Hardness vs Randomness: Impagliazzo's Hardcore Lemma

CS 6810—David Steurer

Spring 2016

We say a distribution H over $\{0,1\}^n$ has density δ if $\mathbb{P}_H(x) \leq 1/(\delta 2^n)$ for all $x \in \{0,1\}^n$. (The set of distribution with density δ turns out to be the convex hull of distributions that are uniform over a subset H' with $|H'| \geq \delta 2^n$.)

A priori, if a function is average-case hard, it could be because every circuit fails to compute f on a different fraction of inputs. The hardcore lemma rules out this possibility. If a function is average-case hard, then every circuit fails essentially on the same fraction of inputs.

Theorem: Let $f\colon \{0,1\}^n \to \{\pm 1\}$ be a mildly average-case hard function, i.e., $\mathbb{E}_{U_n} f \cdot C \leq 1 - \delta$ for every circuit $C\colon \{0,1\}^n \to \{\pm 1\}$ with $\mathrm{size}(C) \leq S$.

Then, there exists a distribution H with density $\Omega(\delta)$ such that $\mathbb{E}_H f \cdot C \leq \varepsilon$ for every circuit $C\colon \{0,1\}^n \to \{\pm 1\}$ with $\mathrm{size}(C) \leq S' := S \cdot O(\varepsilon^2/\log(1/\delta))$.

Proof:

The proof is by linear programming duality and uses that the set of functions with small circuit complexity is *approximately convex*, in the sense that a distribution over circuits can be approximated by a single circuit (of slightly larger size).

We will prove the contrapositive of the theorem statement: Suppose that for every density-δ distribution, there is an S'-sized circuit that computes f with advantage ε. Then there exists a circuit of size $S = S' \cdot n/\varepsilon^2$ that computes f on a $1 - O(\delta)$ fraction of the inputs.

Linear programming duality[1] allows us to exchange quantifiers of our assumption if we pass to distributions over circuits: There exists a distribution D over S'-sized circuits C such that for all density-δ distributions H, D "computes" f with advantage ε. Hence,

$$\min_{H \subseteq \{0,1\}^n,\ |H| \geq \delta 2^n} \mathop{\mathbb{E}}_{H} f \cdot \mathop{\mathbb{E}}_{C \sim D} C \geq \varepsilon \tag{1}$$

What does this condition mean? Let us say that D "computes" f on an input $x \in \{0,1\}^n$ if $f(x) \cdot \mathbb{E}_{C \sim D} C(x) \geq \varepsilon$. On how many inputs can D fail to compute f? We can see that the condition above implies D can fail on at most $\delta 2^n$ inputs. (Otherwise, there would be a set H violating the condition.) Thus D computes f on all but a δ fraction of the inputs.

How can we go from a distribution over circuits to a single circuit? (Note that this distribution could in principle be over an exponential number of circuits.) The idea is to use sample circuits $C_1, \ldots, C_r$ from the distribution and use their outputs to estimate $f(x) \cdot \mathbb{E}_{C \sim D} C(x)$. By the Chernoff bound, every input $x \in \{0,1\}^n$ with $f(x)\, \mathbb{E}_{C \sim D} C(x) \geq \varepsilon$ satisfies

$$\mathop{\mathbb{P}}_{C_1, \ldots, C_r} \Big\{ \mathrm{maj}(C_1(x), \ldots, C_r(x)) \neq f(x) \Big\} \leq 2^{-\Omega(\varepsilon^2 r)}. \tag{2}$$

Hence, for $r = O(\log(1/\delta)/\varepsilon^2)$, the failure probability due to sampling is at most δ. In particular, there exists a choice for $C_1, \ldots, C_r$ such that $\mathrm{maj}(C_1, \ldots, C_r)$ computes f on all

but a δ of the inputs that D computes f on. Therefore, $\mathrm{maj}(C_1, \ldots, C_r)$ is the desired circuit (of size $S = S' \cdot O(r)$).

Footnotes

1. Formally, we define a payoff matrix $M_{H,C} = \mathbb{E}_H f \cdot C$ with rows indexed by distributions H over $\{0,1\}^n$ of density δ and columns indexed by circuits $C\colon \{0,1\}^n \to \{\pm 1\}$ with $\mathrm{size}(S) \leq S'$. Our assumption implies that $\min_p \max_C (p^T M)_C \geq \varepsilon$ (where p is a distribution over density-δ distributions, which is again a density-δ distribution). Hence, by linear programming duality, we also have $\min_H (Mq)_H \geq \varepsilon$ for some distribution q over S'-sized circuits C.