# Additional Problems

Topics in Complexity Theory and Pseudorandomness (Spring 2013)
Rutgers University
Swastik Kopparty

1. Look up the definition of decision tree and depth of a decision tree.

   (a) Suppose $f : \{0,1\}^n \to \{0,1\}$ can be represented as a decision tree of depth $k$. Show that $f$ can be represented as a $k$-CNF and also as a $k$-DNF.

   (b) Suppose $f : \{0,1\}^n \to \{0,1\}$ is a function that can be written as both a $t$-CNF and a $t$-DNF.
   Show that $f$ has a decision tree of depth $t^2$.

   (c) Suppose $f : \{0,1\}^n \to \{0,1\}$ can be represented as a decision tree of depth $\leq n-1$. Show that
   $$\Pr[f(x) = PARITY(x)] = 1/2.$$

   (d) Now suppose $f : \{0,1\}^n \to \{0,1\}$ can be written as a $t$-DNF.
   Show that
   $$\Pr_{x \in \{0,1\}^n}[f(x) = PARITY(x)] \leq \frac{1}{2} + 2^{-\Omega(n/t)}.$$
   Hint: Use the switching lemma and the earlier parts of this problem.

   (e) Complete the proof of Håstad's theorem that if $C$ is an $\mathsf{AC}^0$ circuit of size $\leq 2^{O(n^{1/d})}$ and depth $d$, then
   $$\Pr_{x \in \{0,1\}^n}[C(x) = PARITY(x)] \leq \frac{1}{2} + 2^{-\Omega(n^{1/d})}.$$
   First use the switching lemma to simplify the circuit $C$ to a $t$-DNF. Then use the earlier parts of this problem.

   (f) Let $e_i \in \{0,1\}^n$ denote the vector with 1 only in the $i$th coordinate.
   Show that if $f$ is a decision tree of depth $k$,
   $$\Pr_{x \in \{0,1\}^n, i \in [n]}[f(x) \neq f(x \oplus e_i)] \leq \frac{k}{n}.$$

   (g) Show that if $C$ is an $\mathsf{AC}^0$ circuit of size $s$, depth $d$, on $n$ variables, then
   $$\Pr_{x \in \{0,1\}^n, i \in [n]}[C(x) \neq C(x \oplus e_i)] \leq \frac{(\log s)^{O(d)}}{n}.$$
   This says that functions computed by $\mathsf{AC}^0$ circuits have low "average sensitivity": at a random point, they are unlikely to change value when a random bit is flipped.

2. (a) A map $f : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$ defines a bipartite graph on $(2^n, 2^n)$ in a natural way. This bipartite graph is $k$-Ramsey if there do not exist $S, T \subseteq \{0,1\}^n$, $|S| = |T| = k$ and $b \in \{0,1\}$, with $f(x,y) = b$ for all $x \in S, y \in T$.

   Note that a random function $f$ defines a $O(n)$-Ramsey biparite graph. Show that if $f$ is computed by a size $s$, depth $d$ $\mathsf{AC}^0$ circuit, then its bipartite graph cannot be $2^{\Omega(n/(\log s)^d)}$-Ramsey.

   (b) Let $0 < \delta < 1$ be a constant. A function $f : \{0,1\}^k \to \{0,1\}^n$ is a $\delta$ error-correcting code if for every distinct $x, y \in \{0,1\}^k$, $f(x)$ and $f(y)$ have Hamming distance at least $\delta n$.

   Note that if $n = (1+\Omega(1))k$, a random function $f$ is an $\Omega(1)$ error-correcting code. Show that if $f$ is computed by a size $2^{n^{O(1/d)}}$, depth $d$ $\mathsf{AC}^0$ circuit, then $f$ cannot compute an $\Omega(1)$ error-correcting code.

   In class we will see that polynomial size $\mathsf{AC}^0(\oplus)$ can compute both bipartite $O(n)$-Ramsey graphs and $\Omega(1)$ error-correcting codes.

3. (a) Show the existence of the set-systems that we needed for the Nisan-Wigderson generator. Specifically, given $m, a, b$, show that for $t = O(m^{1/b} \cdot \frac{a^2}{b})$, there exists a family of $m$ subsets of the universe $[t]$, each of size $a$, such that all pairwise intersections are of size at most $b$.

   (b) What is the smallest $\delta(n,s)$ for which you can show the existence of a function $f : \{0,1\}^n \to \{0,1\}$, such that for all circuits $C$ of size $s$,

$$\Pr_{x \in \{0,1\}^n}[C(x) = f(x)] \leq \frac{1}{2} + \delta.$$

   (c) Using such an $f$ in the NW generator, show that one can get pseudorandom generators for size $s$ circuits with seed length $O(\log s)$.

4. Below is a collection of facts/problems related to finite fields. Try to verify them yourself or look them up.

   (a) Let $p$ be prime. Let $\mathbb{F}_p = \{0, 1, \ldots, p-1\}$ along with operations addition and multiplication mod $p$. Every integer can be treated as an element of $\mathbb{F}_p$ (by taking the remainder after dividing by $p$).

   All of $\mathbb{F}_p$ forms a group under addition. The nonzero elements of $\mathbb{F}_p$, denoted $\mathbb{F}_p^*$ form a group under multiplication. Both groups are commutative.

   (b) For each $a \in \mathbb{F}_p$, we have $a^p = a$. If $a \neq 0$, then $a^{p-1} = 1$.

   **Highly recommended:** Following the Razborov-Smolensky proof strategy we saw in class, show that PARITY and MAJORITY do not have polynomial size $\mathsf{AC}^0(\mathrm{Mod}_3)$ circuits. Here $\mathrm{Mod}_3$ gates have unbounded fan-in, and output 0 iff the number of 1's on the input wires is divisible by 3.

   (c) Let $\mathbb{F}_p[X]$ be the set of polynomials with $\mathbb{F}_p$ coefficients. Then the division theorem holds in $\mathbb{F}_p[X]$, and thus every element of $\mathbb{F}_p[X]$ can be uniquely factorized into irreducible polynomials.

   (d) The remainder theorem holds in $\mathbb{F}_p[X]$. Thus $X^p - X = \prod_{\alpha \in \mathbb{F}_p}(X - \alpha)$.

(e) For each integer $d$, the number of $a \in \mathbb{F}_p^*$ satisfying $a^d = 1$ is at most $d$. Combining this with the fact that $\mathbb{F}_p^*$ is commutative, this implies that $\mathbb{F}_p^*$ is cyclic (i.e., there is an element $g \in \mathbb{F}_p^*$ such that $\mathbb{F}_p^* = \{1, g, g^2, \ldots, g^{p-2}\}$.

Not every element of $\mathbb{F}_p^*$ generates $\mathbb{F}_p^*$. Look at the cases $p = 7, 13$ and find a generator for $\mathbb{F}_p^*$ in each case.

(f) Suppose $p$ is an odd prime. Then exactly $1/2$ the elements of $\mathbb{F}_p^*$ are perfect squares. If $a \in \mathbb{F}_p^*$, then $a^{(p-1)/2}$ equals either 1 or $-1$, depending on whether $a$ is a perfect square or not.

(g) Generalize the above to perfect $d$th powers. Note that if $d$ is relatively prime to $p-1$ then every element of $\mathbb{F}_p^*$ is a perfect $d$th power.

(h) Let $f(X)$ be an irreducible polynomial of degree $d$ in $\mathbb{F}_p[X]$. We can consider the set $\mathbb{F}_p[X]/f(X)$ of polynomials modulo $f(X)$. Every polynomial is equivalent modulo $f(X)$ to a unique polynomial of degree $< d$. Thus there are $p^d$ residue classes. Addition and multiplication of polynomials is compatible with reducing mod $f(X)$. Every nonzero element of $\mathbb{F}_p[X]/f(X)$ has a multiplicative inverse (this is where irreducibility of $f(X)$ is used). Thus $\mathbb{F}_p[X]/f(X)$ is a field of cardinality $p^d$.

The relationship between $\mathbb{Z}$, the prime $p$ and the field $\mathbb{Z}/p$ is entirely analogous to the relationship between $\mathbb{F}_p[X]$, the irreducible $f(X)$ and the field $\mathbb{F}_p[X]/f(X)$.

(i) The field $\mathbb{F}_p[X]/f(X)$ is a $d$-dimensional vector space over the field $\mathbb{F}_p$. We denote this field $\mathbb{F}_{p^d}$. It is tricky to prove but true that any two fields of cardinality $p^d$ are isomorphic fields. Thus there is a unique such field. If $n$ is an integer not of the form $p^d$ for $p$ prime, then there does not exist a finite field of cardinality $n$. Thus whenever we talk of the finite field $\mathbb{F}_q$, we will insist that $q$ be a prime power.

(j) Note that the above construction of $\mathbb{F}_{p^d}$ required the existence of an irreducible polynomial of degree $d$ over $\mathbb{F}_p$. Such polynomials exist for every $d$! Try to show this.

(k) Construct the fields $\mathbb{F}_8$ and $\mathbb{F}_9$.

(l) Note that the field $\mathbb{F}_{p^d}$ is not isomorphic to the ring $\mathbb{Z}/p^d$.

(m) Many of the facts you proved about the field $\mathbb{F}_p$ also hold for $\mathbb{F}_{p^d}$. Polynomials over $\mathbb{F}_{p^d}$ can be defined, and they have nice properties. The multiplicative group $\mathbb{F}_{p^d} \setminus \{0\}$ is cyclic. Etc. To prove all these properties, you need not use the explicit construction of $\mathbb{F}_{p^d}$ described above. It suffices to just use the fact that $\mathbb{F}_{p^d}$ is a field of cardinality $p^d$.

(n) $X^{p^d} - X = \prod_{\alpha \in \mathbb{F}_{p^d}} (X - \alpha)$.

5. Let $q$ be a prime power. For each $\alpha \in \mathbb{F}_q$, let $v_\alpha \in \mathbb{F}_q^k$ be the vector $(1, \alpha, \alpha^2, \ldots, \alpha^{k-1})$.

(a) Show that for any $k$ distinct $\alpha_1, \ldots, \alpha_k \in \mathbb{F}_q$, the vectors $v_{\alpha_1}, v_{\alpha_2}, \ldots, v_{\alpha_k}$ are linearly independent.

(b) Now suppose $q = 2^t$. Using the fact that $\mathbb{F}_q$ a vector space of dimension $t$ over $\mathbb{F}_2$, we get a $\mathbb{F}_2$-linear isomorphism $E : \mathbb{F}_q^k \to \mathbb{F}_2^{tk}$. Show that the vectors $\tilde{v}_\alpha = E(v_\alpha) \in \mathbb{F}_2^{tk}$ are such that any $k$ of them are linearly independent over $\mathbb{F}_2$.

Let $n = 2^t$. Show that the $k$-wise independent distribution over $\mathbb{F}_2^n$ that we get from these vectors has seed length $k \log n$.

(c) Again suppose $q = 2^t$, and let $k$ be even. Let $u_\alpha \in \mathbb{F}_q^{k/2}$ be the vector $(\alpha, \alpha^3, \alpha^5 \ldots, \alpha^{k-1})$.

Let $\tilde{u}_\alpha = E(u_\alpha) \in \mathbb{F}_2^{tk/2}$. Show that if $\alpha_1, \ldots, \alpha_k$ are such that $\tilde{u}_{\alpha_1}, \ldots, \tilde{u}_{\alpha_k}$ are linearly dependent over $\mathbb{F}_2$, then $\tilde{v}_{\alpha_1}, \ldots, \tilde{v}_{\alpha_k}$ are linearly dependent over $\mathbb{F}_2$. Thus conclude that every set of $k$ vectors from the collection $\{\tilde{u}_\alpha \mid \alpha \in \mathbb{F}_q\}$ are linearly independent over $\mathbb{F}_2$.

Let $n = 2^t$. Show that the $k$-wise independent distribution over $\mathbb{F}_2^n$ that we get from these vectors has seed length $\frac{k}{2} \log n$.

This construction is also known as the "BCH code", after R. C. Bose, D. Ray-Chaudhuri and Hocquenghem.

(d) Let $k$ be a constant. Suppose $s < \frac{k}{2} \log n - \Omega(1)$. Show that if we take any collection of $n$ vectors in $\mathbb{F}_2^s$, then some $k$ of them are linearly dependent.

Thus the above construction of vectors is essentially as large as possible.

6. In this exercise we will prove some basic properties of the Fourier transform, and use it to prove lower bounds on the seed length required for generating $k$-wise independent random bits and $\epsilon$-biased bits.

Recall that for a function $f : \mathbb{F}_2^n \to \mathbb{R}$, for a set $S \subseteq [n]$, we define:

$$\hat{f}(S) = \langle f, \chi_S \rangle = \sum_{x \in \mathbb{F}_2^n} f(x)\chi_S(x),$$

where $\chi_S(x) = \prod_{i \in S} (-1)^{x_i}$. We then have the basic relation:

$$f(x) = \frac{1}{2^n} \sum_{S \subseteq [n]} \hat{f}(S)\chi_S(x).$$

Sometimes it is more natural to treat $S$ as an element of $\mathbb{F}_2^n$ (viewing it as an indicator vector). In this notation, $\chi_S(x) = (-1)^{\langle x, S \rangle}$ (here the inner product is over $\mathbb{F}_2$).

(a) Let $f, g : \mathbb{F}_2^n \to \mathbb{R}$. Show that

$$\langle f, g \rangle = \frac{1}{2^n} \langle \hat{f}, \hat{g} \rangle.$$

(b) Let $f, g : \mathbb{F}_2^n \to \mathbb{R}$. Let $h : \mathbb{F}_2^n \to \mathbb{R}$ be the function with $h(x) = f(x)g(x)$ for each $x$. Show that $\hat{h}(S) = \sum_{T_1, T_2 \in \mathbb{F}_2^n | T_1 + T_2 = S} \hat{f}(T_1)\hat{g}(T_2)$.

(c) Let $f, g : \mathbb{F}_2^n \to \mathbb{R}$. Let $h : \mathbb{F}_2^n \to \mathbb{R}$ be the function with

$$h(x) = \sum_{y, z \in \mathbb{F}_2^n | y + z = x} f(y)g(z).$$

This is denoted by $h = f \star g$ ($h$ is the *convolution* of $f$ and $g$). This operation shows up naturally in the following situation: if $\mu_1$ and $\mu_2$ are distributions on $\mathbb{F}_2^n$, then $\mu_1 \star \mu_2$ is the distribution of $y + z$ when $y \in \mu_1$ and $z \in \mu_2$ are picked independently.

Show that $\hat{h}(S) = \hat{f}(S)\hat{g}(S)$.

(d) Recall that a distribution $\mu$ over $\mathbb{F}_2^n$ is $k$-wise independent if $\hat{\mu}(S) = 0$ for all $S \subseteq [n]$ with $1 \leq |S| \leq k$.

Let $k$ be a constant. Let $X \subseteq \mathbb{F}_2^n$ with $|X| = o(n^{\lfloor k/2 \rfloor})$. Show that for $d = \lfloor k/2 \rfloor$, there is a function $f : \mathbb{F}_2^n \to \mathbb{R}$ which satisfies:

i. $f$ is not identically 0.

ii. $\hat{f}(S) = 0$ for each $|S| > d$.

iii. $f(x) = 0$ for each $x \in X$.

Use this to show that any $k$-wise independent distribution over $\mathbb{F}_2^n$ has support at least $n^{k/2}$. Thus the BCH construction of a $k$-wise independent distribution has essentially optimal seed length.

(e) Recall that a distribution $\mu$ over $\mathbb{F}_2^n$ is $\epsilon$-biased if $|\hat{\mu}(S)| < \epsilon$ for each $S \neq \emptyset$.
Show that for any $\epsilon$-biased distribution $\mu$, $\sum_{x \in \mathbb{F}_2^n} \mu(x)^2 \leq \epsilon^2 + \frac{1}{2^n}$. Conclude that if $\epsilon < \frac{1}{2^{n/2}}$, then $\mu$ must have support at least $\frac{1}{2}2^n$.

(f) Let $\mu$ be an $\epsilon$-biased distribtion. Let $t$ be a parameter to be chosen later. Let $\nu = \mu \star \mu \star \ldots \star \mu$ ($t$ times).

i. Show that $\nu$ is an $\epsilon^t$-biased distribution.

ii. Show that $|\mathsf{Support}(\nu)| \leq \binom{|\mathsf{Support}(\mu)|+t}{t}$.

Combine these with the previous part to show that

$$|\mathsf{Support}(\mu)| \geq \Omega(\frac{n}{\epsilon^2 \log \frac{1}{\epsilon}}).$$

Thus the seed length required to generate an $\epsilon$-biased distribution is at least

$$\log n + 2 \log \frac{1}{\epsilon} - \log \log \frac{1}{\epsilon} - O(1).$$

This theorem and its proof are due to Noga Alon. Alon's proof is actually even simpler, using only ranks and eigenvalues of matrices, and does not use any Fourier analysis. The original paper (Perturbed Identity Matrices have High Rank) is highly recommended.