

Derandomization: New Results and Applications

Emanuele Viola
Harvard University

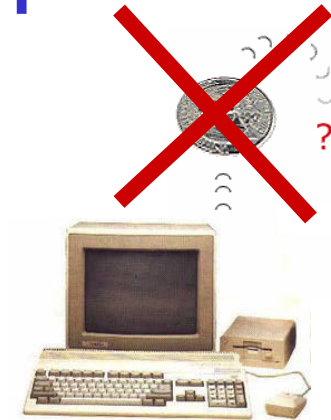
March 2006

Randomness in Computation



- Useful throughout Computer Science
 - Algorithms
 - 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 ($PRIMES \in P$)

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

$$\text{BPP} \subseteq \text{Time}(2^{\text{poly}(n)})$$

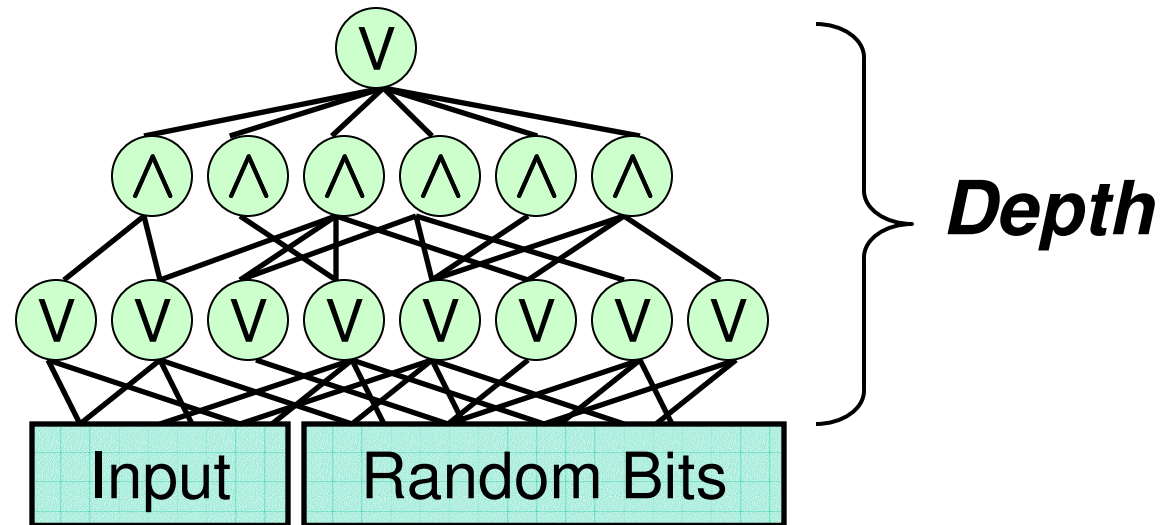
- Strong Belief: $\text{BPP} = \text{P} \quad (\text{Time}(\text{poly}(n)))$
Complexity Assumptions $\Rightarrow \text{BPP} = \text{P}$ [BFNW,NW,IW,...]

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 $BP P \subseteq Time(2^{\text{poly}(n)})$

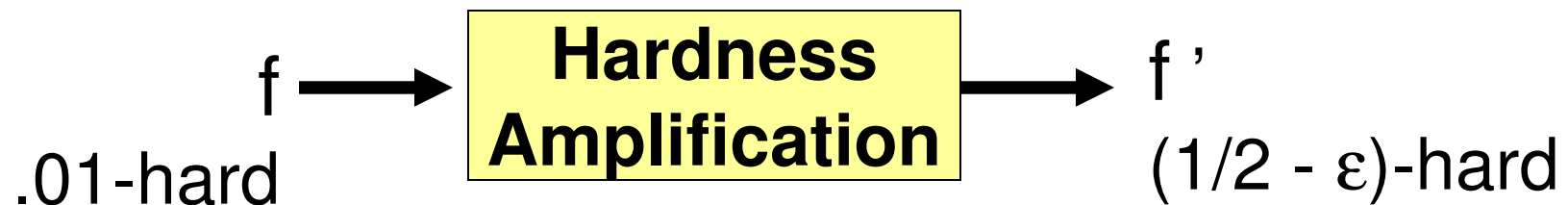
Application: Avg-Case Hardness of NP

- 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 δ -hard if

for every efficient algorithm $M : \Pr_x[M(x) \neq f(x)] \geq \delta$



Previous Results

- **Yao's XOR Lemma:** $f'(x_1, \dots, x_n) := f(x_1) \oplus \dots \oplus f(x_n)$
 $f' \approx (1/2 - 2^{-n})$ -hard, almost optimal
- **Cannot use XOR in NP:** $f \in \text{NP} \not\Rightarrow f' \in \text{NP}$
- **Idea:** $f'(x_1, \dots, x_n) = C(f(x_1), \dots, f(x_n))$, C **monotone**
– e.g. $f(x_1) \wedge (f(x_2) \vee f(x_3))$. $f \in \text{NP} \Rightarrow f' \in \text{NP}$
- **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** [HV \mathbf{V}]: Amplification in NP up to $\approx 1/2 - 2^{-n}$
 - Matches the XOR Lemma

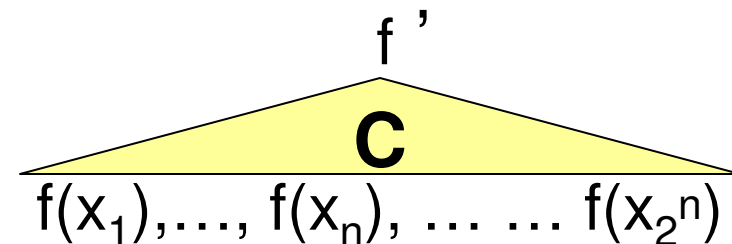
- **Technique**: Derandomize!

Intuitively, $f' := C(f(x_1), \dots, f(x_n), \dots \dots f(x_{2^n}))$

f' $(1/2 - 1/2^n)$ -hard by previous result

Problem: Input length = 2^n

Note C is constant-depth



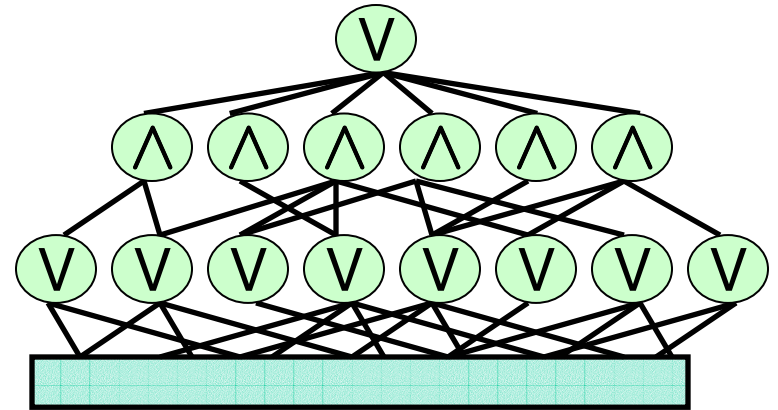
Derandomize: input length $\rightarrow n$, keep hardness

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

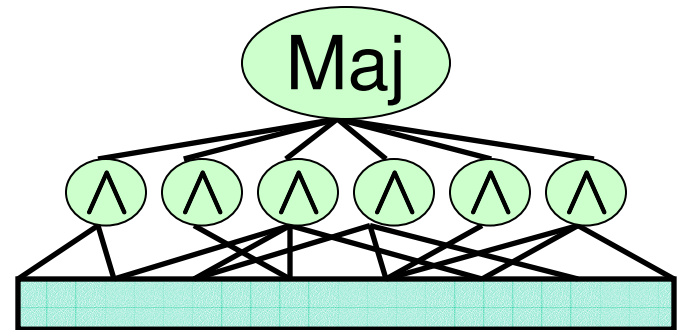
Previous Results

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



- But AC^0 is weak: Majority $\notin \text{AC}^0$
 - Majority(x_1, \dots, x_n) := $\sum_i x_i > n/2$?

- Theorem** [LVW]:
 $\text{BP Maj AND} \subseteq \text{Time}(2^{n^\epsilon})$

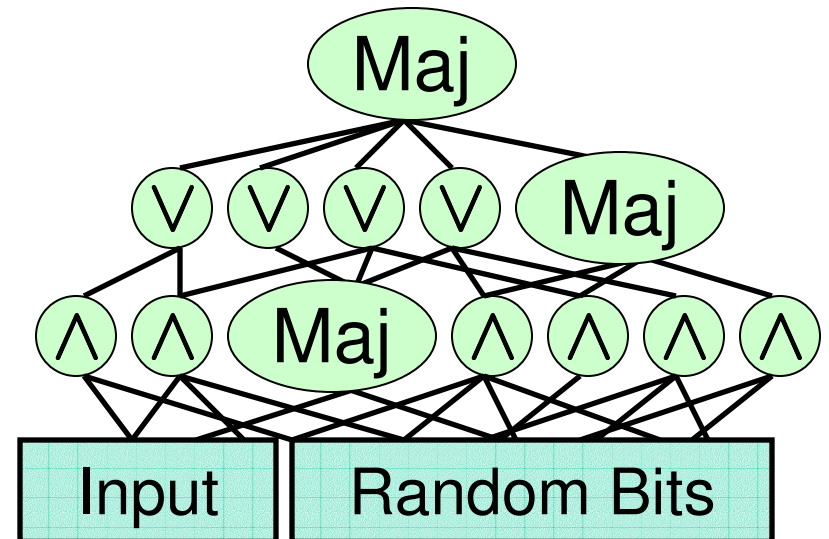


- Derandomize incomparable classes

Our New Derandomization

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

Derandomize
constant-depth circuits
with few Majority gates =



- Improves on [LVW]. Slower than [N] but richer
richest probabilistic circuit class in $\text{Time}(2^{n^\epsilon})$
- **Techniques**: Communication complexity +
switching lemma [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: $BPP = P$ [NW,BFNW,IW,...]
Still open: $BPP \subseteq NP$?
- **Theorem** [SG,L; '83]: $BPP \subseteq \Sigma_2 P$
- Recall
$$\begin{array}{l} NP = \Sigma_1 P \quad \rightarrow \quad \exists y M(x,y) \\ \Sigma_2 P \quad \quad \rightarrow \quad \exists y \forall z M(x,y,z) \end{array}$$

The Problem we Study

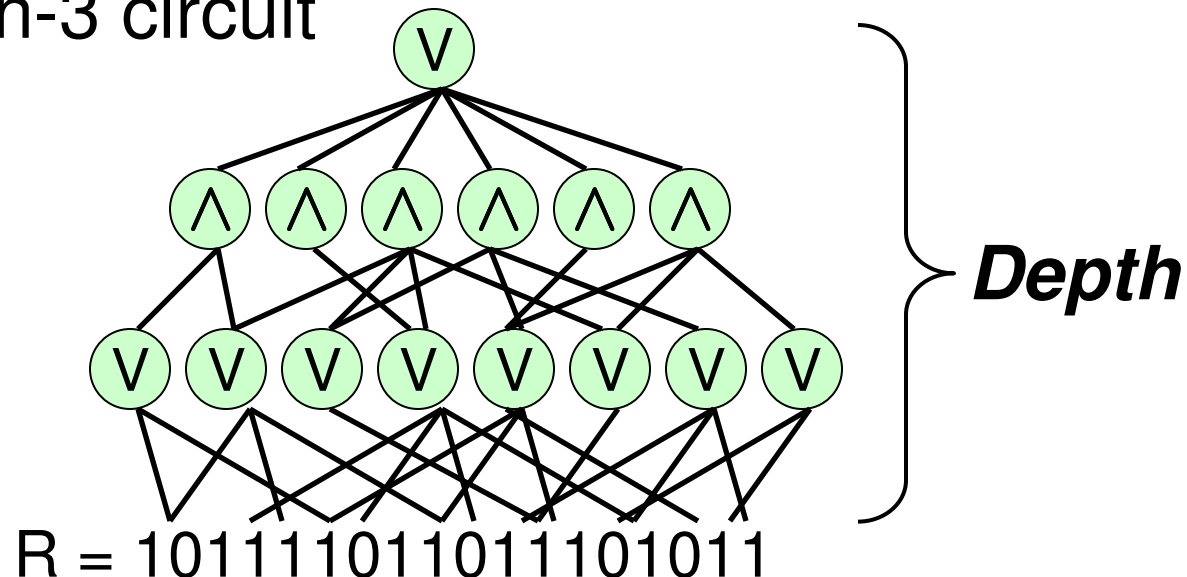
- More precisely [SG,L] give
$$\text{BPTime}(t) \subseteq \Sigma_2\text{Time}(t^2)$$
- Question[Rest of this Talk]:
Is **quadratic slow-down** necessary?
- Motivation: Lower bounds
Know $\text{NTime} \neq \text{Time}$ on some models [P+,F+,...]
Technique: *speed-up* computation with quantifiers
To prove $\text{NTime} \neq \text{BPTime}$ cannot afford $\Sigma_2\text{Time}(t^2)$

Approximate Majority

- Input: $R = 101111011011101011$
- Task: Tell $\Pr_i [R_i = 1] \geq 99\%$ from $\Pr_i [R_i = 1] \leq 1\%$

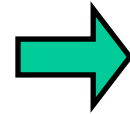
Do not care if $\Pr_i [R_i = 1] \sim 50\%$ (approximate)

- Model: Depth-3 circuit



The connection [FSS]

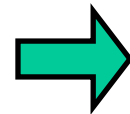
$M(x;u) \in \text{BPTime}(t)$



$R = 11011011101011$
 $|R| = 2^t \quad \curvearrowright \quad R_i = M(x;i)$

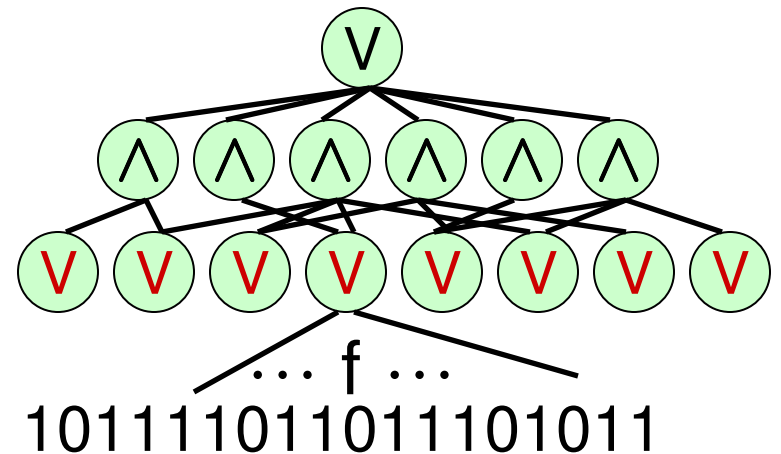
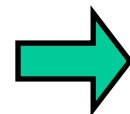
Compute $M(x)$:

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



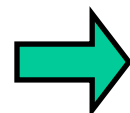
Compute Appr-Maj

$\text{BPTime}(t) \subseteq \Sigma_2 \text{Time}(t')$
 $= \exists \forall \text{Time}(t')$



Running time t'

– run M at most t'/t times



Bottom fan-in $f = t' / t$

Our Results

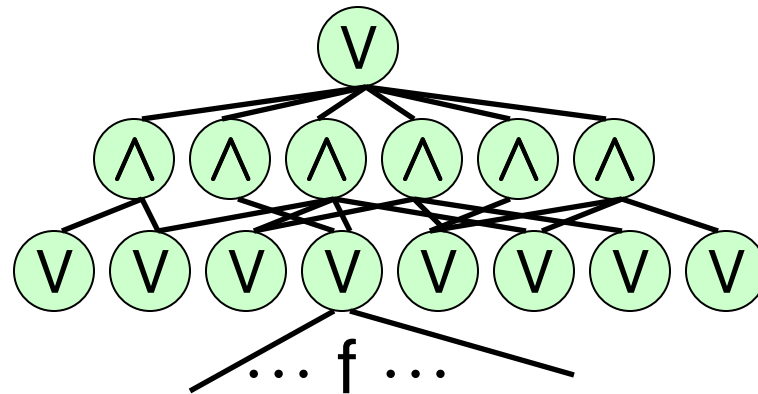
- **Theorem[V]** : Small depth-3 circuits for Approximate Majority on N bits have bottom fan-in $\Omega(\log N)$
- **Corollary**: Quadratic slow-down necessary for relativizing techniques:
$$\text{BPTime}^A(t) \not\subseteq \Sigma_2 \text{Time}^A(t^{1.99})$$
- **Theorem[DvM, V]**: $\text{BPTime}(t) \subseteq \Sigma_3 \text{Time}(t \cdot \log^5 t)$
 - Previous result [A]: $\text{BPTime}(t) \subseteq \Sigma_{O(1)} \text{Time}(t)$
- 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^ϵ} -size depth-3 circuits for Approximate Majority on N bits have bottom fan-in $f = \Omega(\log N)$
- Recall:



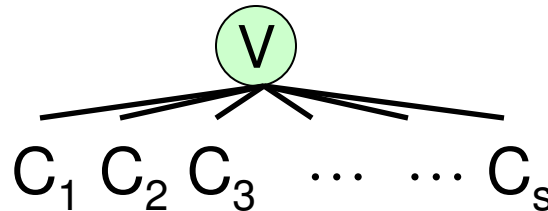
$R = 101111011011101011 \quad |R| = N$

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



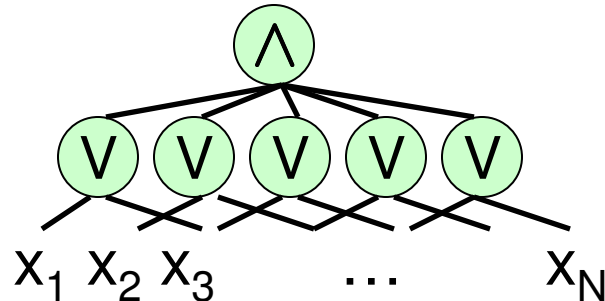
- By definition of OR :
 $R \in \text{YES} \Rightarrow \text{some } C_i (R) = 1$
 $R \in \text{NO} \Rightarrow \text{all } C_i (R) = 0$
- By averaging, fix $C = C_i$ s.t.

$$\begin{array}{l} \Pr_{R \in \text{YES}} [C (x) = 1] \geq 1/s \\ \forall R \in \text{NO} \quad \Rightarrow \quad C (R) = 0 \end{array}$$

- **Claim:** Impossible if C has bottom fan-in $\leq \varepsilon \log N$

CNF Claim

- Depth-2 circuit \Rightarrow CNF



$$(x_1 \vee x_2 \vee \neg x_3) \wedge (\neg x_4) \wedge (x_5 \vee x_3)$$

bottom fan-in \Rightarrow **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 \Rightarrow Theorem

Either $\Pr_{R \in \text{YES}} [C(x)=1] \leq 1/2^{N^\epsilon}$ or $\exists R \in \text{NO} : C(x) = 1$

Proof Outline

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

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

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

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

Case $|S|$ tiny : Fix few variables and repeat

Either $\Pr_{R \in \text{YES}} [C(x)=1] \leq 1/2^{N^\epsilon}$ or $\exists R \in \text{NO} : C(x) = 1$

Case $|S|$ BIG

- $|S| \geq N^\delta \Rightarrow$ have $N^\delta / (\epsilon \cdot \log N)$ **disjoint** clauses Γ_i
 - Can find Γ_i greedily
- $\Pr_{R \in \text{YES}} [C(R) = 1] \leq \Pr [\forall i, \Gamma_i(R) = 1]$
 - $= \prod_i \Pr[\Gamma_i(R) = 1]$ (independence)
 - $\leq \prod_i (1 - 1/100^{\epsilon \log N}) = \prod_i (1 - 1/N^{O(\epsilon)})$
 - $= (1 - 1/N^{O(\epsilon)})^{|S|} \leq e^{-N^{\Omega(1)}} \quad \checkmark$

Either $\Pr_{R \in \text{YES}} [C(x)=1] \leq 1/2^{N^\epsilon}$ or $\exists R \in \text{NO} : C(x) = 1$

Case $|S|$ tiny

- $|S| < N^\delta \Rightarrow$ Fix variables in S
 - Maximize $\Pr_{R \in \text{YES}} [C(x)=1]$
- Note: S **covering** \Rightarrow clauses shrink

Example

$$(x_1 \vee x_2 \vee x_3) \wedge (\neg x_3) \wedge (x_5 \vee \neg x_4) \quad \begin{array}{l} x_3 \leftarrow 0 \\ x_4 \leftarrow 1 \end{array} \Rightarrow (x_1 \vee x_2) \wedge (x_5)$$

- Repeat
Consider smallest covering S' , etc.

Either $\Pr_{R \in \text{YES}} [C(x)=1] \leq 1/2^{N^\epsilon}$ or $\exists R \in \text{NO} : C(x) = 1$

Finish up

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

- When you stop:

Either smallest covering size $\geq N^\delta$



Or $C = 1$

Fixed $\leq (\epsilon \cdot \log N) N^\delta \ll N$ 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 (AC^0)
 - Derandomization of AC^0 [N]
 - Application: Hardness Amplification in NP [HV \checkmark]
 - Derandomization of AC^0 with few Maj gates [\checkmark]
- General models: BPP vs. PH
 - $BPTIME(t) \subseteq \Sigma_2 TIME(t^2)$ [SG,L]
 - $BPTIME(t) \not\subseteq \Sigma_2 TIME(t^{1.99})$ (w.r.t. oracle) [\checkmark]
Lower Bound for Approximate Majority

Thank you!