IITM-CS6840: Advanced Complexity Theory Feb 2, 2012

Lecture 16 : Valiant-Vazirani Lemma

Lecturer: Jayalal Sarma *Scribe: Rahul CS*

THEME: NP $\subseteq$ BP.($\oplus$P)

In the last lecture, we outlined an approach to prove Toda's theorem. One of the key ingredients was a partial answer to the question that we had in one of the earlier lectures. Is NP contained in $\oplus$P? In the last lecture we viewed $\exists$, $\oplus$, BP as operators on complexity classes and stated that NP is contained in BP.($\oplus$P). We also interpreted this in the following way; there is a randomized reduction from SAT to a language in $\oplus$P.

1 Randomized Reduction from SAT to $\oplus$SAT

Define languages,

$$\text{USAT}=\{\phi \mid \#\phi = 1\}$$

$$\oplus\text{SAT}=\{\phi \mid \exists k \in \mathbb{N}, \#\phi = 2k+1\}$$

The following Lemma, famously known as the Valiant-Vazirani Lemma, was proved by Valiant and Vazirani in 1986. It proved instrumental for many results later, including Toda's theorem which we will be taking up in the next lecture.

The lemma states the following reduction from SAT to $USAT$.

Lemma 1 (Valiant-Vazirani Lemma). *There exists a randomized polynomial time algorithm that takes input ϕ and produces a formula $\psi_{\phi,y}$ (say ψ) such that,*

$$\begin{aligned} \phi \in \mathsf{SAT} \Rightarrow Pr(\psi \in USAT) &\geq \frac{1}{8n} \\ \phi \notin \mathsf{SAT} \Rightarrow Pr(\psi' \notin \oplus\mathsf{SAT}] &= 1 \end{aligned}$$

Clearly, $\psi \in$ USAT $\Rightarrow \psi \in \oplus$ SAT. In the other case since $\psi \notin$ SAT, it is also case that $\psi \notin \oplus$SAT. Thus as a corollary to the lemma, we get:

Corollary 2. *There exists a randomized polynomial time algorithm that takes input ϕ and produces a formula ψ such that,*

$$\begin{aligned} \phi \in \mathsf{SAT} \Rightarrow Pr(\psi \in \oplus\mathsf{SAT}) &\geq \frac{1}{8n} \\ \phi \notin \mathsf{SAT} \Rightarrow Pr(\psi \notin \oplus\mathsf{SAT}] &= 1 \end{aligned}$$

Proof. We present a high-level idea first. Given ϕ, a natural approach to produce a formula with unique satisfying assignment is to add another conjunction to produce $\psi = \phi \wedge (\omega)$ such that ω filters out the satisfying assignments using the clause ω such that only one of them will satisfy the resulting formula. Clearly, if ϕ is not satisfiable, then by construction, ψ is also not satisfiable, no matter what ω is. Consider the case when ϕ is satisfiable. Thus, we want ω to state a property which only one of the satisfying assignments have.

We use hashing to achieve this. That is, we will make ω state that $h(x) = 0^k$ where 0^k is in range of the hash family and x is an assignment. The probability that there is a collision at 0^k for two xs that satisfy ϕ (that is, probability that two assignments that satisfy ϕ also satisfies ω) can be controlled by choosing a nice hash family. Thus our filter, with high probability, filters out a unique satisfying assignment for ψ from the set of satisfying assignments of ϕ.

Usually, we design hash families with a size parameter k; the size up to which we want to guarantee collision-free property. But here we do not know the size of the subset (the set of satisfying assignments) a priori for which we are trying to achieve unique mapping(collision-free). The smaller the subset the smaller the range (of the hash functions) that we can work with. Let us say we choose randomly the number k such that the number of satisfying assignments is between 2^k and 2^{k+1} and then decide hash function for all subsets of that size. Since the number of satisfying assignments could be any number from 0 to 2^n, we have already lost out a bit on the probability but by a factor of at most $\frac{1}{n}$.

Now we will formally address this intuition. Let $T \subseteq \{0,1\}^n$ be the set of satisfying assignments of ϕ. Select $k \in \{0, 1, ..., n-1\}$ such that $2^k \leq |T| \leq 2^{k+1}$. Let $H_{n,k}$ be a collection of functions $h : \{0,1\}^n \mapsto \{0,1\}^k$. $H_{n,k}$ is said to be pairwise independent if for every x,$x' \in \{0,1\}^n$ with $x \neq x'$ and for every y,$y' \in \{0,1\}^k$,

$$\Pr_{h \in_R H_{n,k}}[h(x) = y \wedge h(x') = y'] = \frac{1}{2^k} . \frac{1}{2^k} = 2^{-2k}$$

Construct a family of pairwise independent hash functions $H_{n,k+2}$. Hence,

$$\Pr_{h \in H, x \in T}[h(x) = 0^{k+2}] = \frac{1}{2^{k+2}}$$

We calculate the probability of existence of a hash function $h \in H$ such that h maps exactly one satisfying assignment $x \in T$ to 0^{k+2}. This parameter is given by,

$$\Pr_{h \in H}[|\{x \mid x \in T, h(x) = 0^{k+2}\}| = 1]$$

Once we have an h satisfying the above condition, the number of satisfying assignments for ϕ will be exactly 1. That is, in the described transformation, if ω is $h(x) = 0^{k+2}$, the number of satisfying assignments for ψ will be exactly 1, if it is satisfiable. Note that we can

consider the computation sequence of the hash function h on a turing machine and convert that to a SAT formula using Cook-Levin reduction. We claim that,

$$\Pr_{h \in H} [|\{x \mid x \in T, h(x) = 0^{k+2}\}| = 1] \geq \frac{1}{8}$$

Consider,

$$\Pr_{h \in H} [\exists x \in T, h(x) = 0^{k+2} \wedge \forall_{x' \in T, x' \neq x} h(x') \neq 0^{k+2}]$$

$$\begin{aligned}
&= \Pr_h [\forall_{x' \in T, x' \neq x} h(x') \neq 0^{k+2} \mid h(x) = 0^{k+2}]. \Pr_h [h(x) = 0^{k+2}] \\
&= (1 - \Pr_h [\exists_{x' \in T, x' \neq x} h(x') = 0^{k+2} | h(x) = 0^{k+2}]). \Pr_h [h(x) = 0^{k+2}] \\
&= (1 - \sum_{\substack{x' \in T \\ x' \neq x}} \Pr_h [h(x') = 0^{k+2} \mid h(x) = 0^{k+2}]). \Pr_h [h(x) = 0^{k+2}]
\end{aligned}$$

Let us calculate the first part of the expression, $(1 - \sum_{\substack{x' \in T \\ x' \neq x}} \Pr_h [h(x') = 0^{k+2} \mid h(x) = 0^{k+2}])$. Since h is sampled from a family of pairwise independent hash functions, h satisfies the condition of simple uniform hashing. That is, the events $h(x) = 0^{k+2}$ and $h(x') = 0^{k+2}$ are independent. Therefore,

$$\sum_{\substack{x' \in T \\ x' \neq x}} \Pr_h [h(x') = 0^{k+2} \mid h(x) = 0^{k+2}]) = \sum_{\substack{x' \in T \\ x' \neq x}} \frac{1}{2^{k+2}} = |T - \{x\}|. \frac{1}{2^{k+2}}$$

We bound the expression from above by bounding $|T - \{x\}|$ from above. That is,

$$\sum_{\substack{x' \in T \\ x' \neq x}} \Pr_h [h(x') = 0^{k+2} \mid h(x) = 0^{k+2}]) \leq (2^{k+1} - 1) \frac{1}{2^{k+2}}$$

Hence,

$$\begin{aligned}
(1 - \sum_{\substack{x' \in T \\ x' \neq x}} \Pr_h [h(x') = 0^{k+2} \mid h(x) = 0^{k+2}]) &\geq 1 - (2^{k+1} - 1) \frac{1}{2^{k+2}} \\
&= 1 - \frac{1}{2} [\frac{(2^{k+1} - 1)}{2^{k+2}}] \\
&\geq \frac{1}{2}
\end{aligned}$$

Therefore,

$$(1 - \sum_{\substack{x' \in T \\ x' \neq x}} \Pr_h [h(x') = 0^{k+2} | h(x) = 0^{k+2}]). \Pr_h [h(x) = 0^{k+2}] \geq \frac{1}{2} \cdot \frac{1}{2^{k+2}}$$

Thus,

$$|T|.(1-\sum_{\substack{x'\in T\\ x'\neq x}} \Pr_h[h(x')=0^{k+2} \mid h(x)=0^{k+2}]).\Pr_h[h(x)=0^{k+2}] \geq |T|.\frac{1}{2}\frac{1}{2^{k+2}}$$

By bounding $|T|$ from below, the probability that there is a $x \in T$ which uniquely gets mapped to 0^{k+2} is given by,

$$|T|.(1-\sum_{\substack{x'\in T\\ x'\neq x}} \Pr_h[h(x')=0^{k+2} \mid h(x)=0^{k+2}]).\Pr_h[h(x)=0^{k+2}] \geq \frac{1}{2^k}\frac{1}{2}\frac{1}{2^{k+2}} \geq \frac{1}{8}$$

Considering the fact that $k \in \{0,1,...,n-1\}$ satisfies $2^k \leq |T| \leq 2^{k+1}$ with probability $\frac{1}{n}$ we get,

$$\phi \in \mathsf{SAT} \Rightarrow \Pr_x[\psi_{\phi,x} \in USAT] \geq \frac{1}{8n}$$
$$\phi \notin \mathsf{SAT} \Rightarrow \Pr_x[\psi_{\phi,x} \notin SAT] = 1$$

□

Now we make comments about amplification of success probability. Since the algorithm is one-sided error, an approach will be to simply repeat the experiment some ℓ times, and take the $\vee$ of the results. This however, causes loss of uniqueness of the assignment since different h may make different xs to go to 0^k. However, we can still preserve the parity with high probability that that results in the following amplification result which we state as a Lemma.

Lemma 3. *There exists a randomized polynomial time algorithm that takes input ϕ and produces a formula ψ' such that,*

$$\begin{aligned} \phi \in \mathsf{SAT} \Rightarrow Pr(\psi' \in \oplus\mathsf{SAT}) &\geq 1-\frac{1}{2^{q(n)}} \\ \phi \notin \mathsf{SAT} \Rightarrow Pr(\psi' \notin \oplus\mathsf{SAT}] &= 1 \end{aligned}$$

The OR amplification says, that we report that $\phi \in \oplus\mathsf{SAT}$ if and only if at least one of the trials produces a formula in $\oplus\mathsf{SAT}$. But then, can we produce a single formula ψ' such that if ϕ is in SAT, then ψ' is in $\oplus\mathsf{SAT}$ with high probability. We will address such a reduction in the next lecture.

Note that it is an open problem to boost the probability of $\frac{1}{8n}$ in the Valiant-Vazirani reduction to USAT.