CS6840: Advanced Complexity Theory April 12^{th}, 2012

Lecture 52 : Monotone Circuits Lower Bounds Contd.

Lecturer: Jayalal Sarma M.N. *Scribe: Princy Lunawat*

THEME: Circuit Complexity
LECTURE PLAN:This lecture forms the continuation of the proof of a monotone circuit lower bound on $CLIQUE_{(k,n)}$. We start with a quick revision of the (m, l)-approximator gadget used to analyze the clique circuit in a given graph and the statement of the sunflower lemma used in the course of the proof (from the last lecture) followed by a detailed proof of the lower bound.

Theorem 1. *Any monotone circuit computing $CLIQUE_{(k,n)}$ requires size $n^{\sqrt{k}}$.*

Proof Idea

The main idea of the proof is to inductively build a circuit for $CLIQUE_{(k,n)}$ using a gadget referred to as an (m, l)-approximator instead of a gate at the cost of incurring some error. We estimate the error at each insertion of an approximator versus the net error incurred at the root of the circuit. We will observe that while the former causes a small error per approximator, in the latter, the error is too high, hence concluding that the no. of gates has to be too high. □

Proof. Let C be a monotone circuit computing $CLIQUE_{(k,n)}$ on the given input graph $G = V, E$.Let the input graph be specified as set of indicators $I_{\{u,v\}}$ which takes a value 1 iff $u, v \in E$. We define a clique indicator as follows: $I_X = 1$ if the vertices indexed by $X \subseteq V$ form a clique. Then the following expression trivially finds whether a clique of size k exists in a graph:

The size of this gadget is too huge($O(n^k)$ where k depends on n). Hence we look for alternatives to design a monotone circuit, where in comes the (m, l)- approximator defined as follows:

$$\bigvee_{i=1}^{r} I_{X_i} \quad |X_i| \leq l \quad r \leq m$$

1 Construction of the Clique circuit C'

Now, we inductively create a new circuit C' for the clique problem using (m, l)-approximators defined above. The construction is by induction.

Basis:

At the leaves we employ the clique indicator of each edge of the graph itself as inputs to our circuit, that is, I_{X_i} where $|X_i| = 1$. Hence no error is incurred at the leaves. We can see that the this is a trivial (m, l)- approximator.

Induction:

- OR Gate:
  Consider an OR gate with k input (m,l)approximators. We can can reduce than fan-in of this gate by taking the composition of OR's at the cost of a $\lg(k)$ increment in depth of the circuit. Let the input to the final OR- gate after the composition be two (m, l)- approximators. The expression at the OR- gate is given by,

$$\bigvee_{i=1}^{r} I_{X_i} \vee \bigvee_{i=1}^{s} I_{Y_i} \qquad |X_i| \leq l, r \leq m, |Y_i| \leq l, s \leq m$$
$$\Rightarrow \bigvee_{i=1}^{r+s} I_{D_i} \quad |D_i| \leq l, r + s \leq 2m \tag{1}$$

  Hence the result is a $(2m, l)$- approximator. Observe that no error is incurred in so far.We know remove some of the D_i's from the approximator input to achieve a (m, l)-approximator, of course incurring some error. The D_i's to be removed are determined by the **sunflower lemma** stated below:

  Lemma 2. *Let F be a family of subsets of $[n]$, each of size $\leq l$. If $|F| > (p-1)^l l!$, then there exists a sunflower with p petals in the family. That is, there exist sets $Z_1, Z_2, \ldots Z_p \in F$ and a core Z such that*

$$Z = \bigcap_{i=1}^{p} Z_i$$

  and

$$\forall 1 \leq i, j \leq p, \quad (Z_i - Z) \cap (Z_j - Z) = \phi$$

  We use the above lemma in our context as follows:

- From (1), we have a family F of $r+s \leq 2m$ subsets D_i, $|D_i| \leq l$. Suppose, $|F| > m$(If $|F| \leq m$, we already have an (m,l)-approximator). By setting the appropriate values of p and l in the lemma, we get $Z_1, Z_2, \ldots Z_p$ among $D_1, D_2 \ldots D_{r+s}$ with a core Z.
- We replace all OR terms $I_{Z_1}, I_{Z_2}, \ldots I_{Z_p}$ in (1) with the single core I_Z, of course incurring some error. Essentially, we pluck the sunflower and replace it with the core.
- We thereby reduce the number of terms to $r+s-p+1$. If the resulting family size is still $\geq m$ we, re-apply again and again, until the no of terms is $\leq m$ and the lemma cannot be applied anymore.

Hence, at the end of the above procedure, we have succeeded in converting an OR-gate into an (m,l)-approximator. Observe that here we had to deal with only the number of D_i's while the size of each D_i was still bounded by l.

- AND- gate
  As before, let the input to the final AND- gate after the composition be two (m,l)-approximators.The expression at the AND- gate is given by,

$$\bigvee_{i=1}^{r} I_{X_i} \wedge \bigvee_{i=1}^{s} I_{Y_i} \qquad |X_i| \leq l, r \leq m, |Y_i| \leq l, s \leq m$$
$$\Rightarrow \bigvee_{i=1}^{r}\bigvee_{j=1}^{s} \left(I_{X_i} \wedge I_{Y_j}\right) \tag{2}$$

  We convert the above into the desired (m,l)-approximator form as follows:

  - Replace $\left(I_{X_i} \wedge I_{Y_j}\right)$ in (2) by $I_{X_i \cup Y_j} = I_{D_k}$. While the former returns a 1 when the vertices of X and Y form a clique each, the latter returns 1 when vertices of X and Y together form a clique, hence, an error is incurred, but the expression has the form of an approximator now:

$$\bigvee_{i=1}^{rs} I_{D_i} \quad rs \leq m^2, |D_i| \leq 2l$$

    Hence we have a $(m^2, 2l)$-approximator with us.
  - To reduce $2l$ to l, we simply drop all the D_i's such that $|D_i| > l$.
  - To reduce $rs \leq m^2$ to m, we apply the sunflower lemma again and again, as before until the size reduces to m.

At the end of this inductive process for all gates in the circuit C computing $CLIQUE_{(k,n)}$, we have a new circuit C' consisting of only (m,l)-approximators computing $CLIQUE_{(k,n)}$ with some error.

2 Analysis of Error bounds in C'

We now analyse the error bounds for C' on special kinds of inputs called **Positive** and **Negative** inputs.

- *Positive Input:* A graph G is a positive input the circuit C deciding $CLIQUE_{(k,n)}$ if it has exactly one clique of size k and $n-k$ isolated vertices. Clearly, there are $\binom{n}{k}$ such graph instances, and they output 1 when given as an input to C.
- *Negative Input:* A graph G is a negative input the circuit C deciding $CLIQUE_{(k,n)}$ if it is a complete $(k-1)$-partite graph. That is, the set of vertices are partitioned into $(k-1)$ disjoint subsets. There is no edge between any to vertices is the same subset, where as between any two subsets, all possible edges are present (A generalization of the complete bipartite graph.). An example is shown below. Clearly, there are $(k-1)^n$ such graph instances and and they output 0 when given as an input to C, since the maximum clique size is only $k-1$.

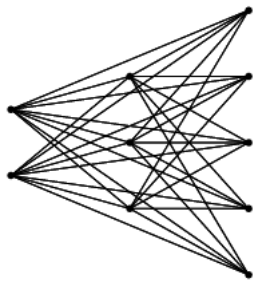

Figure 1: A complete k-partite graph with k = 3

We now prove an important lowerbound on the number of inputs to the approximator circuit C' with respect to the positive and negative inputs, for which the circuit makes an error.

Lemma 3. *Either C' is 0 on all positive inputs $\binom{n}{k}$ or C' is 1 on at least $\left(1 - \frac{\binom{l}{2}}{k-1}\right)(k-1)^n$ negative inputs.*

Proof. The output of C' is of the following form:

$$C' = \bigvee_{i=1}^{r} I_{X_i}$$

If C' is 1 on a positive input then there exists at least one $I_{X_i} = 1$. Hence, we calculate the probability of C' evaluating th 1 conditioned on the existence of such an I_{X_i}:
Choose a negative input randomly,

$$\begin{aligned}\Pr(C' = 1) &\leq \Pr(I_{X_i} = 1) \\ &= 1 - \Pr(I_{X_i} = 0)\end{aligned}$$

On a negative input, the probability of $I_{X_i} = 0$ is the same as the probability that at least of the vertices of X_i fall into the same partition in the negative input graph.

$$\Pr(I_{X_i} = 0) \leq \binom{l}{2}\frac{1}{k-1}$$

$$\Pr(C' = 1) \leq \left(1 - \frac{\binom{l}{2}}{k-1}\right)$$

Hence, the number of negative inputs where $C' = 1$ is given by the above probability multiplied by the number of negative input instances:

$$\left(1 - \frac{\binom{l}{2}}{k-1}\right)(k-1)^n \tag{3}$$

□

The next two lemmas prove an upper bound of the number of errors incurred by our construction of C' in terms of the circuit size C. As mentioned in the proof idea, to meet the lowerbound on (3), the size of C will be very high (We will show this formally.)

Lemma 4. *On positive inputs C fails to evaluate to 1 on almost $size(C).m^2\binom{n-(l+1)}{k-(l+1)}$ inputs.*

Proof. We will analyze the error incurred on both OR and AND gates. Recall, by our construction of C', each OR-gate is converted to an (m, n)-approximator by plucking a sunflower $Z_1, Z_2, \ldots Z_p$ from the inputs to the OR-gate and replacing it by the core Z. If the final output of the circuit is 1 due to a $I_{Z_i} = 1$, then it remains the same when replaced by I_Z, because, if the vertices of Z_i form a clique then the vertices of any set $Z \subseteq Z_i$ also forms a clique. Hence, no error is incurred .
For an AND- gate, we discard all inputs to the gate with length $|D_i| > l$ and reducing number of inputs from rs to m by plucking a sunflower. The latter doesn't cause any error as argued before.Discarding an input with $|D_i| > l$ will only cause an error in graphs which have a clique of size greater than l. The number of inputs to an AND gate is $rs \leq m^2$. Hence the total number of such **positive** input graph instances where an error is incurred at an AND-gate is given by:

$$m^2 \binom{n-(l+1)}{k-(l+1)}^1$$

Now, since the number of AND-gates in a circuit C is bounded by size(C), we upperbound the number of positive input graph instances where C' makes an error to:

$$\leq size(C).m^2 \binom{n-(l+1)}{k-(l+1)}$$

□

Lemma 5. *On negative inputs C fails to evaluate to 0 on almost $size(C).m^2 \left(\frac{\binom{l}{2}}{k-1}\right)^p (k-1)^n$ inputs.*

Proof. OR-gate: Contrary to the previous situation, applying sunflower lemma does lead to an error in the negative input case because if we replace an $I_{Z_i} = 0$ with I_Z, the latter might not evaluate to 0,(the absence of a clique on a set of Z_i vertices doesn't guarantee the same on a subset $Z \subseteq Z_i$).We analyze the probability of this error:

$$\begin{aligned}
&\Pr(\bigvee_{i=1}^{p} I_{Z_i} = 0 \wedge I_Z = 1) \\
&\leq \Pr(\bigvee_{i=1}^{p} I_{Z_i} = 0 | I_Z = 1) \\
&= \prod_{i=1}^{p} \Pr(I_{Z_i} = 0 | I_Z = 1) \\
&= \prod_{i=1}^{p} \Pr(I_{Z_i} = 0)
\end{aligned}$$

[1] l+1 vertices are already in the clique, choose the remaining k-(l+1) to form a k-clique

The above expression is the probability that none of the sets $Z_i, |Z_i| \leq l$ form a clique on the negative instance, that is, at least 2 of the vertices fall into the same partition in the $(k-1)$- partite graph. This probability is given by,

$$\left(\frac{\binom{l}{2}}{k-1}\right)^p \tag{4}$$

AND-gate: In this case, as per our construction, we discard all D_i's with $|D_i| > l$ and apply sunflower lemma to bring down $rs \leq m^2$ to m. The former does not incur any error on a negative input, since this can almost cause the output of the AND to go from 1 to 0, which still cannot cause an error on the desired output for the negative input graph, that is , 0. Now, the sunflower lemma might be applied almost $O(m^2)$ times to reduces the number of inputs from almost m^2 to m, each time incurring an error with probability given by (4). Hence the total number of such **negative** input graph instances where an error is incurred at an AND-gate is given by:

$$m^2 \left(\frac{\binom{l}{2}}{k-1}\right)^p (k-1)^n$$

/footnote$(k-1)^n$ is the total number of negative graph instances. Now, since the number of AND-gates in a circuit C is bounded by size(C), we upperbound the number of negative input graph instances where C' makes an error to:

$$size(C).m^2 \left(\frac{\binom{l}{2}}{k-1}\right)^p (k-1)^n$$

The task now is to correlate the lower and upper bounds on the number of erroneous inputs to C'. For positive inputs, lemma (3) and (4), and for negative inputs, lemma (3) and (5) show the following:

$$\binom{n}{k} \leq \text{Error+}C' \leq\leq size(C).m^2 \binom{n-(l+1)}{k-(l+1)}$$

$$\left(1 - \frac{\binom{l}{2}}{k-1}\right)(k-1)^n \leq \text{Error-}C' \leq size(C).m^2 \left(\frac{\binom{l}{2}}{k-1}\right)^p (k-1)^n$$

We set the values for the parameters in the sunflower lemma as $l = \sqrt{k}$ and $p = \sqrt{k}\log n$ and $m = (p-1)^l l!$ for application of the lemma. Substituting these values in the above expressions, we get the following bound on the size of the circuit C:

$$size(C) \leq n^{\Omega(\sqrt{k})}$$

Hence we have shown that any monotone circuit deciding the $\mathsf{NP} - complete$ problem $CLIQUE_{(k,n)}$ where k is dependent on n cannot be of size polynomial in n. □

□