PSEUDORANDOMNESS

AN OVERVIEW

by

Oded GOLDREICH

WEIZMANN INST.

http://www.weizmann.AC.IL/~oded/PS/prg-no6.ps

RANDOMNESS & COMPUTATION

- **RANDOMNESS as a TOOL used in COMPUTATION**

Essential uses include
+ CRYPTOGRAPHY & DISTRIBUTED COMPUTING
+ PROB. PROOF SYSTEMS (IP, ZK, PCP)
+ Sampling & PROPERTY TESTING

(Omitted: use in standard ALGORITHMS)

- **RANDOMNESS as an OBJECT viewed by COMPUTATION**

$\Rightarrow$ Computational Indistinguishability

$\Rightarrow$ Different objects viewed as equiv. by resource-bounded computations.

$\Rightarrow$ Potential saving/elimination of RANDOMNESS in COMPUT.

(because corresp. applic. cannot tell...)

COMPUTATIONAL VIEW OF RANDOMNESS

COMPUTATIONAL INDISTINGUISHABILITY

$X \equiv Y$ $\qquad Z = \{Z_n\}$

RELAX

$X \overset{s}{\equiv} Y \triangleq \sum_{\alpha} \left| Prob[X_n = \alpha] - Prob[Y_n = \alpha] \right|$
is negl(n)

$X \overset{c}{\equiv} Y \triangleq$ Efficient algorithms (and/or ALG of certain class) "can not tell these apart".

Classes to consider

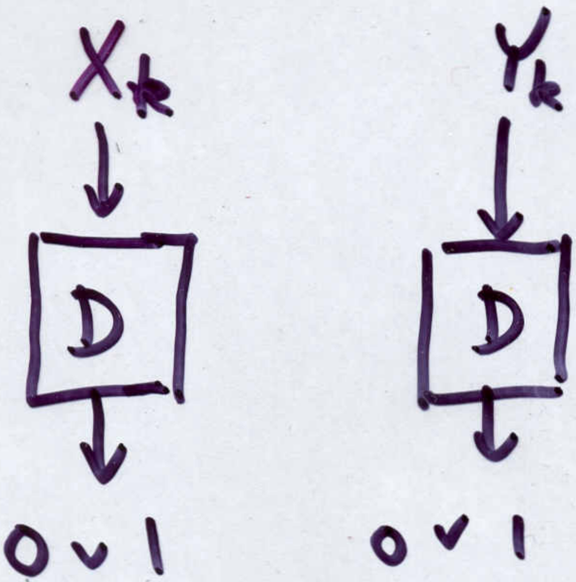
- Poly-time alg.
  poly-size circuits (non-uniform)
- Quad-size circuits
- Space-bounded alg.
- Syn. restricted alg.
  (projection, linear, hitting)

COMPUTATIONAL INDISTINGUISHABILITY

$Z = \{Z_k\}$, where $Z_k \in \{0,1\}^k$ or $\{0,1\}^{\ell(k)}$

DEF: X and Y are ϵ-distinguishable by D

(potential distinguisher)



D's verdict is INSIGNIFICANT

$$\left| \text{Prob}[D(X_k)=1] - \text{Prob}[D(Y_k)=1] \right| \le \epsilon(k)$$

Typically, ϵ is NEGLIGIBLE

$= 1/\text{complexity}(D)$

When class of ALG is understood, we say that X & Y are COMPUT. INDISTING.

Notions of PSEUDORANDOM GENERATORS

$G:\{0,1\}^k \to \{0,1\}^{\ell(k)}$ is a PRG (generic) if

(1) STRETCH $\ell(k) > k$ ($\ell(k) >>> k$)

- ℓ is a polynomial
- ℓ is an exponential [$\ell(k) = 2^{\Theta(k)}$]

(2) EFFICIENT GENERATION

- each bit produced in POLY-TIME
- each bit produced in EXP-TIME

(3) PSEUDORANDOMNESS ≡ Computational Indist. from the uniform (i.e. $\{U_{\ell(k)}\}$)

- by (PROB.) POLY-TIME ALG.
- by QUAD-SIZE CIRCUITS.

GENERAL PURPOSE PRG

canonical DERANDOMIZER

GEN more complex than D

Two popular notions of PRG

- **GENERAL-PURPOSE PRG**

  Can be used to save RANDOMNESS in ANY (efficient) application.*

  Output looks RANDOM also to OBSERVER that uses more RESOURCES THAN the PRG.

- **CANONICAL DERANDOMIZER**

  May (& typically does) use more RESOURCES THAN the OBSERVER.

  Suffices for derandomization of ALGs of specific complexity.

---

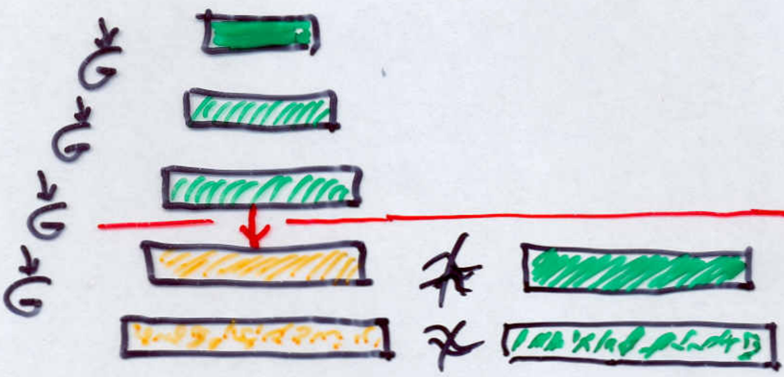
*) Essential to CRYPTO/ADVERSARY APPLICATIONS.

AMPLIFYING THE STRETCH of GENERAL-PURPOSE PRGs

Suppose $G:\{0,1\}^k \to \{0,1\}^{k+1}$ is a PRG, and $\ell:\mathbb{N}\to\mathbb{N}$ is a polynomial (s.t. $\ell(k)>k$)

NAIVE ITERATION METHOD

$G'(s) \triangleq G^{\ell(|s|)-|s|}(s)$, where $G^{i+1}(x)\triangleq G(G^i(x))$ and $G^0(x)\triangleq x$

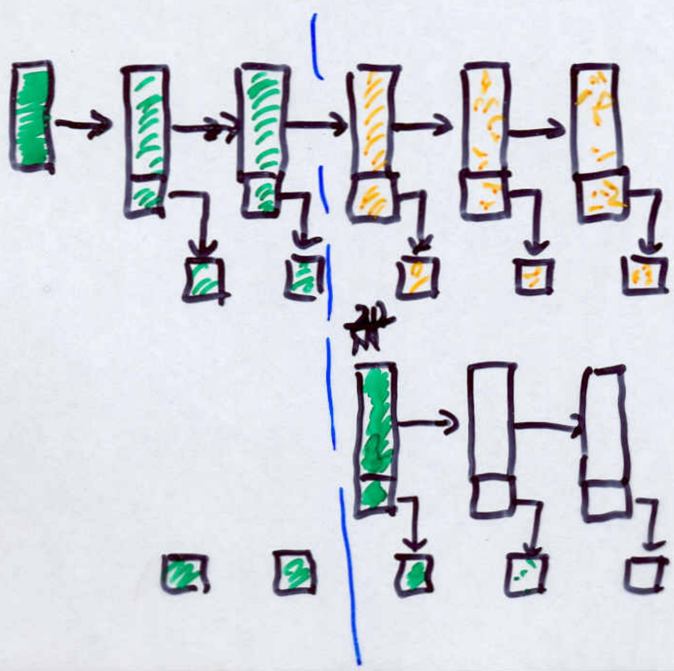


"Deterioration Detected" "at this move"

DISTINGUISH TOP LINE BY APPLYING G

FIXED-LENGTH ITERATION METHOD

$G'(s) \triangleq \sigma_1\cdot\sigma_2\cdots\sigma_{\ell(|s|)}$, where $s_0\triangleq s$ and $\sigma_i s_i \triangleq G(s_{i-1})$



CONSTRUCTING GENERAL-PURPOSE PRG

HARDNESS VS RANDOMNESS

DEF: $f:\{0,1\}^k \to \{0,1\}^k$ is a **OWF** if

(1) poly-time computable

(2) HARD to INVERT on AVERAGE-CASE

$\forall$ ppt A, $\mathrm{PROB}_{x\in\{0,1\}^k}\left[A(f(x)) \in f^{-1}f(x)\right] = \mathrm{negl}(k)$

$$\boxed{\text{THM: PRG exist iff OWF exist.}}$$

PRG $\Rightarrow$ OWF: $G:\{0,1\}^k \to \{0,1\}^{2k}$ PRG

$\Longrightarrow$ $f(x,y) \triangleq G(x)$ for $|x|=|y|$

Inverting f on $f(U_{2k}) \equiv G(U_k)$ implies DIST. $G(U_k)$ FROM U_{2k} (since the latter has f-preimage w. NEGL. PROB.).

OWF $\Rightarrow$ PRG:

We'll only show a special case.

OWF implies PRG

DEF: $b : \{0,1\}^* \to \{0,1\}$ is a HARDCORE of f if

(1) $x \to b(x)$ is POLY-TIME COMPUTABLE

(2) $f(x) \to b(x)$ is HARD to PREDICT on AVERAGE-CASE

$\forall$ ppt A, $\mathrm{Prob}_x[A(f(x)) = b(x)] < \frac{1}{2} + \mathrm{NEGL}(k)$

Note: $b(x)$ HARD TO PREDICT from $f(x)$

$\sim$ $\{f(U_k) \circ \underbrace{b(U_k)}_{\text{same}}\} \stackrel{c}{=} \{f(U_k) \circ \underbrace{U_1}_{\text{independent}}\}$

- Indivi. bits may not be HARDCORE; e.g., $f(x,y) = (f'(x), y)$
- If f is 1-1 & easy to invert then it has NO HARDCORE. $\boxed{f(x) \to x \to b(x)}$

For a 1-1 OWF f, any HARDCORE b yields a PRG $G(s) = \underbrace{f(s)}_{\text{uniform}} \circ \underbrace{b(s)}_{\text{unpredict.}}$.

$\boxed{\text{Amplifying STRETCH}}$

OWF $\Rightarrow$ HARDCORE

OWF f_0 $\Longrightarrow$ OWF $f(x,r) = (f_0(x), r)$

+ HC $b(x,r) = \sum_{i=1}^{k} x_i r_i \mod 2$

LEMMA: Suppose, given $B: \{0,1\}^k \to \{0,1\}$ s.t. $\exists x \in \{0,1\}^k$

$$\text{Prob}_{r \in \{0,1\}^k}\left[B(r) = b(x,r)\right] \geq \frac{1}{2} + \epsilon$$

Then, in poly(k/ϵ)-time, can guess x correctly w.p. $\geq$ poly(ϵ/k)

$\boxed{B_x(r) \triangleq A(f(x), r)}$

Warm-up: Suppose $p_x \triangleq \text{Prob}[B(r) = b(x,r)] \geq 3/4 + \epsilon$

$\Rightarrow$ Recover x_j w.p $1 - 2 \cdot (1 - p_x) \geq \frac{1}{2} + 2\epsilon$

by $r \in_R \{0,1\}^k$ & output $B(r) \oplus B(r \oplus e^j)$

$$\left[b(x,r) \oplus b(x, r \oplus e^j) = \left(\sum_{i=1}^{k} x_i r_i\right) + \left(x_j + \sum_{i=1}^{k} x_i r_i\right) = x_j \right]$$

$\boxed{\text{+ Majority rule } O(n/\epsilon^2) \text{ PAIRWISE IND. VOTES}}$

Eliminate error-doubling

$m \triangleq 5$

Suppose $r^{(1)}, \ldots, r^{(m)} \in \{0,1\}^k$ PAIRWISE INDEP.

and we KNOW $b(x, r^{(1)}), \ldots, b(x, r^{(m)})$.

Then $\text{MAJ}_{i \in [m]} \left\{ b(x, r^{(i)}) \oplus B(r^{(i)} \oplus e^j) \right\} = x_j$

with prob. $\geq 1 - \frac{1}{2k}$

How?

[single call to B, per "vote"]

Generating PAIRWISE IND. samples in $\{0,1\}^k$ with known $b(x,\cdot)$-values

Select $s^{(1)},\dots,s^{(\ell)} \in_R \{0,1\}^k$, where $\ell = \log_2(m+1)$

Guess $b(x,s^{(1)}),\dots,b(x,s^{(\ell)}) \in \{0,1\}$

[correct w.p $2^{-\ell} = \frac{1}{m+1} = \frac{1}{\mathrm{poly}(k/\varepsilon)}$]

Generate $\langle r^{(I)} \rangle_{I \subseteq [\ell],\, I \neq \emptyset}$ s.t. $r^{(I)} = \bigoplus_{i \in I} s^{(i)}$

and note that

$$b(x, r^{(I)}) = b(x, \bigoplus_{i \in I} s^{(i)})$$

$$= \bigoplus_{i \in I} b(x, s^{(i)})$$

Thus, w.p $\frac{1}{m+1}$, we obtain the correct values for <u>all</u> $b(x, r^{(I)})$'s.

+ <u>Note</u>: $r^{(I)}$'s are PAIRWISE IND and uniformly dist. in $\{0,1\}^k$.

HARDNESS VS. RANDOMNESS, ACT 2

$G: \{0,1\}^k \to \{0,1\}^{\ell(k)}$ is a CANONICAL DERANDOMIZER

if G is EXP-TIME Comput. & $\{G(U_k)\} \overset{\text{R.Si.}}{\equiv} \{U_{\ell(k)}\}$.

DERANDOMIZATION of A, where $A(x,r)$ with $|r| = t_A(|x|) = \text{poly}(|x|)$

- $A'(x,s) = A(x, G(s))$ where $|s| = \ell^{-1}(t_A(|x|)) \doteq k$

$$\boxed{\forall x \quad \left|\text{Prob}[A(x, G(U_k)) = 1] - \text{Prob}[A(x, U_{\ell(k)}) = 1]\right| < \frac{1}{10}}$$

- $A''(x) = \text{maj}_{s \in \{0,1\}^k} \{A'(x,s)\}$

$$\boxed{\text{running time} = 2^k \cdot (t_A(|x|) + t_G(k))}$$

THM: If $\exists$ can. derand. with $\ell(k) = 2^{\Omega(k)}$ then BPP = P.

$$\boxed{\begin{array}{l} k = \ell^{-1}(\text{poly}(n)) = O(\log n) \\ \Longrightarrow \; 2^k = \text{poly}(n) \end{array}}$$

THM: If $E = \text{Dtime}(2^{O(n)})$ contains a problem of circuit complexity $2^{\Omega(n)}$ [in worst-case sense but a.e.] then $\exists$ can. derand. with $\ell(k) = 2^{\Omega(k)}$

$$\boxed{\exists c>0 \text{ s.t. } \text{Dtime}(2^n) \not\subseteq \text{Size}(2^{c \cdot n})}$$

"Hierarchy" & "advice"

Constructing a CANON. DERANDOMIZER

Omitted: Worst-case Hardness $\Rightarrow$ Average-Case HARDNESS

$$\exists f \in E \text{ s.t. } \forall\, 2^{\Omega(m)}\text{-size circuit } C_m$$
$$\text{Prob}_{x \in \{0,1\}^m}\left[C_m(x) = f(x)\right] < \frac{1}{2} + 2^{-\Omega(m)}$$

constr.

$$G(\underbrace{s}_{k}) = f(s|_{I_1}) \cdot f(s|_{I_2}) \cdots f(s|_{I_{\ell(k)}}) \quad \text{where}$$

- compute $I_1, I_2, \dots, I_{\ell(k)}$
- evaluate f on $\ell(k)$ points

$I_j \subseteq [k]$

$|I_j| = m$

$\forall j \neq j' \quad |I_j \cap I_{j'}| \le m' \ll m$

time $\sim 2^{O(m)} \gg \ell(k) = 2^{\Theta(k)} \sim (\text{circuit-size})^{1/2}$

$= \exp(O(k))$

Pseudorandomness $\Leftrightarrow$ Unpredictability

$\rightarrow$ Obvious (since UNIFORM is unpredict.)

$\leftarrow$ see next & use this!

WARM-UP: Suppose (I_j)'s are Disjoint.

INTUITION TO REAL CASE:

Small intersections "bound" the gain from $f(s|_{I_1}) \cdots f(s|_{I_j})$ towards predicting $f(s|_{I_{j+1}})$

"limited" depend. on $s|_{I_{j+1}}$

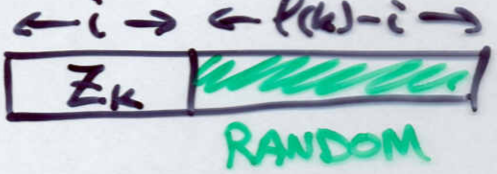
Unpredictability $\Rightarrow$ Pseudorandomness

Suppose $\{Z_k\}$ is <u>not</u> pseudorandom; i.e.

$\exists$eff A s.t. $\{Z_k\} \overset{A}{\not\equiv} \{U_{\ell(k)}\}$.

Consider HYBRIDS

$H_k^{(i)} =$



i^{th} hybrid

$\Downarrow$

(extreme HYBRIDS differ)

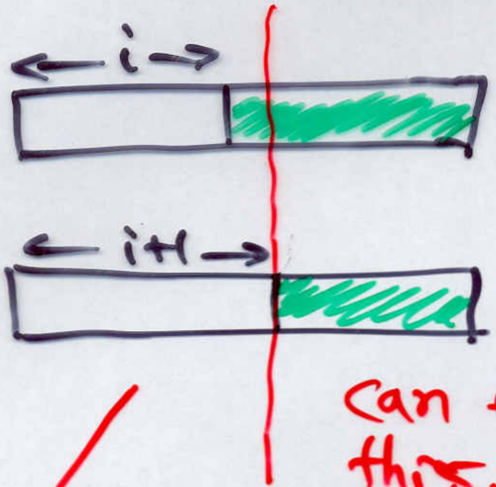
Note $H_k^{(0)} \equiv U_{\ell(k)}$ & $H_k^{(\ell(k))} \equiv Z_k$

$\Downarrow$

$1/\ell(k)$ gap at Neighbor. HYBRIDS

$H_k^{(i)} =$

$H_k^{(i+1)} =$



can emulate this...

can distinguish $i{+}1^{st}$ bit from a random value (when given i-prefix (of Z_k))

$\Downarrow$

can predict $i{+}1^{st}$ bit

PRGs FOR SPACE-BOUNDED DISTINGUISHERS

THM: Every Prob. Poly-Time algorithm can be emulated by a PPT algorithm of RANDOMNESS $= O(|\text{input}| + \boxed{\text{original space complex.}})$

Conj: Similar with RANDOMNESS $= O(\log|\text{input}| + \boxed{\text{original space complex}})$

Support

(1) a PRG with $|\text{seed}| = (\text{space complex})^2$

(2) $BPL \subseteq SC \triangleq TiSp(\text{poly}, \text{polylog})$

(3) UCONN $\in$ L [2005]

RL [1979]

(2') $BPL \subseteq DSPACE((\log)^{1.5})$

SPECIAL-PURPOSE PRGs

PROJECTION TESTS $\Rightarrow$ t-wise indep. PRG

$$G(s_0, \ldots, s_{t-1}) = \left(\sum_{j=0}^{t-1} s_j \cdot \alpha_i^j \right)_{i=1,\ldots \ell(k)}$$

$s_0, \ldots s_{t-1}, \alpha_i$ etc are field elements

• APPLICATIONS

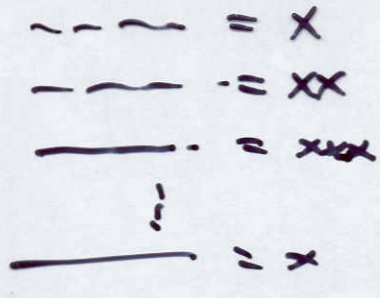
LINEAR TESTS $\Rightarrow$ small-bias PRG

$$G(s,f) = \mathrm{LFSR}_f(s)$$

FEEDBACK RULE $\to f$

s $\leftarrow$ START SEQ.

Application: "PCP of linear system" (ditto quad. sys.)



$\Longrightarrow$ (few) linear combinations of the rows

HITTING TESTS $\Rightarrow$ EXPANDER WALK PRG

$$G(s, \sigma_1, \ldots, \sigma_{\ell-1}) = \underbrace{(v_0, v_1, \ldots, v_{\ell-1})}_{\text{seq. of vertices}}$$

$$v_{i-1} \xrightarrow{\sigma_i \in [d]} v_i$$

$\forall S \subseteq$ VERTEXSET of density $\geq 1/2$

$$\mathrm{Prob}\left[\begin{array}{l} \text{sequence does not} \\ \text{hit the set } S \end{array} \right] < 2^{-\Omega(\ell)}$$

Some Credits

Comput. Indist. ~ [Goldwasser + Micali] + [Yao]

Gen. Pur. PRG ~ [Blum + Micali] + [Yao]

Construction of gen. pur. PRG

- hardcore + iterations [BM]
- hardcore for any OWF [Goldreich + Levin]
- The Char. THM. [Hastad, Impagliazzo, Levin + Luby]

Canonical Derandomizers [Nisan + Wigderson]

$E \notin \mathrm{Size}(2^{\Omega(n)}) \Rightarrow BPP = P$ [Impag. + Wigderson]

PRG for Space-Bounded Disting.
[Nisan + Zuckerman], [Nisan]^2, [Reingold]

Special-Purpose PRGs

- k-wise [Chor + Goldreich] + [Alon, Babai + Itai]
- small-bias [Naor^2] + [Alon, Goldreich, Hastad + Peralta]
- Expander Walk [Ajtai, Komlos + Szemeredi]

More details/material @

http://www.weizmann.ac.il/~oded/pp_pseudo.html