The Complexity of Hardness Amplification and Derandomization

Emanuele Viola
Harvard University
Ph.D. Defense

May 2006
Randomness in Computation

• Useful throughout Computer Science
  – Cryptography
  – Learning Theory
  – Complexity Theory

• Question: Is Randomness necessary?
Derandomization

• **Goal**: remove randomness

• Why study derandomization?

• **Breakthrough** [R ‘04]:
  Connectivity in logarithmic space (SL = L)

• **Breakthrough** [AKS ‘02]:
  Primality in polynomial time (PRIMES ∈ P)
Randomness vs. Time

• Goal:
simulate randomized computation deterministically

• Trivial Derandomization:
If A uses n random bits, enumerate all $2^n$ possibilities

Probabilistic polynomial-time $\subseteq$ exponential time
\[ \text{BPP} \subseteq \text{Time}(2^{\text{poly}(n)}) \]

• Strong Belief: BPP = P ( Time(poly(n)) )
Complexity Assumptions $\Rightarrow$ BPP = P [BFNW,NW,IW,...]
Outline

• Overview of derandomization

• Derandomization of restricted models
  – Application: Hardness Amplification in NP
  – New derandomization

• Derandomization of general models
  – BPP vs. PH
  – Proof of Lower Bound
Constant-Depth Circuits

- Probabilistic constant-depth circuit (BP AC$^0$)

- Theorem [N ‘91]: BP AC$^0$ $\subseteq$ Time($n^{\text{polylog } n}$)
  - Compare to BP P $\subseteq$ Time($2^{\text{poly}(n)}$)
Application: Avg-Case Hardness of NP

- Study hardness of NP on random instances
  - Natural question, essential for cryptography

- Currently cannot relate to $P \neq NP$ [FF,BT,V]

- Hardness amplification
  Definition: $f : \{0,1\}^n \rightarrow \{0,1\}$ is $\delta$-hard if for every efficient algorithm $M : \Pr_x[M(x) \neq f(x)] \geq \delta$

\[ f \rightarrow \text{Hardness Amplification} \rightarrow f' \]
\[ .01\text{-hard} \rightarrow (1/2 - \varepsilon)\text{-hard} \]
Previous Results

• Yao’s XOR Lemma: \( f'(x_1, \ldots, x_n) := f(x_1) \oplus \cdots \oplus f(x_n) \)
  \( f' \approx (1/2 - 2^{-n})\)-hard, almost optimal

• Cannot use XOR in NP: \( f \in \text{NP} \not\Rightarrow f' \in \text{NP} \)

• Idea: \( f'(x_1, \ldots, x_n) = C( f(x_1), \ldots, f(x_n) ) \), \( C \) monotone
  – e.g. \( f(x_1) \land ( f(x_2) \lor f(x_3) ) \). \( f \in \text{NP} \Rightarrow f' \in \text{NP} \)

• Theorem [O’D]: There is \( C \) s.t. \( f' \approx (1/2 - 1/n)\)-hard

• Barrier: No monotone \( C \) can do better!
Our Result on Hardness Amplification

• **Theorem [HVV]**: Amplification in NP up to $\approx 1/2 - 2^{-n}$
  – Matches the XOR Lemma

• **Technique**: Derandomize!
  Intuitively, $f' := C(f(x_1), \ldots, f(x_n), \ldots \ldots f(x_{2^n}))$
  $f'$ $(1/2 - 1/2^n)$-hard by previous result

**Problem**: Input length $= 2^n$

Note $C$ is constant-depth

Derandomize: input length $\rightarrow n$, keep hardness
Outline

• Overview of derandomization

• Derandomization of restricted models
  – Application: Hardness Amplification in NP
  – New derandomization

• Derandomization of general models
  – BPP vs. PH
  – Proof of Lower Bound
Previous Results

• Recall Theorem [N]:
  \( \text{BP AC}^0 \subseteq \text{Time}(n^{\text{polylog } n}) \)

• But \( \text{AC}^0 \) is weak: \( \text{Majority} \not\in \text{AC}^0 \)
  – \( \text{Majority}(x_1, \ldots, x_n) := \sum_i x_i > n/2 \) ?

• Theorem [LVW]:
  \( \text{BP Maj AND} \subseteq \text{Time}(2^{n^\varepsilon}) \)

• Derandomize incomparable classes
Our New Derandomization

- **Theorem** $[\mathcal{V}]$ : $\text{BP Maj AC}^0 \subseteq \text{Time}(2^{n^\epsilon})$

  Derandomize constant-depth circuits with few Majority gates =

- Improves on $[\text{LVW}]$. Slower than $[\text{N}]$ but richer richest probabilistic circuit class in $\text{Time}(2^{n^\epsilon})$

- **Techniques**: Communication complexity + switching lemma $[\text{BNS, HG, H, HM, CH}]$
Outline

• Overview of derandomization

• Derandomization of restricted models
  – Application: Hardness Amplification in NP
  – New derandomization

• Derandomization of general models
  – BPP vs. PH
  – Proof of Lower Bound
BPP vs. POLY-TIME HIERARCHY

- Probabilistic Polynomial Time (BPP):
  for every $x$, $\Pr[M(x) \text{ errs}] \leq 1\%$

- Strong belief: $\text{BPP } = \text{ P }$ [NW,BFNW,IW,…]
  Still open: $\text{BPP } \subseteq \text{ NP }$?

- Theorem [SG,L; ‘83]: $\text{BPP } \subseteq \Sigma_2 \text{ P}$

- Recall
  \[
  \begin{align*}
  \text{NP } &= \Sigma_1 \text{ P } \quad \rightarrow \quad \exists \ y \ M(x,y) \\
  \Sigma_2 \text{ P } \quad &\rightarrow \quad \exists \ y \ \forall \ z \ M(x,y,z)
  \end{align*}
  \]
The Problem we Study

• More precisely [SG,L] give
  \[ \text{BPTime}(t) \subseteq \Sigma_2\text{Time}(t^2) \]

• Question[Rest of this Talk]:
  Is quadratic slow-down necessary?

• Motivation: Lower bounds
  Know NTime \neq \text{Time} on some models [P+, F+, …]
  Technique: speed-up computation with quantifiers
  To prove NTime \neq \text{BPTime} cannot afford \Sigma_2\text{Time}(t^2)
Approximate Majority

• Input: \( R = 101111011011101011 \)

• Task: Tell \( \Pr_i[ R_i = 1] \geq 99\% \) from \( \Pr_i[ R_i = 1] \leq 1\% \)

Do not care if \( \Pr_i[ R_i = 1] \sim 50\% \) (approximate)

• Model: Depth-3 circuit

![Depth-3 Circuit Diagram]
The connection [FSS]

\[ M(x;u) \in \text{BPTime}(t) \]

Compute \( M(x) \):
- Tell \( \Pr_u[M(x) = 1] \geq 99\% \)
- from \( \Pr_u[M(x) = 1] \leq 1\% \)

\[ \text{BPTime}(t) \subseteq \sum_2 \text{Time}(t') \]
\[ = \exists \forall \text{Time}(t') \]

Running time \( t' \)
- run \( M \) at most \( t'/t \) times

\[ R = 11011011101011 \]
\[ |R| = 2^t \]
\[ R_i = M(x;i) \]

Compute Appr-Maj

Bottom fan-in \( f = t'/t \)
Our Results

- **Theorem**[V]: Small depth-3 circuits for Approximate Majority on N bits have bottom fan-in $\Omega(\log N)$

- **Corollary**: Quadratic slow-down necessary for relativizing techniques:
  
  \[
  \text{BPTime}^A(t) \subseteq \Sigma_2 \text{Time}^A(t^{1.99})
  \]

- **Theorem**[DvM,V]: BPTime $^A(t) \subseteq \Sigma_3 \text{Time}^A(t \cdot \log^5 t)$
  - Previous result [A]: BPTime $^A(t) \subseteq \Sigma_{O(1)} \text{Time}^A(t)$

- For time, the level is the third!
Outline

• Overview of derandomization

• Derandomization of restricted models
  – Application: Hardness Amplification in NP
  – New derandomization

• Derandomization of general models
  – BPP vs. PH
  – Proof of Lower Bound
Our Negative Result

- **Theorem [V]:** $2^{N^\epsilon}$-size depth-3 circuits for Approximate Majority on $N$ bits have bottom fan-in $f = \Omega(\log N)$

- Recall:

  $$R = 101111011011101011 \quad |R| = N$$

  Tells $R \in \text{YES} := \{ R : \Pr_i[R_i = 1] \geq 99\% \}$

  from $R \in \text{NO} := \{ R : \Pr_i[R_i = 1] \leq 1\% \}$
Proof

• Circuit is OR of s depth-2 circuits

• By definition of OR:
  \[ R \in \text{YES} \Rightarrow \text{some } C_i (R) = 1 \]
  \[ R \in \text{NO} \Rightarrow \text{all } C_i (R) = 0 \]

• By averaging, fix \( C = C_i \) s.t.
  \[
  \Pr_{R \in \text{YES}} [C(x) = 1] \geq \frac{1}{s} \\
  \forall R \in \text{NO} \Rightarrow C(R) = 0
  \]

• **Claim**: Impossible if \( C \) has bottom fan-in \( \leq \varepsilon \log N \)
CNF Claim

- **Depth-2 circuit** ⇒ **CNF**

\[(x_1 \lor x_2 \lor \neg x_3) \land (\neg x_4) \land (x_5 \lor x_3)\]

**bottom fan-in** ⇒ **clause size**

- **Claim**: All CNF C with clauses of size $\varepsilon \cdot \log N$

  Either $\Pr_{R \in \text{YES}} [C(x) = 1] \leq 1 / 2^{N\varepsilon}$

  or there is $R \in \text{NO} : C(x) = 1$

- **Note**: Claim ⇒ Theorem
Either $\Pr_{R \in \text{YES}} [C(x) = 1] \leq 1/2^{N^\varepsilon}$ or $\exists R \in \text{NO} : C(x) = 1$

**Proof Outline**

- **Definition:** $S \subseteq \{x_1, x_2, \ldots, x_N\}$ is a covering if every clause has a variable in $S$

  E.g.: $S = \{x_3, x_4\}$ $C = (x_1 \lor x_2 \lor \neg x_3) \land (\neg x_4) \land (x_5 \lor x_3)$

- **Proof idea:** Consider **smallest** covering $S$

  \[
  \text{Case } |S| \text{ BIG } : \Pr_{R \in \text{YES}} [C(x) = 1] \leq 1/2^{N^\varepsilon}
  \]

  \[
  \text{Case } |S| \text{ tiny } : \text{Fix few variables and repeat}
  \]
Either $\Pr_{R \in \text{YES}} [C(x) = 1] \leq 1/2^{N^{\varepsilon}}$ or $\exists R \in \text{NO} : C(x) = 1$

**Case $|S|$ BIG**

- $|S| \geq N^{\delta} \Rightarrow$ have $N^{\delta} / (\varepsilon \cdot \log N)$ disjoint clauses $\Gamma_i$
  - Can find $\Gamma_i$ greedily

- $\Pr_{R \in \text{YES}} [C(R) = 1] \leq \Pr \left[ \forall i, \Gamma_i(R) = 1 \right]$
  
  \[
  = \prod_i \Pr[ \Gamma_i(R) = 1] \quad \text{(independence)}
  \]

  \[
  \leq \prod_i \left( 1 - 1/100^{\varepsilon \log N} \right) = \prod_i \left( 1 - 1/N^{O(\varepsilon)} \right)
  \]

  \[
  = \left( 1 - 1/N^{O(\varepsilon)} \right)^{|S|} \leq e^{-N^{\Omega(1)}} \quad \checkmark
  \]
Either \( \Pr_R \in \text{YES} [C(x)=1] \leq 1/2^{N^\varepsilon} \) or \( \exists R \in \text{NO} : C(x) = 1 \)

**Case \(|S|\) tiny**

- \(|S| < N^\delta \) \(\implies\) Fix variables in \(S\)
  - Maximize \( \Pr_R \in \text{YES} [C(x)=1] \)

- Note: \(S\) covering \(\implies\) clauses shrink

Example

\[
(x_1 \vee x_2 \vee x_3) \land (\neg x_3) \land (x_5 \vee \neg x_4)
\]

\[
\begin{align*}
x_3 &\leftarrow 0 \\
x_4 &\leftarrow 1
\end{align*}
\]

\[
(x_1 \vee x_2) \land (x_5)
\]

- Repeat
  Consider smallest covering \(S'\), etc.
Either $\Pr_{R \in \text{YES}} [C(x)=1] \leq 1/2^{N^\varepsilon}$ or $\exists R \in \text{NO} : C(x) = 1$

Finish up

• Recall: Repeat $\Rightarrow$ shrink clauses
  So repeat at most $\varepsilon \cdot \log N$ times

• When you stop:
  Either smallest covering size $\geq N^\delta$
  Or $C = 1$
  Fixed $\leq (\varepsilon \cdot \log N) N^\delta \ll N$ vars.
  Set rest to 0 $\Rightarrow R \in \text{NO} : C(R) = 1$

Q.E.D.
Conclusion

• Derandomization: powerful technique

• Restricted models: Constant-depth circuits (AC⁰)
  – Derandomization of AC⁰  [N]
  – Application: Hardness Amplification in NP  [HHV]
  – Derandomization of AC⁰ with few Maj gates  [V]

• General models: BPP vs. PH
  – BPTime(t) ⊆ Σ₂Time(t²)  [SG,L]
  – BPTime(t) ⊆ Σ₂Time(t^{1.99}) (w.r.t. oracle)  [V]
    Lower Bound for Approximate Majority
Thank you!