CS 6810: Theory of Computing **Spring 2026**

Lecture 20: April 9, 2026

Lecturer: Eshan Chattopadhyay *Scribe: Daniel Cao*

1 More on Arthur–Merlin protocols

In this lecture we continue the study of Arthur–Merlin protocols. Recall that in an AM protocol Arthur first sends a uniformly random string r, Merlin replies with a message m, and Arthur decides whether to accept by evaluating a deterministic polynomial-time predicate $V(x, r, m)$. In an MA protocol the order is reversed: Merlin first sends m, Arthur then samples random coins r, and Arthur accepts or rejects according to $V(x, m, r)$.

From the previous lecture, we already know that $\mathsf{GNI} \in \mathsf{AM}$. The main goals of this lecture are:

1. recall perfect completeness for MA,
2. prove $\mathsf{MA} \subseteq \mathsf{AM}$,
3. deduce $\mathsf{AM}[O(1)] = \mathsf{AM}$,
4. prove perfect completeness for AM,
5. show that $\mathsf{AM} \subseteq \Pi_2^p$,
6. explain why if GI is NP-complete, then PH collapses to level 2.

2 Recap: $\mathsf{GNI} \in \mathsf{AM}$

We briefly recall that graph non-isomorphism has an Arthur–Merlin protocol. If the input graphs G_1, G_2 are non-isomorphic, then Merlin can successfully distinguish a random permutation of one of the two graphs. If they are isomorphic, then no prover can do substantially better than guessing. In the previous lecture this was implemented through the set lower bound protocol.

We will use this important fact later when discussing why graph isomorphism is unlikely to be NP-complete.

3 Recap: perfect completeness for MA

We also recall the perfect completeness theorem for MA.

Theorem 3.1. *For every language $L \in \mathsf{MA}$, there is an MA protocol with perfect completeness. That is, there exists a probabilistic polynomial-time verifier V such that:*

1. *if $x \in L$, then there exists a message m such that*
$$\mathbb{P}_r[V(x, m, r) = 1] = 1,$$
2. *if $x \notin L$, then for every m,*
$$\mathbb{P}_r[V(x, m, r) = 1] < \frac{1}{3}.$$

Proof sketch. For any fixed input x and Merlin message m, let

$$I_{x,m} := \{r : V(x, m, r) = 1\} \subseteq \{0,1\}^\ell.$$

After standard error reduction, we may assume:

1. if $x \in L$, then for some m we have
$$|I_{x,m}| \geq (1-\varepsilon)2^\ell,$$
2. if $x \notin L$, then for every m we have
$$|I_{x,m}| \leq \varepsilon 2^\ell,$$

where $\varepsilon = 2^{-n}$.

Now use the covering trick from the earlier lecture on placing BPP inside PH: there exist $v_1, \ldots, v_t \in \{0,1\}^\ell$, with $t = \text{poly}(n)$, such that whenever $S \subseteq \{0,1\}^\ell$ has size at least $(1-\varepsilon)2^\ell$, then

$$\bigcup_{i=1}^{t} (S \oplus v_i) = \{0,1\}^\ell.$$

Applying this to the large set $I_{x,m}$ from the yes case yields a new verifier that accepts for every random string r, and hence has perfect completeness. In the no case, soundness follows from the union bound. □

4 MA ⊆ AM

We now prove that any Merlin–Arthur protocol can be converted into an Arthur–Merlin protocol.

Theorem 4.1. MA ⊆ AM.

Proof. Let $L \in \mathsf{MA}$. By perfect completeness of MA, there is a verifier V and polynomials f, g such that:

1. if $x \in L$, then there exists some $m \in \{0,1\}^{g(|x|)}$ such that for all $r \in \{0,1\}^{f(|x|)}$,
$$V(x, m, r) = 1,$$
2. if $x \notin L$, then for all $m \in \{0,1\}^{g(|x|)}$,
$$\mathbb{P}_r[V(x, m, r) = 1] < \frac{1}{3}.$$

We first amplify the soundness error while preserving perfect completeness. Define a new verifier $\widehat{V}(x, m, r)$ as follows: interpret r as a tuple

$$r = \langle r_1, \ldots, r_{g(|x|)+2} \rangle,$$

where each $r_i \in \{0,1\}^{f(|x|)}$, and accept iff

$$\bigwedge_{i=1}^{g(|x|)+2} V(x, m, r_i) = 1.$$

Then:

1. if $x \in L$, the same message m causes $\widehat{V}$ to accept for every random string;
2. if $x \notin L$, then for every m,

$$\mathbb{P}_r[\widehat{V}(x, m, r) = 1] < 2^{-g(|x|)-2}.$$

Now turn this into an AM protocol by having Arthur send the full random string r first. Merlin then responds with some message $m(r)$. Let $\widehat{V}'$ denote this public-coin verifier.

If $x \in L$, Merlin simply sends the fixed perfect-completeness witness m, regardless of r. Thus $\widehat{V}'$ accepts with probability 1.

If $x \notin L$, then

$$\mathbb{P}_r[\widehat{V}'(x, m(r), r) = 1] = \frac{1}{2^{|r|}} \sum_{r \in \{0,1\}^{|r|}} \mathbf{1}[\widehat{V}'(x, m(r), r) = 1].$$

We upper bound this by summing over all possible Merlin messages:

$$\mathbb{P}_r[\widehat{V}'(x, m(r), r) = 1] \leq \sum_{m \in \{0,1\}^{g(|x|)}} \mathbb{P}_r[\widehat{V}(x, m, r) = 1] < 2^{g(|x|)} \cdot 2^{-g(|x|)-2} = \frac{1}{4} < \frac{1}{3}.$$

Therefore the resulting public-coin protocol is an AM protocol for L. □

Corollary 4.2. $\mathsf{MAM} = \mathsf{AM}$.

Proof. A MAM protocol begins with a Merlin message and then runs an AM protocol. The initial Merlin message together with Arthur's next random message form an MA subprotocol. By the theorem above, this can be replaced by an AM protocol. Merging consecutive Arthur messages and consecutive Merlin messages gives an AM protocol. □

Corollary 4.3. $\mathsf{AM}[O(1)] = \mathsf{AM}$.

Proof. We repeatedly apply the previous corollary a constant number of times. □

Observation 4.4. *The above argument is specific to a constant number of rounds. The proof does not show that* $\mathsf{AM}[\mathrm{poly}(n)] = \mathsf{AM}$.

5 Perfect completeness for AM

We now prove that AM also admits perfect completeness.

Theorem 5.1. *For every language* $L \in \mathsf{AM}$*, there is a perfectly complete* AM *protocol.*

Proof. Fix $L \in \mathsf{AM}$. After amplification, we may assume that there exists a verifier V using $\ell = \ell(|x|)$ random bits such that:

1. if $x \in L$, then
$$\mathbb{P}_r[\exists m \text{ such that } V(x, r, m) = 1] > 1 - 2^{-n},$$
2. if $x \notin L$, then
$$\mathbb{P}_r[\exists m \text{ such that } V(x, r, m) = 1] < 2^{-n}.$$

Define

$$I_x := \{r \in \{0,1\}^\ell : \exists m \text{ such that } V(x,r,m) = 1\}.$$

Then:

1. if $x \in L$, then
$$|I_x| > (1 - 2^{-n})2^\ell,$$
2. if $x \notin L$, then
$$|I_x| < 2^{-n}2^\ell.$$

By the same covering lemma used for MA, there exist $\omega_1, \ldots, \omega_k \in \{0,1\}^\ell$, with $k = \text{poly}(n)$, such that if $|S| \geq (1 - 2^{-n})2^\ell$, then

$$\bigcup_{i=1}^{k} (S \oplus \omega_i) = \{0,1\}^\ell.$$

Apply this to $S = I_x$ in the yes case. Then for every $r \in \{0,1\}^\ell$, there exists $i \in [k]$ such that $r \oplus \omega_i \in I_x$, meaning there exists a message m for which

$$V(x, r \oplus \omega_i, m) = 1.$$

This yields a perfectly complete MAM protocol:

- Merlin first sends $\omega_1, \ldots, \omega_k$.
- Arthur samples uniform $r \in \{0,1\}^\ell$ and sends it.
- Merlin responds with $i \in [k]$ and a witness message m.
- Arthur accepts iff $V(x, r \oplus \omega_i, m) = 1$.

If $x \in L$, Merlin can choose shifts $\omega_1, \ldots, \omega_k$ so that for every r there is some i with $r \oplus \omega_i \in I_x$, and then send an appropriate witness m. Hence completeness is perfect.

If $x \notin L$, then for any fixed first Merlin message $\omega_1, \ldots, \omega_k$, the acceptance probability is at most

$$\sum_{i=1}^{k} \mathbb{P}_r[r \oplus \omega_i \in I_x] = k \cdot \frac{|I_x|}{2^\ell} < k2^{-n}.$$

After sufficient amplification we may assume $k2^{-n} < 1/3$, so this is a valid MAM protocol.

Finally, since $\mathsf{MAM} = \mathsf{AM}$, we conclude that L has a perfectly complete AM protocol. □

6 $\mathsf{AM} \subseteq \Pi_2^p$

We now show that AM is contained in the second level of the polynomial hierarchy.

Fact 6.1. $\mathsf{AM} \subseteq \Pi_2^p$.

Proof. Let $L \in \mathsf{AM}$. By Theorem 5.1 we may assume that L has a perfectly complete Arthur–Merlin protocol. Thus there is a polynomial-time verifier V such that:

1. if $x \in L$, then for every random string r there exists a message m such that

$$V(x, r, m) = 1,$$

2. if $x \notin L$, then for every prover strategy P,

$$\mathbb{P}_r[V(x, r, P(r)) = 1] < \frac{1}{3}.$$

We claim that

$$x \in L \iff \forall r \in \{0,1\}^{\ell(|x|)} \; \exists m \;\; V(x, r, m) = 1,$$

where $\ell(|x|)$ is the number of random bits used by Arthur.

The forward direction is immediate from perfect completeness.

For the converse, suppose

$$\forall r \in \{0,1\}^{\ell(|x|)} \; \exists m \;\; V(x, r, m) = 1.$$

Then Merlin can use the following prover strategy: on input r, send any message m for which $V(x, r, m) = 1$ holds. Under this strategy Arthur accepts on every random string, and hence with probability 1. This contradicts the soundness condition for $x \notin L$.

Therefore, when $x \notin L$, there must exist some random string r such that for every message m,

$$V(x, r, m) = 0.$$

Equivalently,

$$x \in L \iff \forall r \, \exists m \, V(x, r, m) = 1.$$

Since V is polynomial-time computable and both r and m have polynomial length, this gives a $\forall\exists$ characterization of L. Hence $L \in \Pi_2^p$. □

7 If GI is NP-complete, then PH collapses

Recall first that $\mathsf{GI} \in \mathsf{NP}$: if two graphs G_1, G_2 are isomorphic, then a certificate is simply a permutation π of the vertices such that $\pi(G_1) = G_2$. Since no polynomial-time algorithm for GI is known, it is natural to ask whether GI might in fact be NP-complete. We now explain why this is considered unlikely.

Theorem 7.1. *If* GI *is NP-complete, then* PH *collapses to level* 2.

Proof. Assume that GI is NP-complete. Then its complement GNI is coNP-complete. Since $\mathsf{GNI} \in \mathsf{AM}$, we obtain

$$\text{coNP} \subseteq \mathsf{AM}.$$

We now show that $\Sigma_2^p \subseteq \mathsf{AM}$. Consider the Σ_2^p-complete problem

$$\Sigma_2\text{-SAT} = \{\varphi : \exists x \; \forall y \; \varphi(x, y) = 1\}.$$

Fix a value of the existential witness x, and define

$$S_x := \{\varphi : \forall y \; \varphi(x, y) = 1\}.$$

For each fixed x, the language S_x is in coNP. Since coNP $\subseteq$ AM, there is an AM protocol for S_x.

Now form a protocol for Σ_2-SAT:

1. Merlin first sends the existential witness x.
2. Arthur and Merlin then execute the AM protocol for S_x.

This is a MAM protocol for Σ_2-SAT. Since $\mathsf{MAM} = \mathsf{AM}$, it follows that

$$\Sigma_2\text{-SAT} \in \mathsf{AM}.$$

Hence

$$\Sigma_2^p \subseteq \mathsf{AM} \subseteq \Pi_2^p.$$

Taking complements gives

$$\Pi_2^p = \mathrm{co}\Sigma_2^p \subseteq \mathrm{co}\Pi_2^p = \Sigma_2^p.$$

Therefore

$$\Sigma_2^p = \Pi_2^p.$$

By the standard collapse theorem for the polynomial hierarchy, if $\Sigma_i^p = \Pi_i^p$ for some i, then PH collapses to level i. Hence PH collapses to level 2. □

Remark 7.2. *This is strong evidence that graph isomorphism is unlikely to be NP-complete. Since* $\mathsf{GNI} \in \mathsf{AM}$*, NP-completeness of* GI *would force an unexpected collapse of* PH.