18.405J/6.841J Advanced Complexity Theory	April 8, 1997
Lecture 15	
Lecturer: Daniel A. Spielman	*Scribe: Max Rozenoer*

Proof of Razborov's Theorem Continued

Recap of Last Lecture

We start with a recap of last lecture. Razborov's Theorem roughly states that no family of small monotone circuits[1] exists for an NP-complete problem.

The approach is the following: Let's say that the problem is the $k - clique$ problem for a graph on n nodes. The method is to approximate with a *clean* function a circuit that is said to decide instances of the *clique* of particular size. A *clean* function is is an $\vee$ of a set of *clique indicators* (at most m), where each clique indicator is less than size l [2](m, l and p (see below) to be assigned values later[3].)

The inputs to the circuit are certainly clean functions (trivially, they are $\vee$ of one clique indicator of size 2). We proceed to approximate the whole circuit with a clean function by performing the following operations:

Change every $\vee$ in the circuit for $\sqcup$, where $\sqcup$ is defined as

$$(f \sqcup g) = weed(f \vee g),$$

where *weed* is performed as follows: repeatedly *pluck* sunflowers with at least p petals until at most m clique indicators are left.

Change every $\wedge$ in the circuit for $\sqcap$, where $\sqcap$ is defined as

$$(f \sqcap g) = weed(trim(cover(f \wedge g)))^4,$$

Proof Continued

We start with definitions of a *minimum positive* set of instances of the $k - clique$ and a *maximum negative* set of instances, defined as follows:

POS = graphs on n nodes with exaclty one k-clique and no other edges.

NEG = complete k-1 partite graph (nodes divided into at most k-1 colors,
an edge exists between any two nodes of different colors.)

The idea behind these definitions is that we are only going to analyze performance of the clean functions on these sets, rather than all the instances of the $k - clique$, which will simplify the analysis. We will show that, for these sets, a clean function mishandles most of the inputs and, thus, cannot compute *clique*.

---

[1] Circuits with only $\vee$ and $\wedge$ gates.

[2] A *clique indicator* is thus an $\wedge$ of at most $\binom{l}{2}$ input edges (we can assume without loss of generality that input is a list of edges)

[3] Just to get a notion of the values, $m = (p-1)^l * l!$.

[4] Refer to the previous lecture for references on the *pluck*, *trim* and *cover* operations, as well as the Sunflower Theorem.

Lemma 1 *If f is a clean function, then one of the following holds:*
a) f rejects all graphs (i.e., it's an $\vee$ of an $\emptyset$ of clique indicators). Then, clearly, f mishandles all of the graphs in POS.
b) f accepts a graph from NEG with probability of $(1 - \frac{\binom{l}{2}}{k-1})$.

Proof

Since $a)$ is trivial, we proceed to the proof of $b)$.

The main idea of the proof is to pick a graph from NEG set (a *negative test graph*) and show that the probability of the clean function accepting it is $\geq (1 - \frac{\binom{l}{2}}{k-1})$.
If the conditions in $a)$ do not hold, then f is an $\vee$ of at least one clique indicator. We assume that it is, in fact, an $\vee$ of exactly one clique indicator, for appearance of more clique indicators would not cause the clean function to accept less instances. Let S be the set of nodes for which the indicator is testing for a clique.

We can choose a negative test graph at random[5] by choosing one of $k-1$ colors for each vertex at random, and then setting up the edges as in the definition of NEG. If each node $\in S$ gets a different color, the graph will be accepted. What is the probability of this?

For 2 particular nodes in S, the probability that they are the same color is clearly $\leq \frac{1}{k-1}$ (fix one node, look at the probability that the other is the same color). Then,

$$\text{Prob}[\exists \text{ 2 nodes in } S \text{ that get the same color}] \leq \frac{\binom{|S|}{2}}{k-1} \leq \frac{\binom{l}{2}}{k-1}.$$

■

Thus, if $\frac{\binom{l}{2}}{k-1}$ is relatively small, say $\frac{1}{2}$ (which can be achieved by choosing $l < \sqrt{k}$), then f will accept a random NEG graph with at least a probability of $1 - \frac{1}{2} = \frac{1}{2}$.

We shall now prove a series of lemmas about the effects of the pluck, trim, and cover operations on the probabilities of accepting POS and NEG graphs. The main idea is that the $\sqcup$ and $\sqcap$ approximations neither decrease the probability of accepting a positive test graph by too much, nor increase the probability of accepting a negative test graph by too much.

$POS(operation(f))$ is defined as the probability of accepting a positive test graph by f after *operation* has been applied to f. NEG is defined likewise.

Lemma 2
a) POS(pluck(f)) $\geq$ POS(f).
b) NEG(pluck(f)) $\leq$ NEG(f) + $(\frac{\binom{l}{2}}{k-1})^p$.[6]

Proof

a) We recall that the pluck operation removes the clique indicators corresponding to the petals of the sunflower, and adds the indicator corresponding to the core of the sunflower. Since for any of the removed indicator I, the core $\subseteq I$, $pluck(f)$ will not accept less graphs than f, as the added clique indicator looks for a clique on a subset of any set of vertices corresponding to a removed clique indicator.

b) Let's say that the sunflower being plucked has petals $Z_1 ..., Z_i, .. Z_p$, and core Z. We consider the probability that a NEG graph is accepted by $pluck(f)$ but not by f. The only way f does not

[5] The distribution produced is not uniform, but neither does it have to be, since all the results will be obtained relative to the same distribution of the NEG graphs.

[6] p is the number of petals of the sunflower being plucked.

accept a graph with a clique on the the vertices in the core of the sunflower is if in each petal of the sunflower there are at least 2 nodes of the same color. For $pluck(f)$ to accept the graph, the core must have all different colors. The probability of these two events is:

$$\begin{aligned}
&\text{Prob}[\forall i,\ (Z_i \text{ has 2 nodes of same color}) \wedge (Z \text{ has all distinct colors})] \leq && (1)\\
&\leq \text{Prob}[\forall i,\ Z_i \text{ has 2 nodes of same color} \mid Z \text{ has all distinct colors}] = && (2)\\
&= \textstyle\prod_{i=1}^{p} \text{Pr}[Z_i \text{ has 2 nodes of same color} \mid Z \text{ has all distinct colors}] \leq && (3)\\
&\leq \textstyle\prod_{i=1}^{p} \text{Pr}[Z_i \text{ has 2 nodes of the same color}] \leq (\frac{\binom{l}{2}}{k-1})^p. && (4)
\end{aligned}$$

(2) $\leq$ (1) follows from the definition of conditional probability. (3) = (2) holds by the virtue of the fact that all petals are independent, since if the 2 nodes have the same color (if there are such nodes at all), then they must lie outside the core (or all the petals will have them). The $\leq$ relation of (4) to (3) holds since the events (Z_i has 2 nodes of same color) and (Z has all distinct colors) are negatively correlated (i.e., the more likely one event, the less likely the other).

Thus, $(\frac{\binom{l}{2}}{k-1})^p$ is at most the probability that $pluck(f)$ accepts a NEG graph not accepted by f. ■

Lemma 3
a) POS(cover($f \wedge g$)) $\geq POS(f \wedge g)$.
b) NEG(cover($f \wedge g$)) $\leq NEG(f \wedge g)$.

Reminder: $f \wedge g = \bigvee_{i,j}(clique(S_i) \wedge clique(T_j))$, while $cover(f \wedge g) = \bigvee_{i,j}(clique(S_i \vee T_j))$ (S_i and T_j here are sets of vertices over which presence of cliques is considered).

Proof
a) A POS instance consists of one k-clique. If the k-clique induces a clique on both S_i and T_j, then it certainly induces a clique on $S_i \vee T_j$. Thus, the probability of accepting a POS graph does not decrease.
b) Since $clique(S_i \vee T_j)$ is more restricitve than $clique(S_i) \wedge clique(T_j)$, $cover(f \wedge g)$ does not accept more graphs than $(f \wedge g)$. Thus the probability of accepting a NEG graph does not go up. ■

Lemma 4
a) POS(trim(cover($f \wedge g$))) $\geq POS(f \wedge g) - \frac{m^2\binom{n-l-1}{k-l-1}}{\binom{n}{k}}$.
b) NEG(trim(cover($f \wedge g$))) $\leq NEG(f \wedge g)$.

Recall: trim excludes clique indicators of size $\geq l+1$, since the indicators produced by the cover operation might be too large.

Proof
a) We count the number of POS graphs that are accepted by a clique indicator of size $\geq l+1$. Since we only need an upper bound, let's suppose that this is a clique indicator of size $l+1$. Now, we know the $l+1$ nodes of the indicator. Pick the rest $k-(l+1)$ nodes needed for a clique in the graph to get at most $\binom{n-l-1}{k-l-1}$ graphs that the indicator might have accepted. This is at most the number by which the number of POS graphs accepted can go down after throwing out a clique of size $\geq l+1$.

For trim(cover($f \wedge g$))) we need to do at most m^2 such throwings out by the virtue of the fact that both f and g are clean, that is, of the form $\bigvee_{i=1}^{m} S_i$, and by definition of cover (see above). Thus, the number of POS graphs accepted can go down by at most $m^2\binom{n-l-1}{k-l-1}$. Now, since unlike NEG

graphs, POS graphs can be generated graphs uniformly, the counting argument result translates to a probabilistic one stating that

$$POS(trim(cover(f \wedge g))) \geq POS(f \wedge g) - \frac{m^2\binom{n-l-1}{k-l-1}}{\binom{n}{k}},$$

since $|POS| = \binom{n}{k}$.

b) Since the trim operation decreases the number of graphs accepted by throwing out clique indicators, so the probability of acccepting a NEG graph does not go up. ■

Lemma 5
a) $POS(f \sqcup g) \geq POS(f \vee g)$.
b) $NEG(f \sqcup g) \leq NEG(f \vee g) + m(\frac{\binom{l}{2}}{k-1})^p$.
c) $POS(f \sqcap g) \geq POS(f \wedge g) - \frac{m^2\binom{n-l-1}{k-l-1}}{\binom{n}{k}}$.
d) $NEG(f \sqcap g) \leq POS(f \wedge g) + m^2(\frac{\binom{l}{2}}{k-1})^p$.

Proof
a) Since $(f \sqcup g)$ is defined as weed$(f \vee g)$, this part of the lemma follows from Lemma 2 and the definition of weed operation.
b) This part follows from Lemma 2 and the fact that during the weed operation we pluck the sunflower at most m times.
c) $(f \sqcap g)$ is defined as weed(trim(cover$(f \wedge g)$)). Since the weed operation does not lower the probability of accepting POS instances, this part follows from Lemma 4.
d) Since trim(cover()) operation does not raise the probability of accepting a NEG graph, we only need to worry about the weed operation here. This part of the lemma follows from Lemma 2 and the fact that we do not need to pluck more than m^2 times to weed after trim(cover()) has been applied. ■

Grand Finale

We now come to the final point of the theorem.

Begin with a circuit C that computes k-clique. Replace it with a clean version of C (clean(C)), by appying the $\vee \to \sqcup$, $\wedge \to \sqcap$ transformations.

Lemma 6
Either a) $Size(C) \times \frac{m^2\binom{n-l-1}{k-l-1}}{\binom{n}{k}} \geq 1$
or b) $Size(C) \times m^2(\frac{\binom{l}{2}}{k-1})^p \geq 1 - \frac{\binom{l}{2}}{k-1}$.

Proof By Lemma 1, one of the following has to hold:
a) clean(C) accepts $\emptyset$. C accepted a random graph from POS with probability 1, while clean(C) accepts such a graph with probability 0. The only way this probability could have gone down is through $\wedge \to \sqcap$ transformations. Thus, by Lemma 5, the relation in a) has to hold.
b) clean(C) accepts a random NEG graph with probability $\geq 1 - \frac{\binom{l}{2}}{k-1}$. Clearly, C accepts such

a graph with prob 0. In this case, the probability could have gone up because of either type of transformation, but the relation in b) still has to hold by Lemma 5. ■

We now plug in the following values:

$$k = n^{\frac{1}{4}}, l = n^{\frac{1}{8}}, p = 10n^{\frac{1}{8}} lg(n), m = (p-1)^l l!.$$

When these values are substituted in, we get a contradiction in the realtions derived above, **unless** $Size(C) > 2^{n^{\frac{1}{8}}}$. Q. E. D.