Formula lower bounds (lecture notes)

Course: Circuit Complexity, Autumn 2024, University of Chicago
Instructor: William Hoza (`williamhoza@uchicago.edu`)

---

1 The formula balancing lemma

Definition 1 (Formulas)**.** A *formula* is a circuit $C\colon \{0,1\}^n \to \{0,1\}$ in which each gate has fan-out (out-degree) at most 1. In other words, the underlying graph structure is a tree. A *De Morgan formula* is a formula in which the gates are AND gates and OR gates with fan-in two, with literals and constants at the leaves. The *leafsize* of a formula is the number of leaves in the underlying tree, excluding constants.

Lemma 1 (Formula balancing lemma)**.** *Let $f\colon \{0,1\}^* \to \{0,1\}$. The following are equivalent.*

1. *$f \in \mathsf{NC}^1$, i.e., f can be computed by circuits of depth $O(\log n)$ and size* $\mathrm{poly}(n)$ *over the full binary basis.*
2. *For every $n \in \mathbb{N}$, there is a De Morgan formula C_n of leafsize* $\mathrm{poly}(n)$ *that computes f restricted to inputs of length n.*

Proof. ($1 \implies 2$) If $f \in \mathsf{NC}^1$, then f can be computed by a "De Morgan circuit" (AND/OR gates of fan-in two, with literals and constants at the bottom) of depth $d = O(\log n)$. It is straightforward to show by induction on d that such a circuit can be simulated by a de Morgan formula of leafsize 2^d.

($2 \implies 1$) Let C be a De Morgan formula of leafsize $S = \mathrm{poly}(n)$. We will show by induction on S that C can be computed by a De Morgan formula of depth $3 \log S$. By starting at the root and always choosing the child with more leaf descendants, we can find a gate u in C with children u_L, u_R such that u has at least $S/2$ leaf descendants, whereas u_L and u_R have fewer than $S/2$ leaf descendants each. Identify u with the function $u(x)$ giving the output value at that gate. Let C_0 and C_1 be the formulas obtained from C by replacing u and all of its descendants with a 0 and a 1 respectively. Then

$$C(x) = (C_1(x) \wedge u(x)) \vee (C_0(x) \wedge \neg u(x)). \tag{1}$$

By induction, the output values of u_L and u_R can be computed by De Morgan formulas of depth at most $3 \log(S/2)$, hence $u(x)$ and $\neg u(x)$ can be computed by de Morgan formulas of depth $1 + 3\log(S/2)$. Furthermore, C_0 and C_1 have leafsize at most $S/2$, so by induction, $C_0(x)$ and $C_1(x)$ can be computed by De Morgan formulas of depth $3 \log(S/2)$. Therefore, C can be computed by a De Morgan formula of depth $3 + 3\log(S/2) = 3 \log S$. Note that such a formula necessarily has at most $O(S^3) = \mathrm{poly}(n)$ gates, hence it shows $f \in \mathsf{NC}^1$. ☐

Thus, the question of whether $\mathsf{NC}^1 = \mathsf{P/poly}$ is the question of whether circuits can be converted into formulas with polynomial overhead. The standard conjecture is "no."

2 Subbotovskaya's lower bound

For a function $f\colon \{0,1\}^n \to \{0,1\}$, let $L(f)$ denote the minimum leafsize of any De Morgan formula computing f. It turns out that $L(\mathsf{PARITY}_n) = \Theta(n^2)$. In this section, we will prove the weaker bound $L(\mathsf{PARITY}_n) \geq n^{1.5}$ via a beautiful and powerful technique called *random restrictions*.

Definition 2 (Restrictions)**.** A *restriction* is a string $\rho \in \{0,1,\star\}^n$. If f is a function on $\{0,1\}^n$, then $f|_\rho$ is another function on $\{0,1\}^n$, defined by the rule $f|_\rho(x) = f(y)$, where

$$y_i = \begin{cases} \rho_i & \text{if } \rho_i \in \{0,1\} \\ x_i & \text{if } \rho_i = \star. \end{cases}$$

Lemma 2 (Assigning a value to a single variable)**.** *Let $C\colon \{0,1\}^n \to \{0,1\}$ be a De Morgan formula of size S, where $n \geq 2$. There exists a restriction $\rho \in \{0,1,\star\}^n$ such that $|\rho^{-1}(\{0,1\})| = 1$ and $L(C|_\rho) \leq \left(1 - \frac{1.5}{n}\right) \cdot S$.*

Proof. If C is equivalent to a constant or a literal, then the lemma is trivial, so assume otherwise. The first step is to perform some simplifications to C before applying any restriction. For each subformula of the form $x_i \wedge g$, we can replace each occurrence of x_i in g with the constant 1, because if $x_i = 0$, then the subformula will evaluate to false regardless of what g does. Similarly, in a subformula of the form $\neg x_i \wedge g$, $x_i \vee g$, or $\neg x_i \wedge g$, we can replace each occurrence of x_i in g with an appropriate constant. Then, afterward, we can remove all constants from the formula, because $0 \wedge g \equiv 0$, $1 \wedge g \equiv g$, $0 \vee g \equiv g$, and $1 \vee g \equiv 1$. After making these simplifications, the new formula C' still has size at most S, and now it has the following property: For each vertex u, if ℓ is a leaf that is a child of u and ℓ' is a distinct leaf that is a descendant of u, then ℓ and ℓ' read distinct variables.

Now we are ready to perform the restriction. Pick ρ uniformly at random among all restrictions such that $|\rho^{-1}(\{0,1\})| = 1$. For each leaf ℓ, we divide into three cases.

- Perhaps ρ does not assign a value to the variable that ℓ reads. In this case, we define $K_\ell = \varnothing$.
- Perhaps ρ assigns a value to the variable that ℓ reads, making ℓ a constant, but the parent u of ℓ remains nonconstant. In this case, we define $K_\ell = \{\ell\}$.
- Perhaps ρ assigns a value to the variable ℓ reads that makes both ℓ and its parent u constant. (Note that $0 \wedge g \equiv 0$ and $1 \vee g \equiv 1$ for any g.) In this case, we define $K_\ell = \{\ell, \ell'\}$, where ℓ' is any other leaf that is a descendant of u.

By construction, the function $C|_\rho$ can be computed by a De Morgan formula constructed from C' by replacing some nodes with constants, thereby eliminating all the leaves in $\bigcup_\ell K_\ell$. Furthermore, because of the way we constructed C', we have $K_\ell \cap K_{\ell'} = \varnothing$ whenever $\ell \neq \ell'$. Therefore,

$$\mathbb{E}[L(C|_\rho)] \leq \mathbb{E}\left[S - \sum_\ell |K_\ell|\right] = S - \sum_\ell \left(1 \cdot \frac{0.5}{n} + 2 \cdot \frac{0.5}{n}\right) = S \cdot \left(1 - \frac{1.5}{n}\right).$$

The best case is at least as good as the average case. $\square$

Lemma 3 (Non-optimal shrinkage of De Morgan formulas)**.** *Let $f\colon \{0,1\}^n \to \{0,1\}$, let $p \in [0,1]$, and assume that pn is an integer. There exists a restriction $\rho \in \{0,1,\star\}^n$ such that $|\rho^{-1}(\star)| = pn$ and $L(f|_\rho) \leq p^{1.5} \cdot L(f)$.*

Proof. Let $k = pn$. If $k = 0$, the lemma is trivial, so assume $k \geq 1$. By applying Lemma 2 $n - k$ times, we construct a restriction ρ such that $|\rho^{-1}(\star)| = pn$ and

$$\begin{aligned}
L(f|_\rho) &\leq L(f) \cdot \prod_{i=k+1}^{n} \left(1 - \frac{1.5}{i}\right) \\
&\leq L(f) \cdot \prod_{i=k+1}^{n} \left(1 - \frac{1}{i}\right)^{1.5} && \text{(Bernoulli's inequality)} \\
&= L(f) \cdot \left(\prod_{i=k+1}^{n} \frac{i-1}{i}\right)^{1.5} \\
&= L(f) \cdot (k/n)^{1.5}.
\end{aligned}$$

$\square$

Theorem 1 (Non-optimal formula lower bound for parity)**.** $L(\mathsf{PARITY}_n) \geq n^{1.5}$.

Proof. By Lemma 3, there exists a restriction $\rho \in \{0,1,\star\}^n$ such that $|\rho^{-1}(\star)| = 1$ and

$$L(\mathsf{PARITY}_n|_\rho) \leq (1/n)^{1.5} \cdot L(\mathsf{PARITY}_n).$$

On the other hand, $\mathsf{PARITY}_n|_\rho$ is non-constant, so $L(\mathsf{PARITY}_n|_\rho) \geq 1$. $\square$

3 Near-cubic formula lower bounds

In the previous section, we used two steps to show that there exists an explicit function f (namely, the parity function) such that $L(f) \geq n^{1.5}$:

1. We showed that small De Morgan formulas simplify under random restrictions.

2. We constructed f such that f does not simplify under random restrictions.

It turns out that both of the steps above can be improved, as we now discuss.

3.1 Optimal shrinkage of De Morgan formulas

Definition 3 (Random restrictions)**.** Let $n \in \mathbb{N}$ and $p \in [0,1]$. We define R_p to be the distribution over $\{0,1,\star\}^n$ defined as follows. To sample $\rho \sim R_p$, for each coordinate $i \in [n]$ independently, set

$$\rho_i = \begin{cases} \star & \text{with probability } p \\ 0 & \text{with probability } (1-p)/2 \\ 1 & \text{with probability } (1-p)/2. \end{cases}$$

Theorem 2 (Optimal shrinkage of De Morgan formulas [Tal14])**.** *For every function $f\colon \{0,1\}^n \to \{0,1\}$ and every $p \in [0,1]$, we have*

$$\mathop{\mathbb{E}}_{\rho \sim R_p}[L(f|_\rho)] \leq O\left(p^2 \cdot L(f) + p \cdot \sqrt{L(f)}\right) \leq O(p^2 \cdot L(f) + 1).$$

The proof of Theorem 2 is omitted.

3.2 Andreev's function

Theorem 3 (Near-cubic formula lower bound)**.** *For every $n \in \mathbb{N}$, there exists a function $A\colon \{0,1\}^{2n} \to \{0,1\}$ ("Andreev's function") such that $A \in \mathsf{P}$ and $L(A) \geq \widetilde{\Omega}(n^3)$.*

Proof. Given $f \in \{0,1\}^n$ and $x^{(1)}, \dots, x^{(\log n)} \in \{0,1\}^{n/\log n}$, we interpret f as the truth table of a function $f\colon \{0,1\}^{\log n} \to \{0,1\}$, and we define

$$A(f, x^{(1)}, \dots, x^{(\log n)}) = f(\mathsf{PARITY}_{n/\log n}(x^{(1)}), \dots, \mathsf{PARITY}_{n/\log n}(x^{(\log n)})).$$

Clearly, $A \in \mathsf{P}$. To prove the formula lower bound, sample a restriction $\rho \sim R_p$, where $p = \Theta((\log^2 n)/n)$. On the one hand, by Theorem 2, we have

$$\mathbb{E}[L(A|_\rho)] \leq O(1 + p^2 \cdot L(A)) = O\left(1 + \frac{L(A) \cdot (\log n)^4}{n^2}\right).$$

On the other hand, let us show that $\mathbb{E}[L(A|_\rho)] \geq \widetilde{\Omega}(n)$.

After applying ρ, let us randomly assign values to the remaining variables in the "f" portion of the input of A. This can only make the formula size smaller. By Shannon's counting argument, with probability at least 0.9, the function f has circuit complexity $\Omega(n/\log n)$, hence it also satisfies $L(f) \geq \Omega(n/\log n)$.[1] Meanwhile, the probability that ρ assigns values to all $n/\log n$ of the variables in some block $x^{(i)}$ is at most $\log n \cdot (1-p)^{n/\log n} \leq \log n \cdot \exp(-pn/\log n) \ll 0.1$. Assuming this does not occur, it is possible to deterministically assign values to all but one variable in each block $x^{(i)}$ such that $\mathsf{PARITY}_n(x^{(i)})$ is simply

[1] In fact, Shannon's counting argument can be improved for the special case of De Morgan formulas, but let's just use the bound that we already proved.

a single variable. Consequently, under the resulting restriction ρ', the restricted function $A|_{\rho'}$ is simply f, applied to a subset of the variables. Thus, we have shown that

$$\Pr[L(A|_\rho) \geq \Omega(n/\log n)] \geq 0.8,$$

and hence $\mathbb{E}[L(A|_\rho)] \geq \Omega(n/\log n)$ by Markov's inequality. Combined with the upper bound on $\mathbb{E}[L(A|_\rho)]$, this implies $L(A) \geq \widetilde{\Omega}(n^3)$. $\square$

It is an open problem to show that some $h \in \mathsf{NP}$ satisfies $L(h) \geq n^{3+\Omega(1)}$.

References

[Tal14] Avishay Tal. "Shrinkage of De Morgan Formulae by Spectral Techniques". In: *Proceedings of the 55th Annual Symposium on Foundations of Computer Science (FOCS)*. 2014, pp. 551–560. DOI: 10.1109/FOCS.2014.65.