U.C. Berkeley
CS294: Pseudorandomness and Combinatorial Constructions
Professor Luca Trevisan

Handout N17
October 27, 2005
Scribe: Sebastien Roch

Notes for Lecture 17

In the previous lecture, we introduced the notion of a *randomness extractor*—a procedure which extracts "uniform" randomness from a weak random source and a small number of truly random bits. In this lecture, we give general impossibility results for randomness extraction. The results are taken from [NZ96, RT00].

1 Shannon's Entropy

We seek to characterize the class of random sources for which extraction is possible. We start with a discussion of Shannon's Entropy. We explain, in particular, why this notion is not appropriate for our purposes.

Definition 1 (Shannon's Entropy) *Let X be a random variable. Then Shannon's Entropy is defined as*

$$H(X) \quad = \quad \sum_{x\,:\,\mathbb{P}[X=x]\neq 0} \mathbb{P}[X=x] \log \frac{1}{\mathbb{P}[X=x]} = \mathbb{E}_{x\sim X}\left[\log \frac{1}{\mathbb{P}[X=x]}\right].$$

An interpretation of $H(X)$ is as follows: it counts the number of truly random bits needed to sample from X. Indeed, if t bits are used to encode a particular value a, then we need

$$\mathbb{P}[X=a] \geq \frac{1}{2^t} \qquad \Rightarrow \qquad t \geq \log \frac{1}{\mathbb{P}[X=a]}.$$

Therefore, if R_a denotes the number of random bits used for a, then we have

$$\mathbb{E}_{a\sim X}[R_a] \geq \mathbb{E}_{a\sim X} \log \frac{1}{\mathbb{P}[X=a]}.$$

However, Shannon's Entropy measures only the amount of randomness needed, not how it is "distributed" over all possible values. To see why this is a problem, consider the following random source X over $\{0,1\}^n$. Suppose X is such that

$$\mathbb{P}[X=\mathbf{0}] = 1 - \frac{1}{2^{100}},$$

where $\mathbf{0} = (0,\ldots,0)$, and for all $a \neq \mathbf{0}$,

$$\mathbb{P}[X=a] = \frac{1}{2^{100}}\left(\frac{1}{2^n-1}\right) \sim \frac{1}{2^{n+100}}.$$

Then the Shannon Entropy is

$$H(X) = \left(1-\frac{1}{2^{100}}\right)\log\frac{1}{1-\frac{1}{2^{100}}} + \sum_{a\neq\mathbf{0}} \frac{1}{2^{100}}\left(\frac{1}{2^n-1}\right)\log\frac{1}{\frac{1}{2^{100}}\left(\frac{1}{2^n-1}\right)} = \Omega(n).$$

According to this measure, the source X has high entropy. However, suppose the number of truly random bits used by an extractor is $t = O(\log n)$. Denote $\text{Ext} : \{0,1\}^n \times \{0,1\}^t \to \{0,1\}^m$ the extractor (with $m \gg t$), and U_m the uniform distribution on $\{0,1\}^m$. Then, because $t = O(\log n)$, the distribution of $Y = \text{Ext}[X, U_m]$ is concentrated on a polynomial number of strings z with

$$\mathbb{P}[Y = z] \geq \frac{1}{\text{poly}(n)}.$$

It is easy to see that Y *cannot* be close to uniform. Indeed, suppose Y is $\frac{1}{2}$-close to uniform. Sort the outputs of Y in increasing order of their probability under Y. Let S be the smallest set of "most likely" outputs such that $\mathbb{P}[Y \in S] \geq .51$. Then we must have $\mathbb{P}[U_m \in S] \geq 0.01$. But that implies $|S| \geq 0.01 \cdot 2^m$, i.e. S has to be exponentially large.

We have described a random source with high Shannon entropy for which randomness extraction is not possible. Therefore, Shannon's Entropy is not the right notion of entropy for our purposes.

2 Min-Entropy

It follows from the previous discussion that for a distribution to be close to uniform, it must output (most of the times) values that have exponentially small probability. Because our seed has rather short length, this kind of small probability has to come *from the source*. To capture this, we introduce the notion of *min-entropy*.

Definition 2 *The* min-entropy *of X is*

$$H_\infty(X) \equiv \min_{x:\mathbb{P}[X=x]} \log \frac{1}{\mathbb{P}[X = x]}.$$

If $H_\infty(X) = k$, then for all a, $\mathbb{P}[X = a] \leq 2^{-k}$.

Let m be the size of the output. We will show that extraction is possible only if the min-entropy of the source is close to m. In fact, we will show that there is a fixed extractor for all sources with min-entropy close to m. In this lecture, we focus on the former result. We first define the notion of extractor.

Definition 3 *The function* $\text{Ext} : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ *is a (k, ε)-extractor if* $\forall X : H_\infty(X) \geq k$,

$$\|\text{Ext}(X, U_d) - U_m\|_{\text{SD}} \leq \varepsilon.$$

Trade-off. For n, k fixed (i.e. fixed physical source), we would like d small (i.e. $O(\log n)$), m large (i.e. $\sim k$), and ε small (i.e. constant).

Existence. It is known (non-constructively) that (k, ε)-extractors exist with

$$m = k + d - 2\log \varepsilon^{-1} - O(1),$$

and

$$d = \log(n - k) + 2\log \varepsilon^{-1} + O(1).$$

3 Lower Bound

In this section, we prove the following lower bound on extraction.

Proposition 4 *For every (k, ε)-extractor with $\varepsilon < 1/2$ and $m > d + 2$, the following hold:*

$$m \leq k + d - 2\log \varepsilon^{-1} + O(1),$$

and

$$d \geq \log(n - k) + 2\log \varepsilon^{-1} - O(1).$$

We actually prove a slightly weaker bound.

PROOF:(Sketch) Let $\mathrm{Ext} : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ be a (k, ε)-extractor. Construct a bipartite multi-graph with 2^n nodes on the left side (corresponding to sequences in $\{0,1\}^n$) and 2^m nodes on the right side (corresponding to sequences in $\{0,1\}^m$). There is an edge between x and y iff $\exists z$ s.t. $\mathrm{Ext}(x, z) = y$.

Let $S \subseteq \{0,1\}^n$, $|S| \geq 2^k$, and $T \subseteq \{0,1\}^m$. Pick v uniformly at random in S and w uniformly at random over the neighbours of v. Then by definition of a (k, ε)-extractor, it holds that

$$\left| \mathbb{P}[w \in T] - \frac{|T|}{2^m} \right| \leq \varepsilon.$$

Now assume that, in fact, we pick $T \subseteq \{0,1\}^m$ of size $\varepsilon 2^m + 1$ uniformly at random, and let $S \subseteq \{0,1\}^n$ be the left neighbours of T.

Claim 5 *It holds that $|S| \geq 2^n - 2^k$.*

PROOF: (Sketch) By contradiction, assume $|S^c| > 2^k$. Pick a uniform random neighbour of S^c. The probability that it falls in T is 0. This contradicts the definition of (k, ε)-extractor. □

Fix $v \in \{0,1\}^n$. We compute the probability that the neighbourhood of v hits T when T is chosen in the following way: each element in $\{0,1\}^m$ is picked with probability ε independently. (Note that this choice of T is not quite what we want. This can be fixed. The details are left out.) We have

$$\mathbb{P}_T[\text{no neighbour of } v \text{ is in } T] \geq (1 - \varepsilon)^{2^d}.$$

Therefore,

$$\begin{aligned} 2^k &\geq \mathbb{E}_T[\text{number of vertices on the left with no neighbour in } T] \\ &= \sum_{v \in \{0,1\}^n} \mathbb{P}_T[\text{no neighbour of } v \text{ is in } T] \\ &\geq 2^n (1 - \varepsilon)^{2^d} \sim 2^n e^{-\varepsilon 2^d}. \end{aligned}$$

This implies

$$k \geq n - \varepsilon 2^d O(1),$$

or

$$d \geq \log \varepsilon^{-1} + \log(n - k) - O(1).$$

□

General Proof. To obtain the factor of 2 in front of $\log \varepsilon^{-1}$, one has to consider the following: look at all left vertices such that

$$\frac{\text{number of edges from } v \text{ to } T}{2^d} < \frac{|T|}{2^m} - \varepsilon. \tag{1}$$

There are at most 2^k such vertices (by contradiction). Pick $T \subseteq \{0,1\}^m$ by including each element in $\{0,1\}^m$ with probability $1/2$ independently. Then by the Central Limit Theorem, it holds that the probability that a fixed vertex in $\{0,1\}^n$ satisfies (1) is at least $1/e^{\Omega(\varepsilon^2 2^d)}$. The bound follows.

In the next class, we will see explicit constructions of extractors.

References

[RT00] J. Radhakrishnan and A. Ta-Shma, Bounds for dispersers, extractors, and depth-two superconcentrators, SIAM J. Discrete Math., 13(1):2–24, 2000.

[NZ96] N. Nisan and D. Zuckerman, Randomness is Linear in Space, JCSS, 52(1):43–52, 1996.