CS 880: Advanced Complexity Theory 2/22/2008

Lecture 13: Average-Case Hardness

Instructor: Dieter van Melkebeek Scribe: Tom Watson

In this lecture and the next two lectures we study hardness amplification, in which the goal is to take a mildly average-case hard function from some class and construct another function in that class that is very average-case hard. Today we prove a lemma that roughly states that every average-case hard function has a set of inputs that encapsulates the hardness of that function in a certain sense. In the next two lectures, we will use this tool to prove hardness amplification results within E and within NP. We will not need harmonic analysis in today's lecture.

1 Worst-Case vs. Average-Case Complexity

Consider the following informal question: If a complexity class contains a problem that is worst-case hard, does it contain a problem that is average-case hard? This question has relevance in several contexts:

- *Complexity theory.* Worst-case complexity is the most common measure of hardness in complexity theory. However, for certain results, average-case complexity plays a crucial role. The basic construction of pseudorandom generators for time-bounded computations requires a problem in E that is average-case hard against nonuniform circuits. These pseudorandom generators are computable in time linear exponential in the seed length, which is fine for derandomization purposes since all the seeds must be cycled through anyway. It is known that if E contains a problem that is worst-case hard for exponential-size circuits, then it contains a problem that is very average-case hard for circuits of roughly the same size, so derandomization can be based on worst-case hardness assumptions.

- *Cryptography.* The security of any nontrivial cryptosystem requires some computational problem to be average-case hard in some sense. For example, it is necessary for the security of the RSA cryptosystem that factoring is average-case hard under a certain distribution. We need hard problem that are in NP, since decryption requires that the problem is easy to solve given a secret key. Unlike the derandomization setting, it is unknown how to obtain average-case hard problems in NP from a worst-case hardness assumption on NP. A major open question is whether any nontrivial form of cryptography can be based on the assumption $P \neq NP$. The weakest assumption known to imply nontrivial cryptography is the existence of a one-way function, which is an average-case hardness assumption.

One approach for obtaining the aforementioned worst-case to average-case transformation within E is to start from a worst-case hard function and encode the characteristic sequence at some input length with a good locally list-decodable error-correcting code. The resulting codeword then forms the characteristic sequence of a very average-case hard problem. This approach is problematic if we want the average-case hard problem to be in NP, since evaluating one bit of the codeword seems to require looking at the entire information word, which takes exponential time. Another approach for obtaining the transformation within E operates in two steps:

1. A mildly average-case hard problem is constructed from a worst-case hard problem.

2. A very average-case hard problem is constructed from a mildly average-case hard problem.

We do not know how to achieve step 1 within NP, but step 2, which is known as hardness amplification, *can* be achieved within NP. In this lecture, we develop our main tool for hardness amplification. In the next two lectures we show how to use this tool for hardness amplification within E and within NP. Harmonic analysis is used in determining how average-case hard the constructed functions are.

1.1 Hardcores

We begin with some definitions, leading to the statement of our main lemma. Throughout this lecture, whenever we choose a random input x from some subset of $\{-1,1\}^n$, it is done uniformly. For $H \subseteq \{-1,1\}^n$, we use the notation $\mu(H) = \frac{|H|}{2^n}$.

Definition 1. *A function $f : \{-1,1\}^n \to \{-1,1\}$ is* ϵ-hard *for size s if for all circuits C of size at most s, $\Pr_x[C(x) \neq f(x)] \geq \epsilon$.*

Clearly, we cannot hope for f to be ϵ-hard if $\epsilon > \frac{1}{2}$, since either the constant -1 function or the constant 1 function agrees with f on at least half the inputs. Our goal is to get ϵ as close to $\frac{1}{2}$ as possible. Specifically, given f that is mildly hard ($\epsilon = \frac{1}{poly(n)}$) for size s, we wish to construct g that is very hard ($\epsilon = \frac{1}{2} - o(1)$) for some size s'. In our proofs, we will have to settle for s' being a little smaller than s. In the E setting, it is known how to get $(\frac{1}{2} - \frac{1}{2^{\Theta(n)}})$-hardness from a worst-case hardness assumption. In the NP setting, the result is weaker in two respects: we start from a mild average-case hardness assumption, and we don't get ϵ as close to $\frac{1}{2}$. In the next two lectures, we will show a $(\frac{1}{2} - o(1))$-hardness result; currently the best known result achieves $(\frac{1}{2} - \frac{1}{2^{\Theta(n^{2/3})}})$-hardness.

A key notion in our proofs is that of a *hardcore*.

Definition 2. *Fix a function $f : \{-1,1\}^n \to \{-1,1\}$. A* δ-hardcore for size s *is a set $H \subseteq \{-1,1\}^n$ such that for all circuits C of size at most s, $\Pr_{x \in H}[C(x) = f(x)] \leq \frac{1}{2} + \delta$.*

Note that if f has a δ-hardcore H for size s with $\mu(H) \geq \epsilon$, then f is $\epsilon(\frac{1}{2} - \delta)$-hard for size s. Surprisingly, the converse also holds, modulo some changes in the parameters. That is, if a function is mildly average-case hard, then there is a set of inputs of relative size the hardness of f on which f is very average-case hard.

Lemma 1 (Hardcore Lemma)**.** *If f is ϵ-hard for size s, then for all $\delta > 0$ there exists a δ-hardcore H with $\mu(H) \geq \epsilon$ for size $s' = \Theta(s)$, where the constant in the Θ depends on ϵ and δ.*

The Hardcore Lemma is our main tool for hardness amplification. A number of proofs of this lemma are known; the one we now present does not give the best dependence of s' on ϵ and δ, but the proof is fairly clean as is good enough for our purposes. We refer to the statement at the end of the lecture notes for the precise dependence of s' on ϵ and δ which our argument yields.

2 Proof of the Hardcore Lemma

Fix a function $f : \{-1,1\}^n \to \{-1,1\}$ that is ϵ-hard for size s, and assume WLOG that $\epsilon 2^n$ is an integer. Consider the following two-player zero-sum game: one player picks a circuit C' of size at most s', the other player picks a set $H \subseteq \{-1,1\}^n$ with $\mu(H) = \epsilon$, and the payoff to the C' player is $\Pr_{x \in H}[C'(x) = f(x)]$. That is, the C' player tries to pick a C' that performs well on H, and the H player tries to pick an H on which C' does not perform well. By the well-known Minimax Theorem (which is equivalent to strong duality of linear programming), the expected payoff does not depend on which player selects his strategy first, provided we allow randomized strategies.

$$\underset{\substack{\text{distributions}\\ \text{over } C' \text{ of}\\ \text{size} \le s'}}{\text{MAX}} \quad \underset{\substack{H \text{ with}\\ \mu(H)=\epsilon}}{\text{MIN}} \quad \mathrm{E}_{C'}\Big[\Pr_{x \in H}[C'(x) = f(x)]\Big] = \underset{\substack{\text{distributions}\\ \text{over } H \text{ with}\\ \mu(H)=\epsilon}}{\text{MIN}} \quad \underset{\substack{C' \text{ of}\\ \text{size} \le s'}}{\text{MAX}} \quad \mathrm{E}_{H}\Big[\Pr_{x \in H}[C'(x) = f(x)]\Big] \tag{1}$$

Note that we assume the player who selects his strategy second picks a pure strategy; this is without loss of generality because for any fixed randomized strategy of the first player, the expected payoff for any distribution on pure strategies of the second player is a convex combination of the expected payoffs for pure strategies of the second player.

If we only considered deterministic strategies for the first player, the left side would be 0, assuming $s' \le s$. This is because, by our hardness assumption on f, the maximum over all circuits C' of size at most s of the minimum over H with $\mu(H) = \epsilon$ of $\Pr_{x \in H}[C'(x) = f(x)]$ is 0. With randomized strategies we can always achieve a value of at least $1/2$. In the first part of the proof of the Hardcore Lemma, we show that the value for randomized strategies cannot grow much larger than $1/2$ when f is ϵ-hard. We do so in Section 2.1.

In the second part of the proof we use this upper bound on the right side of (1) to construct a δ-hardcore H with $\mu(H) \ge \epsilon$. We show that if there is a distribution on Hs against which no small circuit can do well, then we can extract a *single* set H' against which no small circuit can do well. This is the contents of Section 2.2.

2.1 Bounding the Value of the Game

We now upper bound the left side of (1) by $\frac{1}{2} + \gamma$ for any γ, when s' is chosen appropriately depending on γ. We do this by showing that if the left side is greater than $\frac{1}{2} + \gamma$ then we can construct a circuit C of size at most s such that $\Pr_x[C(x) \ne f(x)] < \epsilon$, contrary to our hardness assumption on f.

What does it mean for the left side of (1) to be greater than $\frac{1}{2} + \gamma$? It means that there exists some distribution over C' of size at most s' such that for all H with $\mu(H) = \epsilon$,

$$\Pr_{\substack{C'\\ x \in H}}[C'(x) = f(x)] > \frac{1}{2} + \gamma. \tag{2}$$

We construct a distribution over circuits C as follows: take t independent samples of C' and let C on input $x \in \{-1,1\}^n$ output the majority vote of these t circuits. Using the bound (2), we will show that if t is chosen appropriately, then

$$\Pr_{\substack{C\\ x \in \{-1,1\}^n}}[C(x) \ne f(x)] < \epsilon.$$

By the probabilistic method, this implies that there is some particular C such that $\Pr_x[C(x) \neq f(x)] < \epsilon$. This contradicts the supposed hardness of f provided we choose s' small enough to ensure that C has size at most s.

Define

$$B = \Big\{ x \ : \ \Pr_{C'}[C'(x) = f(x)] \leq \frac{1}{2} + \tau \Big\}$$

to be the set of "bad" inputs, where τ is some threshold to be set later. We use the fact that

$$\begin{aligned} \Pr_{C,x}[C(x) \neq f(x)] \ &= \ \mu(B) \cdot \Pr_{C,x}\big[C(x) \neq f(x) \ \big| \ x \in B\big] + \big(1 - \mu(B)\big) \cdot \Pr_{C,x}\big[C(x) \neq f(x) \ \big| \ x \notin B\big] \\ &\leq \ \mu(B) + \Pr_{C,x}\big[C(x) \neq f(x) \ \big| \ x \notin B\big]. \end{aligned}$$

We first upper bound $\mu(B)$. If $\tau \leq \gamma$ then we must have $\mu(B) < \epsilon$ since otherwise we could take $H \subseteq B$ with $\mu(H) = \epsilon$ and have

$$\Pr_{\substack{C' \\ x \in H}}[C'(x) = f(x)] \ = \ \mathop{\mathrm{E}}_{x \in H}\Big[\Pr_{C'}[C'(x) = f(x)]\Big] \ \leq \ \mathop{\mathrm{E}}_{x \in H}\Big[\frac{1}{2} + \tau\Big] \ = \ \frac{1}{2} + \tau \ \leq \ \frac{1}{2} + \gamma,$$

contrary to (2). However, since we want to show that $\mu(B) + \Pr_{C,x}[C(x) \neq f(x) \mid x \notin B] < \epsilon$, we're going to have to do a bit better. Since we know $\mu(B) < \epsilon$, let's extend B to a superset H with $\mu(H) = \epsilon$ in an arbitrary way. Then applying (2), we have

$$\begin{aligned} \frac{1}{2} + \gamma \ &< \ \Pr_{\substack{C' \\ x \in H}}[C'(x) = f(x)] \\ &= \frac{\mu(B)}{\epsilon} \cdot \Pr_{\substack{C' \\ x \in H}}\big[C'(x) = f(x) \mid x \in B\big] \ + \ \Big(1 - \frac{\mu(B)}{\epsilon}\Big) \cdot \Pr_{\substack{C' \\ x \in H}}\big[C'(x) = f(x) \mid x \in H \backslash B\big] \\ &\leq \ \frac{\mu(B)}{\epsilon} \cdot \Big(\frac{1}{2} + \tau\Big) \ + \ \Big(1 - \frac{\mu(B)}{\epsilon}\Big) \cdot 1, \end{aligned}$$

which implies that

$$\mu(B) \ < \ \frac{\frac{1}{2} - \gamma}{\frac{1}{2} - \tau} \cdot \epsilon \ \leq \ \big(1 - (\gamma - \tau)\big)\epsilon.$$

Thus if we pick $\tau < \gamma$ then we get a bound on $\mu(B)$ that is better than ϵ.

We now turn to upper bounding $\Pr_{C,x}[C(x) \neq f(x) \ \big| \ x \notin B]$. For $x \notin B$, each circuit C' chosen for C satisfies $C'(x) = f(x)$ independently with probability greater than $\frac{1}{2} + \tau$, and for $C(x) \neq f(x)$ to hold it would have to be the case that $C'(x) = f(x)$ for at most half of the t circuits C'. Thus by a standard Chernoff bound, we have $\Pr_C[C(x) \neq f(x)] \leq \exp\big(-\Omega(\tau^2 t)\big)$ for each $x \notin B$. It follows that

$$\Pr_{C,x}\big[C(x) \neq f(x) \ \big| \ x \notin B\big] \leq \exp\big(-\Omega(\tau^2 t)\big).$$

Putting everything together, we have

$$\Pr_{C,x}[C(x) \neq f(x)] \ \leq \ \big(1 - (\gamma - \tau)\big)\epsilon \ + \ \exp\big(-\Omega(\tau^2 t)\big).$$

As we noted before, this implies that there exists a particular circuit C such that $\Pr_x[C(x) \neq f(x)]$ satisfies the same bound. Picking $\tau = \frac{\gamma}{2}$ and $t \geq \Theta(\frac{1}{\gamma^2} \log \frac{1}{\gamma\epsilon})$ ensures that this bound is less than

ϵ. This contradicts the supposed hardness of f provided C has size at most s. How big is C? It consists of t circuits C' each of size at most s' and an $O(t)$ size majority circuit, so it has size at most $O(\frac{1}{\gamma^2}\log\frac{1}{\gamma\epsilon}\cdot s')$. This bound at most s if

$$s' \leq O\left(\frac{\gamma^2}{\log\frac{1}{\gamma\epsilon}}\cdot s\right).$$

We summarize what we have shown in the following lemma.

Lemma 2. *If f is ϵ-hard for circuits of size s, then for all $\gamma > 0$ the value of equation (1) is at most $\frac{1}{2}+\gamma$ when $s' \leq O\left(\frac{\gamma^2}{\log\frac{1}{\gamma\epsilon}}\cdot s\right)$.*

2.2 Obtaining the Hardcore

By (1) and Lemma 2 we know that for all $\gamma > 0$, the right side of (1) is at most $\frac{1}{2}+\gamma$ when s' is as in Lemma 2. What does it mean for the right side of (1) to be at most $\frac{1}{2}+\gamma$? It means that there exists some distribution over H with $\mu(H) = \epsilon$ such that for all C' of size at most s',

$$\mathop{\mathrm{E}}_{H}\left[\Pr_{x\in H}[C'(x) = f(x)]\right] \leq \frac{1}{2}+\gamma.$$

What we would like is one particular H such that for all C' of size at most s', $\Pr_{x\in H}[C'(x) = f(x)] \leq \frac{1}{2}+\gamma$; then we could take $\gamma = \delta$ and be done. We may not be able to find such an H, but suppose we could show that for each C', the random variable $\Pr_{x\in H}[C'(x) = f(x)]$ over the choice of H were highly concentrated about its mean. Then for each C', $\Pr_H\left[\Pr_{x\in H}[C'(x) = f(x)] > \frac{1}{2}+2\gamma\right]$ would be extremely low, so by a union bound over C', $\Pr_{x\in H}[C'(x) = f(x)] \leq \frac{1}{2}+2\gamma$ would hold for all C' simultaneously with positive probability over the choice of H, in which case we could take $\gamma = \frac{\delta}{2}$ and be done. We can't guarantee this concentration result for the given distribution over Hs, but we can do a similar thing for a somewhat different distribution, which we now define.

Construct a distribution over sets $H' \subseteq \{-1,1\}^n$ by putting each x in H' independently with probability $\Pr_H[x \in H]$, where the latter probability is over the distribution on Hs with $\mu(H) = \epsilon$ guaranteed by the Minimax Theorem. Note that the H's in the support of this new distribution do not all have $\mu(H') \geq \epsilon$. However, note that

$$\mathop{\mathrm{E}}_{H'}[\mu(H')] = \frac{1}{2^n}\sum_x \Pr_{H'}[x \in H'] = \frac{1}{2^n}\sum_x \Pr_H[x \in H] = \mathop{\mathrm{E}}_H[\mu(H)] = \epsilon.$$

Since x's are placed in H' independently, the distribution of $\mu(H')$ is close to a scaled normal distribution and so we have

$$\Pr_{H'}[\mu(H') \geq \epsilon] = \Omega(1).$$

We will see shortly that this is good enough.

For a fixed C', we intuitively expect the random variable $\Pr_{x\in H'}[C'(x) = f(x)]$ over the choice of H' to be highly concentrated because the x's are chosen independently in H'. We argue something slightly different, namely that the *number* of $x \in H'$ such that $C'(x) = f(x)$ is highly concentrated about its mean. We will then combine this with our above observation that $\Pr_{H'}[\mu(H') \geq \epsilon] = \Omega(1)$ to finish the proof of the Hardcore Lemma.

Let $R_{C'} = \{x \,:\, C'(x) = f(x)\}$ be the set of inputs on which C' is "right", and let $A_{C'} = |H' \cap R_{C'}|$ be the "agreement" random variable over the choice of H'. Note that $A_{C'} = \sum_{x \in R_{C'}} I_{H'}(x \in H')$, where $I_{H'}(x \in H')$ is the indicator random variable over the choice of H' for the event $x \in H'$. We have

$$\begin{aligned}
\mathop{\mathrm{E}}_{H'}[A_{C'}] &= \sum_{x \in R_{C'}} \Pr_{H'}[x \in H'] \\
&= \sum_{x \in R_{C'}} \Pr_{H}[x \in H] \\
&= \mathop{\mathrm{E}}_{H}\left[\sum_{x \in R_{C'}} I_H(x \in H)\right] \\
&\leq \left(\frac{1}{2} + \gamma\right) \cdot \epsilon 2^n.
\end{aligned}$$

Since the indicators $I_{H'}(x \in H')$ for $x \in \{-1,1\}^n$ are fully independent, a standard Chernoff bound (namely, that a sum of fully independent indicators differs from its expectation E by at least ΔE with probability at most $\exp\left(-\Omega(\Delta^2 E)\right)$) yields

$$\Pr_{H'}\left[A_{C'} > \left(\frac{1}{2} + 2\gamma\right) \cdot \epsilon 2^n\right] \leq \exp\big(-\Omega(\gamma^2 \epsilon 2^n)\big).$$

Since this holds for all C' of size at most s', and we can bound the number of circuits of size s' by $(n+s')^{s'}$, a union bound shows that with probability at least

$$1 - (n+s')^{s'} \exp\big(-\Omega(\gamma^2 \epsilon 2^n)\big)$$

over the choice of H', all C' of size at most s' satisfy $A_{C'} \leq \left(\frac{1}{2} + 2\gamma\right) \cdot \epsilon 2^n$. Assuming this probability is sufficiently close to 1, it follows from the fact that $\Pr_{H'}[\mu(H') \geq \epsilon] = \Omega(1)$ that with positive probability H' satisfies both $\mu(H') \geq \epsilon$ and $A_{C'} \leq \left(\frac{1}{2} + 2\gamma\right) \cdot \epsilon 2^n$ for all C' of size at most s'. Thus, there exists an H' such that $\mu(H') \geq \epsilon$ and $\Pr_{x \in H'}[C'(x) = f(x)] \leq \frac{1}{2} + 2\gamma$ for all C' of size at most s'. Taking $\gamma = \frac{\delta}{2}$ and using the fact that every function on n bits can be computed by circuits of size $(1 + o(1))2^n/n$, we have arrived at the following precise formulation of the Hardcore Lemma.

Lemma 3. *There exists a universal constant c such that for all functions $\epsilon(n) > 0$ and $\delta(n) > 0$ there exists an n_0 such that the following holds. If $f : \{-1,1\}^n \to \{-1,1\}$ is $\epsilon(n)$-hard for size $s \leq \epsilon(n)2^n$ and $n \geq n_0$ then there exists a set $H \subseteq \{-1,1\}^n$ such that $\mu(H) \geq \epsilon(n)$ and H is a $\delta(n)$-hardcore for size $s' = c \cdot \frac{\delta(n)^2}{\log \frac{1}{\delta(n)\epsilon(n)}} \cdot s$.*