Boolean function complexity

Lecturer: *Nitin Saurabh* Meeting: 4
Scribe: *Nitin Saurabh* 15.05.2019

Consider the set $\mathcal{R}_k$ of restrictions $\rho\colon [n] \to \{0,1,*\}$ that leaves k variables unset. That is,

$$\mathcal{R}_k := \{\rho\colon [n] \to \{0,1,*\} \mid |\rho^{-1}(*)| = k\}.$$

In the last class we argued that there exists at least one restriction in $\mathcal{R}_k$ that led to a "non-trivial" reduction in formula size. We will now see that, in fact, this is true for a large fraction of them. To this end, we study the expected reduction in formula size when restricted with a restriction in $\mathcal{R}_k$ chosen *uniformly at random.*

Theorem 1 (Subbotovskaya's Theorem (restated))**.** *Let f be a Boolean function on n variables and $\rho \in \mathcal{R}_k$ be chosen uniformly at random. Then,*

$$\mathbb{E}_\rho[\mathsf{L}(f_\rho)] \leq \left(\frac{k}{n}\right)^{3/2} \mathsf{L}(f).$$

Proof. Similar to the proof of Theorem 11, we saw in the last lecture. We sample the restriction ρ in $n-k$ steps as follows: At any step choose a variable uniformly at random from the remaining ones and set it to 0 or 1 again uniformly at random. Clearly this process is equivalent to sampling uniformly at random from $\mathcal{R}_k$.

We now estimate the expected decrease in the formula size after the first step of this random restriction process. Let F be an optimal formula for f, and let ℓ_i be the number of leaves labeled by the variables x_i in F. Then, $\sum_i \ell_i = \mathsf{L}(f)$. We also know from Lemmas 8 and 9 in the previous lecture that for each variable x_i there are ℓ_i distinct siblings, and each sibling gets killed (removed from the formula) on exactly one of the settings to x_i. Let s_{i0} and s_{i1} be the number of siblings that gets killed on setting x_i to 0 and 1 respectively. Then, $s_{i0} + s_{i1} = \ell_i$. Thus, the expected decrease in formula size after the first step of the random process is,

$$\mathbb{E}[\text{decrease in formula size}] \geq \sum_{i=1}^{n} \frac{1}{n}\left[\frac{1}{2}\left(\ell_i + s_{i0}\right) + \frac{1}{2}\left(\ell_i + s_{i1}\right)\right] \geq \frac{3\cdot\mathsf{L}(f)}{2n}.$$

Therefore, the expected formula size after the first step is at most

$$\mathsf{L}(f) - \frac{3\cdot\mathsf{L}(f)}{2n} \leq \left(1-\frac{1}{n}\right)^{3/2}\mathsf{L}(f).$$

Analyzing the subsequent steps recursively as before, we obtain the theorem. □

Having obtained a bound on the expected formula size under the random restriction we can now use Markov's inequality to argue that with high probability the formula size will decrease.

Theorem 2 (Concentrated version)**.** *Let f be a Boolean function on n variables and $\rho \in \mathcal{R}_k$ be chosen uniformly at random. Then, with probability at least $3/4$,*

$$\mathsf{L}(f_\rho) \leq 4 \cdot \left(\frac{k}{n}\right)^{3/2} \mathsf{L}(f).$$

Proof. Use the previous theorem with Markov's inequality. □

The above theorems in fact hold for more general random restrictions. For $p \in [0,1]$, define a *p-random restriction* ρ to be a random restriction that independently decides to leave a variable unfixed with probability p, and sets it 0 or 1 with equal probabilities $(1-p)/2$. We denote this set of random restrictions by $\mathcal{R}_p$. Subbotovskaya basically studied the following question:

> What is the expected formula size of the restricted function when we apply a p-random restriction?

The easy answer to this question is $p \cdot \mathsf{L}(f)$. She showed that in fact formulas shrink more. That is,

Theorem 3 (Subbotovskaya)**.** *For any function f and $p \in [0,1]$, $\mathbb{E}_{\rho \in \mathcal{R}_p}[\mathsf{L}(f_\rho)] = O(p^{3/2}\mathsf{L}(f))$.*

This raises a natural question: how much more can the formula shrink? That is, can we improve the exponent on p in the above theorem? This exponent is known as *shrinkage exponent* in the literature.

Definition 4 (Shrinkage exponent)**.** *The shrinkage exponent of De Morgan formulas is the largest number Γ such that $\mathbb{E}_{\rho \in \mathcal{R}_p}[\mathsf{L}(f|_\rho)] = O(p^\Gamma \mathsf{L}(f))$ for any function f.*

It is easily seen (similar to the arguments in previous lecture) that whatever be the Γ, we obtain a lower bound of n^Γ for the Parity_n function. Therefore, we have that $\Gamma \leq 2$. In a long line of work it has been shown that $\Gamma = 2$: Impagliazzo and Nisan – $\Gamma \geq 1.55$ [IN93], Paterson and Zwick – $\Gamma \geq 1.63$ [PZ93], and Håstad – $\Gamma \geq 2$ [Hås98].

Theorem 5 ([Hås98, Tal14])**.** *For any function f and $p \in [0,1]$, $\mathbb{E}_{\rho \in \mathcal{R}_p}[\mathsf{L}(f|_\rho)] = O(p^2\mathsf{L}(f) + p\sqrt{\mathsf{L}(f)})$.*

For read-once formulas it was shown by Håstad, Razborov and Yao [HRY95] that $\Gamma_{\text{read-once}} = \frac{1}{\log(\sqrt{5}-1)} \approx 3.27$ (see also [DZ94]). For the monotone formulas it is conjectured that $\Gamma_{\text{monotone}} = \Gamma_{\text{read-once}}$.

1 Andreev's function : cubic lower bound

We now prove super-quadratic lower bounds against De Morgan formulas.

Let $n = 2^r$ and $m = n/r$. Define the function $U_n^{\oplus}\colon \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$ as follows. It's a Boolean function on $2n$ variables x and y. Let $x \in \{0,1\}^n$ be represented as

$$x = \begin{pmatrix} x_{1,1} & x_{1,2} & \cdots & x_{1,m} \\ x_{2,1} & x_{2,2} & \cdots & x_{2,m} \\ \vdots & \vdots & \ddots & \vdots \\ x_{r,1} & x_{r,2} & \cdots & x_{r,m} \end{pmatrix}.$$

Let $z_i = x_{i,1} \oplus \cdots \oplus x_{i,m}$ be the parity of the variables in the i-th row. Let $\#(z)$ be the integer represented by bit vector $(z_1, \ldots, z_r)$. Then,

$$U_n^{\oplus}(x,y) = y_{\#(z)}.$$

Theorem 6 (Andreev 1987)**.** $\mathsf{L}(U_n^{\oplus}(x,y)) \geq n^{5/2-o(1)}$.

Proof. Let $h(z)$ be a function on r variables that requires the largest De Morgan formula. Using Shannon-Riordan's lower bound (see Lecture 1), we have

$$\mathsf{L}(h) \geq \frac{2^r}{2\log r}.$$

Let $b \in \{0,1\}^n$ be the truth table of h. Consider the function $f(x) := U_n^{\oplus}(x,b)$. Then,

$$f(x) = h\left(\oplus_{j=1}^m x_{1,j}, \ldots, \oplus_{j=1}^m x_{r,j}\right).$$

We will analyze f when restricted by a random restriction from $\mathcal{R}_k$ for an appropriate choice of k. Our goal is to show that there exists a restriction ρ in $\mathcal{R}_k$ such that the following two properties hold simultaneously.

1. for all $i \in [r]$, $\oplus_{j=1}^m x_{i,j}$ is not a constant function when restricted by ρ. That is, at least one variable in each row of x remains unfixed. This ensures that f_ρ remains as hard as h.

2. The formula computing the restricted function f_ρ is much smaller than the formula computing f. In particular, $\mathsf{L}(f_\rho) \leq 4(k/n)^{3/2}\mathsf{L}(f)$.

Let us now see how the lower bound proof proceeds assuming that we have a restriction ρ satisfying the above properties. We have

$$\frac{2^r}{2\log r} \leq \mathsf{L}(h) \leq \mathsf{L}(f_\rho) \leq 4\left(\frac{k}{n}\right)^{3/2}\mathsf{L}(f) \leq 4\left(\frac{k}{n}\right)^{3/2}\mathsf{L}(U_n^{\oplus}), \tag{1}$$

where the first inequality follows from the choice of h, the second inequality follows from Property 1, the third inequality follows from Property 2, and the last one because f is a subfunction of $U_n^{\oplus}$. Plugging the appropriate value of k gives the lower bound.

We now proceed to show that for an appropriate choice of k we can find a restriction in $\mathcal{R}_k$ that satisfies the required properties. Let us compute the probability that a random restriction ρ in $\mathcal{R}_k$ does not satisfy the Property 1. This happens when all variables in some row are fixed. The probability that a random $\rho \in \mathcal{R}_k$ leaves a variable unfixed is exactly $\frac{k}{n}$. Therefore, the probability that all variables in a particular row are fixed is at most

$$\left(1 - \frac{k}{n}\right)^m.$$

Thus, by the union bound, the probability that some row is completely fixed is at most

$$r\left(1 - \frac{k}{n}\right)^m \leq r \cdot e^{-\frac{km}{n}} = r \cdot e^{-\frac{k}{r}}.$$

Hence, choosing $k = \lceil r \ln(4r) \rceil$, we obtain that with probability at least 3/4, a random $\rho \in \mathcal{R}_k$ leaves at least one variable in each row of x unfixed. Moreover, we know from Theorem 2 that for any k, with probability at least 3/4, a random $\rho \in \mathcal{R}_k$ satisfies Property 2. Therefore, there exists some $\rho \in \mathcal{R}_{\lceil r \ln(4r) \rceil}$ that satisfies both the properties.

Now plugging in $k = \lceil r \ln(4r) \rceil$ in Eq. (1), we obtain the theorem. □

Observe that, in fact, the proof shows a lower bound of $\Omega(n^{\Gamma+1-o(1)})$ for $\mathsf{L}(U_n^{\oplus})$ where Γ is the shrinkage exponent. Therefore, using Håstad's bound of 2 on the shrinkage exponent, we have $\Omega(n^3)$ lower bound for the same function.

2 Nechiporuk's method for formulas over $\mathcal{B}_2$

We now see a method due to Nechiporuck that gives lower bound for formulas over the basis $\mathcal{B}_2$. Recall $\mathcal{B}_2$ denotes the set of all Boolean functions over 2 variables.

Let f be Boolean function over $X = \{x_1, \ldots, x_n\}$. A *subfunction* of f over $Y \subseteq X$ is a function obtained from f by setting all variables in $X \setminus Y$ to constants. Nechiporuk's idea is based on the observation that a small formula can not compute a function with many distinct subfunctions.

Theorem 7 (Nechiporuk 1966)**.** *Let f be a Boolean function over X, and let Y_1, Y_2, ..., and Y_m be a partition of X. Let s_i be the number of distinct subfunctions of f on Y_i. Then,*

$$\mathsf{L}_{\mathcal{B}_2}(f) \geq \frac{1}{4}\sum_{i=1}^{m} \log s_i.$$

Proof. Let F be an optimal formula for f over $\mathcal{B}_2$ and let ℓ_i be the number of leaves labeled by the variables in Y_i. Clearly it suffices to prove that $\ell_i \geq (1/4)\log s_i$.

Consider the subtree T_i of F consisting of all leaves labelled by variables in Y_i and all paths from these leaves to the output of F. The indegree of the nodes in T_i is 0, 1, or 2. Let W_i be the set of nodes in T_i of indegree 2. Since $|W_i| = \ell_i - 1$, it suffices to lower bound $|W_i|$.

Let P_i be the set of paths in T_i starting from a leaf or a node in W_i and ending at a node in W_i or at the root of T_i and containing no node in W_i as a inner node. Since the number of edges in a binary tree with k internal nodes is at most $2k$, we obtain that

$$|P_i| \leq 2|W_i| + 1.$$

We have an extra one because the root of F may not be in W_i. We now count the number of possible distinct subfunctions on Y_i using the above structure of T_i. Let us fix an assignment ρ to the variables in $X \setminus Y_i$. Let p be a path in P_i. We claim that if h is the function computed at the first gate of p, then the function computed at the last edge of p (under the assignment ρ) is either 0, 1, h, or $\neg h$. This is because all (inner) gates on this path have indegree 1. Therefore, the possible number of subfunctions on Y_i is at most $4^{|P_i|}$. We thus have $s_i \leq 4^{|P_i|}$. This implies

$$\frac{1}{2} \log s_i \leq |P_i| \leq 2|W_i| + 1 = 2\ell_i - 1.$$

□

Using Nechiporuk's theorem we can show a quadratic lower bound for formulas over $\mathcal{B}_2$. Consider the following function, known as the *element distinctness function.*

Definition 8 (Element distinctness function)**.** *Let* $\mathsf{ED}_n \colon \underbrace{[m^2] \times \cdots \times [m^2]}_{m \text{ times}} \to \{0,1\}$*, where* $n = 2m \log m$ *and* m *is assumed to be a power of* 2*. Each of the* m *blocks encode a number in* $[m^2]$*. The function accepts an input* $x \in \{0,1\}^n$ *iff all these numbers are distinct.*

Theorem 9. $\mathsf{L}_{\mathbb{B}_2}(\mathsf{ED}_n) = \Omega(n^2/\log n)$*.*

Proof. Consider the partition of variable set X into m disjoint sets $Y_1, \ldots, Y_m$ corresponding to the variables in each m block. Since ED_n is symmetric with respect to blocks we only need to count the number of distinct subfunctions with respect to one of the blocks. Let N be the number of subfunctions of ED_n on Y_1. We now lower bound N.

Consider the set of all $(m-1)$-sized subsets of $[m^2]$. Note that the size of this set is $\binom{m^2}{m-1}$. For every element $\{a_2, \ldots, a_m\}$ of this set we obtain a subfunction $\mathsf{ED}_n(x, a_2, \ldots, a_m)$ over Y_1. We now claim that any two elements of this set give two distinct subfunctions. Let $\{b_2, \ldots, b_m\}$ be another element of this set. Then, there must be an $a_i \notin \{b_2, \ldots, b_m\}$. We thus have $\mathsf{ED}_n(a_i, a_2, \ldots, a_m) = 0$ whereas $\mathsf{ED}_n(a_i, b_2, \ldots, b_m) = 1$. Hence, the two subfunctions on Y_1 are distinct. Therefore, $N \geq \binom{m^2}{m-1}$. Using Nechiporuk's theorem we thus obtain the following lower bound

$$\mathsf{L}_{\mathcal{B}_2}(f) \geq \frac{1}{4} \cdot m \cdot \log \binom{m^2}{m-1} = \Omega(m^2 \log m) = \Omega\left(\frac{n^2}{\log n}\right).$$

□

We note that ED_n also has a matching upper bound.

Remark 2.1. *Nechiporuk's method can not prove better than quadratic lower bound.*

References

[DZ94] Moshe Dubiner and Uri Zwick. How do read-once formulae shrink? *Combinatorics, Probability and Computing*, 3(4):455469, 1994.

[Hås98] Johan Håstad. The shrinkage exponent of de morgan formulas is 2. *SIAM J. Comput.*, 27(1):48–64, 1998.

[HRY95] Johan Håstad, Alexander A. Razborov, and Andrew Chi-Chih Yao. On the shrinkage exponent for read-once formulae. *Theor. Comput. Sci.*, 141(1&2):269–282, 1995.

[IN93] Russell Impagliazzo and Noam Nisan. The effect of random restrictions on formula size. *Random Struct. Algorithms*, 4(2):121–134, 1993.

[PZ93] Mike Paterson and Uri Zwick. Shrinkage of de morgan formulae under restriction. *Random Struct. Algorithms*, 4(2):135–150, 1993.

[Tal14] Avishay Tal. Shrinkage of de morgan formulae by spectral techniques. In *55th IEEE Annual Symposium on Foundations of Computer Science, FOCS 2014, Philadelphia, PA, USA, October 18-21, 2014*, pages 551–560, 2014.