Today:

Worst Case vs. Average Case Hardness

Last time:

Boosting:

if can weakly learn

$\forall f \in \mathcal{C}$ + $\underline{\forall}$ dists $\mathcal{D}$ $\exists \gamma > 0$ s.t. given examples of f can output h s.t. $\Pr_{\mathcal{D}}[h(x) \neq f(x)] \leq \frac{1}{2} - \frac{\gamma}{2}$

then

$\Downarrow$

can strongly learn

given ε, $\mathcal{D}_0$ + as above can output h s.t. $\Pr_{\mathcal{D}_0}[h(x) \neq f(x)] \leq \varepsilon$

Important: description of output hypothesis is $\underbrace{\text{poly}\left(\frac{1}{\varepsilon}, \frac{1}{\delta}, \log|\mathcal{C}|, n\right)}_{\text{Call this fctn } S(\varepsilon, \delta, |\mathcal{C}|, n)} \times$ size of WL hypothesis

[Impagliazzo] $\Rightarrow$ $S(\cdot;\cdot;\cdot) \leq \frac{1}{\varepsilon^2 \delta^2}$

Yao's XOR lemma:

any hard fctn $f \to f'$ hard on ave

intuition:

δ- biased coin

predict correctly with prob $\geq 1-\delta$

k tosses:

predict parity

$\approx \frac{1}{2} + (1-2\delta)^k$ $\leftarrow$ need to guess each bit correctly

$\to \frac{1}{2}$ as $k \to \infty$

is k indep copies of f, k times harder?

??

matrix - vec mult $O(n^2)$

matrix - matrix <u>not</u> $\Omega(n^3)$

Plan

δ-hard fcfn f

boosting $\Rightarrow$ $\Downarrow$

$\delta'(\varepsilon, \delta)$-hardcore measure

$\Downarrow$

$2\delta'$-hardcore set

$\Downarrow$

$2\delta' + (1-\delta)^k$ hardcore on whole domain

any (poly time / fast / efficient) f wrong on some δ fraction of inputs

$\Downarrow$

any efficient program for f wrong on almost $\frac{1}{2}$ in this measure

$\Downarrow$

efficient program wrong on almost $\frac{1}{2}$ inputs in this set of inputs

$\Downarrow$

f^* hard on almost $\frac{1}{2}$ inputs (in whole domain)

More details

[will show hardness for ckts of size g as opposed to run times of Turing machines]

def $f : \{\pm 1\}^n \Rightarrow \{\pm 1\}$ is δ-hard distribution D for size g if $\forall$ Boolean ckt C with $\leq g$ gates

$$\Pr_{x \in_D \{\pm 1\}^n}[C(x) = f(x)] \leq 1 - \delta$$

need to get hard distribution with "enough" inputs, so need to quantify "size" of distribution

<u>def</u>, M measure $\leftarrow$ each x assigned wt $M(x)$

total wt of M

$\mu(M) = \sum_x M(x)$

$D_M(x) = \frac{M(x)}{\mu(M)}$

if $\Pr_{x \in D_M}[C(x) = f(x)] \leq \frac{1}{2} + \frac{\varepsilon}{2}$

$\forall$ ckts C of size $\leq g$

then f is ε- hardcore on M for size g } hardcore measure

need set of inputs on which can't do better than guessing:

<u>def</u> S set

f is "ε- hardcore on S for size g"

if $\forall$ ckts C of size $\leq g$

$\underbrace{\Pr_{x \in_u S}}_{D_M = U_S}[C(x) = f(x)] \leq \frac{1}{2} + \frac{\varepsilon}{2}$

"hard fctns have hard core measures"

Thm f δ-hard for size g on uniform dist } weakly ave case hard

let $0<\varepsilon<1$

then $\exists M$ s.t. $\mu(M) \geq \delta$ s.t.

f is ε-h.c. on M for size

wrong $\geq \frac{1}{2} - \frac{\varepsilon}{2}$ fraction of inputs

$$g' = \underbrace{\tfrac{1}{4}\varepsilon^2\delta^2}\cdot g$$ } ave case hard

Smaller than g
(parameters from [Impagliazzo])

Pf assume not

$\Rightarrow \forall M$ s.t. $\mu(M) \geq \delta$, f not ε-h.c. for g'

$\Rightarrow \exists$ "weak learner" i.e. ckt which predicts $\geq 1/2 + \varepsilon/2$ } def of h.c.

+ size $\leq g'$ on all M s.t. $\mu(M) \geq \delta$

W.L. from last lecture

need to check that WL never called on M with $\mu(M)<\delta$

$\Rightarrow \exists$ ckt of size g' predicts with error $\leq \delta$

total size $\leq s(\varepsilon,\delta,|C|,n) \leq \frac{1}{\varepsilon^2\delta^2}\cdot g' < g$

$\Rightarrow f$ not δ-hard for size $\geq g$ ∎

Hardcore measures $\Rightarrow$ hardcore sets

<u>Thm</u> μ is ε-h.c. measure for size

$$2n < g' < \frac{\varepsilon^2\delta^2}{8}\,\frac{2^n}{n}$$

lose factor 2

then $\exists$ $2\cdot\varepsilon$-h.c for set S for f

& size $\underbrace{g'}_{\text{lose nothing}}$ with $|S| \geq \delta\cdot 2^n$

<u>Pf.</u> # ckts of size $g' << \frac{1}{4}e^{2^n\cdot\varepsilon^2\delta^2}$

Pick S randomly according to D_μ

Show Pr[any C of size g' has $> 2\varepsilon$-advantage] small

using Chernoff & union bnd. $\blacksquare$

Yao's Xor lemma

f hard on hard core set $\Rightarrow$ fcfn f^k which is hard on whole domain given f

$f^{\oplus k}(x_1 \ldots x_k) = f(x_1) \oplus f(x_2) \oplus f(x_3) \ldots \oplus f(x_k)$

f is ε-H.C. on some set H

of size $\geq \delta 2^n$ for size $g+1$

$\Rightarrow f^{\oplus k}$ is $\underbrace{\varepsilon + 2(1-\delta)^k}_{\text{lose a bit here}}$-h.c. for size $\underbrace{g}_{\text{lose very little}}$

Proof. for contradiction

Assume ckt C s.t. $\leq g$ gates

& $\Pr_{x_1 \ldots x_k}[C(x_1 \ldots x_k) = f^{\oplus k}(x_1 \ldots x_k)] \geq \frac{1}{2} + \frac{\varepsilon}{2} + (1-\delta)^k$

Plan $\forall H$ s.t. $|H| \geq \delta 2^n$

will get ckt C' s.t. $|C'| \leq g+1$

which guesses f with prob $\underbrace{\geq \frac{1}{2} + \frac{\varepsilon}{2}}_{\text{so not } \varepsilon\text{-h.c.}}$ on H

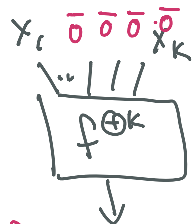
$\Rightarrow\Leftarrow$

a: plan that doesn't work:

assume $f^{\oplus k}$ not $\varepsilon + 2(1-\delta)^k$-h.c. for size g

give ckt for f with size $g+1$

<u>idea</u>



does pretty well

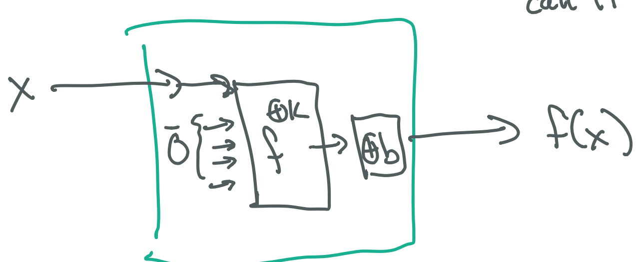
$f(x_1) \oplus \underbrace{f(\bar{0}) \oplus f(\bar{0}) \cdots \oplus f(\bar{0})}_{k-1}$

either $\underbrace{0 \text{ or } 1}_{\text{call it } b}$

hardwire $x_2 = \dots = x_k = \bar{0}$ ← or any arbitrary input

answer is $f(x_1) \oplus 1$ if $\bigoplus_{i=2}^{k} f(\bar{0}) = 1$

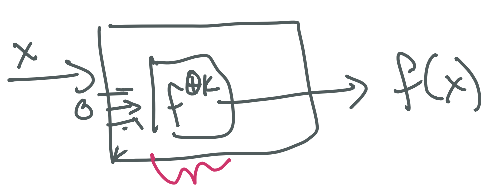
$f(x_1)$ if $= 0$

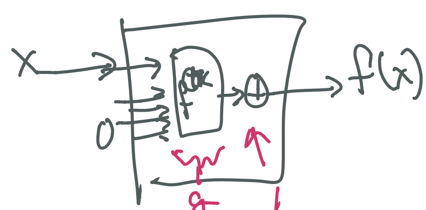


ckt for f

we don't know b,

but <u>one</u> of these works





$\Longrightarrow \exists$ ckt of size $\leq g+1$

PROBLEM!

$f^{\oplus k}$ might be really bad when

$x_2 = x_3 = \ldots = x_k = \bar{0}$

or any other fixed choice

only known to do well when all $x_1 \ldots x_k$ chosen randomly

Going back to proof: assume you know H (will do for all big enough H)

$A_m \equiv$ event that exactly m of $x_1 \ldots x_k$ in H

$\Pr_{x_1 \ldots x_k}[A_0] \leq (1-\delta)^k$ bad event: all x_i's easy unlikely

so $\Pr_{x_1 \ldots x_k}[C(x_1 \ldots x_k) = f^{\oplus k}(x_1 \ldots x_k) \mid \underbrace{\cup A_m \text{ for } m>0}_{\text{not } A_0}]$

$\geq \frac{1}{2} + \frac{\varepsilon}{2}$

so $\exists\ 1 \leq m \leq k$ by averaging s.t.

$\Pr_{x_1 \ldots x_k}[C(x_1 \ldots x_k) = f^{\oplus}(x_1 \ldots x_k) \mid A_m] \geq \frac{1}{2} + \frac{\varepsilon}{2}$ (*)

Construct Idealized ckt: for $x \in_u H$

Compute $f(x)$ as:

1. pick $x_1 \cdots x_{m-1} \in_R H$
2. pick $y_{m+1} \cdots y_k \in_R \overline{H}$
3. randomly permute

$(x_1 \cdots x_{m-1}, x, y_{m+1} \cdots y_k)$ via random permutation π

$$\Pr_{x_1 \cdots x_{m-1}\, x\, y_{m+1} \cdots y_k\, \pi}\left[C(\pi(x_i\text{'s}, x, y_i\text{'s})) = f^{\oplus k}(\pi(x_i\text{'s}\, x\, y_i\text{'s}))\right]$$

$$\geq \frac{1}{2} + \frac{\varepsilon}{2}$$

(same stmt as $*$)

by averaging, $\exists$ choice of $x_1 \cdots x_{m-1}, y_{m+1} \cdots y_k, \pi$

s.t. $\Pr_x\left[C(\pi(x_i\text{'s}, x, y_i\text{'s})) = f^{\oplus k}(\pi(x_i\text{'s}\, x\, y_i\text{'s}))\right]$

$$\geq \frac{1}{2} + \frac{\varepsilon}{2}$$

$f(x) \cdot \underbrace{\bigoplus_i f(x_i) \cdot \bigoplus_j f(y_j)}_{\text{fixed bit: either 1 or 0}}$

b

Construct many ckts!

for each choice of i

x_j's

y_j's

π

b

at least one is good

call it $\tilde{C}$ $\leftarrow$ g gates

$$\Pr_x[\tilde{C}(x) = f(x) \oplus b] \geq \frac{1}{2} + \frac{\varepsilon}{2}$$

given $x \in H$

use $\tilde{C}$ on x to get w $\Big\}$ size $|\tilde{C}| + 1$ $\leq g+1$ gates

output $w \oplus b$

$\Rightarrow$ f is not ε-h.c. for $g+1$

$\rightarrow \Leftarrow$

$\blacksquare$