightness uses the fact that $(d-1)/2 = d-2$, which only happens for $d = 3$.)

(d) Smaller sample space - 1: (To get a code achieving the Gilbert-Varshamov bound, we thus need to only search in a space of size $\min\{q^{nk}, q^{n(n-k)}\}$. But we can do better than this, as shown in this exercise and the next one.) A matrix $G \in \mathbb{F}_q k \times n$ is a Toeplitz matrix if for every i, j, $G_{i,j} = G_{i-1,j-1}$. Show that there exists a Toeplitz matrix G such that the code generated by G achieves $q^n \leq q^k \cdot \mathrm{Vol}_q(d-1, n)$. Conclude there is a small sample space of codes which includes a code meeting the G-V bound in the following sense:[1] Show that there exists a polynomial time algorithm A and a string x of length $O(n)$ such that $A(x, n, k, q)$ produces an $[n, k, d]_q$ code satisfying $q^n \leq q^k \cdot \mathrm{Vol}_q(d-1, n)$.

(e) Smaller sample space - 2: This result shows an alternate way to get a small sample space. It is not as general as the method above (works only when k divides n, and our exercise only covers the case where $k = n/2$). However it is a good way to explore the concept of finite fields and their extensions. The resulting collection of codes is called the Wozencraft ensemble.

i. Given positive integer k, show there exists a bijection $f : \mathbb{F}_{2^k} \to \mathbb{F}_2^k$ such that for every $\alpha, \beta \in \mathbb{F}_{2^k}$, $f(\alpha) + f(\beta) = f(\alpha + \beta)$.

ii. For $\alpha \in \mathbb{F}_{2^k}$, define the encoding function $E_\alpha : \mathbb{F}_2^k \to \mathbb{F}_2^{2k}$ which maps x to $(x, f(\alpha \cdot f^{-1}(x)))$.

- Show that the image of E_α is a linear code.
- For integer d, say that α is *bad* if there exists a non-zero x such that $E_\alpha(x)$ has weight less than d (i.e., has fewer than d 1s). Prove that there are at most $\mathrm{Vol}_2(d-1, n)$ bad α's. Conclude that if $2^k > \mathrm{Vol}_2(d-1, n)$ then there exists an $[2k, k, d]_2$ code among the images of E_α.
- How does this code compare with the other codes constructed in this problem (in terms of performance)? What is the sample size?

4. Food for thought (i.e., not to be turned in either). (Context/Reference: A recent paper of Tao Jiang and Alex Vardy shows how to improve on the Gilbert-Varshamov bound asymptotically (on the number of codewords, but not the rate). The paper is based on simple combinatorial ideas, though working out the full proof is beyond the scope of this class. Here we describe the main steps. Warning: The hard steps are *really* hard. The idea is for you to understand the

[1] The reason for this elaboration is important. Note that once we prove there exists a code meeting the G-V bound, we have also proved there exists a sample space of size *one* which contains a code achieving the G-V bound! To articulate the fact that we have done better than this trivial fact, we need to find a mathematical way to express our enhanced knowledge. We do so by making explicit the fact that we can generate points from the sample space efficiently (since A runs in polynomial time), and that one point in the sample space has the desired properties.

steps, not necessarily to prove them.) Consider the following graph $G = G_{n,d} = (V, E)$. The vertex set $V = \{0,1\}^n$. The edge set E contains all pairs (x, y) for $x, y \in V$ if $\Delta(x, y) < d$. Say that a set $I \subset V$ is independent in G if for every pair of distinct vertices $x, y \in I$, $(x, y) \notin E$. Let $\alpha(G)$, called the independence number of G, denote the size of the largest independent set I in G.

- Show that there is an $(n, k, d)_2$ code if and only if $k \leq \log_2 \alpha(G)$.
- For a vertex $v \in G$, define its degree $\deg(v)$ to be the number of $u \in V$ such that $(u, v) \in E$. Define $\deg(G)$ to be the maximum over $v \in V$ of $\deg(v)$. Prove that every graph G has an independent set of cardinality $|V|/(deg(G) + 1)$. How does this relate to other things you've done above.
- (Hard) A triangle in a graph G is a triple of vertices x, y, z such that $(x, y), (y, z), (x, z) \in E$. Let $T(G)$ denote the number of triangles in G.
  - Prove that if G has no triangles, then $\alpha(G) \geq |V|/(8 \cdot \deg(G)) \cdot \log_2 \deg(G)$.
  - For general G, show that $\alpha(G) \geq |V|/(8 \cdot \deg(G)) \cdot (\log_2 \deg(G) - \frac{1}{2} \log_2(T(G)/|V|))$.
- (Medium) Bound the number of triangles in $G = G_{n,d}$. Prove that this number is of the form $|V| \cdot \deg(G)^{2-\epsilon}$ for some $\epsilon > 0$. Conclude that there exists an $(n, k, d)_2$ code satisfying $2^k = \Omega(d2^n/\operatorname{Vol}_2(d-1, n))$. (When is this bound better than the Varshmov Bound - 2 above?)