Fall 2006
Coding Theory

Homework 1
Due 2006/11/2

Oded Regev
Dept. 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. Make sure you are comfortable with some basic algebraic notions, such as rings, fields, polynomial rings, finite fields, and vector spaces. A list of some of these basic facts can be found under the "Notes on Algebra" in Madhu Sudan's 2001 course.

2. **Parity check matrix:** (no need to turn in)

   (a) Let G be a $k \times n$ matrix with $0, 1$ entries of rank k over $\mathbb{F}_2$, and let $C = \{\mathbf{x}G \mid \mathbf{x} \in \mathbb{F}_2^k\}$ (C is the linear code generated by G). Show that there exists an $m \times n$ matrix H (known as the parity check matrix of G) such that $C = \{\mathbf{y} \in \mathbb{F}_2^n \mid H\mathbf{y} = \mathbf{0}\}$. What is the relationship between m, n and k above?

   (b) Give an efficient algorithm that given G computes such an H, and vice versa.

3. **Source coding:**

   (a) Prove the converse to the source coding theorem (for $\{0, 1\}$ distributions).

   (b) Extend the source coding theorem and its converse to arbitrary distributions (no need to turn in).

4. **Entropies:**

   (a) Show that $H(X|Y) \leq H(X)$.

   (b) Show that if Y is a (deterministic) function of X then $H(Y) \leq H(X)$.

   (c) Show that $H(Y|X) = 0$ iff Y is a (deterministic) function of X.

   (d) Assume $f : \mathcal{X} \times \mathcal{Y} \to \mathcal{Z}$ is such that for any fixed $y \in \mathcal{Y}$, $f(\cdot, y)$ is one-to-one. Show that $H(f(X, Y)|Y) = H(X|Y)$.

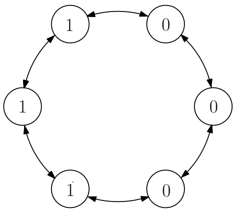
   (e) Let X be a uniformly distributed Boolean random variable, and let Y be a Boolean random variable satisfying $\Pr[X = Y] = p$. Show that $I(X : Y) \geq 1 - H(p)$. Hint: you can use (d),(a)

5. **The importance of admitting ignorance:** Let X be a binary random variable uniformly distributed on $\{0, 1\}$. Let Y be a binary random variable obtained from X by the following process: with probability ε it equals X, and with probability $1 - \varepsilon$ it is a uniformly random bit. Let Z be a random variable on $\{0, 1, 2\}$ obtained from X by the following process: with

probability ε it equals X, and with probability $1-\varepsilon$ it equals 2. Compute $I(X:Y)$ and $I(X:Z)$ and find the rough asymptotic behavior as $\varepsilon \to 0$. How do you explain this?

6. **Conditional vs. unconditional information:**
   (a) Show that if Z is a randomized function of Y then $I(X:Y|Z) \leq I(X:Y)$.
   (b) Show that if Z is independent of Y then $I(X:Y|Z) \geq I(X:Y)$.
   (c) For both inequalities, give an example of a strict inequality where one side is zero. Hint: $\oplus$

7. **Rate of a Markovian source:**[1] Consider a Markovian source of bits, where the source consists of a 6-cycle with three successive vertices outputting 0, and three successive vertices outputting 1, with probability $1/2$ to go left or right from each vertex. Compute the rate of this source, i.e., show matching upper and lower bounds on the ability to compress strings originating from this source. Your proof should be from first principles.



8. **Coins:** Let $X = (X_1, \ldots, X_n)$ be the result of n independent tosses of a biased coin with bias p. We are interested in functions $f : \{0,1\}^n \to \{0,1\}^*$ such that $f(X)$ is a sequence of unbiased coins. More precisely, we want that for any length k, all 2^k sequences of length k are equally likely to appear as $f(X)$. For example, for $n = 2$ we can take the function that maps $00 \to \phi$, $11 \to \phi$, $01 \to 0$, and $10 \to 1$. Show that in any such function, the average number of unbiased coins in $f(X)$ is at most $H(p)n$, i.e., $E[|f(X)|] \leq H(p)n$. Hint: consider the entropy of $f(X)$

9. **Random access code:** An (n, m, p) random access code is a randomized function f from $\{0,1\}^n$ to $\{0,1\}^m$ and randomized functions $g_1, \ldots, g_n : \{0,1\}^m \to \{0,1\}$ such that

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

Show that there is no $(2, 1, p)$ random access code for $p > \frac{1}{2}$.

[1] Taken from Madhu Sudan's problems