Lecture 23

Limitations of IPs

Foundations of Probabilistic Proofs
Alessandro Chiesa

IPs with Bounded Communication

Let $\mathsf{IP}[pc=1]$ = "languages decidable via IPs where prover sends 1 bit".

Q: Is $\mathsf{IP}[pc=1]$ trivial (contained in BPP)?

A: Unlikely, because $GNI \in \mathsf{IP}[pc=1]$ and GNI is not known to be in BPP.

→ Even IPs with small communication can decide non-trivial languages.

Q: Can we hope for $SAT \in \mathsf{IP}[pc = o(n)]$? (pc is sublinear in number of variables)

Note that $SAT \in NP \subseteq IP$ so we are asking if there is an IP for SAT that provides an improvement in communication over the trivial IP.
(send the candidate assignment)

Today we study IPs with BOUNDED COMMUNICATION:

- $\mathsf{IP}[pc, vc, vr]$ = "languages decidable via IPs where { prover sends pc bits; verifier sends vc bits; verifier uses vr random bits }"
- $\mathsf{AM}[pc, vc, vr]$ = "same as above but via public-coin IPs"

Warmup: Short Proof → Easy Language [1/2]

Define $NP[pc]$ = "NP languages where the proof string (aka witness) is pc bits"

lemma: $NP[pc] \subseteq DTIME(2^{O(pc)} \cdot poly(n))$

proof: Try every possible proof string.

$A(x) :=$ 1. For every NP proof $\tilde{\pi} \in \{0,1\}^{pc}$: if $V_{NP}(x,\tilde{\pi})=1$ then output 1.
2. Output 0.

■

Now we consider proof strings checked with randomness.

Define $MA[\varepsilon_c, \varepsilon_s, pc, vr]$ = " languages decidable via pc-bit proof strings with completeness error ε_c and soundness error ε_s, using vr bits of randomness"

lemma: ① $MA[\varepsilon_c, \varepsilon_s, pc, vr] \subseteq DTIME(2^{O(pc+vr)} \cdot poly(n))$

② $MA[\varepsilon_c, \varepsilon_s, pc, vr] \subseteq BPTIME\left(2^{O(pc)} \cdot poly\left(\frac{1}{1-\varepsilon_c-\varepsilon_s}, n\right)\right)$

Proof of ①: try all possible proof strings and randomness strings

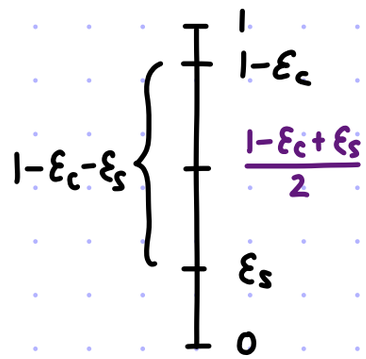
Proof of ②: we need a new idea because we CANNOT afford trying all randomness strings

Warmup: Short Proof → Easy Language [2/2]

$$MA[\varepsilon_c, \varepsilon_s, pc, vr] \subseteq \text{BPTIME}\left(2^{O(pc)} \cdot \text{poly}\left(\frac{1}{1-\varepsilon_c-\varepsilon_s}, n\right)\right)$$

Idea: APPROXIMATE the acceptance probability for every possible MA proof.

$A(x) :=$
1. Sample $\rho_1, \ldots, \rho_t \in \{0,1\}^{vr}$.
2. For every MA proof $\tilde{\pi} \in \{0,1\}^{pc}$:
   - Compute $N(\tilde{\pi}) := \left|\{i \in [t] \mid V_{MA}(x, \tilde{\pi}; \rho_i) = 1\}\right|$.
   - If $N(\tilde{\pi})/t > \frac{1-\varepsilon_c+\varepsilon_s}{2}$ then output 1.
3. Output 0.



For $\tilde{\pi}$ and ρ, $Z(\tilde{\pi}, \rho) :=$ "indicator that $V_{MA}(x, \tilde{\pi}; \rho) = 1$".

$Z(\tilde{\pi}, \rho_1), \ldots, Z(\tilde{\pi}, \rho_t)$ are i.i.d. samples from the Bernoulli distribution with bias $\delta(\tilde{\pi}) := \Pr_\rho[V_{MA}(x, \tilde{\pi}; \rho) = 1]$.

By an additive Chernoff bound, $\Pr_{\rho_1, \ldots, \rho_t}\left[\left|\frac{1}{t}\sum_{i=1}^{t} Z(\tilde{\pi}, \rho_i) - \delta(\tilde{\pi})\right| > \alpha\right] \leq \exp(-t \cdot \alpha^2)$.

$$\begin{cases} x \in L \to \exists \pi \text{ s.t. } \delta(\pi) \geq 1-\varepsilon_c \\ x \notin L \to \forall \tilde{\pi} \;\; \delta(\tilde{\pi}) \leq \varepsilon_s \end{cases} \to \text{We need } \alpha < \frac{1}{2} \cdot ((1-\varepsilon_c)-\varepsilon_s) \text{ to distinguish between these.}$$

If we set $t := O\left(\frac{1}{\alpha^2} \cdot pc\right)$ then each estimation is within $\pm\alpha$ except w.p. $\exp(-pc)$.

By a union bound across the 2^{pc} estimations, A has constant two-sided error. ■

The Case of Interactive Proofs

A similar statement holds for any IP:

Theorem: ① $IP[\varepsilon_c, \varepsilon_s, pc, vc, vr] \subseteq DTIME(2^{O(pc+vc+vr)} \cdot poly(n))$

② $IP[\varepsilon_c, \varepsilon_s, pc, vc, vr] \subseteq BPTIME\left(2^{O(pc+vc)} \cdot poly\left(\frac{1}{1-\varepsilon_c-\varepsilon_s}, n\right)\right)$

①: if we bound (two-way) communication and randomness
then we can decide the language in deterministic exponential time

②: if we bound (two-way) communication only
then we can decide the language in probabilistic exponential time

→ relation between communication complexity of IP
and the time complexity of the language it decides

Example for 3SAT: it is unlikely that $3SAT \in IP[pc = o(n), vc = o(n)]$
It would imply that $3SAT \in BPTIME(2^{o(n)})$, contradicting RETH.

Randomized Exponential-Time Hypothesis: $\exists c > 0 : 3SAT \notin BPTIME(2^{c \cdot n})$

Q: What about one-way communication? We discuss this later.

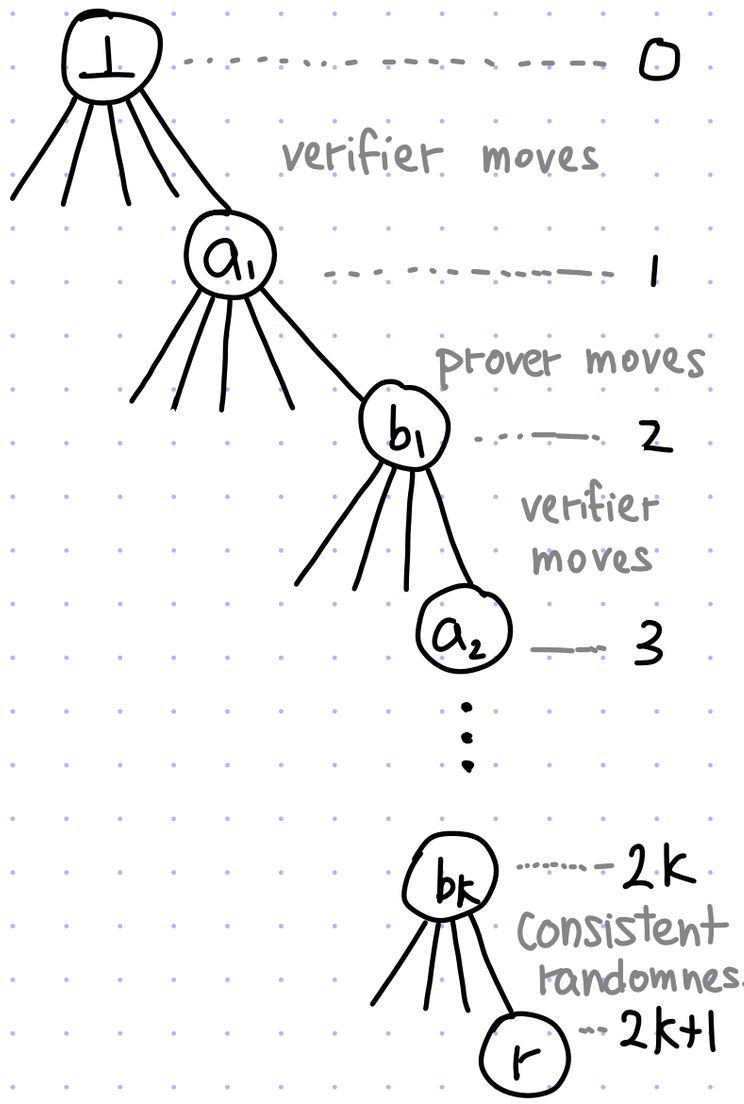
Game Tree for an IP

A transcript (of interaction) is a tuple $(a_1, b_1, \ldots, a_k, b_k)$.

An augmented transcript is $(a_1, b_1, \ldots, a_k, b_k, r)$ where r is verifier randomness.

Fix an IP verifier V and instance x.

The game tree $T = T(V,x)$ of $V(x)$ is the tree of all possible augmented transcripts →



for $i = 0, 1, \ldots, k-1$:

- verifier moves at level $2i$
- prover moves at level $2i+1$

Edges from $2i$ to $2i+1$ are possible moves by verifier.

Edges from $2i+1$ to $2(i+1)$ are possible moves by prover.

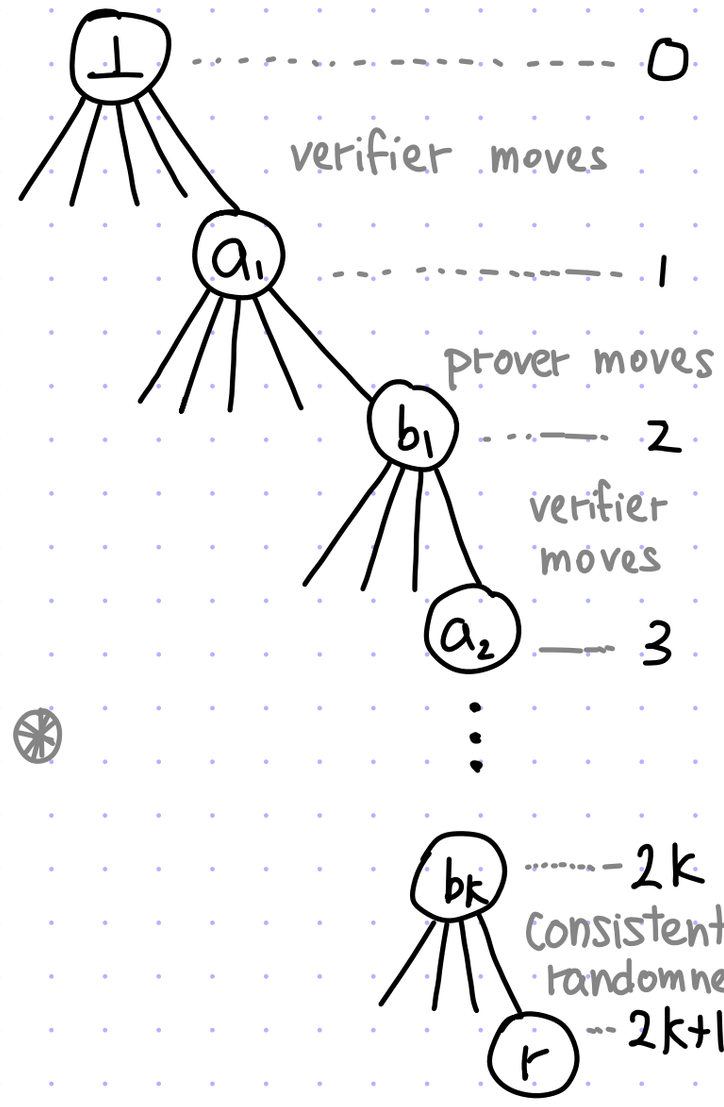
Edges from $2k$ to $2k+1$ are possible random strings consistent with transcript.

Approximating the Value Suffices

The value of the tree is $\mathrm{val}(T) := \mathrm{val}(\text{root})$.

The value of a node is recursively defined:

- $\mathrm{val}\big(\text{leaf node } tr=(a_1,b_1,\dots,a_k,b_k,r)\big) := V(x,b_1,\dots,b_k;r)$.
- $\mathrm{val}\big(\text{internal node } tr=(a_1,b_1,\dots,a_i,b_i) \text{ at level } 2i\big)$
  $:= \mathop{\mathbb{E}}_{a_{i+1}} \mathrm{val}\big(\text{child node } (a_1,b_1,\dots,a_i,b_i,a_{i+1}) \text{ at level } 2i+1\big)$
  a_{i+1} ← sampled according to this verifier message's probability
- $\mathrm{val}\big(\text{internal node } tr=(a_1,b_1,\dots,a_i,b_i,a_{i+1}) \text{ at level } 2i+1\big)$
  $:= \max_{b_{i+1}} \mathrm{val}\big(\text{child node } (a_1,b_1,\dots,a_{i+1},b_{i+1}) \text{ at level } 2i+2\big)$.



✳ This includes the special layer 2k, where the randomness r can be viewed as a fictitious final verifier message.

Observe: $x \in L \to \mathrm{val}(T) \geq 2/3$
$x \notin L \to \mathrm{val}(T) \leq 1/3$

} to decide $x \stackrel{?}{\in} L$ it suffices to approximate $\mathrm{val}(T)$ to within $\pm 1/6$

We showed that $\mathrm{val}(T)$ is computable in space $\mathrm{poly}(n)$ (and thus time $\exp(\mathrm{poly}(n))$).

TODAY: We analyze the time complexity to approximate $\mathrm{val}(T)$.

Bounded Randomness & Two-Way Communication

Theorem: $IP[\varepsilon_c, \varepsilon_s, pc, vc, vr] \subseteq DTIME(2^{O(pc+vc+vr)} \cdot poly(n))$

Let $c = pc + vc + vr$ be a bound on communication AND randomness.

The number of nodes in the tree $T(V,x)$ is $2^{O(c)}$ because:

- the number of possible transcripts is $\leq 2^{pc+vc}$,
- each transcript has $\leq 2^{vr}$ possible augmentations.

Hence, we can compute $val(T(V,x))$ exactly in time $2^{O(c)} \cdot poly(n)$, by writing down the tree $T(V,x)$ and following the recursive computation.

Q: How to compute the probabilities of verifier messages?

Associate to each verifier node the set of random strings consistent with transcript so far.

Iterating over this set partitions it by the next verifier message.

Note that no partitioning occurs at prover nodes, so the same randomness s may appear in multiple leaves.

NOTE: can set $c = pc + vr$ because the number of augmented transcripts is $\leq 2^{pc+vr}$

Bounded Two-Way Communication

Theorem: $IP[\varepsilon_c, \varepsilon_s, pc, vc, vr] \subseteq BPTIME\left(2^{O(pc+vc)} \cdot poly\left(\frac{1}{1-\varepsilon_c-\varepsilon_s}, n\right)\right)$

Let $c := pc + vc$ be a bound on communication ONLY.

There are $\leq 2^c$ possible transcripts. (Hence $\leq 2^{O(c)}$ internal nodes.)

PROBLEM: each transcript may have $2^{poly(n)}$ augmentations,
so we cannot construct the tree T in time $2^{O(c)} \cdot poly\left(\frac{1}{1-\varepsilon_c-\varepsilon_s}, n\right)$
nor compute the probabilities of verifier messages in T.

IDEA: use randomness to APPROXIMATE $val(T)$ in time $2^{O(c)} \cdot poly\left(\frac{1}{1-\varepsilon_c-\varepsilon_s}, n\right)$
with bounded probability of error.

Let (P,V) be an IP for $L \in IP[\varepsilon_c, \varepsilon_s, pc, vc, vr]$.

We design a $2^{O(c)} \cdot poly\left(\frac{1}{\varepsilon}, n\right)$-time probabilistic algorithm A s.t.

$$Pr\left[\,|A(x) - val(T(V,x))| > \varepsilon\,\right] \leq \frac{1}{100}.$$

Setting $\varepsilon := \frac{1-\varepsilon_c-\varepsilon_s}{2}$, this suffices because: $\begin{cases} x \in L \rightarrow val(T(V,x)) \geq 1-\varepsilon_c \\ x \notin L \rightarrow val(T(V,x)) \leq \varepsilon_s \end{cases}$

Proof [1/4]

$A(x)$:

1. Set $t := \Theta\left(\frac{2^c \cdot c}{\varepsilon^2}\right)$.
2. Sample $g_1, \dots, g_t$ independently at random in $\{0,1\}^{vr}$.
3. Construct $T(V,x)[R]$, the residual game tree obtained by omitting nodes in $T(V,x)$ inconsistent with $R = \{g_1, \dots, g_t\}$ (and adjusting weights).
4. Compute and output $\mathrm{val}(T(V,x)[R])$.

The algorithm runs in time $2^{O(c)} \cdot \mathrm{poly}(\frac{1}{\varepsilon}, n)$ because:

- $T(V,x)[R]$ has size $2^{O(c)} \cdot |R| = 2^{O(c)} \cdot t = 2^{O(c)} \cdot \frac{1}{\varepsilon^2}$
- the running time is $\mathrm{poly}(|T(V,x)[R]|) \cdot \mathrm{poly}(n)$.

We are left to argue CORRECTNESS.

We prove that:

<u>lemma</u>: $\Pr_R\left[\, |\mathrm{val}(T(V,x)[R]) - \mathrm{val}(T(V,x))| > \varepsilon \,\right] \leq \frac{1}{100}$

Henceforth we write $T := T(V,x)$.

Proof [2/4]

<u>lemma</u>: $\Pr_R\left[\,|\mathrm{val}(T[R]) - \mathrm{val}(T)| > \varepsilon\,\right] \leq \frac{1}{100}$

Define V_R to be the verifier V restricted to sample randomness in R rather than $\{0,1\}^{vr}$.

Obsere that $\mathrm{val}(T[R]) = \max_{\tilde{P}} \Pr\left[\langle \tilde{P}, V_R(x)\rangle = 1\right]$.

Fix a prover strategy $\tilde{P}$ and define:

$$\Delta(\tilde{P}, R) := \Pr_{s\leftarrow R}\left[\langle \tilde{P}, V(x;s)\rangle = 1\right] - \Pr_{s\leftarrow\{0,1\}^{vr}}\left[\langle \tilde{P}, V(x;s)\rangle = 1\right]$$

$$= \underbrace{\Pr\left[\langle \tilde{P}, V_R(x)\rangle = 1\right]}_{\text{depends on R}} - \underbrace{\Pr\left[\langle \tilde{P}, V(x)\rangle = 1\right]}_{\text{independent of R}}$$

<u>claim</u>: $\Pr_R\left[\,\exists\,\tilde{P} : |\Delta(\tilde{P}, R)| > \varepsilon\,\right] \leq \frac{1}{100}$

This implies the lemma: $\Pr_R\left[\,|\mathrm{val}(T[R]) - \mathrm{val}(T)| > \varepsilon\,\right] \leq \Pr_R\left[\,\exists\,\tilde{P} : |\Delta(\tilde{P}, R)| > \varepsilon\,\right] \leq \frac{1}{100}$.

Indeed, for every choice of R, the event on the left implies the event on the right:

- $\mathrm{val}(T[R]) > \mathrm{val}(T) + \varepsilon \rightarrow \Pr\left[\langle P_R^*, V_R(x)\rangle = 1\right] > \Pr\left[\langle P^*, V(x) = 1\right] + \varepsilon \geq \Pr\left[\langle P_R^*, V(x)\rangle = 1\right] + \varepsilon$
- $\mathrm{val}(T) > \mathrm{val}(T[R]) + \varepsilon \rightarrow \Pr\left[\langle P^*, V(x) = 1\right] > \Pr\left[\langle P_R^*, V_R(x)\rangle = 1\right] + \varepsilon \geq \Pr\left[\langle P^*, V_R(x) = 1\right] + \varepsilon$

Proof [3/4]

We are left to prove claim: $\Pr_R\left[\exists\, \tilde{P} : |\Delta(\tilde{P},R)| > \varepsilon\right] \leq \frac{1}{100}$

① We use a concentration argument to show that $\Delta(\tilde{P},R)$ is small w.h.p. over the choice of R:

$$\forall\, \tilde{P},\ \Pr_R\left[|\Delta(\tilde{P},R)| > \varepsilon\right] \leq 2\cdot e^{-2\cdot\varepsilon^2\cdot t}$$

Define $z_i := \langle \tilde{P}, V(x;\rho_i)\rangle$ where ρ_i is the i-th random string in R.

The random variables $z_1,\dots,z_t$ are i.i.d. because $\rho_1,\dots,\rho_t$ are i.i.d..

Observe that:
- $\mathbb{E}[z_i] = \Pr[\langle \tilde{P}, V(x)\rangle = 1]$ because each ρ_i is random in $\{0,1\}^{vr}$
- $\frac{z_1+\cdots+z_t}{t} = \Pr[\langle \tilde{P}, V_R(x)\rangle = 1]$

Hence:

$$\Pr_R\left[|\Delta(\tilde{P},R)| > \varepsilon\right]$$

$$= \Pr_R\left[\left|\Pr[\langle \tilde{P}, V_R(x)\rangle = 1] - \Pr[\langle \tilde{P}, V(x)\rangle = 1]\right| > \varepsilon\right]$$

$$= \Pr\left[\left|\frac{z_1+\cdots+z_t}{t} - \mathbb{E}[z_1]\right| > \varepsilon\right] \leq 2\cdot e^{-2\cdot\varepsilon^2\cdot t}$$

↑ additive Chernoff bound (for $z_1,\dots,z_t$ i.i.d. in $[0,1]$)

Proof [4/4]

We are left to prove claim: $\Pr_R\left[\exists \tilde{P} : |\Delta(\tilde{P},R)| > \varepsilon\right] \leq \frac{1}{100}$

① We use a concentration argument to show that $\Delta(\tilde{P},R)$ is small w.h.p. over the choice of R:

$$\forall \tilde{P},\ \Pr_R\left[|\Delta(\tilde{P},R)| > \varepsilon\right] \leq 2\cdot e^{-2\cdot\varepsilon^2\cdot t} \quad \checkmark$$

② We use a union bound to conclude the claim's proof.

Any prover $\tilde{P}$ is a function from transcript so far to next message.
There are at most $(2^c)^{2^c} = 2^{c\cdot 2^c}$ provers (because input and output are at most c bits).

By a union bound on all such provers, and taking $t = \Theta\left(\frac{c\cdot 2^c}{\varepsilon^2}\right)$ large enough,

$$\Pr_R\left[\exists \tilde{P} : |\Delta(\tilde{P},R)| > \varepsilon\right] \leq \sum_{\tilde{P}} \Pr_R\left[|\Delta(\tilde{P},R)| > \varepsilon\right] \leq 2^{c\cdot 2^c}\cdot 2\cdot e^{-2\cdot\varepsilon^2\cdot t} \leq \frac{1}{100}.$$

Bounded One-Way Communication

Handling the case where we bound ONLY prover communication is HARDER.

- With perfect completeness, we can non-deterministically decide the complement.

  theorem: $IP[\varepsilon_c = 0, pc] \subseteq coNTIME(2^{O(pc)} \cdot poly(n))$

- Without perfect completeness, it is more complicated.

  theorem: $AM[k, pc] \subseteq BPTIME(2^{O(pc + k \cdot \log k)} \cdot poly(n))$

We do not prove these.

theorem: $IP[pc] \subseteq BPTIME(2^{O(pc \cdot \log pc)} \cdot poly(n))^{NP}$

theorem: $IP[k, pc] \subseteq coAM(rounds = O(k), pc' = 2^{pc} \cdot poly(k^k, n))$

Example for GNI:

prover sends pre-image $(H, \pi) \in \{0,1\}^{n^2 + n \cdot \log n}$ & isomorphism $\phi : [n] \to [n]$

We know that $GNI \in IP[pc = 1]$ and $GNI \in AM[pc = O(n^2)]$.

But we should NOT expect that $GNI \in AM[pc = o(\frac{\log n}{\log\log n})]$ unless $GNI \in P$.