CSC2429 / MAT1304: Circuit Complexity 17 January 2019

Lecture 2: Stuff

Instructor: Benjamin Rossman *Scribe: Scribe*

1 Circuit Size Hierarchy

Theorem 1. ***(Circuit Size Hierarchy Theorem.)*** *If* $n \leq s(n) \leq \frac{2^{n-2}}{n}$, *then* $\mathsf{SIZE}[s] \subsetneq \mathsf{SIZE}[4s]$.

Proof. Combination of Shannon and Lupanov. Pick[1] $m < n$ such that

$$s(n) \leq \frac{2^m}{m} \leq 2s(n).$$

By Shannon, there exists a function $f : \{0,1\}^m \to \{0,1\}$ such that

$$\mathcal{C}(f) > \frac{2^m}{m} \geq s(n).$$

Thus $f \notin \mathsf{SIZE}[s]$. By the tight bound from Lupanov's theorem, $\mathcal{C}(f) \leq 2^m/m + o(2^m/m)$ so

$$\mathcal{C}(f) \leq \frac{2 \cdot 2^m}{m} \leq 4s(n)$$

and $f \in \mathsf{SIZE}[4s]$. □

2 Khrapchenko and Koutsoupias Lower Bound

Let us define $PARITY_n$ as $\bigoplus_n$ and $1 - PARITY_n$ as $\overline{\bigoplus}_n$. Recall[2] that $\mathcal{C}(\bigoplus_n) \leq 3(n-1)$ and $\mathcal{L}(\bigoplus_n) \leq 2^{2\lceil \log n \rceil}$. We will show that these bounds are tight.

Notation: $\lambda(\mathbf{P})$ is the largest eigenvalue of a symmetric matrix $\mathbf{P}$. Recall that

$$\lambda(\mathbf{P} + \mathbf{Q}) \leq \lambda(\mathbf{P}) + \lambda(\mathbf{Q}).$$

For non-empty $A, B \subseteq \{0,1\}^n$, the matrix $\mathbf{M} \subseteq \{0,1\}^{A \times B}$ is the matrix

$$\mathbf{M}_{a,b} = \begin{cases} 1 & \text{if } a_i \neq b_i \text{ for exactly one } i \\ 0 & \text{otherwise} \end{cases}$$

[1] Such an m must exists. When $m = 1$, $2^m/m \leq s(n)$ and when $m = n-1$, $2^m/m \geq s(n)$ so there must be some m such that $2^m/m \leq s(n)$ and $2^{m+1}/(m+1) \geq s(n)$. If $2^{m+1}/(m+1) \geq 2 \cdot s(n)$ then

$$s(n) \leq \frac{2^m}{m+1} \leq \frac{2^m}{m}$$

which contradicts our original choice of m.

[2] Construct a circuit with $n-1$ $\oplus$-gates and substituting three DeMorgan gates $(x \wedge \neg y) \vee (\neg x \wedge y)$ for each $x \oplus y$.

you can read this as "the hamming distance of $\mathbf{a}$ and $\mathbf{b}$ differs by exactly one". Note that $\mathbf{M}^\intercal\mathbf{M} \in \mathbb{N}^{B\times B}$ with entry (i,j) interpreted as "the number of vectors $\mathbf{a} \in A$ such that both $\mathbf{b}_i$ and $\mathbf{b}_j$ are one away from $\mathbf{a}$". Similarly $\mathbf{M}\mathbf{M}^\intercal \in \mathbb{N}^{A\times A}$ with entry (i,j) interpreted as "the number of vectors $\mathbf{b} \in B$ such that both $\mathbf{a}_i$ and $\mathbf{a}_j$ are one away from $\mathbf{b}$". It is a fact from linear algebra that $\mathbf{M}^\top\mathbf{M}$ and $\mathbf{M}\mathbf{M}^\top$ have the same non-zero eigen-values. In particular, $\lambda(\mathbf{M}^\intercal\mathbf{M}) = \lambda(\mathbf{M}\mathbf{M}^\intercal)$.

Theorem 2. *(Koutsoupias '93) For any $f : \{0,1\}^n \to \{0,1\}$ and $A \subseteq f^{-1}(0)$ and $B \subseteq f^{-1}(1)$,*

$$\mathcal{L}(f) \geq \lambda(\mathbf{M}^\intercal\mathbf{M}).$$

Proof. By induction on $\mathcal{L}(f)$. The base case occurs when $\mathcal{L}(f) = 1$. Thus the circuit only reads in one out of the n variables of the input. W.l.o.g assume that the input to the leaf is x_1. Then $f(\mathbf{x}) = x_1$ or $f(\mathbf{x}) = 1 - x_1$; w.l.o.g assume the former for ease of exposition. Let $A = f^{-1}(0)$ and $B = f^{-1}(1)$. Then $A = \{0s : s \in \{0,1\}^{n-1}\}$ and $B = \{1s : s \in \{0,1\}^{n-1}\}$. Recall that entry (i,j) of $\mathbf{M}^\top\mathbf{M}$ is the number of elements $\mathbf{a} \in A$ such that both $\mathbf{b}_i$ and $\mathbf{b}_j$ differ from $\mathbf{a}$ by one. Notice that $\mathbf{a} = 0s$ and $\mathbf{b} = 1s'$ differ by exactly one if and only if $s = s'$ thus $\mathbf{M}^\top\mathbf{M}$ is exactly the identity matrix of dimension $|B|$. Thus $\lambda(\mathbf{M}^\top\mathbf{M}) = 1$ satisfying the theorem.

In the inductive step, let F be a formula which computes f of size $\mathcal{L}(f)$. Suppose that $F = F_1 \wedge F_2$ for some circuits F_1 and F_2. Let f_1 and f_2 be the functions computed by F_1 and F_2 respectively. Notices that $\mathcal{L}(f) = \mathcal{L}(f_1) + \mathcal{L}(f_2)$. □

The following is only for the DeMorgan basis.

Corollary 3. *(Khrapchenko 1971)*

$$\mathcal{L}(f) \geq \frac{\left(\sum_{a\in A}\sum_{b\in B}\mathbf{M}_{a,b}\right)^2}{|A|\cdot|B|}$$

Proof. [3]. The idea is to use Rayleigh quotient to write out $\lambda(\mathbf{M}^\intercal\mathbf{M})$ and then find a particular vector, namely $\mathbf{z} = \mathbb{1}$, to plug into the inequality to get that lower bound. □

The above technique can achieve a gap of at most n^2. An example of its application would be to show that $\mathcal{L}(\bigoplus_n) \geq n^2$. Take A and B to be the set of even and odd[4] strings of n respectively. Then, by the above corollary,

$$\mathcal{L}(f) \geq \frac{\left(\sum_{a\in A}\sum_{b\in B}\mathbf{M}_{a,b}\right)^2}{|A|\cdot|B|} = \frac{\left(n2^{n-1}\right)^2}{2^{n-1}\cdot 2^{n-1}} = n^2.$$

Exercise: (1)[5] prove lower-bound $\mathcal{L}(MAJ_n) \geq \Omega(n^2)$ and (2) can you devise an upper bound of $\mathcal{L}(n^{O(1)})$.

[3]:) I like this

[4]Here the parity of the string corresponds to the parity of the sum of ones.

[5]Hint: Take $A = \{s \in \{0,1\}^n : s \text{ has exactly } \lceil n/2\rceil - 1 \text{ ones}\}$ and $B = \{t \in \{0,1\}^n : t \text{ has exactly } \lceil n/2\rceil \text{ ones}\}$.

3 Restrictions and Gate Elimination

For $i \in [n]$ and $b \in \{0,1\}$ the *1-bit restriction*, $x_i \leftarrow b$ is the n-ary function $f^{(x_i \leftarrow b)}$. The same can be done for circuits C, namely, $C^{(x_i \leftarrow b)}$. The technique is to substitute $x_i \leftarrow b$ and $\overline{x}_i \leftarrow 1 - b$ and performing the relevant simplifications.

Theorem 4. *(Schnorr 1979)* $C(XOR_n) \geq 3(n-1)$.

Proof. By induction. The base case where $n = 1$ is trivial. The crucial observation is as follows. If a literal is below k $\wedge/\vee$ gates (of the same type), then there is a setting of the literal such that you can knock out at least k gates. Just think about the different settings of the literal.

Consider any circuit C which calculates the XOR_n function. Identify three gates in C:

1. A gate whose inputs are two literals. Let these be x_i and x_j.
2. Pick a literal of the previous gate, say x_i. Find another gate with x_i as an input. Suppose such a gate does not exist. Then, by setting x_j appropriately, we could knock out the gate in step 1 and the output would not depend on x_i. This would not calculate the XOR_n function.
3. The gate above the one in step 2. Such a gate exists if the gate from step 2 is not the output of the circuit. Suppose for a contradiction that it was. Then a setting of x_i would fix the output. This would also not calculate the XOR_n function.

By setting x_i appropriately, we can kills all three gates above. See Figure.

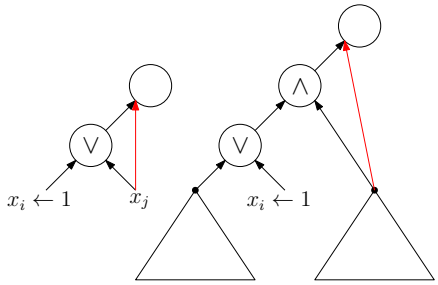

Figure 1: The three gates that get eliminated when we restrict x_i. The actual setting of x_i depends on the actual gate type.

By the induction hypothesis, $C^{(x_i \leftarrow b)}$ has at least $3(n-2)$ gates. Since we were able to eliminate three gates by setting x_i, we know that C has to have $3(n-1)$ gates. □

4 Subbotovskaya's Method (1961)

Definition 5. *A formula F is* ***nice*** *if for every sub-formula of the form $x_i \wedge F'$, $\overline{x}_i \wedge F'$, $x_i \vee F'$, $\overline{x}_i \vee F'$, the variable x_i does* not *occur in F'.*

Lemma 6. *Every formula is equivalent to a* nice *formula of the same (or less) leaf size.*

Proof. Repeated apply $x_i \wedge F' \leftarrow x_i \wedge F'^{(x_i \leftarrow 1)}$ (similarly for the other three subformulas). □

Lemma 7. *For every $f : \{0,1\}^n \to \{0,1\}$,*

$$\mathop{\mathbb{E}}_{i \in [n], b \in \{0,1\}} \left[\mathcal{L}\left(f^{(x_i \leftarrow b)} \right) \right] \leq \left(1 - \frac{1}{n} \right)^{1.5} \mathcal{L}(f).$$

Proof. □

Definition 8. *A **restriction** ρ is a function $\rho : [n] \to \{0,1,*\}$. Further ρ is a k**-start restriction** if $\rho^{-1}(*) = k$.*

*Let $p \in [0,1]$. The p**-random restriction** where you set*

$$R_p(i) = \begin{cases} * & \text{with probability } p \\ 0 & \text{with probability } \frac{1-p}{2} \\ 1 & \text{with probability } \frac{1-p}{2} \end{cases}$$

Open problem: shrinkage exponent of monotone formulas? (this is known to be between 2 and $\left(\log(\sqrt{5}) - 1\right)^{-1} = 3.27$).

5 Composition of Boolean Functions

Definition 9. *Let $f : \{0,1\}^k \to \{0,1\}$ and $g : \{0,1\}^m \to \{0,1\}$. Let $f \otimes g : (\{0,1\}^m)^k \to \{0,1\}$ is defined as*

$$(f \otimes g)(\mathbf{x}_1, ..., \mathbf{x}_k) = f\left(g(\mathbf{x}_1), ..., g(\mathbf{x})\right)$$

so basically the composition is of the form $f \otimes g = f \circ g^k$.

Think of the input of the composition as a matrix $\mathbf{X} \in \{0,1\}^{k \times m}$ with rows $\mathbf{x}_1, ..., \mathbf{x}_k$. Apply g to each row, then apply f to the resulting column vector. Observe that $\mathcal{L}(f \otimes g) \leq \mathcal{L}(f) \cdot \mathcal{L}(g)$.

The following is an explicit n-ary Boolean function for which the converse is true.

Lemma 10. *For all $k, m \geq 1$ and $f : \{0,1\}^k \to \{0,1\}$,*

$$\mathcal{L}\left(f \otimes \bigoplus\nolimits_m \right) \geq \mathcal{L}(f) \cdot \Omega\left(\frac{m}{\log k} \right)^2.$$

Proof. Let $p = \frac{2 \ln k}{m}$. Apply R_p on $k \times m$ variables of $f \otimes \bigoplus_m$. If R_p has a $*$ in every row then

$$\mathcal{L}\left(\left(f \otimes \bigoplus\nolimits_m \right) \rightharpoonup R_p \right) \geq \mathcal{L}(f)$$

since a formula which calculates the LHS can be used to calculate the RHS. It remains to bound the probability that the above. This is just your standard combination of probability, union bound, Markov inequality, and the shrinkage formula from before. Just rearrange and do not forget $1 - \frac{1}{k} \in O(1)$ and drop the one. □

This is a special case of a general conjecture on the leaf size of composed functions.

Conjecture 11. ***(KRW).*** *For all functions f and g,*

$$\mathcal{L}\, f \otimes g = \tilde{\Omega}\left(\mathcal{L}(f) \cdot \mathcal{L}(g)\right)$$

where $\tilde{\Omega}(t(n)) = \Omega(t(n))/\left(\log t(n)\right)^{O(1)}$ for any function $t(n)$.

Definition 12. *For parameters $k, m \in \mathbb{N}$. $ANDREEV_{k,m} : \{k\text{-ary Boolean function}\} \times \{0,1\}^{k \times m} \leftarrow \{0,1\}$ such that $ANDREEV(f, X) = \left(f \otimes \bigoplus_m\right)(\mathbf{X})$.*

Observe that when $m = 1$, then you get the multiplexor function. Further it should be noted that $\mathcal{C}(ANDREEV_{k,m}) = O(n)$.

Lemma 13. *For every $f : \{0,1\}^k \leftarrow \{0,1\}$ we have*

$$\mathcal{L}(ANDREEV_{k,n}) \geq \mathcal{L}(f \otimes \bigoplus\nolimits_m).$$

This result uses the Shannon theorem idea. Also this guy is tight.

6 General Binary Basis

Let B_2 be the full binary basis (all Binary functions are acceptable gates). Unfortunately, the random restriction idea above no longer works.

Definition 14. *For $f : \{0,1\}^n \leftarrow \{0,1\}$ and $V \subset [n]$,*

$$sub_V(f) = \{f \rightharpoonup \rho : \rho : [n] \to \{0,1,*\} \text{ such that } \rho^{-1}(*) = V\}$$

For an example, consider $MAJ_3(x_1, x_2, x_3)$ and $V = \{1,2\}$ Note that $sub_V(MAJ_3)$ is .

Define further $sub^*_V(f) = \{f', 1 - f', \mathbf{0}, \mathbf{1} : f' \in sub_V(f)\}$

There is also a definition for ℓ_v

Suppose $F = gate(G, H)$ for some $gate : \{0,1\}^2 \to \{0,1\}$. Then

$$sub_V(f) \subseteq \{gate(G, H), g \in sub_V(G), h \in sub_V(H)\}$$

Suppose that $F = gate(G, H)$ and $\ell_V(H) = 0$. Then $sub^*_V(F) \subset sub^*_V(G)$. Ok, the notation is awful, but this should actually make quite a bit of sense!

Lemma 15. *If F is andy n-ary formula and $V \subseteq [n]$ and $\ell_v(F) \geq 1$, then $|sub^*_V(F)| \leq 4 \cdot 16^{\ell_V(F)-1}$.*

Proof. By induction on the leaf size of F. □

Theorem 16. ***(Nechiparuk's Bound).*** *For any $f : \{0,1\}^n \leftarrow \{0,1\}$ and any partition $[n] : V_1 \cup \cdots \cup V_t$ (these are dijoint decomp so the 'u' with a + inside)*

$$\underset{B_2}{\mathcal{L}}(f) \geq \frac{1}{4} \sum_{i=1}^{t} \log |sub_{V_i}(f)|.$$

Next we are going to get a lower bound for an explicit function in the full binary basis. The function that we are going to consider is the *Element Distinctness function.*

Definition 17. *For $k \in \mathbb{N}$, let $n = 2^k \cdot 2k$. The **Element Distinctness** function ED_n is*

$$ED_n : \{0,1\}^{2^k \times 2k} \to \{0,1\}$$

where

$$ED_n(X_1, ..., X_{2^k}) = \begin{cases} 1 & \text{if } X_1, ..., X_{2k} \text{ are distinct elements of } \{0,1\}^{2k} \\ 0 & \text{otherwise} \end{cases}$$

7 Upper Bound

Definition 18. *A **randomized circuit** for a function $f : \{0,1\}^n \to \{0,1\}$ is a circuit C with $n+m$ variables $x_1, ..., x_n$ and $y_1, ..., y_m$ (think of this as a random seed) such that for every $\mathbf{x} \in \{0,1\}^n$*

$$\Pr_{y \in \{0,1\}^m} [C(\mathbf{x}, \mathbf{y}) = 1] = \begin{cases} \geq \frac{2}{3} & \text{if } f(x) = 1 \\ \leq \frac{1}{3} & \text{if } f(x) = 0 \end{cases}$$

Theorem 19. ***(Adelman, 1978).*** *If f is computable by polynomial size randomized circuit, then its computable by poly-sized (deterministic) circuits.*

Proof. You want to improve the probability of success by doing repeated trials, taking the majority, then use the probabilistic method to show that there was a good choice of the randomization which we could have just wired into our circuit initially. □

Welp, that is all the bounds that we could get out of the general case. Time to consider some restricted settings.

8 Monotone and Others

Definition 20. *A **monotone circuit***

Theorem 21. *MAJ_n has poly-sized monotone circuits. (Also has monotone Formulas!)*

Proof. □