Hardness implies derandomization

Amnon Ta-Shma and Dean Doron

1 Designs and Weak Designs

Definition 1 (A design [1]). *A family of sets $Z_1, Z_2, \ldots, Z_m \subseteq [t]$ is a (ℓ, a) design if*

1. *For all $i \in [t]$, $|Z_i| = \ell$, and*
2. *For all $i \neq j$, $|Z_i \cap Z_j| \leq a$.*

Claim 2. *For every ℓ, m there exists a $(\ell, a = \log m)$ design $Z_1, \ldots, Z_m \subseteq [t]$ where $t = O(\ell^2)$.*

Proof. Assume w.l.o.g. that ℓ is a prime power. Consider the numbers in $[t]$ as pairs of elements in $\mathbb{F}_\ell$. I.e., identify $[t]$ with $\{(x, y) \mid x, y \in \mathbb{F}_\ell\}$.

For every polynomial $p \in \mathbb{F}_\ell[X]$ of degree at most a, define the set of all evaluations $S_p = \{(x, p(x)) \mid x \in \mathbb{F}_\ell\}$. There are at least $\ell^{a+1} \geq m$ such polynomials, so all that is left is to observe that:

1. For every p, $|S_p| = \ell$.
2. For every $p_1 \neq p_2$, $|S_{p_1} \cap S_{p_2}| \leq a$.

Therefore, every m sets from $\{S_p\}_p$ is a (ℓ, a) design. □

In fact, a slightly more refined notion that already suffices is of a weak design:

Definition 3 (Weak design [2]). *A family of sets $Z_1, \ldots, Z_m \subseteq [t]$ is a weak (ℓ, ρ) design if*

1. *For all $i \in [t]$, $|Z_i| = \ell$, and*
2. *For all $i \neq j$, $\sum_{j<i} 2^{|Z_i \cap Z_j|} \leq \rho \cdot (m-1)$.*

We cite without a proof:

Lemma 4 ([2]). *For every ℓ, m and $\rho > 1$, there exists a weak (ℓ, ρ) design $Z_1, \ldots Z_m \subseteq [t]$ with $t = \left\lceil \frac{\ell}{\ln \rho} \right\rceil \cdot \ell$. Such a family can be found in time* $\mathrm{poly}(m, t)$.

2 The Nisan-Wigderson generator

We would like to construct a pseudo-random generator (PRG) fooling a class of circuits (such as AC^0, $\mathsf{P/poly} = \mathsf{SIZE}(\mathrm{poly}(n))$ or even $\mathsf{SIZE}(2^{\sqrt{n}})$). A PRG against a class of functions F is a function $G : \{0,1\}^\ell \to \{0,1\}^n$ such that no function $f \in F$ ε-distinguishes $G(U_\ell)$ from the uniform distribution.

Throughout the lectures, given $f : \{0,1\}^n \to \{0,1\}$, we say that $\mathsf{Size}(f) > s$ if no family of circuits of size $s = s(n)$ computes f correctly, and that $\mathsf{Size}_\varepsilon(f) > s$ if no family of circuits of size $s = s(n)$ computes f correctly on more than ε-fraction of the inputs.

The existence of a PRG implies the existence of a hard function:

Theorem 5. *If there exists a PRG $G : \{0,1\}^\ell \to \{0,1\}^{s(\ell^c)}$ against circuits of size $s(\ell^c)$ for some constant c and $\ell \le s(\ell) \le 2^\ell$ running in time exponential in ℓ then there exists a function f in* EXP *that is average-case hard for circuits of size $s(\ell^c)$.*

The proof of this lemma is left as an exercise.

The converse is much more difficult to achieve and is our goal today. Nisan and Wigderson described such a black-box reduction. We are given some $f : \{0,1\}^\ell \to \{0,1\}$ (and we think of f as a "hard" function for some computation class). Let $S_1, \ldots, S_m \subseteq [t]$ be a $(\ell, 2)$ weak design that is guaranteed by Lemma 4. The generator $G^f_{\ell,m} : \{0,1\}^t \to \{0,1\}^m$ is given by:

$$G^f_{\ell,m}(y) \;=\; f(y|_{S_1}), \ldots, f(y|_{S_m}).$$

We prove that:

Theorem 6 ([1])**.** *Suppose $f : \{0,1\}^\ell \to \{0,1\}$ is a function such that no circuit of size s can compute f correctly on more than a $\frac{1}{2} + \frac{\varepsilon}{m}$ fraction of the inputs. Then, $G^f_{\ell,m}$ is a PRG against circuits of size $s - m^2$ with error ε.*

The NW construction and also the later improvements are black-box constructions in the following sense: They start with an explicit function $f : \{0,1\}^\ell \to \{0,1\}$ and construct from it a new function $G^f : \{0,1\}^t \to \{0,1\}^m$ (where the notation is meant to indicate that G makes black-box oracle calls to f).

Moreover, the proof of Theorem 6 will be by "black-box reconstruction", namely, the proof describes an efficient "reconstruction" oracle Turing Machine R such that for every boolean function $f : \{0,1\}^\ell \to \{0,1\}$, if there is a small circuit C that ε-distinguishes $G^f(U_t)$ from uniform, then there exists a *short* advice string $z = A(f)$ such that $R^C(z,i)$ computes $f(i)$. Formally,

Definition 7 (Reconstructive PRG)**.** *We say the NW generator $G^f_{\ell,m}$ has (p,q) reconstruction with:*

- *Advice function $A = A(D,f)$, and,*
- *Reconstruction oracle circuit R,*

if for every $f : \{0,1\}^\ell \to \{0,1\}$ and every distinguisher $D : \{0,1\}^m \to \{0,1\}$ for $G^f_{\ell,m}$ with advantage p, we have that

$$\Pr_{y \in \{0,1\}^\ell}[R^D(A(D,f), y) = f(y)] \;\ge\; q.$$

One thing to notice is that the advice function does not depend on the input x. Thus, given f and D we can hardwire the value $A(D,f)$ and it is not counted in the circuit complexity.

Theorem 8. *For every $p > 0$, $G^f_{\ell,m}$ has $(p, q = \frac{1}{2} + \frac{p}{m})$ reconstruction with a reconstruction circuit $R \in \mathsf{SIZE}(O(m))^D$.*

Proof. (Sketch). Suppose $D : \{0,1\}^m \to \{0,1\}$ distinguishes $G^f_{\ell,m}$ with advantage p. By a hybrid argument there is an $1 \le i \le m$ where we get $\frac{p}{m}$ advantage. There is a way to fix the bits of $y|_{\overline{S_i}}$ so the advantage is preserved. Similarly, there is a way to fix the output bits $j > i$ so that the advantage is preserved. The advice functions $A(f)$ contains:

- The index $1 \le i \le m$,
- The fixing of $y|_{\overline{S_i}}$, i.e., the fixing of the seed y outside S_i,
- A string $w \in \{0,1\}^{m-i-1}$ that fixes all the bits after i so that the distinguishing gap is preserved,
- For every $j < i$, and every string $w \in \{0,1\}^{|S_j \cap S_i|}$, the values $f(\sigma)$, where σ is the restriction of y to S_i, when y outside S_i is fixed as before, and y restricted to S_i is σ.

Notice that the advice function contains $\log m + t + m + \sum_{j<i} 2^{|S_i \cap S_j|} = O(m)$ bits. We remark that some parts of the advice can be chosen at random (instead of being given as advice), e.g., the second and third items above.

We now describe the circuit R. On input $x \in \{0,1\}^m$, the first bits are fed with $f(x|_{S_1}), \ldots, f(x|_{S_{i-1}})$ computed by circuits implementing their truth tables, where the bits outside S_i are fixed according to the advice. The $m - i - 1$ bits following the i-th bits are fixed according to the advice as well. D is then applied, where the i-th bit is the input to the circuit. By the discussion above we have that $\Pr_{x|_{S_i} \in \{0,1\}^\ell}[R(x|_{S_i}) = f(x|_{S_i})] > \frac{1}{2} + \frac{p}{m}$. Also, the size of R is $O(m) + |D|$. □

With that we can prove Theorem 6:

Proof. Assume towards contradiction that $G^f_{\ell,m}$ is not a PRG against circuits of size $s - m^2$ with error ε. Hence, there exists a size $s - m^2$ circuit C such that $|C(U_m) - C(G_{\ell,m}(U_t))| > \varepsilon$. This implies that C is a distinguisher for $G_{\ell,m}$ with advantage ε, and by Theorem 8 we have a circuit R that computes f correctly on more than $\frac{1}{2} + \frac{\varepsilon}{m}$ of the inputs. As $|R| = |C| + O(m) \le s$, this is a contradiction to the fact that $\mathsf{Size}_{\frac{1}{2}+\frac{\varepsilon}{m}}(f) > s$. □

3 Conditional derandomization of BPP

We want to show that if there is a language decidable in time $2^{O(n)}$ and requires circuits of size $2^{\Omega(n)}$ then $\mathsf{P} = \mathsf{BPP}$. So far we have seen:

- How to derandomize BPP assuming a language that is hard *on average* for a family of circuits.
- A worst-case to average-case reduction for PSPACE (or higher classes).

Together this gives:

Theorem 9. *If there exists a language $L \in \mathsf{E}$ and a constant $c > 0$ such that L cannot be computed by circuits of size $2^{c \cdot n}$ then $\mathsf{P} = \mathsf{BPP}$.*

Proof. Consider $f : \{0,1\}^\ell \to \{0,1\}$ to be the truth table of L, and let $f' = C(f) : \{0,1\}^{\ell'=O(\ell)} \to \{0,1\}$. As discussed earlier, f' (as a language) is in E as well. We saw that no circuit of size $s' = 2^{(c/2)\ell}$ can compute f' correctly with advantage greater than $\varepsilon' = 2^{-c'\ell}$ for some constant c'.

We have that $G^{f'}_{\ell',m} : \{0,1\}^d \to \{0,1\}^n$ is a PRG against circuits of size $2^{(c/2)\ell} - n^2$ with error $\varepsilon' n$ and seed length $O(\frac{\ell^2}{\log n})$. Choosing ℓ to be a large enough multiplication of $\log n$, we have that the PRG's seed length is logarithmic in the output length and the error.

To show that $\mathsf{BPP} \subseteq \mathsf{P}$, let $A \in \mathsf{BPP}$ and let $C(x,y)$ be a circuit of polynomial size. Given an input x to A, let $C_x(y)$ be the circuit whose input bits are fixed. Setting the parameters according to the size of C, we have that $G^{f'}_{\ell',m}$ fools C_x. As the seed length is logarithmic in the output length and the error, we can set the error to be polynomially-small so by going over all the outputs of the generator and taking the majority vote we can decide whether $x \in A$ with high probability and in polynomial time. $\square$

References

[1] Noam Nisan and Avi Wigderson. Hardness vs randomness. *Journal of computer and System Sciences*, 49(2):149–167, 1994.

[2] Ran Raz, Omer Reingold, and Salil Vadhan. Extracting all the randomness and reducing the error in trevisan's extractors. In *Proceedings of the thirty-first annual ACM symposium on Theory of computing*, pages 149–158. ACM, 1999.