Lecture 3: Boolean Function Complexity Measures

Topics in Algorithms and Complexity Theory (Spring 2020)
Rutgers University
Swastik Kopparty
Scribes: Erica Cai and Zach Langley

In the previous lecture, we studied two important Boolean function complexity measures: *decision tree depth* and *certificate complexity*. In this lecture, we introduce four more measures—*sensitivity*, *block sensitivity*, *degree*, and *approximate degree*—and study how they relate to each other. Ultimately, we will show that all of these measures are equivalent up to polynomial factors, but in this lecture, we omit a lower bound for sensitivity.

1 Decision tree complexity versus certificate complexity

Recall that the *decision tree complexity* $D(f)$ of a Boolean function f is the depth of the smallest decision tree computing f. The decision tree complexity is also called the *deterministic query complexity*, since it is equivalently the fewest number of queries a deterministic algorithm needs to make in order to compute f.

Certificate complexity captures the non-deterministic analog of decision tree complexity. Recall that a *certificate* for an input $x \in \{0,1\}^n$ to a Boolean function f is a set $S \subseteq [n]$ of indices such that f is constant on all inputs that match x on S. Notationally, S is a certificate for x if $y|_S = x|_S$ implies $f(y) = f(x)$. The *certificate complexity of f at x* is the size of the smallest certificate for x. The *1-certificate complexity* $C_1(f)$ is then defined as $C_1(f) := \max_{x:f(x)=1} C(f,x)$, and the *0-certificate complexity* $C_0(f)$ is defined analogously as $C_0(f) := \max_{x:f(x)=0} C(f,x)$. Finally, the *certificate complexity* $C(f)$ of f is defined as $C(f) := \max\{C_0(f), C_1(f)\}$.

It should be clear that $C(f) \leq D(f)$, since the root-to-leaf path in the decision tree computation of an input x immediately gives a certificate for x. It is natural to wonder how large the gap between certificate and decision tree complexity can be. Could it be, for example, that $D(f) = \Omega(\sqrt{n})$ and $C(f) = O(\log n)$ for some function f? The next theorem says the answer is no: $D(f) \leq C(f)^2$ for all Boolean functions f.

Theorem 1. *For every Boolean function $f : \{0,1\}^n \to \{0,1\}$,*

$$D(f) \leq C_0(f) \cdot C_1(f).$$

Proof. Let $k = C_0(f)$ and let $\ell = C_1(f)$. As we saw last time, we may write f as a k-CNF $f = \bigwedge C_i$ and also as an ℓ-DNF $f = \bigvee T_i$.

Observe that every clause C_i must share at least one variable with every term T_j. Indeed, if $C_i(x) = 0$ for some x, then $\bigvee T_i(x) = f(x) = 0$. In other words, any clause being zero implies that all terms are zero.

We now demonstrate how to evaluate f in at most ℓk adaptive queries to x. First, we query the at-most-k variables in C_1 leaving us with a function f' on fewer variables after substituting the values for the variables. Since every term T_i shares at least one variable with C_1, after simplifying the terms knowing the values of the variables in C_1, every term is now is a conjunction of at most $\ell - 1$ literals. Thus, $C_1(f') \leq \ell - 1$. Inductively, we can evaluate f' in at most $(\ell - 1)k$ additional queries, and thus we use ℓk queries in total. □

Since decision tree depth captures the number of queries needed by a deterministic algorithm for f and certificate complexity captures the number of queries needed by a nondeterministic algorithm for f, the previous theorem may be considered a "$\mathsf{P} = \mathsf{NP} \cap \mathsf{co\text{-}NP}$" type of result for query complexity.

2 Sensitivity

Definition 2. *Let $f : \{0,1\}^n \to \{0,1\}$ be a Boolean function and let $x \in \{0,1\}^n$. The* sensitivity *of f at x is the number of positions $i \in [n]$ such that flipping the ith bit of x changes the output $f(x)$. The* sensitivity *$s(f)$ of f is the maximum sensitivity $s(f,x)$ over all points x.*

Informally, the sensitivity of a function f measures how unstable the output of f is to small perturbations to the input. Here are some examples:

- The sensitivity of AND_n is n, since at $x = 1^n$, flipping any bit in x changes the value of $\mathrm{AND}_n(x)$.
- Similarly, the sensitivity of OR_n is n, since at $x = 0^n$, flipping any bit in x changes the value of $\mathrm{OR}_n(x)$.
- For all $x \in \{0,1\}^n$, the sensitivity of PARITY_n at x is n.

We now prove a simple relationship between sensitivity and certificate complexity. In the proof, e_i is the Boolean vector which is 1 at position i and 0 elsewhere.

Proposition 3. *For every Boolean function f,*

$$s(f) \leq C(f).$$

Proof. If $f(x \oplus e_i) \neq f(x)$ for $i \in [n]$, then i must be present in every certificate for x. The set of all such i has size $s(f,x)$, and so $s(f,x) \leq C(f,x)$. □

3 Degree of polynomial representations

A *polynomial representation* of a Boolean function $f : \{0,1\}^n \to \{0,1\}$ is a polynomial $P : \mathbb{R}^n \to \mathbb{R}$ which agrees with f on Boolean inputs. For example, we can represent AND_3 by

$$(x_1, x_2, x_3) \mapsto x_1 x_2 x_3$$

and OR_3 by

$$(x_1, x_2, x_3) \mapsto x_1 + x_2 + x_3 - x_1x_2 - x_1x_3 - x_2x_3 + x_1x_2x_3.$$

Note that both examples are *multilinear*: no variable appears with an exponent larger than one. In general, a polynomial representation of a Boolean function never needs exponents larger than one since $x_i^k = x_i$ whenever $x_i \in \{0,1\}$ and $k \neq 0$. Moreover, this multilinear representation is unique, which we will now prove.

Proposition 4. *Every Boolean function $f : \{0,1\} \to \{0,1\}$ has a unique multilinear representation.*

Proof. First we show that there exists a multilinear representation of f. Inductively, there are multilinear polynomial representations Q and R for $(x_1, \ldots, x_{n-1}) \mapsto f(x_1, \ldots, x_{n-1}, 0)$ and $(x_1, \ldots, x_{n-1}) \mapsto f(x_1, \ldots, x_{n-1}, 1)$, respectively. Thus, we may write

$$P(x_1, \ldots, x_n) = x_n R(x_1, \ldots, x_{n-1}) + (1 - x_n)Q(x_1, \ldots, x_{n-1}),$$

which is multilinear.

We now show uniqueness. Suppose P and P' are both multilinear representations of f so that $(P - P')(x) = 0$ for all $x \in \{0,1\}^n$. And suppose to obtain a contradiction that $P - P'$ is not identically zero and therefore contains some monomial. Let $S \subseteq [n]$ be a minimal set of indices such that the monomial $\prod_{i \in S} x_i$ appears in $P - P'$ with nonzero coefficient. Notice that $(P - P')(\chi_S) \neq 0$, a contradiction, and therefore, $P - P'$ is identically zero. □

(Appealing to linear algebra, we obtain an even simpler proof: The set of multilinear polynomials over $\mathbb{R}$ forms a 2^n dimensional vector space, and there is an obvious orthonormal basis of degree 2^n which is just the set of monomials.)

That Boolean functions admit unique multilinear representations gives rise to another natural complexity measure, namely, the degree of the multilinear representation.

Definition 5. *The* degree $\deg(f)$ *of f is the degree of the unique multilinear polynomial representation of f.*

We now show that the degree of a Boolean function is at most its decision tree complexity.

Theorem 6. *For every Boolean function f,*

$$\deg(f) \leq D(f).$$

Proof. Fix a decision tree for f. For each leaf ℓ of the decision tree, define P_ℓ the polynomial which equals one on inputs that reach ℓ and 0 otherwise. That is, to form P_ℓ we multiply the terms x_i for each internal node i for which we went "right" on the way to ℓ and $(1 - x_i)$ for each internal node i for which we went "left" on the way to ℓ.

Now define $P(x) := \sum_\ell T_\ell P_\ell(x)$, where T_ℓ is the label of the leaf ℓ. This multilinear polynomial P represents f and has degree $\max_\ell\{\deg(P_\ell(x))\} = D(f)$. □

The following theorem is was proved by Nisan and Szegedy [4] in 1994.

Theorem 7 (Nisan and Szegedy [4]). *For every Boolean function f,*

$$\deg(f) \geq \sqrt{s(f)/2}.$$

Before proving Theorem 7, we will a few ingredients. First, we state an inequality by Markov.[1]

Theorem 8 (Markov's Inequality). *If $q : \mathbb{R} \to \mathbb{R}$ is a polynomial of degree d such that $b_1 \leq q(x) \leq b_2$ whenever $a_1 \leq x \leq a_2$, then the derivative of Q satisfies $|q'(x)| \leq d^2(b_2 - b_1)/(a_2 - a_1)$ whenever $a_1 \leq x \leq a_2$.* □

Markov's inequality will not be immediately useful to us as stated because its hypothesis requires us to know something about all real x within a certain range, whereas we will only be able to reason about *integer* x within a certain range. However, it is not hard to use Markov's inequality to prove the following corollary pertaining to integer values.

Corollary 9. *If $q : \mathbb{R} \to \mathbb{R}$ satisfies*

1. *$b_1 \leq q(n) \leq b_2$ for all $n \in \{0, 1, \ldots, m\}$, and*
2. *for some real $x \in [0, m]$, the derivative of q satisfies $|q'(x)| \geq c$,*

then $\deg(q) \geq \sqrt{cm/(c + b_2 - b_1)}$. □

Another key component in our proof is the "method of symmetrization," attributed to Minksy and Papert [2]. For a polynomial $p : \mathbb{R}^n \to \mathbb{R}$, the *symmetrization* $p^{sym} : \mathbb{R}^n \to \mathbb{R}$ averages p over all inputs:

$$p^{sym}(x_1, \ldots, x_n) := \frac{1}{n!} \sum_{\pi \in S_n} p(x_{\pi(1)}, \ldots, x_{\pi(n)}).$$

Note that $\deg(p^{sym}) \leq \deg(p)$.

The polynomial p^{sym} is still a multivariable polynomial, which prevents us from applying Markov's inequality to it directly. Thus, the final ingredient allows us to reason about the degree of a univariate polynomial rather than a multivariate one.

Lemma 10 (Minsky and Papert [2]). *For every multilinear polynomial $p : \mathbb{R}^n \to \mathbb{R}$ there is a univariate polynomial $q : \mathbb{R} \to \mathbb{R}$ such that* $\deg(q) \leq \deg(p)$ *and*

$$p^{sym}(x_1, \ldots, x_n) = q\left(x_1 + \cdots + x_n\right)$$

for all $x \in \{0, 1\}^n$.

[1] This inequality is not to be confused with the more famous "Markov's inequality" about probabilities of non-negative random variables. The Markov is the same—Andrey Markov—although his lesser-known brother Vladimir Markov generalized the inequality to higher order derivatives, and so the result is sometimes called the "Markov brothers' inequality."

Proof. Define

$$P_k(x) := \sum_{S \in \binom{[n]}{k}} \prod_{i \in S} x_i$$

to be the sum of all monomials of degree k. Letting d be the degree of p, by induction we may write

$$p^{sym}(x) = c_0 + c_1 P_1(x) + \cdots + c_d P_d(x),$$

with $c_i \in \mathbb{R}$, leveraging the symmetry of p^{sym}. Now notice that $P_k(x) = \binom{|x|}{k}$ and define the univariate polynomial q by

$$q(t) := c_0 + c_1 \binom{t}{1} + \cdots + c_d \binom{t}{d},$$

which satisfies $p^{sym}(x) = q(|x|)$.

□

We are now ready to prove Theorem 7.

Proof of Theorem 7. We may assume that $\arg\max_x s(f,x) = 0$ and that $f(0) = 0$. Let g be the restriction of f to its m sensitive bits on input 0. That is, $g = f|_S$, where $S := \{i : f(e_i) = 1\}$ and $|S| = m$. Let p be the multilinear representation of g and let $q : \mathbb{R} \to \mathbb{R}$ be the univariate polynomial promised by Lemma 10 with respect to p^{sym}.

Notice that $q(1) = 1$ since g is sensitive on every input. Also, $q(0) = 0$, and so by the mean value theorem there is some $x \in [0,1]$ such that $q'(x) \geq 1$. Finally, by Corollary 9, we have $\deg(f) \geq \deg(q) \geq \sqrt{m/2} = \sqrt{s(f)/2}$. □

4 Approximate Degree

Definition 11. *Let f be a Boolean function. A real-valued polynomial p* approximates *f if for every $x \in \{0,1\}^n$, it holds that $|p(x) - f(x)| \leq 1/3$. The* approximate degree $\widetilde{\deg}(f)$ *is the minimum degree of p over all polynomials p that approximate f.*

The constant $1/3$ is inconsequential; we may replace it with any other constant in $(0, 1/2)$ and the results will only change by a constant factor.

Clearly $\deg(f) \leq \widetilde{\deg}(f)$ since the multilinear representation approximates f. However, $\widetilde{\deg}(f)$ is also not too small with respect to the other measures.

Theorem 12 (Nisan and Szegedy [4])**.** *For every Boolean function f,*

$$\widetilde{\deg}(f) \geq \sqrt{s(f)/6}.$$

The proof follows along exactly the same lines as that of Theorem 7.

5 Block sensitivity

Here we use the notation x^B for some $B \subset [n]$ to mean x with the coordinates in B flipped. A *sensitive block for f at x* is a set $B \subseteq [n]$ such that $f(x) \neq f(x^B)$.

Definition 13. *The* block sensitivity $bs(f, x)$ *of f at x is the maximum number of disjoint sensitive blocks for f at x.*

Block sensitivity was first studied by Nisan [3] in 1991, where he proved the following theorem.

Theorem 14 (Nisan [3])**.** *For every Boolean function f,*

$$bs(f) \geq \sqrt{C(f)}.$$

Proof. First we show that if B is a minimal sensitive block of f at some x, then $|B| \leq s(f)$. Indeed, if B is a minimal sensitive block, then $f(x^B) \neq f(x)$ and $f(x^C) = f(x)$ for all proper subsets $C \subsetneq B$. Thus, $s(f, x^B) \geq |B|$, since any flipping any bit in B of x^B is equivalent to flipping some $|B| - 1$ bits of x. It follows that $s(f) \geq s(f, x^B) \geq |B|$.

Let x be any input and let $B_1, \ldots, B_m$ be disjoint sensitive blocks such that $m = bs(f, x)$. For each $i \in [m]$, replace B_i with a minimal subset $\widetilde{B}_i \subseteq B_i$ such that $\widetilde{B}_i$ is a sensitive block. We claim that $S := \bigcup_{i=1}^m \widetilde{B}_i$ is a certificate for f at x.

Suppose to obtain a contradiction that for some input y we have $x|_S = y|_S$ but $f(x) \neq f(y)$. Let $B^* = \{i : x_i \neq y_i\}$. Notice that $f(x^{B^*}) = f(y) \neq f(x)$, and so B^* is a sensitive block. Also, by definition, $B^* \cap S = \emptyset$. Thus $\widetilde{B}_1, \ldots, \widetilde{B}_m, B^*$ are disjoint sensitive blocks for f at x, contradicting that $m = bs(f, x)$.

Now if $m \geq \sqrt{C(f)}$, we are done. Otherwise, there is some i such that $|\widetilde{B}_i| \geq \sqrt{C(f)}$, in which case, $bs(f, x^{B_i}) \geq \sqrt{C(f)}$. □

We also have the following theorem, similar to Theorem 7, which we state without proof. (The proof uses similar ideas to those we have already seen.)

Theorem 15 (Nisan and Szegedy [4])**.** *For every Boolean function f,*

$$\deg(f) \geq \sqrt{bs(f)/2},$$

and

$$\widetilde{\deg}(f) \geq \sqrt{bs(f)/6}.$$

6 Conclusion and the Sensitivity Conjecture

At this point, we have almost shown that decision tree complexity, certificate complexity, sensitivity, block sensitivity, degree and approximate degree are all within polynomial factors of each other. For example, we have shown that $\deg(f) \geq \sqrt{bs(f)/2}$ in Theorem 15, and also that $\deg(f) \leq D(f) \leq (C(f))^2 \leq bs(f)$ by Theorems 1 and 14. However, we cannot yet show, to take another

example, that sensitivity is bounded below by a polynomial in decision tree complexity. The circle would be complete by showing that sensitivity is bounded below by a polynomial in any of the other measures, a statement known as the "Sensitivity Conjecture." Nisan and Szegedy [4] put forth this conjecture in 1994, and it remained open until 2019, when Hao Huang [1] settled it with a remarkable proof.

Theorem 16. *For every Boolean function f,*

$$s(f) \geq \sqrt{\deg(f)}.$$

Thus, in fact, every pair of Boolean function complexity measures in this lecture are within a polynomial factor of each other for all Boolean functions f.

References

[1] Hao Huang. Induced subgraphs of hypercubes and a proof of the sensitivity conjecture. *Ann. of Math. (2)*, 190(3):949–955, 2019.

[2] Marvin Minsky and Seymour A Papert. *Perceptrons: Expanded Edition.* The MIT Press, Cambridge, MA, 1988. First edition appeared in 1968.

[3] Noam Nisan. CREW PRAMs and decision trees. *SIAM J. Comput.*, 20(6):999–1007, 1991.

[4] Noam Nisan and Márió Szegedy. On the degree of Boolean functions as real polynomials. *Comput. Complexity*, 4(4):301–313, 1994. Special issue on circuit complexity (Barbados, 1992).