Lecture 26
Computational Complexity Theory

Abhishek Shetty
Undergraduate Department
Indian Institute of Science

Raghav Malhotra
Undergraduate Department
Indian Institute of Science

Instructor
Chandan Saha
Computer Science and Automation
Indian Institute of Science

November 5 2015

1 Leftover Hash Lemma

Let $\mathcal{H}_{n,m}$ be a family of pairwise independent hash functions from $\{0,1\}^n$ to $\{0,1\}^m$ and $\epsilon \in (0,1)$ be a constant.

Theorem 1.1. *Let $S \subseteq \{0,1\}^n$ be such that $|S| \geq \frac{2^{m+2}}{\epsilon^2}$. Let $A = \{a \in S : h(a) = 0\}$ and $n_A = |A|$*

$$\Pr_{h \in \mathcal{H}_{n,m}} \{|n_A - \frac{|S|}{2^m}| \geq \epsilon \frac{|S|}{2^m}\} \leq \frac{1}{4}$$

Proof. Let χ_i be the indicator variable defined as follows

$$\chi_i = \begin{cases} 1 & h(a_i) = 0 \\ 0 & \text{otherwise} \end{cases}$$

We note that $n_A = \sum_{i=1}^{k} \chi_i$. Let X denote the sum of random variables χ_i and μ denote $\mathbb{E}[X] = \frac{|S|}{2^m}$. From Chebyshev's Inequality, we have,

$$Pr\{|X - \mu| \geq \epsilon\mu\} \leq \frac{Var(X)}{\epsilon^2\mu^2} \tag{1}$$

As the family $\mathcal{H}$ is pairwise independent,

$$Var\left(\sum_{i=1}^{k} \chi_i\right) = \sum_{i=1}^{k} Var(\chi_i)$$

By definition, $Var(\chi_i) = E[\chi_i^2] - (E[\chi_i])^2$ but as $\chi_i^2 = \chi_i$, we get $Var(\chi_i) = E[\chi_i] - (E[\chi_i])^2$. Thus, we get

$$Var(\chi_i) \leq E[\chi_i] = Pr\{\chi_i = 1\} = \frac{1}{2^m}$$

Plugging this in (1), we get

$$Pr\{|X - \mu| \geq \epsilon\mu\} \leq \frac{|S|/2^m}{\epsilon^2(|S|/2^m)^2} = \frac{2^m}{\epsilon^2|S|} \leq \frac{2^m\epsilon^2}{2^{m+2}\epsilon^2} = \frac{1}{4}$$

as required. □

2 Toda's theorem

Theorem 2.1. $\mathsf{NP} \cup \mathsf{co} - \mathsf{NP} \subseteq \mathsf{P}^{\#\mathtt{SAT}}$

Proof. This is clear since $\mathtt{SAT}$ and $\mathtt{co\text{-}SAT}$ can be decided by querying the oracle for the number of satisfying assignments and then checking whether the returned answer is zero. □

The following theorem strengthens the above theorem considerably stating the whole of the polynomial hierarchy can be decided by a polynomial machine with access to an $\#\mathtt{SAT}$ oracle.

Theorem 2.2 ([Tod91], [AB09])**.** $\mathsf{PH} \subseteq \mathsf{P}^{\#\mathtt{SAT}}$

We prove this theorem in two steps. We show that any problem in the polynomial hierarchy can be randomly reduced to a problem in $\bigoplus\mathsf{P}$ (defined in section 2.1) , that is $\mathsf{PH} \subseteq \mathsf{BPP}^{\bigoplus\mathsf{P}}$. This is captured in the following theorem.

Theorem 2.3. *There exists a probabalisitic poly-time algorithm* $\boldsymbol{A}$*, which on input* ψ*, a quantified boolean formula with* c *quantifiers and a parameter* m*, outputs a boolean formula* ϕ *such that*

$$\begin{aligned} \psi \textit{ is true} &\implies Pr\{\#\phi \textit{ is odd }\} \geq 1 - 2^{-m} \\ \psi \textit{ is false} &\implies Pr\{\#\phi \textit{ is odd }\} \leq 2^{-m} \end{aligned}$$

$\boldsymbol{A}$ *runs in time poly(*m*,*$|\psi|$*).*

Then we derandomize this by using a more powerful $\#\mathsf{P}$ oracle, proving that $\mathsf{BPP}^{\bigoplus\mathsf{P}} \subseteq \mathsf{P}^{\#\mathsf{P}}$. Towards proving the theorems we first look at the properties of the class $\bigoplus\mathsf{P}$.

2.1 Class $\bigoplus\mathsf{P}$

Definition 2.1. A language $\mathtt{L} \subseteq \{0,1\}^*$ is in $\bigoplus\mathsf{P}$ if there is a poly-time DTM $\mathbf{M}$ and a poly-time computable function q such that

$$x \in \mathtt{L} \iff |\{u \in \{0,1\}^{q(|x|)} : \mathbf{M}(x,u) = 1\}| \text{ is odd.}$$

Definition 2.2. $\bigoplus\mathtt{SAT} := \{\phi : \phi \text{ is a boolean formula and } \#\phi \text{ is odd}\}$

Theorem 2.4. $\bigoplus\mathtt{SAT}$ *is complete for* $\bigoplus\mathsf{P}$ *under polynomial time Karp reductions.*

Proof. The theorem follows from the fact that $\forall \mathtt{L} \in \mathsf{NP}$, $\mathtt{L}$ reduces to $\mathtt{SAT}$ parsimoniously, due to the Cook-Levin Theorem. □

Definition 2.3. The $\bigoplus$ quantifier is defined to be such that $\bigoplus_x \phi(x)$ is true if the number of satisfying assignments of ϕ is odd.

Remark. We note the following properties of the parity quantifier.

- Identifying true with 1 and false with 0 we have
$$\bigoplus_{x\in\{0,1\}^n} \phi(x) \equiv \sum_{x\in\{0,1\}^n} \phi(x) \pmod 2$$
- Define $(\phi\cdot\psi)(x,y) = \phi(x) \wedge \psi(y)$. Note that, $\#(\phi\cdot\psi) = \#\phi\cdot\#\psi$. This gives us,
$$\bigoplus_x \phi(x) \bigwedge \bigoplus_y \psi(y) = \bigoplus_{x,y} (\phi\cdot\psi)(x,y)$$
- Let $\#\phi(x) = m$ and $\#\psi(y) = n$. We would like to construct γ such that $\#\gamma = m + n$. Consider
$$\gamma(x,y,z) = ((z=0) \bigwedge \phi(x) \bigwedge (y=0)) \bigvee ((z=1) \bigwedge \psi(y) \bigwedge (x=0))$$
Note that γ satisfies the requirement. We denote $\gamma = \phi + \psi$.
- Let $1(y) = y_1 \wedge \cdots \wedge y_n$. Define $(\phi+1)(x,y) = \phi(x) + 1(y)$. Note that $\#(\phi+1)(x,y) = \#\phi + 1$. Thus, we have
$$\neg \bigoplus_z \phi(x) = \bigoplus_{x,y} (\phi+1)(x,y)$$
.

Theorem 2.5. $\mathsf{co}\text{-}\bigoplus\mathsf{P} = \bigoplus\mathsf{P}$

Proof. This is clear from the fact that $\phi \in \overline{\bigoplus\mathtt{SAT}} \implies (\phi+1) \in \bigoplus\mathtt{SAT}$. □

2.2 Valiant-Vazirani Theorem

Definition 2.4. $\mathtt{USAT} := \{\psi : \psi \text{ is a CNF formula and } \#\psi = 1\}$

Theorem 2.6 ([VV85])**.** *There is a poly-time PTM* $\boldsymbol{M}$ *such that on input* ϕ*(boolean formula in* n *variables),* $\boldsymbol{M}$ *outputs another formula* ψ *such that*

$$\phi \in \mathtt{SAT} \implies Pr\{\psi \in \mathtt{USAT}\} \geq \frac{1}{8n}$$
$$\phi \notin \mathtt{SAT} \implies Pr\{\psi \notin \mathtt{SAT}\} = 1$$

Proof. Let S be a the set of all satisfying assignments of input ϕ. Note that $0 \leq |S| \leq 2^n$. Say that $2^{k-2} \leq |S| \leq 2^{k-1}$ where $k \in [2, n+1]$. $\mathbf{M}$ picks a k randomly from the set $\{1, ..., n+1\}$. With probability n^{-1}, $\mathbf{M}$ will pick k for which $2^{k-2} \leq |S| \leq 2^{k-1}$. $\mathbf{M}$ then picks a hash function h uniformly at random from a pairwise independant family $\mathcal{H}_{n,k}$. Let $X = |\{a \in S : h(a) = 0^k\}|$ and note that $E(X) = \frac{|S|}{2^k}$. Thus, we have $\frac{1}{4} \leq E[X] \leq \frac{1}{2}$. From the inclusion-exclusion principle we have,

$$Pr[X = 1] = Pr[X \geq 1] - Pr[X \geq 2]$$

Denoting S as $\{a_1, \ldots a_{|S|}\}$, we defined the following sets as $\epsilon_i = \{h : h(a_i) = 0\}$. We bound the above probabilities as follows.

$$\begin{aligned} Pr[X \geq 1] &= Pr\left[\bigcup_{i=1}^{|S|} \epsilon_i\right] \\ &\geq \sum_{i=1}^{|S|} Pr[\epsilon_i] - \sum_{i \neq j} Pr[\epsilon_i \cap \epsilon_j] \\ &= \frac{|S|}{2^k} - \binom{|S|}{2} \frac{1}{2^{2k}} \end{aligned}$$

Continuing

$$\begin{aligned} Pr[X \geq 2] &\leq \sum_{i \neq j} Pr[\epsilon_i \cap \epsilon_j] \\ &= \binom{|S|}{2} \frac{1}{2^{2k}} \end{aligned}$$

Thus, we get

$$\begin{aligned} Pr[X = 1] &\geq \frac{|S|}{2^k} - 2\binom{|S|}{2} \frac{1}{2^{2k}} \\ &\geq \frac{|S|}{2^k} - \frac{|S|^2}{2^{2k}} \\ &= \frac{|S|}{2^k}\left(1 - \frac{|S|}{2^k}\right) \\ &\geq \frac{1}{8} \end{aligned}$$

Consider $\psi(x) = \phi(x) \wedge (h(x) = 0)$. We can find a CNF formula corresponding to $h(x) = 0$, say $\tau(x)$. Thus, we have a CNF formula $\psi(x)$ such that

$$\phi \in \mathtt{SAT} \implies Pr\{\psi \in \mathtt{USAT}\} \geq \frac{1}{8n}$$
$$\phi \notin \mathtt{SAT} \implies Pr\{\psi \notin \mathtt{SAT}\} = 1$$

as required. ∎

References

[AB09] Sanjeev Arora and Boaz Barak. *Computational complexity: a modern approach.* Cambridge University Press, 2009.

[Tod91] Seinosuke Toda. PP is as hard as the polynomial-time hierarchy. *SIAM Journal on Computing*, 20(5):865–877, 1991.

[VV85] Leslie G Valiant and Vijay V Vazirani. NP is as easy as detecting unique solutions. In *Proceedings of the seventeenth annual ACM symposium on Theory of computing*, pages 458–463. ACM, 1985.