CS 810: Introduction to Computational Complexity Theory 04/22/2003

Lecture 25: Lautemann's Proof, Lemma on Fat and Thin Sets

Instructor: Jin-Yi Cai Scribe: Michael DeCoster, Vinod Ganapathy, and Charles Kahn

1 Approximate Counting

We continue our discussion of approximate counting schemes. For $T \subseteq \{0,1\}^n$, choose k such that $1 \leq k \leq n \cdot m = l$. We can repeat the process m times, so define $S = T^m \subseteq \{0,1\}^{n \cdot m}$.

From the results shown in the last lecture, we know that if $|S| \leq \frac{2^k}{2}$, then for $\forall x \in S$ with high probability there exists a hash function h_i that isolates x. Conversely, if $|S| > k \cdot 2^k$, then no hash function h_i isolates x.

There are only polynomially many k to evaluate, so it's a question in Σ_2^p. We claim that there is a value of k that represents the last "flip" from a "no" answer for $k-1$ to a "yes" answer (i.e., "Yes, there is a hash function that isolates any x.") for all values at least k.
For $k = s+1$,

$$(s - \log s) \cdot 2^{s - \log s} < |S| \leq \frac{2^{s+1}}{2}$$

Then $2^{k-2} < |S| \leq k \cdot 2^k$, so we can estimate $|S|$ with an error of approximate magnitude $4k$. We estimate $|S| = 2^k$, which yields $|T| = 2^{k/m}$.

$$\frac{1}{4} < \frac{|S|}{2^k} \leq k$$

$$(\frac{1}{4})^{\frac{1}{m}} < \frac{|S|^{\frac{1}{m}}}{2^{\frac{k}{m}}} \leq k^{\frac{1}{m}}$$

This formulation gives us:

$$k^{\frac{1}{m}} = e^{\frac{1}{m}\log k} \leq e^{\frac{1}{m}(\log m + \log n)}$$

$$= 1 + \frac{1}{m}(\log m + \log n) < 1 + \frac{1}{n^c}$$

for arbitrarily large c, that is, c as a polynomial in n.

Thus our process yields an estimate

$$|T| = |S|^{\frac{1}{m}} \leq (2^{\frac{k}{m}}) \cdot (1.00...0xxx...)$$

where there are $O(\log n)$ zeroes after the decimal point, which, when multiplied by some power of 2 gives us an approximation of the $O(\log n)$ leading bits. This is Stockmeyer's approximate counting scheme. We can state this as a theorem.

Theorem 1. *In $P^{\Sigma_2^p}$ we can generate $O(\log n)$ leading bits of any counting problem.*

For instance, we can approximate $|S|$ for $S \subseteq \{0,1\}^n$, where we can decide whether $x \in S$ in deterministic polynomial time.

This can be done "lower" in the polynomial hierarchy, and we'll use randomization – that is, a probabilistic algorithm – to do so. To achieve accuracy $c \log n$, choose a polynomial m such that $l = n \cdot m$, randomly pick $h_1, h_2, \cdots, h_k : \{0,1\}^l \mapsto \{0,1\}^k$ and determine if this random assignment works.

$$\forall x \in S, \exists h_i \forall y \in S, y \neq x [h_i \text{ isolates } x]$$

$$\forall x \in S, \forall y_1, \ldots, y_k \in S[y_1 \neq x, \ldots, y_k \neq x][h_1(y_1) \neq h_1(x) \vee \cdots \vee h_k(y_k) \neq h_k(x)]$$

Thus, it's a coNP question, so it can be decided in BPP^{NP}. Note that the original statement by Stockmeyer required $P^{\Sigma_2^p}$, and our revised approach can be done in BPP^{NP}.

The algorithm can be made to err only in saying "yes"; the last "no" is at worst an overestimate.

Lautemann's approach is simpler and more elegant, but not as versatile as Sipser's.

2 Fat and Thin Sets

Definition 1 (Fat Set). *For set $S \subseteq \{0,1\}^n$, we say S is $\frac{1}{m}$-fat if $|S| > (1 - \frac{1}{m}) \cdot 2^n$.*

Definition 2 (Thin Set). *Set S is $\frac{1}{m}$-thin if S^c is fat.*

We can think of $\{0,1\}^n$ as an n-dimensional space, where we are interested in the effect of "shifting" a set S by adding various n-dimensional vectors u_i. As defined here, addition is the bitwise exclusive-or operation, where for any $u, v \in \{0,1\}^n$, $(S \oplus v) \subseteq \{0,1\}^n$ and $u \oplus v \oplus v = u$.

Before proving our lemmas about fat and thin sets, we state a simple lemma that aids in the proofs.

Lemma 1. *$(E \oplus u)^c = E^c \oplus u$.*

Proof. We'll show one direction of the proof. We know $x = (x \oplus u) \oplus u$ for $\forall x \in E, \forall u \in \{0,1\}^n$. Take $x \in (E \oplus u)^c$. We claim $(x \oplus u) \notin E$. If $(x \oplus u) \in E$, then $x = (x \oplus u) \oplus u$, which makes $(x \oplus u) \in E$, which is a contradiction. The other direction is proved in similar fashion. □

We will randomly and uniformly pick "shift vectors" $u_1, u_2, \cdots, u_m \in \{0,1\}^n$. If S is "fat", then it's likely that for any "point" in $\{0,1\}^n$, some shifts will cover it. We also can state that if S is fat, then the intersection of the "shifted" sets $(S \oplus u_i)$ will be non-empty, that is,

$$\Pr_{u_1, \cdots, u_m} [\forall v \in \{0,1\}^n, \exists u_i, 1 \leq i \leq m, v \in S \oplus u_i] \geq 1 - \frac{2^n}{m^m}$$

Take $m > n$, so that $\frac{2^n}{m^m} \to 0$.

$$\Pr_{u_1,\cdots,u_m}\left[\bigcap_{i=1}^{m}(S \oplus u_i) \neq \emptyset\right] = 1$$

$$\left[\bigcap_{i=1}^{m}(S \oplus u_i)\right]^c = \bigcup_{i=1}^{m}(S \oplus u_i)^c = \bigcup_{i=1}^{m}(S^c \oplus u_i)$$

$$\left|\left[\bigcap_{i=1}^{m}(S \oplus u_i)\right]^c\right| < \frac{1}{m} \cdot m \cdot 2^n = 2^n$$

$$\Pr_{u_1,\cdots,u_m}\left[\forall v \in \{0,1\}^n, \forall u_i, 1 \leq i \leq m, v \notin S \oplus u_i\right] \leq 2^n \left(\Pr_u\left[v \notin S \oplus u\right]\right)^m$$

$$v \notin S \oplus u \Rightarrow v \oplus u \notin S \Rightarrow u \notin S \oplus v$$

$$|S \oplus v| = |S| > \left(1 - \frac{1}{m}\right) \cdot 2^n$$

So, if S is thin, then the probability that the "shifted" sets $(S \oplus u)$ cover all of the points in $\{0,1\}^n$ is zero. That is,

$$\Pr_{u_1,\cdots,u_m}\left[\forall v \in \{0,1\}^n, \forall u_i, 1 \leq i \leq m, v \notin S \oplus u_i\right] = 0$$

$$\Pr_{u_1,\cdots,u_m}\left[\bigcup_{i=1}^{m}(S \oplus u_i) = \{0,1\}^n\right] = 0$$

The proof is the complement of the proof for "fat" sets. Also, for "thin" sets,

$$\Pr_{u_1,\cdots,u_m}\left[\bigcap_{i=1}^{m}(S \oplus u_i) \neq \emptyset\right] \leq \frac{2^n}{m^m}$$

3 BPP

Theorem 2. $BPP \subseteq \Sigma_2^P$.

If $x \in BPP$, the witness set Y is either thin or fat in $\{0,1\}^m$, where $m = \text{poly}(n)$. For $L \in BPP$, $L = L(M)$:

If $x \in L$, then the witness set is fat in $\{0,1\}^m$ (for appropriate m, some large polynomial in n), so

$$\exists u_1, \cdots, u_m \; s.t. \; \forall v, \bigvee_{i=1}^{m} M(x, u_i \oplus v) = 1$$

Conversely, if $x \notin L$, the witness set is thin.