CS277: Concrete Complexity **Fall 1996**

Lecture 15: October 22, 1996

Lecturer: Uri Zwick *Scribe: Philip Chong*

15.1 The Method of Approximations

15.1.1 Introduction

We will now see how the *method of approximations* (Razborov, 1985) can be used to obtain exponential lower bounds for the monotone circuit size of monotone functions.

Consider a monotone circuit for the function f, and consider a series of functions

$$f_1, f_2, f_3, \ldots, f_t = f$$

where each f_i is one of:

- x_j (an input of the circuit)
- $f_{j_1} \vee f_{j_2}$
- $f_{j_1} \wedge f_{j_2}$

where $j_1, j_2 < i$. Note that any circuit can be described this way using a topological sorting of the nodes. Now construct a series of approximations

$$c_1, c_2, c_3, \ldots, c_t$$

where each *approximator* c_i is of a certain restricted form. That is,

- if $f_i = x_j$ then $c_i = x_j$
- if $f_i = f_{j_1} \vee f_{j_2}$ then $c_i = c_{j_1} \sqcup c_{j_2}$
- if $f_i = f_{j_1} \wedge f_{j_2}$ then $c_i = c_{j_1} \sqcap c_{j_2}$

where $\sqcup$ and $\sqcap$ denote "approximate" *or* and *and* operations.

To obtain a lower bound for a function f, it is enough to satisfy the following two goals:

1. Every approximator is a very *bad* approximation of f

2. For every two approximators c_1 and c_2, $c_1 \sqcup c_2$ and $c_1 \sqcap c_2$ are relatively *good* approximations of $c_1 \vee c_2$ and $c_1 \wedge c_2$, respectively.

The following example will illustrate what is meant by "bad" and "good" approximators. The general reasoning behind this method is that, with proper selection of approximators, each step in the series of approximations only adds a small number of errors to the approximation. If it can be shown that the end result (c_t) of the series differs greatly from the original function, then there must be a large number of steps in the series, and hence the circuit must be large.

15.1.2 Example

The following example is due to Alon and Boppana (1986), and is based on the work of Razborov.

Consider the function $CLIQUE_3(n)$, defined as follows: $CLIQUE_3(n)$ takes as input $\binom{n}{2}$ variables $x_{ij}, 1 \leq i < j \leq n$, which correspond to the edges of the complete graph on n vertices. The output is a 1 if and only if there is a triangle of edges in the underlying graph whose edges correspond to inputs with value 1. That is, the input is a labeling of the edges with the values 0 and 1, and the output indicates the existence of a clique of size 3 whose edges are all labeled with 1s.

Note that the monotone circuit complexity of $CLIQUE_3(n)$ is $O(n^3)$.

An approximator for this example will be an *or* of up to m variables, where $m = 2(p-1)^2$ and p is a parameter to be chosen later. (We will see later that we require $p \approx 3 \log n$, so that $m \approx 18 \log^2 n$.) We will also allow approximators to be a constant (0 or 1).

We define the *approximate-or* function as:

$$\left(\bigvee_{\{i,j\} \in A} x_{ij} \right) \sqcup \left(\bigvee_{\{i,j\} \in B} x_{ij} \right) = \begin{cases} \left(\bigvee_{\{i,j\} \in A \cup B} x_{ij} \right) & \text{if } |A \cup B| \leq m \\ 1 & \text{otherwise} \end{cases}$$

where A and B are sets each of at most m edges in the underlying graph of $CLIQUE_3(n)$. Similarly, we define the *approximate-and* function as:

$$\left(\bigvee_{\{i,j\} \in A} x_{ij} \right) \sqcap \left(\bigvee_{\{i,j\} \in B} x_{ij} \right) = \left(\bigvee_{\{i,j\} \in A \cap B} x_{ij} \right)$$

It is easy to see that $c_{j_1} \sqcup c_{j_2} \geq c_{j_1} \vee c_{j_2}$, and $c_{j_1} \sqcap c_{j_2} \leq c_{j_1} \wedge c_{j_2}$.

We define a *positive test graph* as an assignment of 1 to the input variables corresponding to a single triangle in the underlying graph; there are thus $\binom{n}{3}$ distinct positive test graphs.

We define a *negative test graph* as an assignment of 1 to the input variables corresponding to a complete bipartite graph on the n vertices of the underlying graph; note that this is equivalent to coloring each of the vertices with one of two colors; there are 2^n such colorings, and we can say there are equally many negative test graphs (for simplicity, we choose to count colorings of the vertices instead; note that any negative test graph gives rise to two different colorings).

15.1.3 Obtaining a Lower Bound

We can make the following claims:

Claim 15.1 *Every approximator either outputs* 0 *on all positive test graphs or outputs* 1 *on at least half of the negative test graphs.*

Proof: If the approximator is identically 0, then the first case is true. Otherwise, we may write an approximator c as:

$$c = \bigvee_{\{i,j\} \in A} x_{ij} \geq x_{i_0 j_0}$$

where $\{i_0, j_0\} \in A$. Since $x_{i_0 j_0} = 1$ for exactly half of the negative test graphs, then c must be 1 for at least half of the negative test graphs. ■

Claim 15.2 *Let $f_1, f_2, \ldots f_t$ be the representation of a monotone circuit for CLIQUE$_3(n)$ and let $c_1, c_2, \ldots c_t$ be the corresponding sequence of approximations. If $f_t(G) = 1$ but $c_t(G) = 0$, then there exists i such that $f_i = f_{j_1} \wedge f_{j_2}$ and $(c_{j_1} \sqcap c_{j_2})(G) < (c_{j_1} \wedge c_{j_2})(G)$. Also, if $f_t(G) = 0$ but $c_t(G) = 1$, then there exists i such that $f_i = f_{j_1} \vee f_{j_2}$ and $(c_{j_1} \sqcup c_{j_2})(G) > (c_{j_1} \vee c_{j_2})(G)$.*

Proof: The proof is left as an unnumbered exercise. ■

Claim 15.3 *$(c_{j_1} \sqcap c_{j_2})(G) < (c_{j_1} \wedge c_{j_2})(G)$ for at most m^2 positive test graphs.*

Claim 15.4 *$(c_{j_1} \sqcup c_{j_2})(G) > (c_{j_1} \vee c_{j_2})(G)$ for at most 2^{n-p} negative test graphs.*

We will in a moment prove these two claims. Note that the following corollary results:

Corollary 15.5 *Every monotone circuit for CLIQUE$_3(n)$ contains either at least $\frac{\binom{n}{3}}{m^2}$* and *gates or at least 2^{p-1}* or *gates.*

A cubic lower bound is thus obtained if $p \approx 3 \log n$ and $m \approx 18 \log^2 n$; we have

$$C_{M_2}(CLIQUE_3(n)) \geq \min\{\frac{\binom{n}{3}}{m^2}, 2^{p-1}\}$$

$$C_{M_2}(CLIQUE_3(n)) = \Omega(\frac{n^3}{\log^4 n}) \quad \text{or} \quad C_{M_2}(CLIQUE_3(n)) = \Omega(n^3)$$

$$C_{M_2}(CLIQUE_3(n)) = \Omega(\frac{n^3}{\log^4 n})$$

Proof: (Claim 15.3)

We have

$$\left(\bigvee_{\{i,j\}\in A} x_{ij}\right) \sqcap \left(\bigvee_{\{i,j\}\in B} x_{ij}\right) < \left(\bigvee_{\{i,j\}\in A} x_{ij}\right) \wedge \left(\bigvee_{\{i,j\}\in B} x_{ij}\right)$$

Note that for inequality the right hand side must be 1. Since each positive test case corresponds to a triangle in the underlying graph, the right hand argument is 1 iff the triangle has an edge in A and an edge in B. There are at most $|A||B| \leq m^2$ such triangles. ■

Proof: (Claim 15.4)

We have

$$\left(\bigvee_{\{i,j\}\in A} x_{ij}\right) \sqcup \left(\bigvee_{\{i,j\}\in B} x_{ij}\right) > \left(\bigvee_{\{i,j\}\in A} x_{ij}\right) \vee \left(\bigvee_{\{i,j\}\in B} x_{ij}\right) \tag{15.1}$$

In this case, the left hand side must be identically equal to 1 for inequality (this can be seen from the definition of the approximate-or function. Thus $|A \cup B| > m$.

We will now show that a negative test graph chosen uniformly at random from the set of all possible negative test graphs has at most a probability of 2^{-p} of satisfying the inequality; if this is the case, then since there are 2^n possible negative test graphs (recall that we chose to count colorings of the graphs), there can be at most 2^{n-p} negative test graphs for which the inequality holds.

Note that the inequality (equation 15.1) is satisfied iff all edges in $A \cup B$ are labeled with the value 0; the equivalent graph coloring is such that for any edge $\{i,j\} \in A \cup B$, the vertices i and j have the same color.

Suppose one of the following were true:

- A *matching* of at least p edges can be found in $A \cup B$.
- A *star* of at least p edges can be found in $A \cup B$.

In the first case, any vertex is covered by at most one edge in the matching, and so the probability that an edge in the matching covers two like-colored vertices is independent of the coloring of any other vertex covered by the matching. Thus, since there are at least p edges in the matching, the inequality (equation 15.1) holds with probability at most 2^{-p}.

In the second case, the probability of an outer vertex of the star having the same color as the center is independent of the coloring of any other vertex in the star, and since there are p edges in the star, the inequality (equation 15.1) holds with probability at most 2^{-p}.

The following lemma shows that indeed at least one of:

- A *matching* of at least p edges can be found in $A \cup B$.
- A *star* of at least p edges can be found in $A \cup B$.

is true. Recall that $|A \cup B| > m = 2(p-1)^2$.

Lemma 15.6 *Any set of more than $m = 2(p-1)^2$ edges either contains a matching of size p or a star of size p.*

Proof: Choose a maximal matching in the given set of edges. If this matching has size at least p, we are done. If the size of the matching is at most $p-1$, then there are at most $2(p-1)$ vertices covered by the matching. Thus, by the pigeonhole principle, some vertex covered by the matching must touch more than $\frac{m}{2(p-1)} = p-1$ edges in the given set. This vertex is thus the center of a star of at least p edges. ■

Thus claim 15.4 must be true.

■

CS277: Concrete Complexity	**Fall 1996**

Lecture 16: October 24, 1996

Lecturer: Uri Zwick *Scribe: Philip Chong*

16.1 Cliques of Larger Sizes

We can generalize the result from the previous section for cliques of larger sizes ($CLIQUE_k(n)$). First we need a lemma.

16.1.1 Sunflowers

A *sunflower* with p petals is a collection of p sets $Z_1, Z_2, \ldots, Z_p$ such that all the intersections $Z_i \cap Z_j, i \neq j$ are all the same. The common intersection is called the *center* of the sunflower.

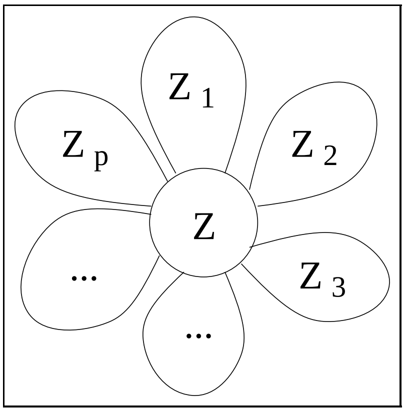

Figure 16.1: A Sunflower with Center Z

Lemma 16.1 *(Erdös, Rado, 1960) If $\mathcal{Z}$ is a collection of sets each of size at most ℓ, and if $|\mathcal{Z}| > (p-1)^\ell \ell!$, then $\mathcal{Z}$ contains at least one sunflower with p petals.*

Proof: We use induction on ℓ. If $\ell = 1$ and $\mathcal{Z}$ contains at least p distinct sets, each of cardinality ≤ 1, then we clearly have a sunflower with p petals.

Suppose this claim holds for $\ell - 1$. Let $\mathcal{M} \subseteq \mathcal{Z}$ be a maximum subcollection of pairwise disjoint sets. If $|\mathcal{M}| \geq p$ then $\mathcal{M}$ is a sunflower with at least p petals and an empty center.

Assume therefore that $|\mathcal{M}| \le p-1$. Every set $Z \in \mathcal{Z}$ intersects at least one set of $\mathcal{M}$. Note that there are at most $(p-1)\ell$ distinct elements contained in the sets in $\mathcal{M}$. By the pigeonhole principle, there must be an element i that is contained in more than $\frac{(p-1)^\ell \ell!}{(p-1)\ell} = (p-1)^{\ell-1}(\ell-1)!$ sets in $\mathcal{Z}$. Let

$$\mathcal{Z}' = \{Z \backslash \{i\} : Z \in \mathcal{Z}, i \in Z\}$$

Thus $|\mathcal{Z}'| > (p-1)^{\ell-1}(\ell-1)!$.

By the induction hypothesis, $\mathcal{Z}'$ contains a sunflower $Z'_1, \ldots, Z'_p$ with p petals. Taking $Z_j = Z'_j \cup \{i\}$, then $Z_1, \ldots, Z_p \in \mathcal{Z}$ is also a sunflower with p petals. ■

16.1.2 A Small Improvement

Using the definition of a sunflower, we can make a small improvement to the lower bound for the size of $CLIQUE_3(n)$. Recall that we defined

$$\left(\bigvee_{\{i,j\}\in A} x_{ij}\right) \sqcup \left(\bigvee_{\{i,j\}\in B} x_{ij}\right) = \begin{cases} \left(\bigvee_{\{i,j\}\in A\cup B} x_{ij}\right) & \text{if } |A\cup B| \le m \\ 1 & \text{otherwise} \end{cases}$$

and for any two approximators c_{j_1} and c_{j_2} we can have

$$(c_{j_1} \sqcup c_{j_2})(G) > (c_{j_1} \vee c_{j_2})(G)$$

for at most 2^{n-p} negative test graphs. We also defined

$$\left(\bigvee_{\{i,j\}\in A} x_{ij}\right) \sqcap \left(\bigvee_{\{i,j\}\in B} x_{ij}\right) = \left(\bigvee_{\{i,j\}\in A\cap B} x_{ij}\right)$$

and for any two approximators c_{j_1} and c_{j_2} we can have

$$(c_{j_1} \sqcap c_{j_2})(G) < (c_{j_1} \wedge c_{j_2})(G)$$

for at most m^2 positive test graphs. This led to the conclusion that every monotone circuit for $CLIQUE_3(n)$ must contain either at least $\frac{\binom{n}{3}}{m^2}$ *and* gates or at least 2^{p-1} *or* gates.

We can make a simple change in the definition for the approximate-or function to improve this lower bound. Take

$$\left(\bigvee_{\{i,j\}\in A} x_{ij}\right) \sqcup \left(\bigvee_{\{i,j\}\in B} x_{ij}\right) = \begin{cases} 1 & \text{if } A\cup B \text{ contains a sunflower of size } p \\ \left(\bigvee_{\{i,j\}\in A\cup B} x_{ij}\right) & \text{otherwise} \end{cases}$$

Here we consider the sets in the sunflower to be the graph edges in $A \cup B$; the center of the sunflower can thus either be a single vertex of the edges or the empty set.

Note that with this change, no approximator can contain a sunflower of size p; that is, no vertex may touch more than p edges in A, and likewise for B. Now consider the inequality

$$\left(\bigvee_{\{i,j\}\in A} x_{ij}\right) \sqcap \left(\bigvee_{\{i,j\}\in B} x_{ij}\right) < \left(\bigvee_{\{i,j\}\in A} x_{ij}\right) \wedge \left(\bigvee_{\{i,j\}\in B} x_{ij}\right)$$

We saw that if this inequality holds for a positive test graph then the triangle represented by the test graph has an edge in A and an edge in B (claim 15.3). Now, since no vertex can touch more than p edges in B, any edge e in A can touch at most $2p$ edges in B, and so there are at most $2p$ distinct triangles which contain both e and an edge in B. Thus, there can be at most $2mp$ triangles which satisfy the above inequality (compared to m^2 using the old definition of approximate-or).

As a result, the lower bound improves; every monotone circuit for $CLIQUE_3(n)$ contains at least $\Omega((\frac{n}{\log n})^3)$ gates.

16.1.3 Exercise 20

Prove the following:

> If f can be computed using a monotone circuit with k *and* gates, then f can also be computed using a monotone circuit with k *and* gates and $O(k(n+k))$ *or* gates.

Note that had we chosen $p \approx 6\log n$ and $m \approx 72\log^2 n$, we would have obtained bounds of either $\Omega(\frac{n^3}{\log^3 n})$ *and* gates or $\Omega(n^6)$ *or* gates. Suppose the latter were true; that is, we have $\Omega(n^6)$ *or* gates. If f can be computed with at most $O(\frac{n^3}{\log^3 n})$ *and* gates, then the above exercise implies that we have at most $O(\frac{n^6}{\log^6 n})$ *or* gates, a contradiction. Thus we must have $\Omega(\frac{n^3}{\log^3 n})$ *and* gates in particular.

16.1.4 Generalized Definitions for $CLIQUE_k(n)$

We now define *approximators*, the *approximate-or* and *approximate-and* functions, and *positive* and *negative test graphs* in a manner suitable for applying the method of approximations to find a lower bound for $CLIQUE_k(n)$.

Let $X \subseteq [n]$. The *clique indicator* of X is defined as:

$$\lceil X \rceil = \bigwedge_{i,j\in X, i<j} x_{ij}$$

That is, the clique indicator has value 1 iff the input graph has a clique on the set of vertices corresponding to X.

We will let an approximator be of the form

$$\bigvee_{i=1}^{m'} \lceil X_i \rceil$$

where $m' \leq m, |X_i| \leq \ell$ and m, ℓ are constants whose values we will later determine. That is, an approximator is the disjunction of at most m clique indicators, each of size at most ℓ.

Consider the "natural" definition of approximate-or:

$$\bigvee_{i=1}^{m_1} \lceil X_i \rceil \sqcup \bigvee_{j=1}^{m_2} \lceil Y_j \rceil \longrightarrow \bigvee_{i=1}^{m_1} \lceil X_i \rceil \vee \bigvee_{j=1}^{m_2} \lceil Y_j \rceil$$

Note that this this definition is slightly flawed; the approximator obtained in this manner (right hand side) may be the disjunction of *too many* ($> m$) clique indicators. To fix this, we can state that if the resulting approximator contains more than m clique indicators, then we can use a process of *plucking* to reduce the number of clique indicators.

Choose $m = (p-1)^\ell \ell!$; note that now we can simply select as constants p and ℓ and thus obtain m. Now if there are more than m clique indicators in a potential approximator, we can find a sunflower with p petals amongst them. These p clique indicators can be replaced with a single clique indicator:

$$\bigvee_{i=1}^{p} \lceil Z_i \rceil \longrightarrow \lceil Z \rceil$$

where Z is the center of the sunflower. This "plucks" the sunflower, and reduces the number of clique indicators in the approximator. This process can be repeated until there are at most m clique indicators in the (now valid) approximator.

Thus, for approximate-or, we can use the construction:

$$\bigvee_{i=1}^{m_1} \lceil X_i \rceil \sqcup \bigvee_{j=1}^{m_2} \lceil Y_j \rceil \longrightarrow \bigvee_{i=1}^{m_1} \lceil X_i \rceil \vee \bigvee_{j=1}^{m_2} \lceil Y_j \rceil \stackrel{\text{plucking}}{\longrightarrow} \bigvee_{i=1}^{m'} \lceil W_i \rceil, \quad (m' \leq m)$$

Note that if a clique exists on a petal of a sunflower, then a clique must exist on the center of the sunflower; that is,

$$\left(\bigvee_{i=1}^{p} \lceil Z_i \rceil = 1 \right) \Rightarrow \lceil Z \rceil = 1$$

Thus, for any two approximators c_1, c_2,

$$c_1 \sqcup c_2 \geq c_1 \vee c_2$$

In the case of approximate-and, we use the formulation

$$\bigvee_{i=1}^{m_1} \lceil X_i \rceil \sqcap \bigvee_{j=1}^{m_2} \lceil Y_i \rceil \longrightarrow \bigvee_{(i,j):|X_i \cup Y_j| \le \ell} \lceil X_i \cup Y_j \rceil \stackrel{\text{plucking}}{\longrightarrow} \bigvee_{i=1}^{m'} \lceil W_i \rceil, \quad (m' \le m)$$

That is, we form the approximate-and in two steps. First, we take the disjunction of $\lceil X_i \cup Y_j \rceil$ over all i, j such that $|X_i \cup Y_j| \le \ell$. Thus, all the clique indicators here are of size at most ℓ. For the second step, we use plucking to reduce the number of clique indicators just as we did for approximate-or. Note that for approximate-and, the resulting approximator can either be larger or smaller than the conjunction of the arguments.

The definitions for positive and negative test graphs here are similar to these definitions used in the case of $CLIQUE_3(n)$. Here, a positive test graph is a graph where the value 1 is only assigned to the edges of a clique of size k. A negative test graph is a $(k-1)$-partite graph; the vertices of the graph are colored using $(k-1)$ colors, and only those edges joining vertices of differing colors are assigned the value 1.

In the next section, we will see how to use these definitions to obtain an exponential lower bound for $CLIQUE_k(n)$.