OWF vs. TFNP: Simpler and Improved

Folwarczny, Goos, Hubacek, Maystre, Yuan

Kashvi Gupta, Tianqi Yang

Overview

Discussions from Previous Classes

- We saw that if NP is hard-on-average, and no OWF → TFNP is hard-on-average
- If NP is hard-on-average → TFNP/poly is hard-on-average

Natural Gap

NP is hard-on-average, and no OWF → TFNP is hard-on-average

NP is hard-on-average → TFNP/poly is hard-on-average

?

OWP exist → PPP is hard-on-average (better than TFNP hard-on-average)

What can we say about TFNP assuming OWF exist?

Goal of the Paper

Either:

1) Construct a problem in TFNP that is hard-on-average assuming OWF exist
2) Show that it cannot be done

This paper: shows that this cannot be done using a specific kind of black-box reductions UNLESS it meets specific conditions

Black Box Reductions

P ⇒ Q is fully black-box if

CONSTRUCTION

F

F is assumed construction for P

G

G is a construction for Q w/ oracle

SECURITY PROOF

F

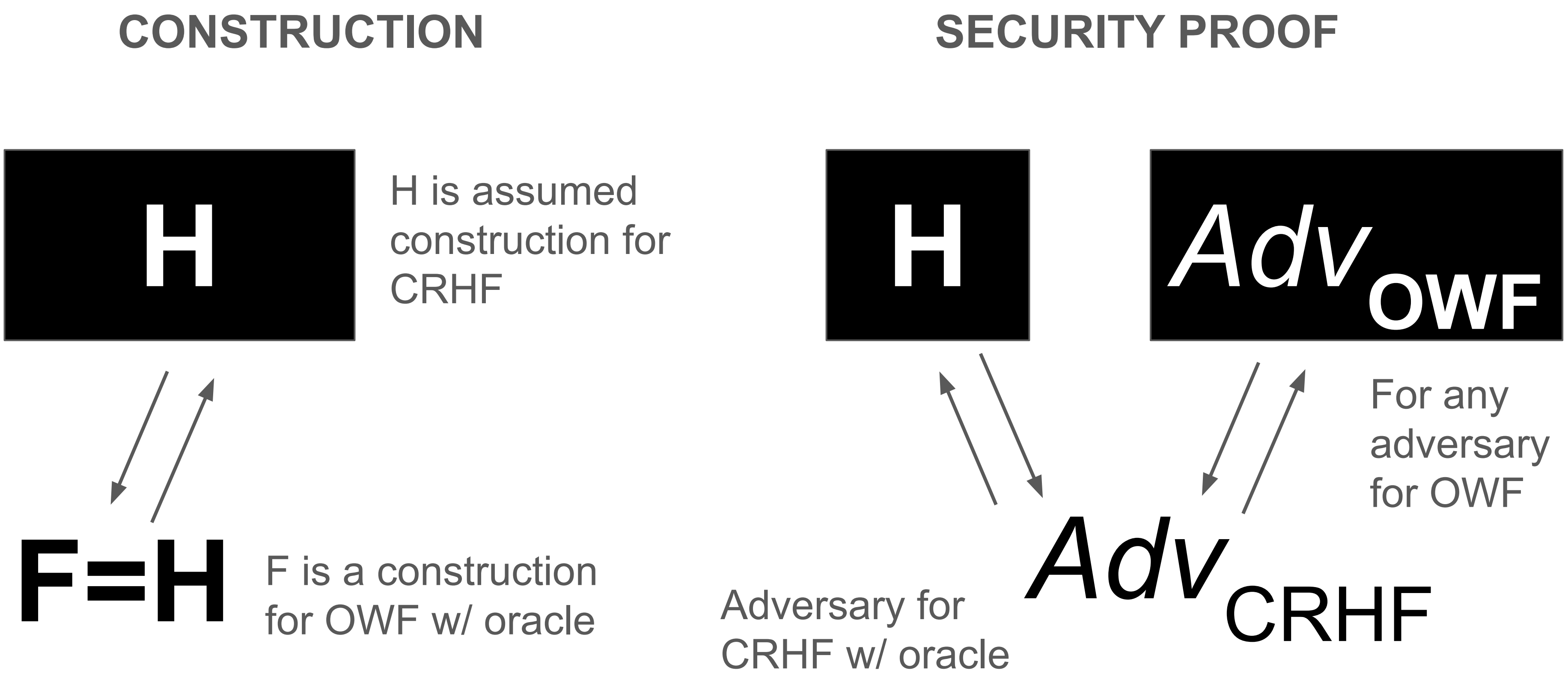
Adv_G

For any adversary for Q

Adv_F

Adversary for P w/ oracle

Example: ∃ CRHF ⇒ ∃ OWF



∃ CRHF ⇒ ∃ OWF (Black Box Security Proof)

Adv_{OWF}: f(x) → preimage x'

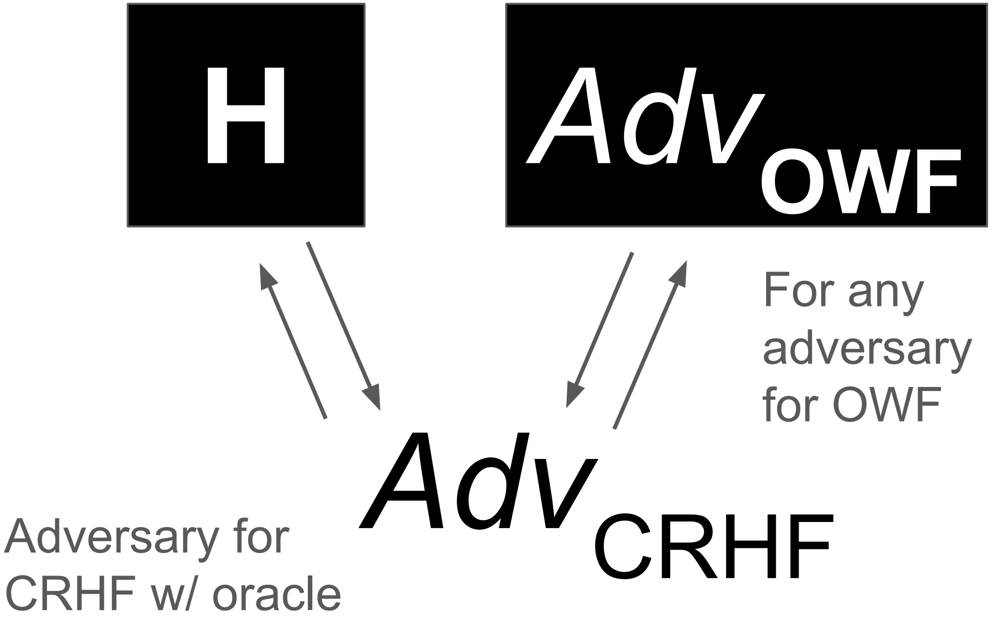
Adv_{CRHF}:

1. Sample $x \in \{0,1\}^n$
2. Query H(x) → y
3. Query Adv_{OWF} → x'
4. Output (x,x')

With non-negligible probability, $x \neq x'$ and $H(x)=H(x')$

(recall H is shrinking, eg $H:\{0,1\}^{2n} \rightarrow \{0,1\}^n$)

SECURITY PROOF



Why do we care about black box reductions?

- Most of the techniques we know in cryptography
- Relate many primitives to each other

- For this paper, gives two options
  - +) Hint to what proof of security looks like
  - -) Important first step to proving no black box construction is possible

? ? ?

$$\exists\ \text{OWF} \Rightarrow \exists\ \text{CRHF}$$

Simon Says …

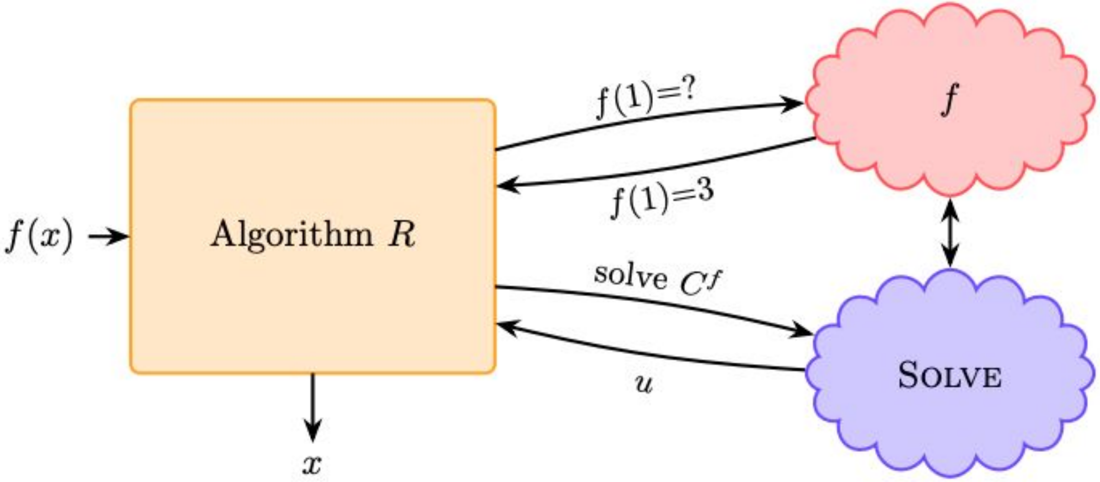
… it’s impossible to construct CRHF from OWF using BBR

- Can we define an Oracle O relative to which OWF exist, but CRHF do not exist?

… it's impossible to construct CRHF from OWF using BBR

- Can we define an Oracle O relative to which OWF exist, but CRHF do not exist?
- Define 2 oracles (f, SOLVE) such that they satisfy
  - **Random Injection**: Oracle f : $\{0,1\}^* \rightarrow \{0,1\}^*$ is an injective black-box function mapping n-bit strings to (n+1)-bit strings
  - **Collision finder**: Oracle SOLVE : $\{0,1\}^* \rightarrow \{0,1\}^*$ is a black box function that can find collisions in any shrinking function
  - **One-wayness**: f is one-way even in the presence of SOLVE. ie: if given f(x) for a randomly chosen x, no ppt algorithm given f(x) and (f, SOLVE) can output x with non-negligible probability

… it's impossible to construct CRHF from OWF using BBR



Generalizing Simon to TFNP

Simon says there is no construction of CRHF (TFNP problem) from OWF via black box reduction $\rightarrow$ maybe this implies there is no black box reduction of any TFNP problem from OWF?

Generalizing Simon to TFNP

- There exists a pair of oracles (f, SOLVE) satisfying
  - **Random injection**: Oracle f
  - **TFNP Solver**: Oracle SOLVE : special oracle that can find solution to any TFNP problem
  - **Single Query One-wayness**: f is one-way if the reduction calls SOLVE one time, before ever calling f

Main Conclusion of this Paper

Can you show in a black-box way that OWF implies TFNP hard on average? This paper says partial no.

What this paper proves: If a black box reduction exists, it must make multiple queries or it has to make a query to OWF before calling TFNP solver (they rule out any reduction that simultaneously satisfies both conditions)

Equivalently: If it calls TFNP solver before OWF and makes only a single query it is not a viable black box reduction.

Stability Lemma

OWF in Random Oracle Model

- Probability[$Adv^f(y)$ finds a preimage of y] ≤ negligible
  - Assuming $y = f(x)$ where f is a random oracle OWF

- What if adversary has access to SOLVE as well?
  - Pr[$Adv^{(f,SOLVE)}(y)$ finds a preimage of y]
  - We can't say if this is also negligible probability
  - This is where stability lemma comes in

Stability Lemma

Lemma 2 (Stability Lemma)**.** *There exists an oracle* SOLVE *satisfying* **(2⁺)** *such that for every* $y \in [2N]$ *and every satisfiable circuit* C^f *of size* t,

$$\Pr_{\substack{f \sim \mathcal{F}^{-y} \\ x \sim [N]}} [\textsc{Solve}(C^f) \neq \textsc{Solve}(C^{f_{x \to y}})] \leq O(t/N^{1/2}).$$

Key Takeaway

- Stability Lemma says SOLVE tells us nothing about how to get the preimage of $y = f(x)$ where f is random oracle OWF
  - Caveat: can only make 1 query to SOLVE, multiple queries could leak information
- "Preserves one-wayness of (f, SOLVE) system"

Main Theorem

Given the oracle (f,SOLVE), no adversary R that can access SOLVE once before any access to f, can invert f.

3 Games

- Sample a random f and a random x.
- Adversary R wins the game if they can correctly invert y = f(x) to get the pre-image x

3 Games

- H_1: R gets the oracle for f, SOLVE and y = f(x)
- H_2: sample y in non-image(f) and give R $f_{x \to y}$, SOLVE and y
- H_3: sample y in non-image(f) and give R f, SOLVE and y
  - Note: probability of winning H_3 is 1/(N) where N is the size of the domain, because there is no y to invert in the image. Ie: for our case, it is negligible

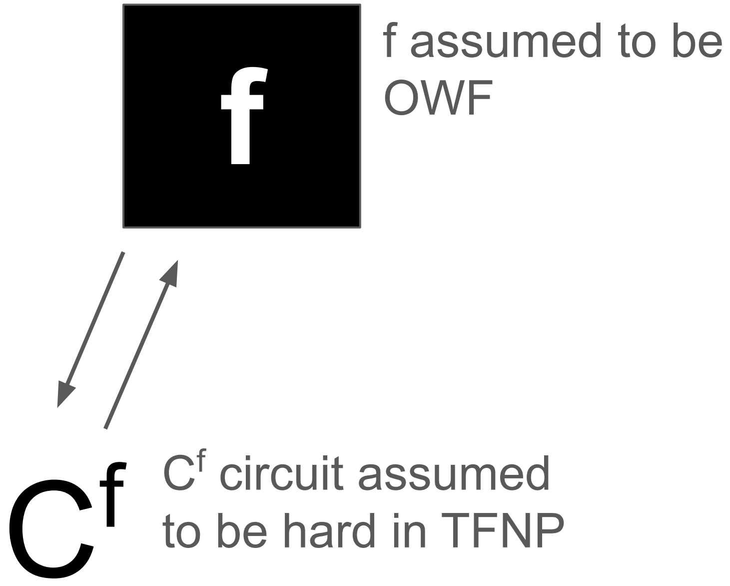
Probability of Winning

- $\Pr[R \text{ wins } H_1] = \Pr[R \text{ wins } H_2]$ b/c they have the same distribution
- $|\Pr_{f,x,y}[R \text{ wins } H_2] - \Pr_{f,x,y}[R \text{ wins } H_3]| \leq \Pr_{f,x,y}[R^{(f\_x\to y,\ SOLVE)}(y) \neq R^{(f,SOLVE)}(y)]$
  - Applying stability lemma gives us
  - $|\Pr_{f,x,y}[R \text{ wins } H_2] - \Pr_{f,x,y}[R \text{ wins } H_3]| \leq$ negligible
    - $\Pr_{f,x,y}[R \text{ wins } H_3] =$ negligible
- $\Pr_{f,x,y}[R \text{ wins } H_2] \leq$ negligible

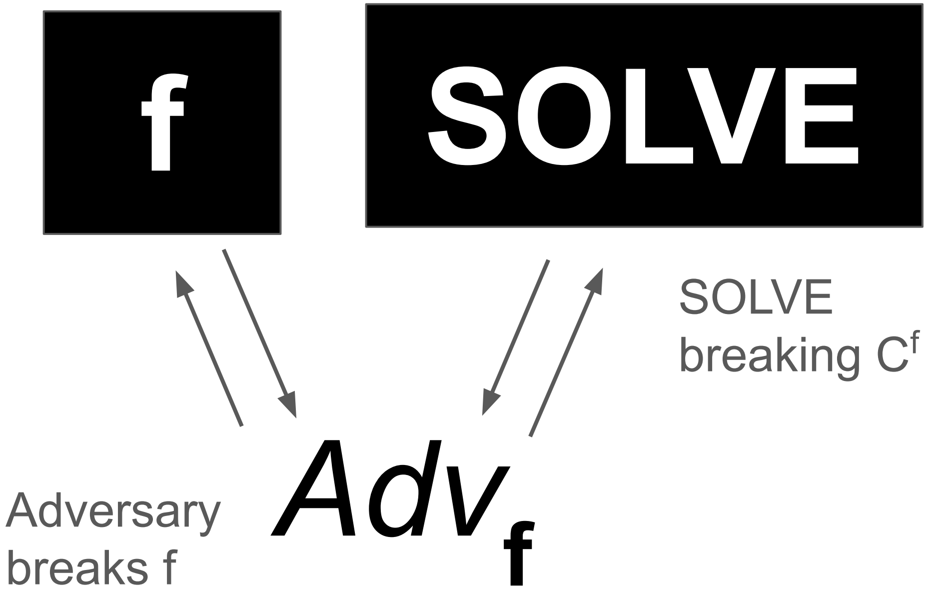
Constructing Black Box Reduction

Constructing the Reduction

CONSTRUCTION



SECURITY PROOF



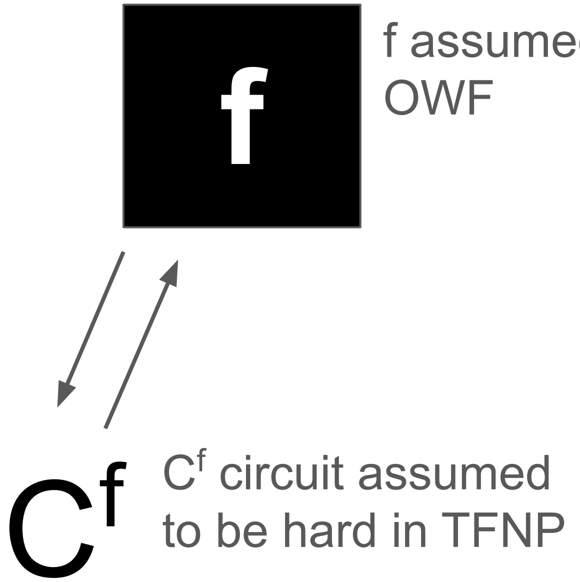
…HOWEVER

We just proved this is impossible via the stability lemma!

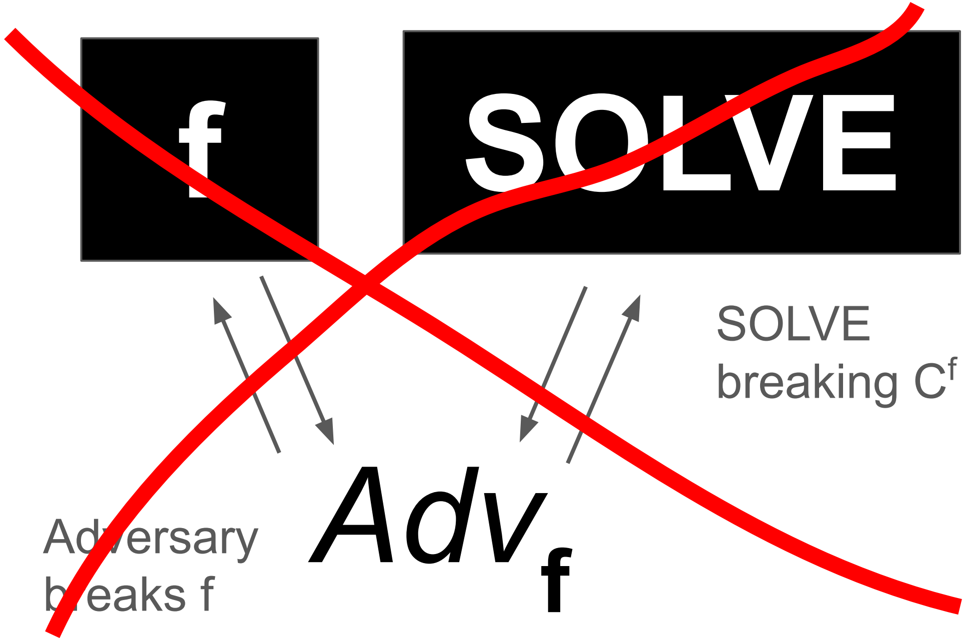
$\Pr[R \text{ wins } H_1] = \Pr_{\text{random } x}[\text{Adversary}^{(f,\ \text{SOLVE})} \text{ inverts } y] \leq \text{negligible}$

…THEREFORE

CONSTRUCTION



SECURITY PROOF



Conclusion

Final Takeaway

If a black box reduction exists, it must make multiple queries or it has to make a query to OWF before calling TFNP solver (they rule out any reduction that simultaneously satisfies both conditions)

Equivalently: If it calls TFNP solver before OWF and makes only a single query it is not a viable black box reduction.

Why is this negative result important?

“We do not conclude that researchers should give up on proving serious lower bounds. Quite the contrary, by classifying a large number of techniques that are unable to do the job we hope to focus research in a more fruitful direction.” ~ Razborov & Rudich