

CMPT 710 - Complexity Theory: Lecture 19

Valentine Kabanets

November 8, 2007

1 Randomized NP

Randomness is a computational resource that can be combined with other resources, e.g., nondeterminism. We will consider such a combination next.

Recall that a language $L \in \mathbf{NP}$ if there is polytime verifier such that, for every $x \in L$, there is a proof (or witness) of small size that makes the verifier accept. (The verifier is the polytime relation $R(x, y)$, where x is the input, and y is supposed to be a “proof” that $x \in L$.) In this definition, the verifier is a *deterministic* polytime algorithm. By letting the verifier be *randomized* polytime algorithm, we obtain an extension of \mathbf{NP} , called \mathbf{MA} (which stands for Merlin-Arthur). Here, Merlin is an all-powerful wizard that tries to convince a probabilistic polytime verifier Arthur that an input string is in the language. If a string is indeed in the language, Merlin can make Arthur accept most of the time. If, on the other hand, the string is not in the language, then no matter what kind of proof Merlin shows to Arthur, Arthur will reject most of the time.

More formally, we say that a language $L \in \mathbf{MA}$ if there is a constant c , and a polytime relation $R(x, y, z)$ such that, for every x of length n , we have

$$\begin{aligned}x \in L &\Rightarrow \exists y \Pr_z[R(x, y, z) = 1] \geq 3/4 \\x \notin L &\Rightarrow \forall y \Pr_z[R(x, y, z) = 1] \leq 1/4,\end{aligned}$$

where $|y| = |z| \leq n^c$. In other words, if $x \in L$, then there is a short proof y that will convince Arthur with probability at least $3/4$, and if $x \notin L$, then every y will be rejected by Arthur with probability at least $3/4$.

2 Example of a language in MA

Recall that an *arithmetic formula* is a tree whose leaves are labeled by variables or constants, and whose inner nodes are labeled by arithmetic operations $+$, $-$, and $*$. The *size* of an arithmetic formula is the number of leaves in the tree representation of the formula. (Note that variables labeling the leaves may repeat, i.e., the same variable may label more than one leaf.)

Each arithmetic formula $f(x_1, \dots, x_n)$ computes some polynomial $p(x_1, \dots, x_n)$ over the integers. The same polynomial may or may not be computable by a smaller formula. We

say that an arithmetic formula f is *optimal* if it is a smallest possible formula that computes the polynomial p , i.e., any smaller arithmetic formula computes a different polynomial than p . We say that a formula is *non-optimal* if there is a smaller formula computing the same polynomial.

Define $L = \{f(x_1, \dots, x_n) \mid f \text{ is a non-optimal arithmetic formula}\}$. We will show that $L \in \mathbf{MA}$.

Let $f(x_1, \dots, x_n)$ be an input arithmetic formula of size s . From the last lecture, we know that the degree of the polynomial computed by f is at most s . Let $S = \{1, 2, \dots, 4s\}$. Merlin will try to convince Arthur that f is non-optimal by providing Arthur with a strictly smaller formula $g(x_1, \dots, x_n)$ that is supposed to compute the same polynomial as f . Upon receiving a formula $g(x_1, \dots, x_n)$, Arthur will pick a sequence of random integers r_1, \dots, r_n from the set S defined above, and will check whether $f(r_1, \dots, r_n) = g(r_1, \dots, r_n)$. If the equality holds, then Arthur accepts; otherwise, Arthur rejects.

Let us analyze the correctness of the described \mathbf{MA} protocol. Suppose f is indeed non-optimal. Then Merlin can send to Arthur a smaller equivalent formula g , and Arthur will accept with probability $1 \geq 3/4$ (since $f \equiv g$ means that f and g agree on all possible inputs). Now suppose that f is optimal. Then whatever smaller formula g is sent to Arthur, the polynomial $f - g$ is non-zero, of degree at most s . Hence, by the Schwartz-Zippel lemma, Arthur will accept in this case with probability at most $s/(4s) = 1/4$.

Remark The reasoning above can be generalized to show that a language of non-optimal arithmetic *circuits* is also in \mathbf{MA} (by using the modular-arithmetic algorithm from the last lecture for testing if a given arithmetic circuit is zero).

Remark Earlier we saw in class that the language of non-optimal *Boolean* circuits is in Σ_2^P . The case of *arithmetic* circuits is easier — the language of non-optimal arithmetic circuits is in \mathbf{MA} , which is a subset of Σ_2^P . (The inclusion $\mathbf{MA} \subseteq \Sigma_2^P$ can be shown using the ideas from the proof that $\mathbf{BPP} \subseteq \Sigma_2^P$ that we saw in class. *Exercise:* Prove that $\mathbf{MA} \subseteq \Sigma_2^P$. Hint: Use the error reduction result from the next section.)

3 Error Reduction in MA

As in the case of \mathbf{BPP} , we can reduce the error probability of any \mathbf{MA} protocol to be less than an inverse exponential in the input size. Here's how.

Let $L \in \mathbf{MA}$ be any language. Let $R(x, y, z)$ be a polytime relation for L such that, for every $x \in L$, there is a y with $\Pr_z[R(x, y, z) = 1] \geq 3/4$; and for every $x \notin L$, for every y , it holds that $\Pr_z[R(x, y, z) = 1] \leq 1/4$.

Consider a new protocol where, upon receiving a string y , Arthur randomly and independently chooses k strings z_1, \dots, z_k , and accepts iff $R(x, y, z_i) = 1$ for more than half of these k strings.

We use Chernoff bounds to analyze the correctness of the described protocol. Suppose first that $x \in L$. Then Merlin can send Arthur a string y such that $\Pr_z[R(x, y, z) = 1] \geq 3/4$. Every string z_i , $1 \leq i \leq k$, randomly chosen by Arthur has probability at least $3/4$ of satisfying R . The expected number of z_i 's that satisfy R is $\mu \geq \frac{3}{4}k$. Let X_i , $1 \leq i \leq k$, be a random variable that is 1 if z_i satisfies R , and 0 otherwise. Let $X = \sum_{i=1}^k X_i$. As we

just argued, the expectation of X is μ . Using Chernoff bounds, we get that $\Pr[X \leq k/2] \leq \Pr[|X - \mu| > k/4] < 2e^{-k/48}$. Thus, Arthur will accept with probability exponentially close to 1.

On the other hand, suppose that $x \notin L$. Then, whatever y is sent to Arthur, $\Pr_z[R(x, y, z) = 1] \leq 1/4$. Let X_i be random variables as defined above, and let $X = \sum_{i=1}^k X_i$. Then the expectation of X is $\leq \frac{1}{4}k$. The probability that Arthur accepts in this case is $\Pr[X > k/2]$, which, by Chernoff bounds, is at most $2e^{-k/48}$. Thus, in this case, Arthur will accept with probability exponentially close to 0.

4 Class AM

Suppose we change the order in which Merlin and Arthur communicate by letting Arthur go first, and Merlin go second. More formally, we say that a language $L \in \mathbf{AM}$ if there is a constant c and a polytime relation $R(x, y, z)$ such that, for every x of length n ,

$$\begin{aligned} x \in L &\Rightarrow \Pr_z[\exists y : R(x, y, z) = 1] \geq 3/4, \\ x \notin L &\Rightarrow \Pr_z[\exists y : R(x, y, z) = 1] \leq 1/4, \end{aligned}$$

where $|y| = |z| \leq n^c$.

In other words, if $x \in L$, then Merlin can successfully answer with y almost every random challenge z from Arthur. If $x \notin L$, then, for most random challenges from Arthur, Merlin does not have a good answer.

How are the classes MA and AM related to each other? We show the following.

Theorem 1. $\mathbf{MA} \subseteq \mathbf{AM}$

We'll give the proof next time.