Recent advances in Complexity CIS 6930/CIS 4930	November 14, 2002
Lecture 23	
Lecturer: Dr. Meera Sitharam	Scribe: Srijit Kamath

1 Extractors

1.1 Motivation

The Nisan-Wigderson theorem draws a relationship between hardness and randomness. It shows that it is possible to use a hard function f to develop a good psuedorandom generator G (and vice versa). The construction of G does not rely on the fact that f is hard. It only relies on the complexity upper bound of f (and that too, only to establish the efficiency of computing G). However, the fact that G is a good psuedorandom generator that results in a bias defined in (3) of Nisan-Wigderson theorem relies on the hardness of approximating f using circuit C. The question now is whether it is possible to develop good pseudorandom generators without attaching any hardness assumptions and for any arbitrary test C. For instance we ask if it is possible to come up with an *extractor* Ext such that

$$|P[C(y) = 1] - P[C(Ext(x)) = 1]| \leq \frac{1}{\epsilon},$$

for all C, for some constant ϵ, as opposed to a rhs that is related to the size of the circuit C.

It turns out that this type of bias is sufficient for determininstic simulations of BPP to settle $Pvs.BPP$. – we will see this type of simulation soon. In fact, we will see that this type of deterministic simulation is a little different from the simulation (i.e, in the proof of the corrolaries of the NW generator, and in the introductory lecture to pseudorandomness). That simulation uses the entire set of a pseudorandom generator's outputs for all possible seeds, which necessitates that the (fully random) seeds need to be small, in order for the deterministic simulation to be efficient.

In the case of the simulation using extractors, we can also allow the input to the extractor Ext to be not necessarily a "small" random seed, but instead, longer strings obtained from a "weak random source." to be defined below. Good extractors could thus result in derandomization results such as $P = BPP$ without any complexity theoretic hardness assumptions.

1.2 Preliminaries

We begin by making a few definitions.

Definition 1 *A distribution D over $\{0,1\}^n$ is said to have min entropy atleast k if*

$$|D(x)| \leq \frac{1}{2^k}, \forall x \in \{0,1\}^n$$

Here is an alternative definition of min entropy.

Definition 2 *A random variable X of range $\{0,1\}^n$ has min entropy atleast k if it holds that*

$$\forall x \in \{0,1\}^n, P[X = x] \leq \frac{1}{2^k}$$

Exercise 1 *Any distribution D of min entropy $\geq k$ can be written as a convex combination $D = \sum \lambda_S U_S$, where $S \subseteq \{0,1\}^n, |S| = 2^k, \sum \lambda_S = 1, 0 \leq \lambda_S \leq 1$, U_S is the uniform distribution over S and the summation is carried out over all such susets S.*

Intuitively a string picked from a distribution over $\{0,1\}^n$ with min entropy $\geq k$ contains (or encodes) k bits of 'true randomness'. A good example is $D = U_S = \{0,1\}^k 1^{n-k}$. This is a subcube of $\{0,1\}^n$. The existence of k true random bits is obvious in this case, but is not so for many other min entropy $\geq k$ distributions. We would like to 'extract' k true bits of randomness from any such distribution. In other words we would like to convert a (potentially) weak random source to a strong random source. That this is always possible has not been proven and remains a conjecture.

Definition 3 *Two random variables X, Y taking values in $\{0,1\}^n$ are $\epsilon-$close if*

$$\forall T : \{0,1\}^n \longrightarrow \{0,1\} |P(T(X) = 1) - P(T(Y) = 1)| \leq \epsilon,$$

i.e., if

$$max_{T:\{0,1\}^n \longrightarrow \{0,1\}} |P(T(X) = 1) - P(T(Y) = 1)| \leq \epsilon$$

Exercise 2 *Show that this is equivalent to showing that*

$$\frac{1}{2} \sum_{v \in \{0,1\}^n} |P(X = v) - P(Y = v)| \leq \epsilon$$

Next we restate the above definitions as applicable to distributions. Note that one can think of these concepts in terms of random variables or distributions, whichever is more convenient. Understanding both may help in developing a better intuition.

Definition 4 *Two distributions D, F are $\epsilon-$close if*

$$max_{T:\{0,1\}^n \longrightarrow \{0,1\}} \left| \sum_{x|T(x)=1} D(x) - \sum_{y|T(y)=1} F(y) \right|$$

$$= \frac{1}{2} \sum_{v \in \{0,1\}^n} |D(v) - F(v)| \leq \epsilon$$

Definition 5 *A (k, ϵ) extractor is a function $EXT : \{0,1\}^k \times \{0,1\}^t \longrightarrow \{0,1\}^m$ if for every random variable X of min entropy $\geq k$, it holds that $EXT(X, U_t)$ is $\epsilon -$ close to U_m, where U_t and U_m are uniform distributions over $\{0,1\}^t$ and $\{0,1\}^m$ respectively.*

What it means is that an extractor has the property that if X is drawn from a source of min entropy $\geq k$ and if the other input to the extractor is drawn from a uniform distribution over $\{0,1\}^t$ then the output is uniformly distributed over $\{0,1\}^m$.