DERANDOMIZATION

AGENDA
1. Derandomization
2. Nisan-Wigderson

Two things we don't know:
(1) Are there hard functions?
(2) Does randomness help computation?

Punchline for the next few days:

$\exists$ hard functions $\Rightarrow$ randomness doesn't help

there are functions that need big circuits to compute exactly $\Rightarrow$ BPP $\subseteq$ ___ ← P or QUASIP or SUBEXP or...

Route:

fns that are hard in the worst case $\Rightarrow$ fns that are hard on average $\Rightarrow$ pseudorandom generators $\Rightarrow$ BPP $\subseteq$ ___

Two options:
(1) local decoding + XOR Lemmas
(2) local list decoding

Nisan-Wigderson

derandomization

Start on this end.

(1) Derandomization. Does BPP = P?

One way to derandomize a randomized algorithm:

Given an alg A w/ $x \in \{0,1\}^m$ → A → output y which is correct w/ prob $2/3$, $r \in \{0,1\}^m$ random

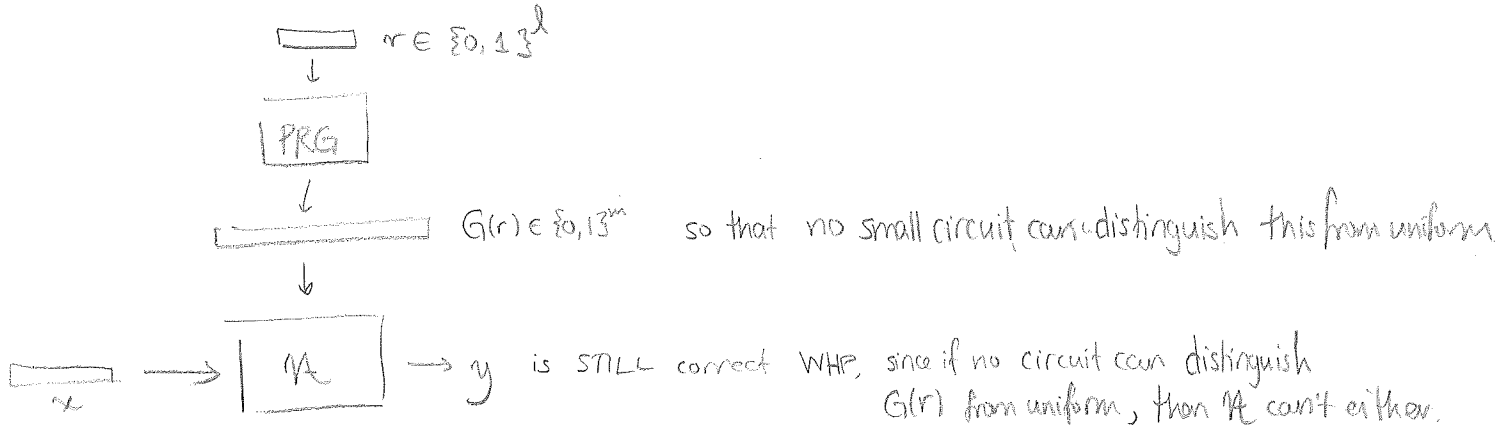
Deterministic algorithm:

try all the seeds $r \in \{0,1\}^m$ and take majority outcome.

Problem:

this takes time 2^m.

[2] Hope to fix this with a Pseudorandom Generator (PRG).



New deterministic algorithm: try all the seeds $r \in \{0,1\}^\ell$ and take majority vote.

Def A distribution $\mathcal{R}$ over $\{0,1\}^m$ is (s, ε) - PSEUDORANDOM if:

$\forall$ circuits C of size $|C| \le s$,

$$\left| \mathbb{P}_{r \sim \mathcal{R}}\{C(r)=1\} - \mathbb{P}_{u \sim U_m}\{C(r)=1\} \right| < \varepsilon.$$

(U_m: uniform on m bits)

Let $s : \mathbb{N} \to \mathbb{N}$ (time constructible, non-decreasing)

$G : \{0,1\}^* \to \{0,1\}^*$ is $s(\ell)$ PRG if $\{0,1\}^* \to \{0,1\}^*$ is [illegible]

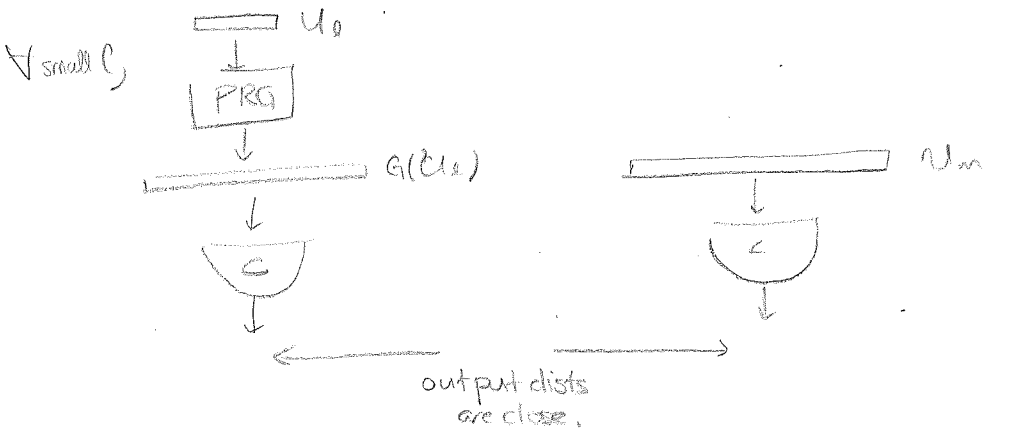
$\forall z \in \{0,1\}^*$, $|G(z)| = s(|z|)$

$\forall \ell \in \mathbb{N}$, $G(U_\ell)$ is $\left(s(\ell)^{10}, \frac{1}{10} \right)$ - pseudorandom

$\forall z \in \{0,1\}^*$, $G(z)$ can be computed in time $\le 2^{O(|z|)}$

(10, $\frac{1}{10}$: arbitrary constants.)

That is, G is PRG if no small circuit can distinguish it from uniform.



3 If we have a PRG we can derandomize.

LEMMA Say $\exists$ $s(\ell)$-PRG. Then $\forall$ polytime computable $\ell : \mathbb{N} \to \mathbb{N}$,

$$BPTIME(s(\ell(n))) \subseteq DTIME(2^{c \cdot \ell(n)})$$ for some constant c (which depends on ℓ)

Proof DTIME alg is: on input $x \in \{0,1\}^n$:

- for all $z \in \{0,1\}^{\ell(n)}$, compute $A(x, G(z))$ ← here $\leq s(\ell(n))$
- output majority answer.

CLAIM. For sufficiently large n, $\forall x$, $P_z\{A(x, G(z)) = L(x)\} \geq \frac{2}{3} - \frac{1}{10} > \frac{1}{2}$.

This is enough. [The "c" in the statement is for "sufficiently large"]

pf of CLAIM Suppose otherwise, $\exists$ only many x s.t.

$$P_z\{A(x, G(z)) = L(x)\} < \frac{2}{3} - \frac{1}{10}.$$

We'll build a circuit that distinguishes $G(z)$ from uniform.

How?

$$A(x, \bullet) \xrightarrow{\text{cook-levin}} C_x(\bullet) \text{ circuit of size } O(s(\ell(n))^2) \leq [s(\ell(n))]^{10}$$

for sufficiently large n.

COR

$2^{\Sigma\ell}$ PRGs $\Rightarrow$ BPP = P

2^{ℓ^ϵ} PRGs $\Rightarrow$ BPP = QUASI-P := D-TIME$(2^{\text{polylog}(n)})$

$\forall c \geq 1$ ℓ^c PRGs $\Rightarrow$ BPP $\subseteq$ SUBEXP := $\bigcap_{\epsilon > 0}$ DTIME(2^{n^ϵ})

② Nisan-Wigderson : Avg-case hardness ⇒ PRGs

DEF $H^{\ell}_{avg}(f) = \max\left\{ s : \forall C, |C| \le s, \; \Pr_{x \sim U_n}\{C(x) = f(x)\} < \ell \right\}$

aka, any circuit that computes f w/ success prob $\ge \ell$ on a random input has to be bigger than $H^{\ell}_{avg}(f)$.

$$H_{avg}(f) = \max\left\{ s : H^{1/2 + 1/s}_{avg}(f) \ge s \right\}$$

$$H_{worst}(f) = H^{1}_{avg}(f) = \max\left\{ s : \text{no circuit of size } s \text{ computes } f \text{ correctly on all inputs} \right\}$$

Thm (Nisan-Wigderson)

Suppose $f \in TIME(2^{O(n)})$, $H_{avg}(f) \ge H(n)$.

Then $\exists$ $s(\ell)$-PRG w/

$$s(\ell) \ge (H(n))^{\delta}$$

for all $n \ge \sqrt{\delta \ln \log(H(n))}$ for some $\delta > 0$.

→ This is what we'll show. | Actually one can show:

Thm $f \in TIME(2^{O(n)})$, $H_{avg}(f) \ge H(n)$.

$\Rightarrow \exists$ $[H(\delta \ell)]^{\delta}$-PRG. for some $\delta > 0$.

COR - $f \in TIME(2^{O(n)})$ and

$H_{avg}(f) \ge 2^{\varepsilon n} \Rightarrow 2^{\varepsilon \cdot \delta^2 \ell}$-PRGs $\Rightarrow P = BPP$

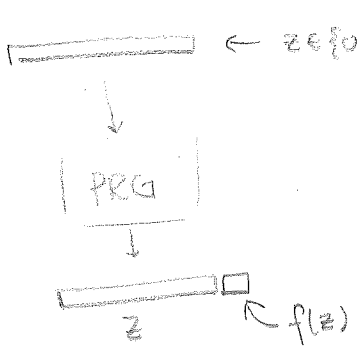
$H_{avg}(f) \ge 2^{n^{\varepsilon}} \Rightarrow 2^{(\delta \ell)^{\varepsilon \cdot \delta}}$-PRGs $\Rightarrow$ BPP = QUASI-P

$H_{avg}(f) \ge n^{c} \Rightarrow (\delta \ell)^{c \cdot \delta}$-PRGs $\Rightarrow$ BPP = SUBEXP

5 Warm-up.

Lemma Say that f is hard on average: $f \in E$ and $H_{avg}(f) > n^{100}$.
Then $\exists$ an $(\ell+1)$-PRG.

Construction is:



Lemma (Yao) Let $\mathcal{R}$ be a dist over $\{0,1\}^m$, and say $\exists$ $S > 10m$ and $\epsilon > 0$ so that $\forall$ C of size $|C| \le 2S$, and $\forall i \in [m]$,

$$\mathbb{P}_{r \sim \mathcal{R}}\{ C(r_1, \dots, r_{i-1}) = r_i \} \le \frac{1}{2} + \frac{\epsilon}{m}$$

Then $\mathcal{R}$ is (S, ϵ)-pseudorandom.

That is, "unpredictability $\Rightarrow$ pseudorandomness"!
We'll prove this later. For now, how do we use it?

Recall, PRG means output is $(s(\ell)^{1/a}, 1/b)$-PR

→ **Pf** It suffices to show that there's no circuit of size $2 \cdot \ell^{10}$ and $i \in [m]$ s.t. C can predict $f(z)$, given z_i.

← because clearly no circuit can predict z_i given $z_1, \dots, z_{i-1}$ for any i.

$$\mathbb{P}_{z \sim \mathcal{U}_\ell}[C(z) = f(z)] > \frac{1}{2} + \frac{1/10}{\ell+1}$$

But by def of hard, there's no circuit of size $< \ell^{100}$ s.t.

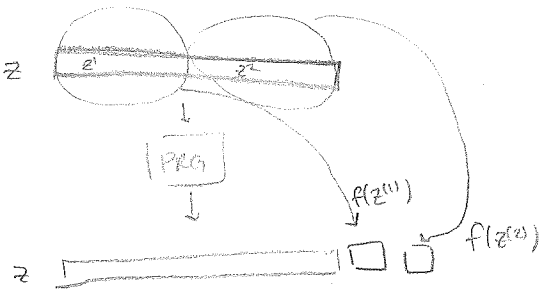
$$\mathbb{P}_{z \sim \mathcal{U}_\ell}[C(z) = f(z)] > \frac{1}{2} + \frac{1}{\ell^{100}}$$

, so that is definitely enough. ∎

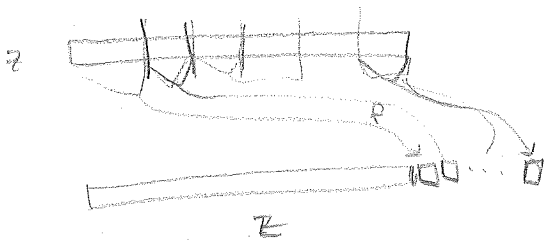
Of course, this is no good for derandomization.

16 How can we extend this?

By one more bit...

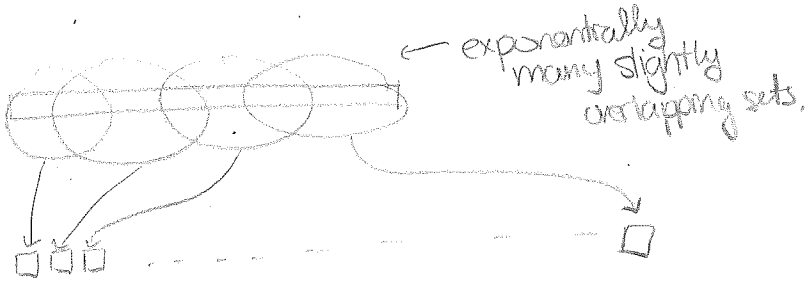


OK... that'll work, but to get a lot of stretch we'll have at best

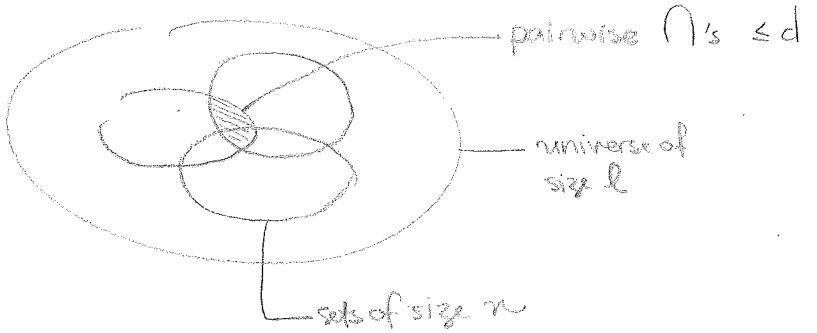


so this idea can never expand z by more than a factor of 2.

How about if we let them slightly overlap?



Def Say $\ell > n > d$. A family $\mathcal{I} = \{I_1, \ldots, I_m\}$, $I_i \subseteq [\ell]$ is an (ℓ, n, d)-design if $|I_j| = n \ \forall j$, and $|I_j \cap I_k| \leq d \ \forall j \neq k$.



Lemma We can construct (ℓ, n, d) designs for $n > d$ and $\ell > \frac{10 \cdot n^2}{d}$ in time $2^{O(\ell)}$ and w/ size $|\mathcal{I}| = 2^{d/10}$

Given that, here's the construction:

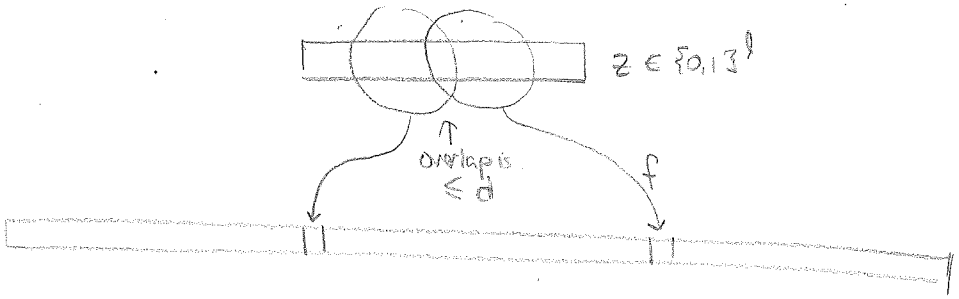
Suppose $f : \{0,1\}^n \longrightarrow \{0,1\}$, $H_{avg}(f) = H$

$\mathcal{I} = \{I_1, \ldots, I_m\}$ is an (ℓ, n, d)-design.

$NW^f_{\mathcal{I}} : \{0,1\}^\ell \longrightarrow \{0,1\}^m$ is given by

$$NW^f_{\mathcal{I}}(z) = f(z|_{I_1}) \circ f(z|_{I_2}) \circ \cdots \circ f(z|_{I_m})$$

$\circ$ ↑ concatenation



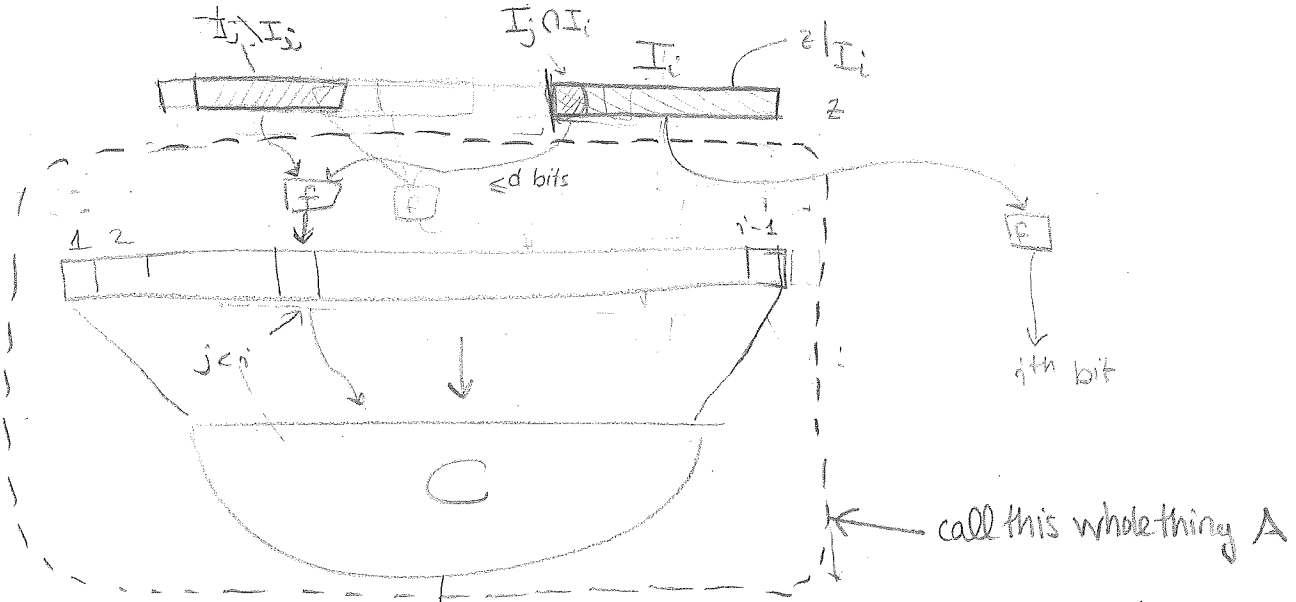
$NW(z) \in \{0,1\}^m$

$m = 2^{d/10}$

8 Pf of NW Thm. → LEMMA: If $f:\{0,1\}^n \to \{0,1\}$ has $H_{avg} > 2^{\delta d}$, then $NW_I^f(U_d)$ is $(2^{\delta d}/4, 1/10)$-PR.

Pf of lemma: Say that $H_{avg}(f) > 2^{\delta d}$.

- Want to show no circuit of size $\frac{2^{\delta d}}{4}$ can distinguish $NW_I^f(U_d)$ from U_m w/ err $1/10$
- Suffices to show that no circuit of size $\frac{2^{\delta d}}{2}$ can predict the i^{th} bit with advantage $\geq 1/10 \cdot m$.
- Say there is such a circuit; C



output $= f(z|_{I_i})$ w/ advantage $\geq \frac{1}{10m}$

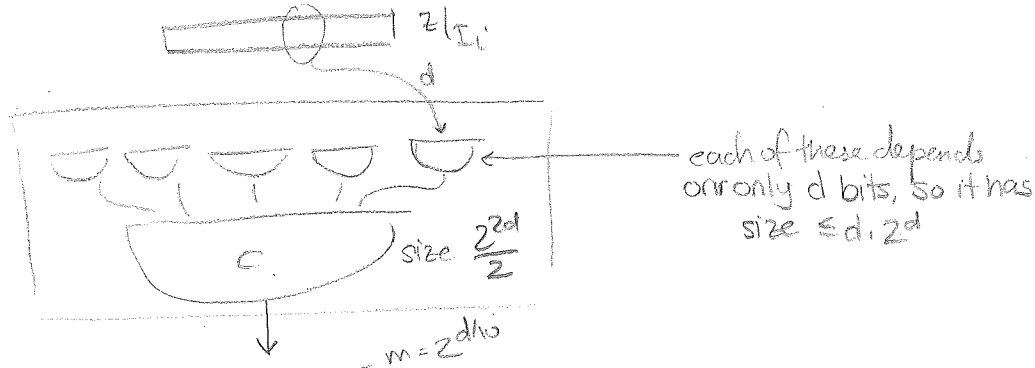
$$\mathbb{P}\{ C(\dots) = f(z|_{I_i}) \} \geq \frac{1}{2} + \frac{1}{10m}.$$

That means $\exists$ some string $z|_{rest} \in \{0,1\}^{\ell - n}$ so that

$$\mathbb{P}_{z|_{I_i}}\{ A(z|_{rest}, z|_{I_i}) = f(z|_{I_i}) \} \geq \frac{1}{2} + \frac{1}{10m}.$$

($z|_{rest}$: fixed; $z|_{I_i}$: random)

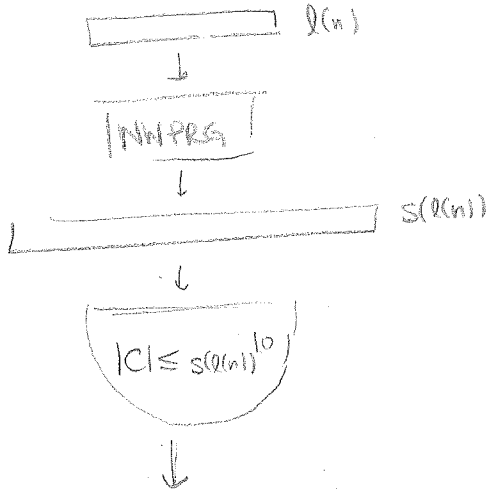
But we can implement A as a small circuit (w/ input $z|_{I_i}$):



$\Rightarrow$ total size is $\leq \frac{2^{\delta d}}{2} + m \cdot d \cdot 2^d < 2^{\delta d}$. ↯ this contradicts the hardness of f.

[9] To use that Lemma, wade through a mess of parameters.

Say $f \in \text{TIME}(2^{O(n)})$, $H_{avg}(f) \geq H(n)$.



Choose $d = \log(H(n))/2$

$$\ell(n) = \frac{10 \cdot n^2}{d} = \frac{20\, n^2}{\log(H(n))}$$

$$s(\ell(n)) = 2^{d/10} = H(n)^{1/20}$$

No circuit of size $\leq \frac{H(n)}{4}$ can distinguish from uniform.

So no circuit of size $s(\ell(n)) = \sqrt{H(n)} \leq \frac{H(n)}{4}$ can do that either.

Notice also NWPRG make $\text{poly}(H(n)) \leq 2^{O(n)}$ calls to f, which also needs time $2^{O(n)}$. So total running time is $2^{O(n)} \subseteq 2^{O(\ell)}$.

This gives the thm w/ $\varepsilon = 1/20$, (probably, maybe a factor of 2...)

Finally, the proof of this lemma.

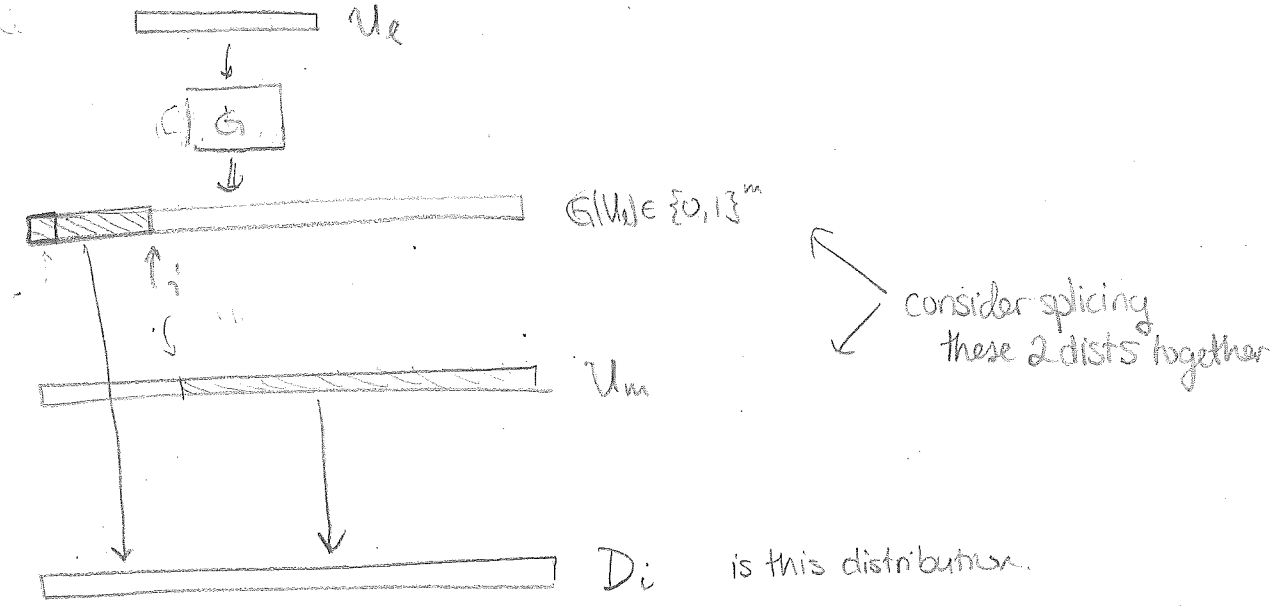
WTS that unpredictability ⇒ pseudorandomness.

Proof (sketch) by contradiction. Given a circuit that distinguishes $G(U_\ell)$ from U_m;

$$\left| \mathbb{P}\{C(G(U_\ell))=1\} - \mathbb{P}\{C(U_m)=1\} \right| > \varepsilon.$$

We'll construct a circuit B that has, for some i,

$$\mathbb{P}_{r \sim G(U_\ell)}\{B(r_1,\dots,r_{i-1}) = r_i\} > \frac{1}{2} + \frac{\varepsilon}{m}.$$



So $D_0 = U_m$, $D_m = G(U_\ell)$.

Let $p_i = \mathbb{P}\{C(D_i)=1\}$.

Then by assm $\varepsilon < |p_0 - p_m| < |p_0 - p_1| + |p_1 - p_2| + \cdots + |p_m - p_{m-1}|$,

so one of these increments (say, i) has $|p_{i-1} - p_i| > \varepsilon/m$. Consider the algorithm:

(say wlog $p_i > p_{i-1}$)

$A(r_1,\dots,r_{i-1})$:

draw $z_i,\dots,z_m$ randomly

If $C(r_1,\dots,r_{i-1},z_i,\dots,z_m)=1$, output z_i

else, output $1-z_i$

11 Now,

$$\mathbb{P}_{z,r}\{A(r_1,\dots,r_{i-1}) = r_i\} = \frac{1}{2}\mathbb{P}\{C(D_{i-1})=1 \mid z_i = r_i\} + \frac{1}{2}\mathbb{P}\{C(D_{i-1})=0 \mid z_i \neq r_i\}$$

correct if either
$z_i = r_i$ and $C(D_{i-1})=1$
or
$z_i \neq r_i$ and $C(D_{i-1})=0$

$$= \mathbb{P}\{C(D_{i-1})=1 \mid z_i = r_i\} = \mathbb{P}\{C(D_i)=1\} = p_i$$

$$\mathbb{P}\{C(D_{i-1})=0 \mid z_i \neq r_i\} = 1 - \mathbb{P}\{C(D_{i-1})=1 \mid z_i \neq r_i\}$$

$$= 1 - 2p_{i-1} + p_i$$

Now,

$$p_{i-1} = \frac{1}{2}\mathbb{P}\{C(D_{i-1})=1 \mid z_i = r_i\} + \frac{1}{2}\mathbb{P}\{C(D_{i-1})=1 \mid z_i \neq r_i\}$$
$$= \frac{1}{2}p_i + \frac{1}{2}\mathbb{P}\{C(D_{i-1})=1 \mid z_i \neq r_i\}$$

$$\Rightarrow \mathbb{P}\{C(D_{i-1})=1 \mid z_i \neq r_i\} = 2p_{i-1} - p_i$$

So

$$\mathbb{P}_{z,r}\{A(r_1,\dots,r_{i-1}) = r_i\} = \frac{1}{2} + (p_i - p_{i-1}) \geq \frac{1}{2} + \varepsilon/m.$$

So then there's some string z so that

$$\mathbb{P}_r\{A(r_1,\dots,r_{i-1}; z) = r_i\} \geq \frac{1}{2} + \varepsilon/m.$$

Let B be the circuit for A w/ that string z hard-coded in. ∎