Randomness is Linear in Space

Nisan and Zuckerman (1994)

Henry Scheible

November 2025

Main Result

Theorem (1)

Any randomized algorithm A that runs in space S and time T and uses $\mathrm{poly}(S)$ random bits can be simulated using only $O(S)$ random bits in space S and time $T + \mathrm{poly}(S)$. The distribution of the output of the simulation is within statistical distance of $\exp(-S^{1-\gamma})$ from the distribution of the output of A. Here $S = S(n) \geq \log(n)$, $T = T(n) \geq n$, and $\gamma > 0$ is an arbitrary constant.

For any polynomial $p(x)$, we need a psuedo-random generator $G : \{0,1\}^{O(S)} \to \{0,1\}^{p(S)}$ for space S with parameter $\exp(-S^{1-\gamma})$ running in time $\mathrm{poly}(S)$ and space $O(S)$.

Agenda

- Psuedo-Random Generators for Space S
- Model of space-bounded computation, motivation for extractors
- Definition and intuition for extractors
- Some ideas of existence proofs for extractors

Psuedo-Random Generators for Space S

For any polynomial $p(x)$, we need a psuedo-random generator $G : \{0,1\}^{O(S)} \to \{0,1\}^{p(S)}$ for space S with parameter $\exp(-S^{1-\gamma})$ running in time $\mathrm{poly}(S)$ and space $O(S)$.

Definition

A generator $G : \{0,1\}^n \to \{0,1\}^m$ is called a psuedo-random generator for space S with parameter ϵ if, for every randomized space S algorithm A and every input to it,

$$|\Pr[A(y) \text{ accepts}] - \Pr[A(G(x)) \text{ accepts}]| \leq \epsilon$$

Psuedo-Random Generators for Space S

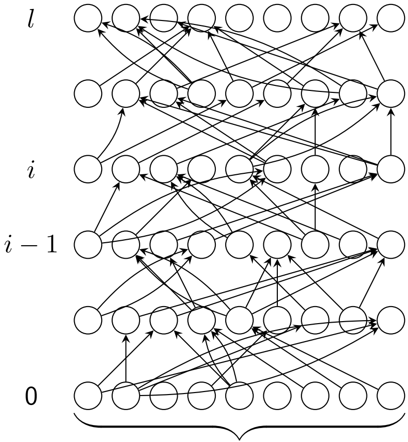
For any polynomial $p(x)$, we need a psuedo-random generator $G : \{0,1\}^{O(S)} \to \{0,1\}^{p(S)}$ for space S with parameter $\exp(-S^{1-\gamma})$ running in time $\text{poly}(S)$ and space $O(S)$.

Lemma

Let $G_1 : \{0,1\}^{R_2} \to \{0,1\}^{R_1}$ be a generator for space S_1 with parameter ϵ_1 running in space S_2. Let $G_2 : \{0,1\}^{R_3} \to \{0,1\}^{R_2}$ be a generator for space $S_1 + S_2$ with parameter ϵ_2 running in space S_3. Then $G_1 \circ G_2 : \{0,1\}^{R_3} \to \{0,1\}^{R_1}$ is a psuedo-random generator for space S_1 with parameter $\epsilon_1 + \epsilon_2$ running in space $S_2 + S_3$.

Assume $p(n) = n^c$. Thus, we only need a psuedo-random generator $G : \{0,1\}^R \to \{0,1\}^{\Omega(RS^\gamma)}$ with parameter $\exp(-S^{1-\gamma})$ running in time $\text{poly}(S)$ that we compose with itself $(c-1)/\gamma$ times. Assume without loss of generality that $R \geq 4S$.

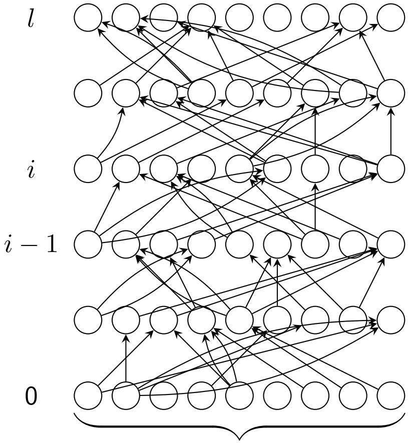
Proof



$O(2^S)$ configurations

- Let M be an arbitrary space(S) machine
- Each layer represents configuration after a set of S random bits.
- Edge $((i, j), (i+1, k))$ labeled by $r \Leftrightarrow S$-bit random string r causes M to go from j to k.
- U_i: Distribution on layer i in response to true randomness
- D_i: Distribution on layer i after running G
- We want $||U_l - D_l|| \leq 2^{-\Omega(S^{1-\gamma})}$. Let $\epsilon = 2^{-\Omega(S^{1-\gamma})}$.

Proof (continued)



$O(2^S)$ configurations

Let $x \in \{0,1\}^R$ be the input for G, and let $r_1, ..., r_l \in \{0,1\}^S$ be the outputs. Let X be the random variable for x (X is uniform).

We want $||U_l - D_l|| \leq \epsilon$.

We prove by induction that $||U_i - D_i|| \leq i\epsilon/l$.
Let D_i^j and U_i^j be U_i and U_j conditioned on the fact that M was in configuration j at step $i-1$.

Dividing up cases

Consider dividing into good configurations (say A) and bad configurations at step $i-1$. Let A be the configurations j which have $D_{i-1}[j] \geq 2^{-2S}$.

- If $j \in A$, then look at $X|(i-1,j)$.
  We know that

  $$\begin{aligned} \Pr[X = x|(i-1,j)] &\leq \Pr[X = x] / \Pr[(i-1,j)] \\ &\leq 2^{-R} 2^{2S} \leq 2^{-R/2} \end{aligned}$$

  We want that $||U_i^j - D_i^j|| \leq \epsilon'$, which would come from the distribution of r_i conditioned on $(i-1,j)$ being quasi-random to within ϵ'.
- For $j \notin A$, $D_{i-1}[j] \leq 2^{-2S}$, and there are at most 2^S possible values of j, so

  $$\sum_{j \notin A} D_{i-1}[j] \leq 2^S 2^{-2S} = 2^{-S}$$

Proof (continued)

$$
\begin{aligned}
||U_i - D_i||_1 &= \left\| \sum_j U_{i-1}[j] U_i^j - \sum_j D_{i-1}[j] D_i^j \right\|_1 \\
&\leq \left\| \sum_j U_{i-1}[j] U_i^j - \sum_j D_{i-1}[j] U_i^j \right\|_1 \\
&\quad + \left\| \sum_j D_{i-1}[j] U_i^j - \sum_j D_{i-1}[j] D_i^j \right\|_1 \\
&\leq \left(\sum_j |U_{i-1}[j] - D_{i-1}[j]| \right) ||U_i^j||_1 \\
&\quad + \left(\sum_{j \in A} |D_{i-1}[j]| \right) ||U_i^j - D_i^j||_1 + \left(\sum_{j \notin A} |D_{i-1}[j]| \right) ||U_i^j - D_i^j||_1 \\
&\leq ||U_{i-1} - D_{i-1}||_1 \cdot 1 + 1 \cdot 2\epsilon' + 2^{-S} \cdot 2.
\end{aligned}
$$

What we need

Distribution of X in $j \in A$ case

If $j \in A$, then look at $X|(i-1,j)$.
We know that

$$\begin{aligned}\Pr[X = x|(i-1,j)] &\leq \Pr[X = x]/\Pr[(i-1,j)] \\ &\leq 2^{-R}2^{2S} \leq 2^{-R/2}\end{aligned}$$

We want that $||U_i^j - D_i^j|| \leq \epsilon'$, which would come from the distribution of r_i conditioned on $(i-1,j)$ being quasi-random to within ϵ'.

Definition (δ-source)

A distribution D on $\{0,1\}^n$ is a *δ-source* if for all $x \in \{0,1\}^n$, $D(x) \leq 2^{-\delta n}$.

Note that we know that $X|(i-1,j)$ is a δ-source.

Extractors

This motivates the need for an extractor!

Definition

Let $E : \{0,1\}^n \times \{0,1\}^t \to \{0,1\}^m$. E is called a (δ, ϵ)-extractor if for every δ-source D, the distribution of $E(x, y) \circ y$ induced by choosing x form D and y uniformly in $\{0,1\}^t$ is within statistical distance of ϵ from the uniform distribution (on $\{0,1\}^m \times \{0,1\}^t$.)

Extractors

This motivates the need for an extractor!

Definition

Let $E : \{0,1\}^n \times \{0,1\}^t \to \{0,1\}^m$. E is called a (δ, ϵ)-extractor if for every δ-source D, the distribution of $E(x,y) \circ y$ induced by choosing x form D and y uniformly in $\{0,1\}^t$ is within statistical distance of ϵ from the uniform distribution (on $\{0,1\}^m \times \{0,1\}^t$.)

Intuition for extractors

Extractors

This motivates the need for an extractor!

Definition

Let $E : \{0,1\}^n \times \{0,1\}^t \to \{0,1\}^m$. E is called a (δ, ϵ)-extractor if for every δ-source D, the distribution of $E(x, y) \circ y$ induced by choosing x form D and y uniformly in $\{0,1\}^t$ is within statistical distance of ϵ from the uniform distribution (on $\{0,1\}^m \times \{0,1\}^t$.)

Intuition for extractors

- Hash families

Extractors

This motivates the need for an extractor!

Definition

Let $E : \{0,1\}^n \times \{0,1\}^t \to \{0,1\}^m$. E is called a (δ, ϵ)-extractor if for every δ-source D, the distribution of $E(x,y) \circ y$ induced by choosing x form D and y uniformly in $\{0,1\}^t$ is within statistical distance of ϵ from the uniform distribution (on $\{0,1\}^m \times \{0,1\}^t$.)

Intuition for extractors

- Hash families
- Bipartite graph on $\{0,1\}^n \times \{0,1\}^m$ with good expansion properties.

More Extractor Intuition

Consider a set $U \subseteq \{0,1\}^n$ with $|U| \geq 2^{\delta n}$, and suppose you have a random element from A.

How can you extract the δn bits of randomness?

An extractor gives you $\Omega(\delta^2 n)$ bits.

Existence of Extractors

Lemma

For any parameters $\delta = \delta(n)$ and $\epsilon = \epsilon(n)$ with $1/n \leq \delta \leq 1/2$ and $2^{-\delta n} \leq \epsilon \leq 1/n$, there exists an easily computable (and explicitly given) (δ, ϵ)-extractor $E : \{0,1\}^n \times \{0,1\}^t \to \{0,1\}^m$, where $t = O(\log \epsilon^{-1} \log^2 n \log \delta^{-1}/\delta)$ and $m = \Omega(\delta^2 n/\log(\delta^{-1}))$.

Proof Idea: Leftover Hash Lemma and converting δ-sources to block-wise δ-sources.

Proof Ideas: Leftover Hash Lemma

Lemma (Leftover Hash Lemma)

Let $X \subset \{0,1\}^n$, $|X| \geq 2^r$. *Let* $k > 0$, *and* H *be a 2-universal family of hash functions mapping* n *bits to* $r - 2k$. *Then the distribution* $(h, h(x))$ *is quasi-random within* $1/2^k$ *(on the set* $H \times \{0,1\}^{r-2k}$*), where* h *is chosen uniformly at random from* H*, and* x *uniformly from* X.

Proof Ideas: Leftover Hash Lemma

Lemma (Leftover Hash Lemma)

Let $X \subset \{0,1\}^n$, $|X| \geq 2^r$. *Let* $k > 0$, *and* H *be a 2-universal family of hash functions mapping* n *bits to* $r - 2k$. *Then the distribution* $(h, h(x))$ *is quasi-random within* $1/2^k$ *(on the set* $H \times \{0,1\}^{r-2k}$*), where* h *is chosen uniformly at random from* H, *and* x *uniformly from* X.

Corollary

Let D *be a distribution on* $\{0,1\}^n$ *such that for all* $x \in \{0,1\}^n$, $D(x) \leq 2^{-r}$. *Let* $k > 0$, *and let* H *be a universal family of hash functions mapping* n *bits to* $r - 2k$ *bits. Then the distribution* $(h, h(x))$ *is quasi-random within* $1/2^k$ *(on the set* $H \times \{0,1\}^{r-2k}$*), where* h *is chosen uniformly at random from* H, *and* x *according to* D.

Block-wise δ-sources

Definition

A distribution D on the space $\{0,1\}^{l_1} \times \{0,1\}^{l_k}$ is called a block-wise δ-source if, for $1 \leq i \leq k$ and for all values $x_1 \in \{0,1\}^{l_1}, ..., x_i \in \{0,1\}^{l_k}$, we have that

$$\Pr[X_i = x_i | X_1 = x_1, ..., X_{i=1} = x_{i-1}] \leq 2^{-\delta l_i}$$