:COS 598: Lecture 1. Pseudo-random functions, Natural Proofs and Matrix Rigidity

Scribe: Simion Filip

Feb 08, 2008

1 Pseudo-random functions

Roughly speaking, a pseudo-random function is one which behaves in a non-deterministic way. While there is no universal definition, various properties can distinguish functions as pseudo-random.

Example 1.1. Let us consider the set

$$H = \{f | \text{ the smallest circuit for } f \text{ has size } \geq 2^{n/2}\}$$

If we denote by $\overline{H}$ the complement of H, it is clear that

$$|\overline{H}| \leq \#\{\text{circuits of size } \leq 2^{n/2}\} \leq O(2^{n/2}n)$$

Clearly then $Pr[f \in \overline{H}] \rightsquigarrow 0$, that is, most functions are in H.

Another good property that is definitely expected of random functions is given in the following

Example 1.2. Consider $f : \{0,1\}^n \to \{0,1\}$. We will say that $f \in H$ if after fixing $n - n^{1/10}$ indices $i_1, \dots i_{n-n^{1/10}}$ and the values of those input bits to $x_{i_j} = a_{i_j}$, then $f|_{\overline{a}} \neq \text{const}$.

We then have the following

Theorem 1.1 (follows from Hastad's switching lemma)**.** *If $f \in H$, for H as in the previous example, f does not have a polynomial-size constant-depth circuit.*

As an example of $f \in H$ with this property, consider $f(x_1, \dots x_n) = x_1 \oplus \dots \oplus x_n$. (Recall that only AND, OR and Negation are allowed in a circuit)

2 Natural Proofs

Suppose we would like to show that a function f has no small circuit, i.e.$f \notin SIZE(n^c)$ for $c >> 0$, where by $SIZE(n^x)$ we denote functions with circuits of size n^c. The paradigm of natural proofs is to find some H with the following properties:

Pseudo-largeness - $Pr[f \in H] \geq 1 - \varepsilon$, with ε very small.

Constructiveness - Can determine if $f \in H$ in $2^{O(n)}$ time.

Useful - $H \bigcap SIZE((n^c) = \emptyset$

With this in mind, we arrive at the following

Definition 2.1. H is "n^c-natural" if

$$f \in H \Rightarrow f \notin SIZE(n^c)$$

Although one might hope to prove various lower bounds by trying to find H's that are n^c-natural, the following theorem of Rudich and Razborov (see [RR97])

Theorem 2.1. *If we assume a conjecture that is weaker than average-case factoring n-bit integers* $\in SIZE(2^{n^{1/10}})$ *then there does not exist a* n^c*-natural* H *for* $c >> 0$.

3 The probabilistic method and constructions

Some of the most famous problems in combinatorics are related to Ramsey numbers. By definition, the Ramsey number $R(n,k)$ is defined to be the smallest number such that whenever one has a graph G with $N \geq R(n,k)$ vertices, the graph G either has an $n-clique$, or k independent vertices. A priori, it is not clear that $R(n,k)$ is finite. However, the following is easy to prove:

Theorem 3.1 (Ramsey). $R(n,k) \leq \binom{n+k}{k} + O(n+k)$

Proof. We use induction on $n+k$. Clearly, for $n = k = 1$, every graph on 2 vertices either has an edge, or 2 independent vertices, and in either case the bound holds. Suppose we are in the situation with $n+1$ and $k+1$. Suppose you have a graph. Pick any vertex v. If it has $R(n,k+1)$ neighbors, then either that subgraph has $k+1$ independent vertices, or it has an n-clique, which together with v gives an $n+1$-clique. If the vertex v has $R(n+1,k)$ vertices with which it is not connected, then that subgraph either has an $n+1$-clique, or it has k independent vertices, which together with v yield $k+1$ independent ones. So, we obtain that $R(n+1,k+1) \leq R(n+1,k) + R(n,k+1) + 1$, from which the theorem follows. □

If we look closer at the asymptotic of this sequence, we see that $R(n, n) \leq O(2^n)$. So, it is interesting to know a good lower bound. The first one to find a non-trivial lower bound was Erdos, but his proof is not constructive. This was probably the first use of the probabilistic method.

Theorem 3.2 (Erdos). $R(n, n) \geq O(2^{n/2})$

Proof. This is a simple matter of counting. We call a graph bad if it contains either an n-clique or n independent vertices. We would like to show that the number of bad graphs on N vertices is less than the total number of graphs on N vertices, for N reasonably large. Now, the number of bad graphs is bounded by $2\binom{N}{n} \cdot 2^{\binom{N}{2} - \binom{n}{2}}$, since $\binom{N}{n}$ is the total number of choices for a sub-graph with n vertices, and $2^{\binom{N}{2} - \binom{n}{2}}$ is the number of choices for the edges outside this fixed graph. The factor 2 comes from considering both cliques and independent sets. Now, we would like this number to be less than the total number of graphs, which is just $2^{\binom{N}{2}}$. So, we need

$$2\binom{N}{n} \cdot 2^{\binom{N}{2} - \binom{n}{2}} < 2^{\binom{N}{2}} \Leftrightarrow 2\binom{N}{n} < 2^{\binom{n}{2}}$$

It is enough to check this for $N = 2^{n/2}$. Let's approximate $\binom{N}{n} \leq \frac{N^n}{n!}$, and since $\binom{n}{2} = \frac{n(n-1)}{2}$, this is equivalent to $2^{n/2+1} < n!$, which is obvious. □

3.1 Hadamard Matrices

A good example, although far from sharp, related to the question of Ramsey numbers, is given by Hadamard matrices. These are defined as follows.

Definition 3.1. The entries of the Hadamard matrix $M_{2^n \times 2^n}$ are given by $M_{\overrightarrow{x}, \overrightarrow{y}} =< \overrightarrow{x}, \overrightarrow{y} >= \sum x_i y_i (\mod 2)$, where we consider the rows and columns indexed by binary vectors.

This matrix clearly gives us a graph on $N = 2^n$ vertices, and remark that the set $S = (v, v)$ has roughly $\sqrt{N}$ vertices, which are independent (a similar construction works for a clique). This example turns out to be sharp, as can be seen from the following theorem:

Theorem 3.3. *A clique or an independent set in the graph given by the Hadamard matrix has size* $\leq O(\sqrt{N})$.

Before proving the theorem, we begin with the following

Lemma 3.4 (The Johnson bound). *Suppose we have* $v_1, \dots v_t \in \mathbb{R}^n$ *vectors, with* $\|v_i\| = 1$, *and* $< v_i, v_j > \leq -\delta$, $\forall i \neq j$. *Then* $t + 1 \leq \frac{1}{\delta}$

Proof. (of the Lemma) We have

$$0 \leq \| \sum v_i \|^2 = \sum \|v_i\|^2 + \sum_{i \neq j} < v_i, v_j > \leq t - \delta(t^2 - t)$$

, hence $t + 1 \leq \dfrac{1}{\delta}$. □

Now, back to the proof of our theorem.

Proof. (of the theorem on Hadamard graphs) Suppose we pick $a = \sqrt{N}$ vertices given by vectors $v_1, \ldots, v_a$. To be able to apply the lemma, we view the vectors as having elements $\{\pm 1\}$ instead of $\{0, 1\}$. Now, we disregard the entries at the intersection of the same numbered columns, and notice that after normalization, the inequality in the lemma is satisfied by the scalar products of the vectors, with $\delta = O(N^{-\frac{1}{2}})$. We then have the desired bound. □

3.2 Unnatural constructions

As we saw in the previous section, it is quite non-trivial to present explicit examples using pseudo-random constructions. Moreover, there is a clear limit to what can be achieved by these methods. However, many interesting and quite strong (in terms of the lower bounds they provide) are constructions from Number Theory. They are often related to modular arithmetic and similar tools. We will work out an example, the Frankl-Wilson graph (see [FW81]). This is an example where for N vertices, we get maximal cliques or independent sets of size $O(\sqrt{\log N})$.

Example 3.1 (The Frankl-Wilson Graph)**.** Fix a prime p. We will construct a graph on $V = \binom{p^3}{p^2 - 1}$ vertices, which are parameterized by subsets of $p^2 - 1$ elements in $[p^3]$. We put an edge between two vertices (S, T) iff $|S \bigcap T| \equiv -1(\text{ mod p})$.

Theorem 3.5 ([FW81])**.** *Any clique or maximally independent set of vertices in the above graph has size at most* $p^{O(p)}$. *Given that* $V \geq (\dfrac{p^3}{p^2 - 1})^{p^2 - 1} \sim p^{p^2}$, *we get the promised bound.*

Proof.

$$(S, T) \in E \Leftrightarrow |S \bigcap T| \in \underbrace{\{p - 1, 2p - 1 \ldots, (p - 1)p - 1\}}_{|L| = p} =: L$$

We shall prove the statement for the cliques, and the statement for independent sets will follow from similar techniques. Consider the sets S_i which for a clique and note that $|S_i| = p^2 - 1$. Remark the following useful

Lemma 3.6. *If $S_1, \dots, S_n \in [n]$ and $\forall i \neq j, |S_i \bigcap S_j| \in L$ then $k \leq n^{O(|L|)}$*

Proof. (of the Lemma) Assign the natural correspondence $S_i \leftrightsquigarrow v_i \in \{0,1\}^n$. Let now

$$P_i(\overrightarrow{x}) = \prod_{l \in L} (< \overrightarrow{x}, v_i > -l)$$

Remark now that clearly $P_i(v_j) = 0$, unless $i = j$. This allows us to conclude that the P_i's are linearly independent as polynomials. Moreover, after reduction of powers of the x_i's, we see that the polynomials in the linear space of polynomials that have at most $|L|$ products of x_i, and hence of dimension at most $n^{|L|}$, whence $k \leq n^{|L|}$. □

Going back to our problem, we immediately obtain the bound we wanted for the case of cliques. For the independent set situation, we use a "modulo p" variant of the above argument to obtain similar bounds. □

4 Matrix Rigidity

Roughly speaking, a matrix is rigid if it cannot be written as the sum of a matrix of a low rank and a sparse matrix. More technically for our purposes, we have

Definition 4.1. A n by n matrix is called *rigid* if $\forall B, C$ with rank $B \leq \frac{n}{1000}$ and C is $n^{1/1000}$-sparse, i.e. every row of C has at most $n^1/1000$ non-zero elements, $A \neq B + C$.

Remark. While the above definition works for any field, one usually considers it over $\mathbb{F}_2$. In this case, notice that being rigid is a pseudo-random property. Indeed, the number of matrices of rank at most $\frac{n}{1000}$ is at most the number of products of two matrices of size $n/1000 \times n$ and $n \times n/1000$, which is $2^{\frac{2n^2}{1000}}$. The number of sparse matrices is even less, from which we see that the ratio of rigid matrices among all matrices is close to 1.

However, even proving that a particular matrix is rigid is a very difficult thing. No good methods are known, although any matrix one thinks of seems to be rigid. The importance of rigid matrices lies in the following theorem of Valiant.

Theorem 4.1 ([Val77])**.** *If A is a rigid matrix then $x \mapsto Ax$ has no $O(n)$-size, $O(\log n)$-depth linear circuit.*

Before sketching the proof, we notice the following general graph-theoretic

Lemma 4.2. *Consider a directed acyclic graph with depth $\leq c \log n$ and n vertices. (in our case, this will be the circuit). Then, $\forall \varepsilon$ there exists a set of edges S with $|S| \leq \frac{d^{\frac{1}{\varepsilon}}}{c \log \log n} < \frac{n}{1000}$ such that every path of length $\geq \varepsilon \log n$ hits S.*

Proof. Label the nodes of the graph with increasing (according to the direction of the graph) labels, and consider them in binary form, as binary sequences with $\lceil \log\log n \rceil$ bits. Let now X_i be the set of edges for which the starting and ending vertices have the most significant difference in bits on the i^{th} position. Removing one X_i, the depth of the modified graph roughly halved, since we can just forget about the i^{th} bit in the labeling and still get a good(w.r.t. the direction) labeling. If we order the sets X_i according to size and remove the smallest k, for k chosen big enough, we get that the bound on the number of removed edges is satisfied, and also that the depth decreased to $\frac{\log n}{2^k}$, which means there are no paths in the initial graph of length greater than that, not passing through the selected edges. □

Proof. (Sketch of Valiant's theorem) The idea is that if we temporarily disregard all but the edges chosen in the lemma and their predecessors, and view the last nodes as outputs, this projects an input into something of size less than, say, $n/1000$. This is clearly given by a matrix of rank at most $n/1000$. Now, notice that the remaining circuit has length less than $\varepsilon \log n$, whence it comes from a sparse matrix. Considering now the composition of these things, we can deduce that $A = B + C$, where B is of low rank and C is sparse. But this is in contradiction to the assumption that A is rigid. □

References

[FW81] P. Frankl and R. M. Wilson. Intersection theorems with geometric consequences. *Combinatorica. An International Journal of the János Bolyai Mathematical Society*, 1(4):357–368, 1981.

[RR97] Alexander A. Razborov and Steven Rudich. Natural proofs. *J. Comput. System Sci.*, 55(1, part 1):24–35, 1997. 26th Annual ACM Symposium on the Theory of Computing (STOC '94) (Montreal, PQ, 1994).

[Val77] Leslie G. Valiant. Graph-theoretic arguments in low-level complexity. In *Mathematical foundations of computer science (Proc. Sixth Sympos., Tatranská Lomnica, 1977)*, pages 162–176. Lecture Notes in Comput. Sci., Vol. 53. Springer, Berlin, 1977.