The Valiant-Vazirani Lemma

This is the proof that was presented in class on April 12 and 14, 2016.

Consider the following "promise problem," which we denote by USAT ("U" for uniquely). The input is a CNF formula ϕ; the output N is a non-negative integer. Recall that $\#(\phi)$ is the number of satisfying assignments of ϕ.

$$\begin{aligned}
\#(\phi) = 0 &\Rightarrow \; N = 0 \\
\#(\phi) = 1 &\Rightarrow \; N = 1 \\
\#(\phi) > 1 &\Rightarrow \; N \text{ can be any non-negative integer.}
\end{aligned}$$

So the "promise" in USAT is that ϕ either is unsatisfiable or has a unique satisfying assignment. If the promise is kept, then the output is correct; if the promise is broken, then the output could be anything.

On the face of it, USAT seems as though it may be easier than the general SAT problem. However, Valiant and Vazirani have shown the following:

Theorem 1: If USAT can be solved in polynomial time, then NP = RP.

Their main technical result is

Theorem 2: There is a probabilistic polynomial-time algorithm A such that

$$\begin{aligned}
\phi \in \text{SAT} &\Rightarrow \; \text{Prob}[A(\phi) \text{ has a unique satisfying assignment}] \geq \frac{1}{8n}, \text{ and} \\
\phi \notin \text{SAT} &\Rightarrow \; \text{Prob}[A(\phi) \text{ has zero satisfying assignments}] = 1.
\end{aligned}$$

Here, the probability is computed over the coin tosses of A.

We will prove Theorem 2 below. First observe that Theorem 2 implies Theorem 1. Suppose that B were a deterministic algorithm that solved USAT in polynomial time. Then $B \circ A$ would be an algorithm that, if given a satisfiable formula as input, would output 1 with probability at least $\frac{1}{8n}$ and, if given an unsatisfiable formula as input, would output 0 with probability 1. Note that the correctness probability of the first case could be made exponentially close to 1 using polynomially many independent trials of $B \circ A$ and outputting 1 if and only if at least one trial outputs 1. The existence of such an algorithm would imply that SAT was in RP and thus that NP = RP.

Proof of Theorem 2: Let $\mathcal{H}_{n,k}$ be a pairwise-independent hash-function family, where $2 \leq k \leq n+1$. Let $p = 2^{-k}$ and $S \subseteq \{0,1\}^n$ be such that $2^{k-2} \leq |S| \leq 2^{k-1}$.

Recall that

$$\text{Prob}_{h \in_R \mathcal{H}_{n,k}}[h(x) = 0^k] = p, \text{ for all } x, \text{ and} \tag{1}$$

$$\text{Prob}_{h \in_R \mathcal{H}_{n,k}}[h(x) = 0^k \wedge h(x') = 0^k] = p^2, \text{ for all } x \neq x'. \tag{2}$$

Claim 3: $\text{Prob}_{h \in_R \mathcal{H}_{n,k}}[|x \in S \text{ such that } h(x) = 0^k| = 1] \geq \frac{1}{8}$.

We will prove Claim 3 below. First, we argue that it can be used to prove Theorem 2. For a given $k \in [2, n+1]$ and $h \in \mathcal{H}_{n,k}$, let M be a polynomial-time machine that accepts x if and only if $h(x) = 0^k$. Note that M can also be regarded as a nondeterministic, polynomial-time machine with the property that, on any input x, M has either one or zero accepting computations.

Now apply the Cook-Levin reduction to M. Because the reduction is parsimonious,[1] this produces, for every x such that $h(x) = 0^k$, a CNF formula $\tau(x, z)$ that has a unique satisfying assignment and, for every x such that $h(x) \neq 0^k$, a CNF formula $\tau(x, z)$ that has no satisfying assignments. Here z is the string of Boolean "snapshot variables" created by the Cook-Levin reduction.

The probabilistic, polynomial-time algorithm A of Theorem 2 proceeds as follows. First choose k uniformly at random from $[2, n + 1]$. Next, choose h uniformly at random from $\mathcal{H}_{n,k}$. For the M that corresponds to h and k, compute the corresponding $\tau(x, z)$. Then $A(\phi) = \phi(x) \wedge \tau(x, z)$. Clearly, if $\phi \notin$ SAT, then $A(\phi) \notin$ SAT as well. If $\phi \in$ SAT, then, with probability at least $\frac{1}{n}$, A chooses a k for which the set S of satisfying assignments of ϕ is such that $2^{k-2} \leq |S| \leq 2^{k-1}$. Conditioned upon A's making such a choice, with probability at least $\frac{1}{8}$, Claim 3 guarantees that there is a unique $x \in S$ such that $h(x) = 0^k$ and hence a unique x, y such that $\tau(x, z) = 1$. Thus, with probability at least $\frac{1}{8n}$, $A(\phi)$ has a unique satisfying assignment.

Proof of Claim 3: Let $N = |\{x \in S \text{ such that } h(x) = 0^k\}|$. Then N is a random variable the distribution of which is determined by the uniformly random choice of $h \in \mathcal{H}_{n,k}$. We are interested in the probability that $N = 1$, which is the difference between the probability that $N \geq 1$ and the probability that $N \geq 2$.

$$\mathrm{Prob}_{h \in_R \mathcal{H}_{n,k}}[N \geq 1] = \mathrm{Prob}_{h \in_R \mathcal{H}_{n,k}}[\exists x \text{ s.t. } h(x) = 0^k] = \mathrm{Prob}_{h \in_R \mathcal{H}_{n,k}}[\bigvee_{x \in S} h(x) = 0^k].$$

Apply equations (1) and (2) above, together with the inclusion-exclusion principle, to obtain

$$\begin{aligned} & \mathrm{Prob}_{h \in_R \mathcal{H}_{n,k}}[\bigvee_{x \in S} h(x) = 0^k] \\ \geq \quad & \sum_{x \in S} \mathrm{Prob}_{h \in_R \mathcal{H}_{n,k}}[h(x) = 0^k] \quad - \sum_{\{x, x'\} \in S,\ x \neq x'} \mathrm{Prob}_{h \in_R \mathcal{H}_{n,k}}[h(x) = h(x') = 0^k] \\ = \quad & |S|p - \binom{|S|}{2} p^2. \end{aligned}$$

[1]Parsimony was not covered in the January 26, 2016, lecture in which we proved the Cook-Levin Theorem. However, it is there implicitly and can be stated as follows. Consider the bitstring $y_1 y_2 \ldots y_n y_{n+1} y_{n+2} \ldots y_{n+p(n)} t_1 t_2 \ldots t_{cT(n)}$ defined in that lecture. There is a one-to-one correspondence between substrings $y_{n+1} y_{n+2} \ldots y_{n+p(n)}$, which encode potential witnesses u that x is in the NP language L, and potential satisfying assignments to the CNF formula produced by the reduction. Note that, if L is in P, then the substring $y_{n+1} y_{n+2} \ldots y_{n+p(n)}$ is empty; there is just one computation of the machine M, and it is an accepting computation if and only if $y_1 y_2 \ldots y_n t_1 t_2 \ldots t_{cT(n)}$ satisfies the formula.

Now apply the union bound to obtain

$$\begin{aligned}
\mathrm{Prob}_{h\in_R \mathcal{H}_{n,k}}[N \geq 2] &= \mathrm{Prob}_{h\in_R \mathcal{H}_{n,k}}[\bigvee_{\{x,x'\}\in S,\ x\neq x'} h(x) = h(x') = 0^k] \\
&\leq \sum_{\{x,x'\}\in S,\ x\neq x'} \mathrm{Prob}_{h\in_R \mathcal{H}_{n,k}}[h(x) = h(x') = 0^k] \\
&= \binom{|S|}{2} p^2.
\end{aligned}$$

Putting together the lower bound on the probability that $N \geq 1$ and the upper bound on the probability that $N \geq 2$, we have

$$\begin{aligned}
\mathrm{Prob}_{h\in_R \mathcal{H}_{n,k}}[N = 1] &= \mathrm{Prob}_{h\in_R \mathcal{H}_{n,k}}[N \geq 1] - \mathrm{Prob}_{h\in_R \mathcal{H}_{n,k}}[N \geq 2] \\
&\geq |S|p - 2\binom{|S|}{2} p^2 \\
&\geq |S|p - |S|^2 p^2 \\
&= |S|p \cdot (1 - |S|p).
\end{aligned}$$

Because $p = 2^k$ and $2^{k-2} \leq |S| \leq 2^{k-1}$, we have $|S|p \geq \frac{1}{4}$ and $(1 - |S|p) \geq \frac{1}{2}$ (the latter because $|S|p \leq \frac{1}{2}$). This implies that $|S|p \cdot (1 - |S|p)$, our lower bound on the probability that $N = 1$, is at least $\frac{1}{8}$ and completes the proof of Claim 3.

We do not know how to amplify the correctness probability of this USAT version of the Valiant-Vazirani lemma. However, it is the $\oplus$SAT version that we need for the proof of Toda's Theorem, and we do know how to amplify *its* correctness probability. To do so, we will treat $\oplus$ and $\#$ as operators on formulas. Let $\phi(x_1, x_2, \ldots, x_n)$ and $\psi(y_1, y_2, \ldots, y_m)$ be formulas on disjoint sets of boolean variables $\{x_1, x_2, \ldots, x_n\}$ and $\{y_1, y_2, \ldots, y_m\}$. Denote by $\phi \cdot \psi$ a formula on $\{x_1, x_2, \ldots, x_n\} \cup \{y_1, y_2, \ldots, y_m\}$ such that $(\phi \cdot \psi)(x_1, x_2, \ldots, x_n, y_1, y_2, \ldots, y_m)$ is satisfied if and only if both $\phi(x_1, x_2, \ldots, x_n)$ and $\psi(y_1, y_2, \ldots, y_m)$ are satisfied. Then

$$\#(\phi \cdot \psi) = \#(\phi) \cdot \#(\psi).$$

Now suppose that $n \geq m$. (The case of $m \geq n$ can be handled analogously.) Denote by $\phi + \psi$ the formula on $\{z_0, z_1, \ldots, z_n\}$ that is satisfied if and only if $(z_0 = 0) \wedge \phi(z_1, \ldots, z_n)$ or $(z_0 = 1) \wedge \psi(z_1, \ldots, z_m) \wedge (z_{m+1} = \cdots = z_n = 0)$. Note that this is, by definition, an exclusive or because of the role of z_0. Then

$$\#(\phi + \psi) = \#(\phi) + \#(\psi).$$

Finally, denote by "1" a formula (on whatever number of variables is required) that has a unique satisfying assignment. For example, we can take it to be $(z_1 = 1) \wedge (z_2 = 1) \wedge \cdots \wedge (z_n = 1)$ if we need a formula on n variables.

Note that multiplying or adding formulas produces a result whose size is the *sum* of the sizes of the original formulas, not the product. This will be needed for Claim 4 below.

We have the following implications for membership in $\oplus$SAT, where the notation "$\oplus_{\overline{x}}\phi(\overline{x})$" means that $\phi \in \oplus$SAT.

$$
\begin{aligned}
(\oplus_{\overline{x}}\phi(\overline{x})) \wedge (\oplus_{\overline{y}}\psi(\overline{y})) &\longleftrightarrow \oplus_{\overline{x},\overline{y}}(\phi \cdot \psi)(\overline{x},\overline{y}) \\
(\oplus_{\overline{x}}\phi(\overline{x})) \vee (\oplus_{\overline{y}}\psi(\overline{y})) &\longleftrightarrow \oplus_{\overline{x},\overline{y},\overline{z}}((\phi+1)\cdot(\psi+1)+1))(\overline{x},\overline{y},\overline{z}) \\
\neg(\oplus_{\overline{x}}\phi(\overline{x})) &\longleftrightarrow \oplus_{\overline{x},\overline{z}}((\phi+1)(\overline{x},\overline{z}))
\end{aligned}
$$

Note that Theorem 2 gives us a probabilistic reduction A with the property that

$$
\begin{aligned}
\phi \in \text{SAT} &\Rightarrow \ \text{Prob}[A(\phi) \in \oplus SAT] \geq \frac{1}{8n} \\
\phi \notin \text{SAT} &\Rightarrow \ \text{Prob}[A(\phi) \in \oplus SAT] = 0,
\end{aligned}
$$

where n is the number of variables in ϕ. To amplify the correctness probability of this reduction to $1-2^{-m}$, where $m = \text{poly}(n)$, first choose an R such that $1-(1-\frac{1}{8n})^R \geq 1-2^{-m}$. Run R independent executions of $A(\phi)$ to produce formulas ψ_1, ψ_2, ..., ψ_R. Compute a single formula ψ as follows:

$\psi \leftarrow \psi_1$;
FOR ($i \leftarrow 1$ TO $R-1$)
  $\psi \leftarrow (\psi+1)\cdot(\psi_{i+1}+1)+1$;
OUTPUT ψ;

Claim 4: $\psi \in \oplus$SAT if and only if at least one of $\psi_i \in \oplus$SAT, for $1 \leq i \leq R$.

Claim 4 gives us a reduction with the properties we want: If $\phi \notin$ SAT, then all of the ψ_i are unsatisfiable (and hence not in $\oplus$SAT), and the probability that $\psi \in \oplus$SAT is 0. If $\phi \in$ SAT, then the ψ_i's are, independently, in $\oplus$SAT with probability $\frac{1}{8n}$; so the probability that at least one of them (and hence ψ) is in $\oplus$SAT is at least $1-(1-\frac{1}{8n})^R \geq 1-2^{-m}$.