Cryptography 10 Years Later, Boot Camp

# Foundations

Iftach Haitner

Stellar Development Foundation
& Tel Aviv University

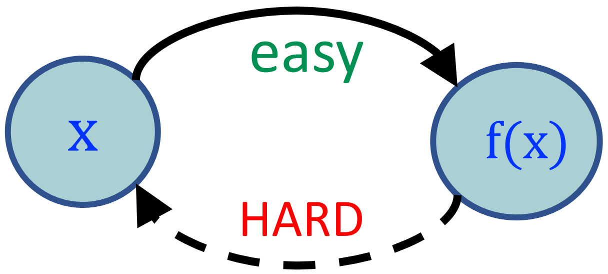
# Talk roadmap

- Minicrypt
- Computational correlation/Public-key world/Cryptomania

Computational analogues of entropy

# Minicrypt

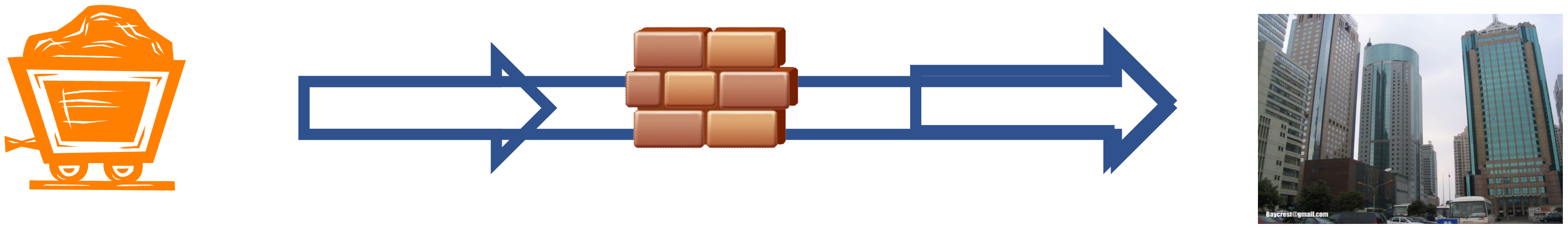
# One-way functions (OWFs)




- Easy to compute
- Hard to invert (even on the average)
- Poly-time $f: \{0,1\}^n \mapsto \{0,1\}^n$ is one-way if ∀PPT $A$:

$$\Pr_{y \leftarrow f(U_n)}[A(y) \in f^{-1}(y)] \le \mathrm{negl}(n)$$

- Unstructured
- Implied by most crypto
- Much of crypto can be based on the existence of OWFs



OWF

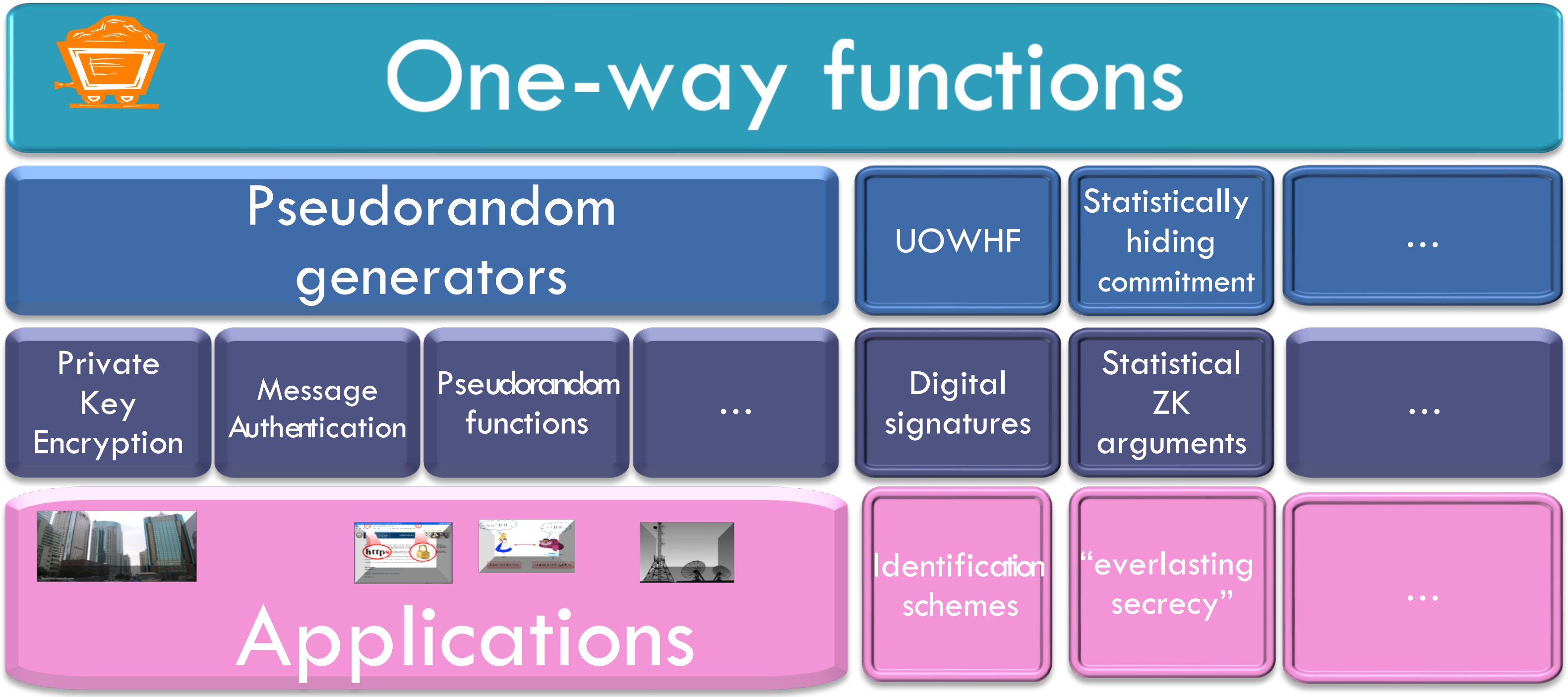
Intermediate primitives

# Pseudorandom generators [BM, Yao 82]

Poly-time function $G: \{0,1\}^s \mapsto \{0,1\}^m$


- Stretching ($m > s$)
- Output is computationally indistinguishable from uniform
  - No PPT distinguishes $G(U_S)$ from $U_m$ (with more than $\mathrm{negl}(m)$ advantage)

# OWF-based cryptography



# OWF → PRG

[BMY 82], [GKL 90], [HILL 91], [Hol 06], [HHR 06], [HRV 10], [VZ 11], [YLW 15], [MZ 22], [MP 22]

Key concepts:

- Leftover hash lemma
- Randomness extractors
- Pseudoentropy
- Next-block pseudoentropy
- KL hardness

# Pseudoentropy generator [HILL 91]

n+|h| bits of entropy

n+|h| + 1 bits of **pseudoentropy**

Looks uniform

g(x,h) := | f(x) | h | $h(x)_{1...d(x)}$ | $+1$ |

?

Goldreich-Levin hardcore bit

- f:$\{0,1\}^n \mapsto \{0,1\}^n$ is OWF
- h is $n \times n$ Boolean matrix, $h(x) := h \times x$ mod
- d(x) := log|$f^{-1}(f(x))$|

Might be inefficient to compute

**Cl**... h,$h(x)_{1..d(x)}$ is (almost) injective

What is the **entropy** of g'(x,h)? (over uniform inputs)

**Claim:** g' is one way

**Pf:** Leftover Hash Lemma

Y is g(x,h) with $h(x)_{d(x)+1}$ replaced with a uniform bit

The (Shannon) entropy of X is
$H(X) := E_{x \leftarrow X}[\log(1/\Pr[X=x])]$
"Unpredictability of X"

X has pseudoentropy k if ∃ Y
1. $X \approx_c Y$
2. H(Y) = k

# Pseudoentropy generator [HILL 91], cont.

$$g(x,h) = f(x), h, h(x)_{1..d(x)+1}$$

$$g(x,h,i) = f(x), h, h(x)_{1..i}$$

Pseudoentropy gap = (output) pseudoentropy – (output) entropy = 1/n

**Disadvantages:**

1. Pseudoentropy gap is small
2. Output pseudoentropy $<$ input entropy
3. Value of output pseudoentropy is **unknown**

Yet, using information theoretic tools (repetitions and extractions) implies PRG, but rather complicated and inefficient

- # of $f$-calls: $n^8$
- Seed length: $n^8$

# But what if we do not truncate?

g(x,h) := | f(x) | h | h(x) |

**Nonsense:** g is invertible and therefore has no pseudoentropy gap

Well yes, but g does have pseudoentropy gap “in the eyes of an online observer”

# Next-block pseudoentropy [HRV '10]

- $H(X) = k \quad \Leftrightarrow \quad \sum_i H(X_i | X_{<i}) = k$
  - $X_{<i} := X_1, \ldots, X_{i-1}$

$H(A|B) := E_{b \leftarrow B}\left[A\big|_{B=b}\right]$

- $X$ has "next-block entropy" $k$ in the eyes of **online** (unbounded) observer

$X = (X_1, \ldots, X_n)$ has next-block pseudoentropy $k$ if $\exists$ (jointly dis.) $Y = (Y_1, \ldots, Y_n)$ s.t:

- $(X_{<i}, X_i) \approx_c (X_{<i}, Y_i)$
- $\sum_i H(Y_i | X_{<i}) \geq k$

I.e., $X_i$ is somewhat hard to predict given $X_{<i}$

- Quantitative variant of Yao's unpredictability
- Might be larger than pseudoentropy!

# Next-block pseudoentropy of g

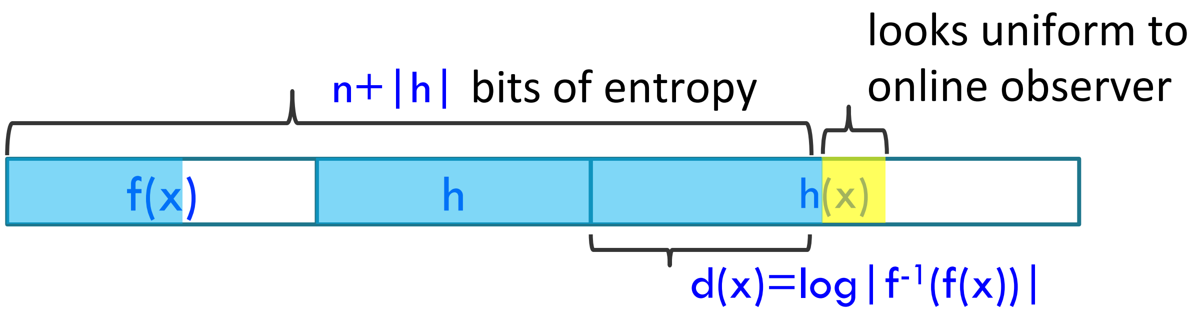
$X$ has NB pseudoentropy $k$ if $\exists Y$ s.t:

- $(X_{<i}, X_i) \approx_c (X_{<i}, Y_i)$
- $\sum_i H(Y_i | X_{<i}) \geq k$

**Claim**: Output of g has next-block pseudoentropy $n + |h| + 1$

**Proof:**

g(x,h) :=

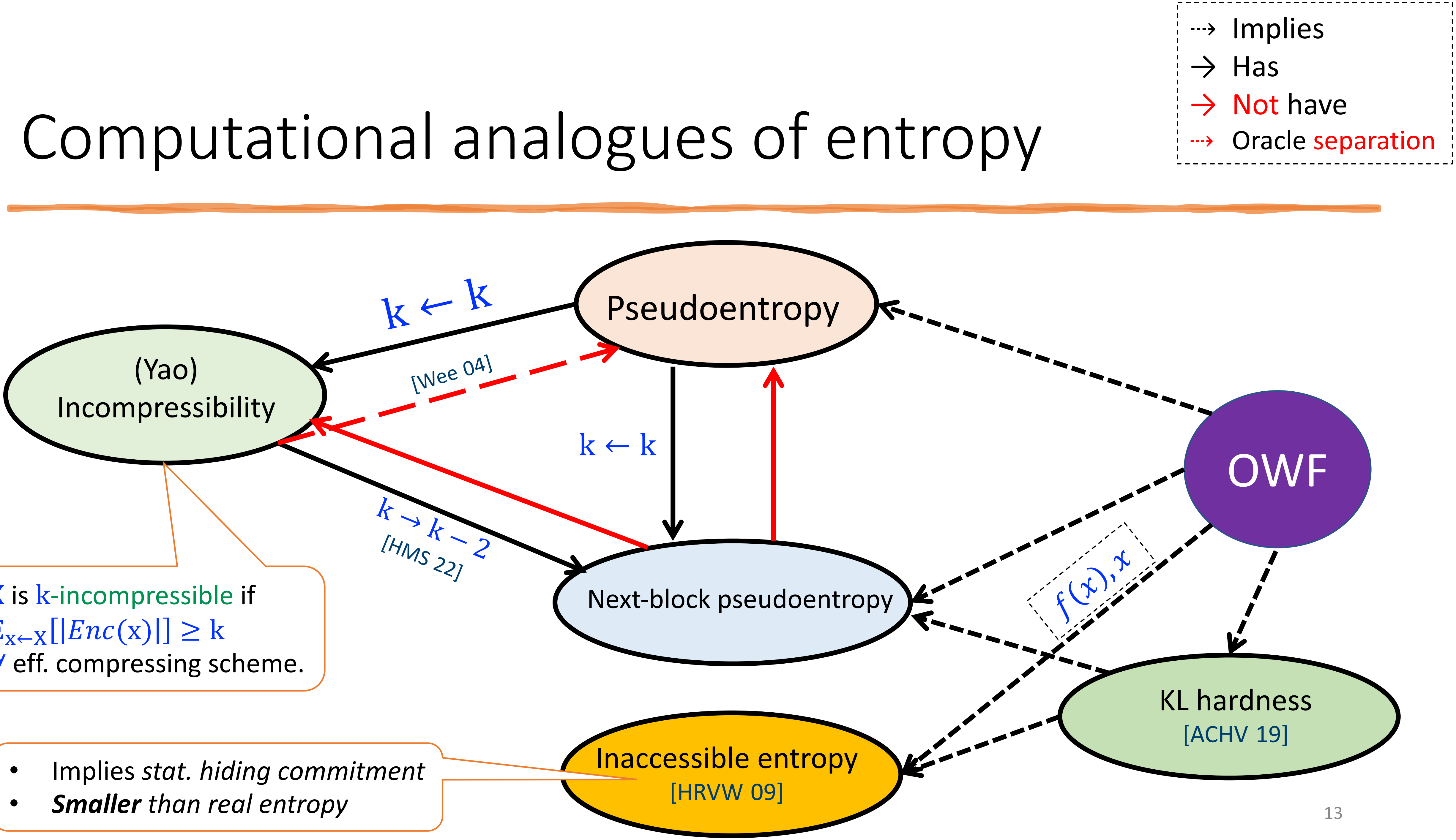



Y is set to g(x,h) with $h(x)_{d(x)+1}$ replaced by a uniform bit

- Jointly distributed with g(x,h)

- Leads to significantly more efficient PRG (seed length and # of $f$ calls $n^3$)
- [VZ 11]: (f(x),x)
- [MP 22]: Simpler, yet useful, notion of next-block pseudoentropy

# Computational analogues of entropy




X is k-incompressible if $E_{x \leftarrow X}[|Enc(x)|] \geq k$ $\forall$ eff. compressing scheme.

- Implies *stat. hiding commitment*
- ***Smaller*** *than real entropy*

# Efficiency lower bounds

The best OWF-based PRG

- Has seed length $n^3$
- Makes $n^3$ calls to $f$

Can we do better?

What does it mean?

# Bounds on **black-box** reductions

"Reductions"

- Construction: for any eff. $f$ exists eff. G
- Security proof: If G is broken then $f$ is not one-way

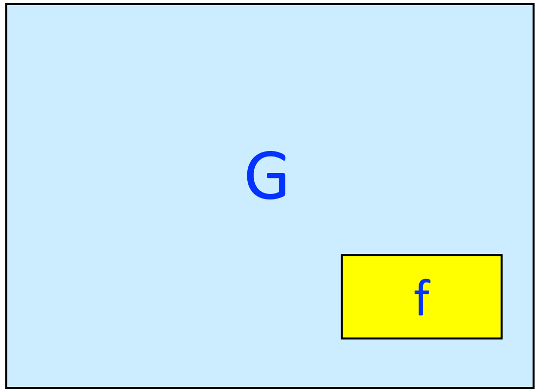
Too general to refute

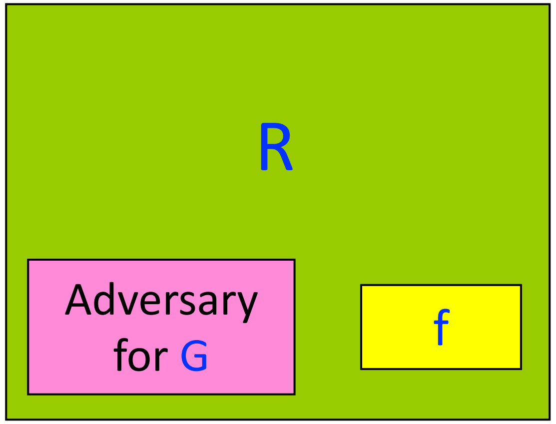
Black-box reductions

- Construction: G makes **oracle** use of $f$
- Security proof: Eff. R that makes **oracle** use of $f$ and the adversary A
- G and R should work for any (even inefficient) $f$ and A

[GT 01]: **Length-doubling** PRG makes $\Omega(n/\log n)$ $f$-calls

- Even if $f$ is one-way permutation



# Random permutations are exp. hard to invert [GT 01]

**Thm.** Whp over permutation $f: \{0,1\}^n \mapsto \{0,1\}^n$, a $2^{o(n)}$-query A inverts $f$ wp $2^{-\Omega(n)}$

**Pf**: Assume A makes no $f$-calls

- How many permutations A inverts w.p. 1?
  - One, since A determines $f^{-1}$
- How many $f$'s algorithm A inverts w.p $\epsilon \gg 2^{-n}$?
  - A partially determines $f^{-1} \rightarrow$ cannot hold for many $f$'s
- Slightly more complicated argument when A does make $f$-calls

# Length-doubling PRG makes $\Omega(\frac{n}{\log n})$ $f$-calls [GT 01]

| $g(x_1, x_2) =$ | $f(x_1)$ | $I(x_2)$ |
|---|---|---|

Let $g: \{0,1\}^n \mapsto \{0,1\}^n$ be a **concatenation** of

- Random permutation $f: \{0,1\}^{\omega(\log n)} \mapsto \{0,1\}^{\omega(\log n)}$
- The identity function $I: \{0,1\}^{n-\omega(\log n)} \mapsto \{0,1\}^{n-\omega(\log n)}$

**Claim**: $g$ is one-way: whp over $g$, a $poly(n)$-query A inverts g wp $negl(n)$.

- Let $G: \{0,1\}^n \mapsto \{0,1\}^{2n}$ be BB PRG that makes $o(\frac{n}{\log n})$ $f$-calls
- Let $G': \{0,1\}^{s<2n} \mapsto \{0,1\}^{2n}$ be variant of $G^g$ that samples the answers of $g$-calls by **itself** (using randomness given as additional input)

**Claim:** ∃ (unbounded) $D$ that tells $G'(U_s)$ from $U_{2n}$

→ $D$ tells $G^g(U_n)$ from $U_{2n}$

→ $R^{g,D}$ inverts $g$

- But $R^{g,D}$ makes $poly(n)$ # of $g$-calls

# Lower bounds on black-box reductions cont.

- [HS 12]: **Any** PRG makes $\Omega(n/\log n)$ calls
  - Even if $f$ is unknown regular
- [CGVZ 18]: Seed length $\Omega(n^3)$ for **certain** PRG constructions
- Many other lower bounds on the (BB) complexity OWF-based UOWHF, commitments schemes, and more
- Many open questions

# Missing lower bound: Weak-OWF amplification

Weak OWF: $\forall$ PPT $A$

$$\Pr_{y \leftarrow f(U_n)}[A(y) \in f^{-1}(y))] \leq 1 - \delta$$

- Can we construct OWF out of f?
- [Yao82]: Yes, $g(x_1, \ldots, x_\ell) := f(x_1) \ldots, f(x_\ell)$ for $\ell = \omega(\log n)/\delta$
- If $f: \{0,1\}^{100} \mapsto \{0,1\}^{100}$ and $\delta = 2^{-10}$, input length of $g$ is about $10^5$
- [GILVZ 90, HHR 06]: Input length $O(n)$ for unknown regular f
- [LTW 05]: $\ell$-queries is needed for BB reductions
- [BCKR 22]: Seed length $\ell$ needed for non-adaptive, non-post-processing, BB reductions

# Necessity of one-way functions

In "most" cryptographic primitives there is a hidden OWF

- What is the OWF in PRG $\mathrm{G}: \{0,1\}^{\mathrm{s}} \mapsto \{0,1\}^{\mathrm{m}}$?
- In commitment schemes, key-agreement, oblivious transfer?
  - In $\mathrm{G}_1, \mathrm{G}_2: \{0,1\}^{\mathrm{m}} \mapsto \{0,1\}^{\mathrm{m}'}$ s.t $\mathrm{G}_1(U_m)$ and $\mathrm{G}_2(U_m)$ are statistically far but comp. indiguishable?
  - Is it $\mathrm{G}(\mathrm{x}, \mathrm{b}) := \mathrm{G}_b(x)$?
  - What if $\mathrm{G}_1$ and $\mathrm{G}_2$ have the same support?
  - [IL '89]: $\forall \mathrm{f}\ \exists \mathrm{f}'$ such that: $\mathrm{f}'$ is not one-way → $\mathrm{f}$ is distributional invertable
- In coin-flipping protocols?
  - No single-attacking-point
  - Attack changes the object

Sampling random preimage is easy

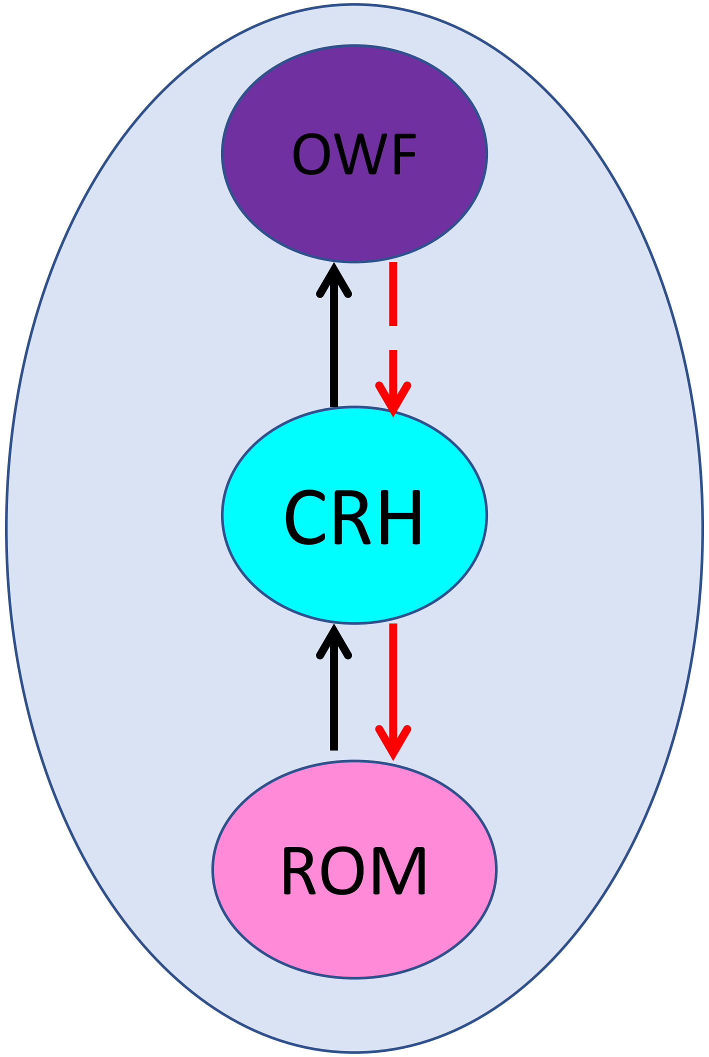
# Additional open questions (for OWFs)

- Simpler constructions
- Matching BB lower bound  for PRG, UOWHF, …

Many other gaps…

# Minicrypt beyond OWFs

- One-way permutations
  - Injective OWF
- Collision resistant hash
  - Assumption of different nature
  - Implies OWF
  - [Simons 98]: Not implied by OWF in a black-box way
- Random Oracle Model (ROM)
  - Parties have oracle access to a random function, adversaries are computationally unbounded
  - Extremely popular (random oracle heuristic)
  - Is it in minicrypt?



# Beyond minicrypt

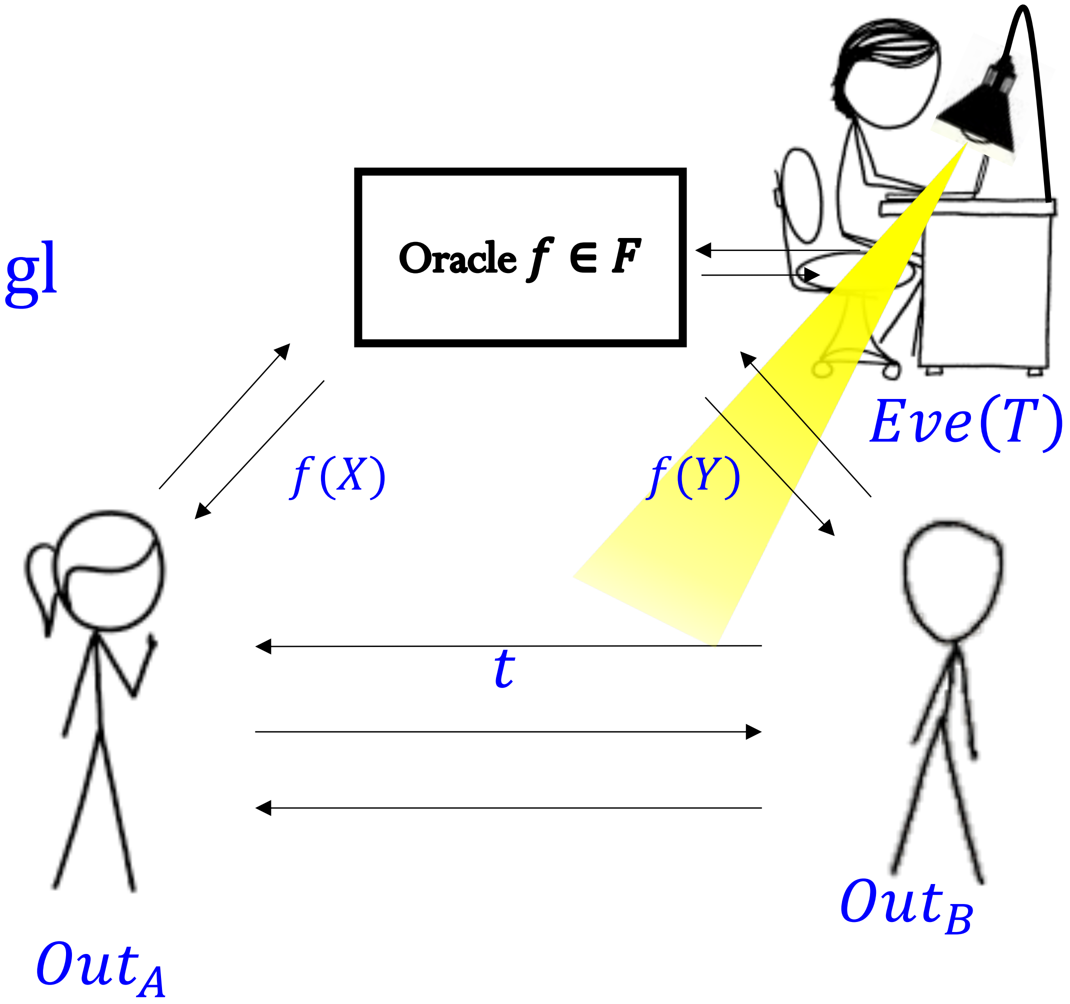
# Key agreement is not in minicrypt

Key agreement

- $\Pr[Out_A = Out_B] \approx 1$
- For any PPT E: $\Pr[E(t) = Out_A] \leq$ ½ + negl

- Can we construct KA from a minicrypt primitive?
- [IR 89, BM 09]: No KA in the ROM

→ No black-box reduction from OWF/CRH to KA



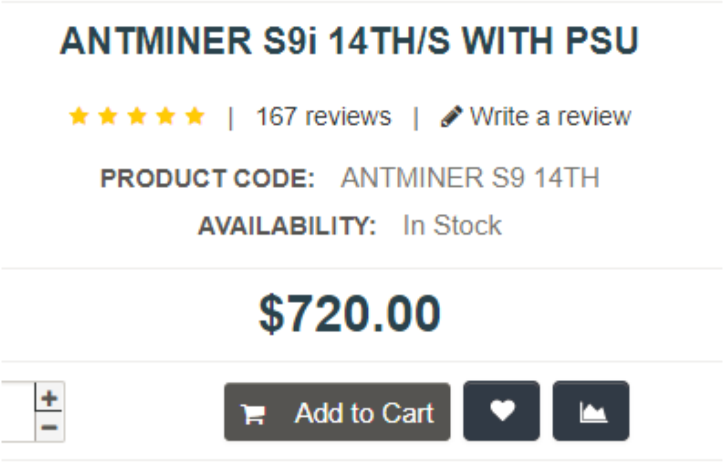
# Key agreement is not in minicrypt?

*Merkle-puzzle: $\ell$-query ROM KA that takes* $\ell^2$ queries to break

Using specialized hardware for computing SHA-256

- $10^{13}$-query to SHA-256 takes one second!
- Eve needs 1,000,000 years to break 1-sec Merkle-puzzle

Seems suffice, but Alice needs to send 100 TB ☹




[HMORY 19]: Communication of Merkle's Puzzle is **optimal** for limited family of protocols: two-message non-adaptive KA

[HMYZ 23]: For non-adaptive perfect KA

Question is still wide open

# Characterizing not-minicrypt

- *"Cannot be implemented in the ROM"* is not very useful…
- "Public-key world" is not the right definition either,
  does not include many important protocols, e.g., key agreement.

In minicrypt [IL 89]: A poly-time $f$

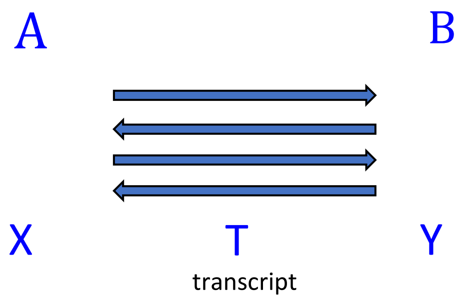
- Is distributionally invertible
- Or can be transformed into OWF

So, either $f$ is useless from cryptographic point of view, or it is as strong as OWF.

Goal: win-win dichotomy for not-minicrypt

# Two-party protocols w/ single-bit output

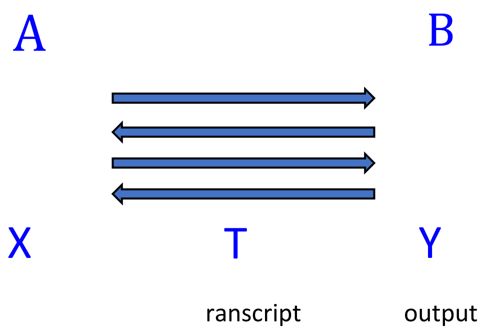
- Two-party protocol $(A, B)$
- Parties interact
- Each party outputs a value




- Can X and Y be correlated given T?
  - $I(X;Y|T) > 0$; $X$ and $Y$ are dependent given $T$
- **No**
- But can they be computationally corrlated?
- What does that mean?

# Key-agreement, revisited

- Eff. two-party protocol $(A, B)$
- Parties interact
- Each party outputs a single bit




Agreement: $X = Y$

- Since I(X;Y|T)= 0, T **determines** X and Y

Secrecy: $\forall$ ppt E: $\Pr[E(T) = X] \leq \frac{1}{2} + \text{negl}$

- X and Y, given T, are highly correlated **in the eyes of efficient observer**

Can we generalize this phenomena?

# Uncorrelated protocols

**Dfn:** protocol $\Pi = (\mathrm{A}, \mathrm{B})$ is uncorrelated if $\exists$ eff. $\mathrm{Decr}$ (decorlator) s.t:

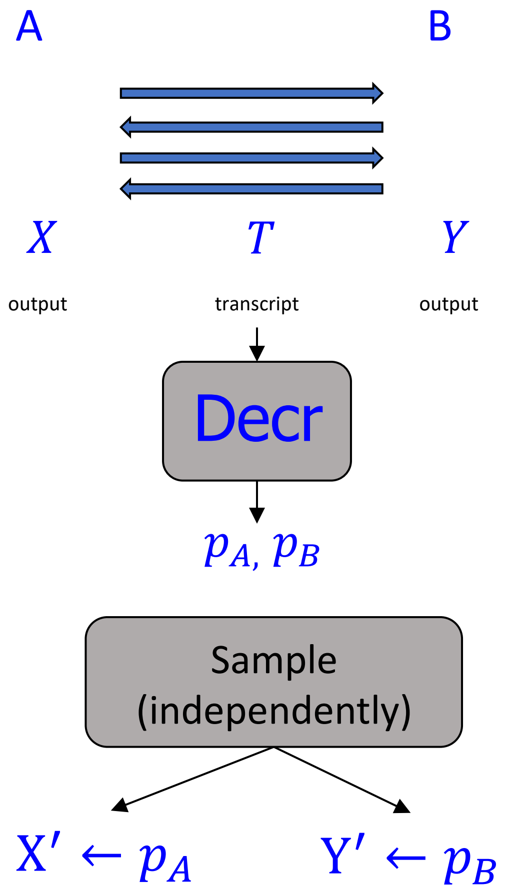
1. $(X, Y, T) \leftarrow \Pi$
2. $(\mathrm{p_A}, \mathrm{p_B}) \leftarrow \mathrm{Decr}(\mathrm{T})$
3. $X'_k \leftarrow p_A$ and $Y'_k \leftarrow p_B$

Then $(X, Y, T) \approx^C (X', Y', T)$

Uncorrelated protocols can be simulated

- (cryptographically) useless

- Key agreement is "highly correlated"
- Are there protocols in between?




$(X, Y, T) \approx^C (X', Y', T)$

*real* *simulated*

# key-agreement dichotomy

[HNOSS 18]: **Every** efficient (single-bit) two-party protocol is either **uncorrelated** or can be transformed into **key-agreement**

No intermediate concept!

😀 Holds in ROM

😢 Only holds for (any) constant distinguishing gap

😢 Only for single-bit output protocols

# Oblivious transfer dichotomy?

Oblivious transfer (OT): *receiver* learns **one** of two strings held by sender, w/o revealing which

- **Complete** functionality for MPC [GMW 87]
- Rich set of theoretic and practical applications
- Can we find dichotomy for OT?
    - Trivial from insider point-of-view: can be simulated using KA
    - Or implies OT
- Barrier: OT is rather **poorly understood** even information theoretically
    - Specifically, **0/1 rule** is proved using the parties' view

# Summary

Foundation of cryptography is about

*Deeply understanding the fundamental primitives and concepts*

Many exciting questions are still open