Razborov's theorem, interpolation method, and lower bounds for Resolution and Cutting Planes.

Alisa Knizel

2009 April

Abstract

The goal of this article is to illustrate the interpolation method of proving the lower bounds for resolution and cutting plane proofs. The idea is to reduce this problem to the problem of proving the lower bounds for monotone circuits which was solved by Razborov.

1 Introduction

The problem of proving lower bounds on the length of propositional proofs is one of the most important problems in logic and complexity theory. It is so important because we know the connection: $NP = co\text{-}NP$ if and only if there is a polynomialy-bounded proof system for the classical tautologies. There is no general technique to proof lower bounds for all propositional proof systems so we are trying to extend current methods to stronger and stronger systems.

We can also formulate a congecture, a strengthering of the $P \neq NP$, which is very close to the problem we have discussed above, that NP-complete problems have no polynomial circuits. The progress in proving this conjecture has been very slow. At present the best lower bounds we have been able to proof for explicit families of functions are of the form $k * n$ for small constants k. So it was decided to try to prove this statement in a weeker circuit model: the monotone circuits. Many NP-complete problems cannot be computed by monotone circuits but some important ones

can be computed (for example, $CLIQUE_{n,k}$). The question is how small these monotone circuits can be?

First of all, in this article we will answer this question. After it we will proof lower bounds for resolutios and for cutting planes proofs with the help of this result.

2 Basic definitions

Definition 2.0.1. Boolean circuit is a directed acyclic graph. All the nodes (gates) are labelled by: inputs, AND, OR, NOT. It computes a function of its n input bit in the natural way.

Definition 2.0.2. Monotone circuits are ones without NOT gates.

Remark 1. *Monotone circuits can only compute monotone functions(* $x \leq y \Rightarrow f(x) \leq f(y)$*), and* $\forall$ *monotone function can be computed by monotone circuit.*

Definition 2.0.3. $CLIQUE_{n,k}$ is the Boolean function. $CLIQUE(G(V,E)) =$ 1 if G has a clique of size k.

Remark 2. $CLIQUE_{n,k}$ *is a monotone function because if the graph has a clique than the graph with an added edge will also have a clique.* $CLIQUE_{n,k}$ *is* NP*-complete.*

Now we are going to construct monotone circuit for $CLIQUE_{n,k}$:

- input gate $g[i,j]$ is set to true $\Leftrightarrow [i,j] \in E$
- $\forall S \subseteq V$ with $|S| = k$ test with AND gates whether S forms a clique
- repeat $\forall S \subseteq V$ with $|S| = k$ and take a big OR of the outcomes

Definition 2.0.4. Crude circuit is a circuit testing whether a family of subsets of V form a clique and returning true $\Leftrightarrow$ one of the sets does. The above circuit is denoted $CC(S_1, ..S_{\binom{n}{k}})$

3 Razborov's Theorem

Theorem 1. *There is a constant* **c** *such that for large enough* n *all monotone circuits for* $CLIQUE_{n,k}$ *with* $k = \sqrt[4]{n}$ *have size at least* $2^{c\sqrt[8]{n}}$

The plan of the proof:

We are going to construct a restricted kind of crude circuit from the given monotone circuit for $CLIQUE_{n,k}$. This process we will call approximation. This curcuit will be very crude but we are going to test it only on a very special kind of inputs. We will call them possitive and negative examples. The approximation is devided into steps. Each step corresponds to each gate of the original curcuit. The errors which each approximation step can introduce we will call false positives and false negatives. We will count how many false negatives and false positives each approximation step can introduce. We will show that each step introduces rather few errors but the resulting crude circuit has exponentially many errors. It means that the approximation takes exponentially many steps. Thus the original monotone circuit has exponentially many gates.

Now we have to proof a lemma which will be very useful for us.

Definition 3.0.5. A *sunflower* is a family of p sets $\{P_1, ..., P_p\}$, called *petals*, each of cardinality at most ℓ, such that all pairs of sets in the family have the same intersection (called *the core* of sunflower).

Lemma 1. *(The Erdös-Rado Lemma) Let* $\mathbf{Z}$ *be a family of more than* $M = (p-1)^\ell \ell!$ *nonempty sets, each of cardinality* ℓ *or less. Then* $\mathbf{Z}$ *must contain a sunflower.*

Proof. Induction on ℓ.

- $\ell = 1 \Leftrightarrow$ different singletons form a sunflower. $\mathbf{D}$ is a maximal subset of $\mathbf{Z}$ of disjoint sets.
- $|D| \geq p$ sets, then it constitutes a sunflower with empty core.
- Let's define $\mathbf{F} := \bigcup H_i, H_i \in \mathbf{D}$. We know: $|\mathbf{F}| \leq (p-1)\ell$ and that $\mathbf{D}$ intersects every set in $\mathbf{Z}$.
- Thus there is an element $d \in \mathbf{D}$ which intersects more than $\frac{M}{(p-1)\ell} = (p-1)^\ell(\ell-1)!$ sets.

- Let's define $\mathbf{G} := \{\mathbf{S} - d : \mathbf{S} \in \mathbf{Z} \text{ and } d \in \mathbf{Z}\}$
- $\mathbf{G}$ has more than $(p-1)^\ell(\ell-1)!$ sets $\Rightarrow$ by induction it contains a sunflower $P_1, ..., P_p$. Then $\{P_1 \cup \{d\}, ..., P_p \cup \{d\}\}$ is a sunflower in $\mathbf{Z}$.

□

Definition 3.0.6. Plucking a sunflower entails replacing the sets in the sunflower by its core.

$$Z_1, .., Z_p \longrightarrow Z$$

Remark 3. *During the whole proof the parameters p and ℓ will be fixed. With the help of plucking we can also fix the maximal number of the sets in the family. If there are more than M sets in a family, we can reduce their number by repeatedly finding a sunflower and plucking it.*

It is time to describe the approximation process. We will do it inductively because any monotone circuit can be considered as the OR or AND of two subcircuits(the induction is easy to stert because each input gate $g_{i,j}$ denoting whether $[i,j] \in E$ can be seen as a crude circuit CC(i,j)). Let's define the step of induction. Suppose there are two circuits CC($\mathbf{X}$), CC($\mathbf{Y}$), where $\mathbf{X}$ and $\mathbf{Y}$ are families of $\leq$ M sets of nodes, each set with $\leq \ell$ $(= \sqrt[8]{n})$ nodes. (Later we will define: $M = (p-1)^\ell \ell!$, $\ell = \sqrt[8]{n}$) and p is about $\sqrt[8]{n}$))

Approximation steps:

- A$[CC(\mathbf{X}) \vee CC(\mathbf{Y})]$ = CC(pluck($\mathbf{X} \cup \mathbf{Y}$))
- A$[CC(\mathbf{X}) \wedge CC(\mathbf{Y})]$ = CC(pluck $(\{U_i \cup V_j : U_i \in \mathbf{X}, V_i \in \mathbf{Y}$, and $|U_i \cup V_j| \leq \ell\})$)

A positive example is simply a graph with $\binom{k}{2}$ edges connecting k nodes in all possible ways. There are $\binom{n}{k}$ such graphs, and they all should elicit the "true". The negative examples are outcomes of following experiment: color the nodes with $k-1$ different colors. Then join by an edge any two nodes that are colored differently. Such a graph has no k-clique. There are $(k-1)^n$ negative examples overall.

Definition 3.0.7. (False positives)

Let N be a negative example. $CC_1(N) = false, CC_2(N) = false, CC = A[CC_1 \vee CC_2]$ and $CC(N) = true \Rightarrow$ the approximation step has introduced a **false positive**.

Let N be a negative example. $CC_1(N) = false, CC = (AND)A[CC_1 \wedge CC_2]$ and $CC(N) = true \Rightarrow$ the approximation step has introduced a **false positive**.

Definition 3.0.8. (False negatives)

Let E be a positive example. $CC_1(E) = true, CC_2(N) = true, CC = A[CC_1 \wedge CC_2]$ and $CC(E) = false \Rightarrow$ the approximation step has introduced a **false negative**. Let E be a positive example. $CC_1(E) = true, CC = A[CC_1 \vee CC_2]$ and $CC(E) = false \Rightarrow$ the approximation step has introduced a **false negative**.

These two lemmas state that each approximation step introduces rather few errors.

Lemma 2. *(about false positives) Each approximation step introduces* $\leq M^2 2^{-p}(k-1)^n$ *false positives.*

Proof. First for an OR.

A false positive introduced by plucking (the replacement of sunflower $\{Z_1, ..., Z_p\}$ by its core $\mathbf{Z}$) is a coloring such that there is a pair of identically colored nodes in each petal, but at least one node from each petal was plucked away. Let's count such colorings.

Let R(X) be the probability of the event that there are repeated colors in set X. So we have:

$$\mathbf{prob}[R(Z_1) \wedge ... \wedge R(Z_p) \wedge \neg R(Z)]$$

$$\leq \mathbf{prob}[R(Z_1) \wedge ... \wedge R(Z_p) | \neg R(Z)] =$$

$$= \prod_{i=1}^{p} \mathbf{prob}[R(Z_i) | \neg R(Z)] \leq \prod_{i=1}^{p} \mathbf{prob}[R(Z_i)]$$

The first inequality holds because the left-hand side is actually equal to the right-hand side devided by $\mathbf{prob}[\neg R(Z)] < 1$ (this is the definition of conditional probability).The second equality is true because the only common verticies the $Z_i's$ have are in Z, and, given that there are no repeated colors in Z, the probabilities of repeated colors in the $Z_i's$ are independent. The last inequality holds because the probability of repetitions in Z_i is descreased if we restrict ourselves to colorings with no repetitions in $Z \subseteq Z_i$.

Let's consider two nodes in Z_i, **prob**[they have the same color]$=\frac{1}{k-1}$. Then

$$\mathbf{prob}[R(Z_i)] \leq \frac{\binom{|Z_i|}{2}}{k-1} \leq \frac{\binom{\ell}{2}}{k-1} \leq \frac{1}{2}$$

Thus the probability that a randomly chosen coloring is a new false negative is at most 2^{-p} There are $(k-1)^n$ different coloring $\Rightarrow$ each plucking introduces $\leq 2^{-p}(k-1)^n$ false positives. The approximation step entails up to $\frac{2M}{p-1}$ pluckings, so the lemma holds for the OR approximation step.

Consider now an AND approximation step. It can be broken down in 3 phases:

- first we form $CC(\{U \cup V : U \in \mathbf{X}, V \in \mathbf{Y}\})$, this introduces no false positives, because any graph in which $U \cup V$ is a clique must have a clique in both U and V.
- we delete the sets with cardinality more than ℓ and it introduces no false positives.
- the third phase entails a sequence of less than M^2 pluckings, during each of which $\leq 2^{-p}(k-1)^n$ false positives are introduced.

□

Lemma 3. *(about false negatives) Each approximation step introduces* $\leq M^2\binom{n-\ell-1}{k-\ell-1}$ *false negatives.*

Proof. Plucking can introduce no false negatives, since replacing a set in a crude circuit by a subset can only increase the number of accepted graphs. As the approximation of an OR entails only pluckings, it introduces no false negatives. Consider now an AND approximation step. When we form $CC(\{U \cup V : U \in \mathbf{X}, V \in \mathbf{Y}\})$ no false negatives can be introduced. If a positive example is accepted by both $CC(\mathbf{X})$ and $CC(\mathbf{Y})$; it must be the case that its clique contains one set in $\mathbf{X}$ and one set in $\mathbf{Y}$, but then it contains the union of these sets(because of the structure of positive example), and thus it is accepted by the new circuit.

Next we have to delete all sets which are larger than ℓ. Each deletion of a set W which is larger than ℓ can introduce several false negatives, namely the cliques that contain $W \Rightarrow$ at most $\binom{n-\ell-1}{k-\ell-1}$ false negatives can be introduced by each deletion. There are at most M^2 sets to be deleted. Thus the lemma is proved. □

So we can make a conclution. Previous lemmas show us that each approximation step introduces "few" false positives and false negatives. We'll next show that the resulting crude circuit must have "a lot".

Lemma 4. *(number of errors) Every crude circuit is not identically false (and thus is wrong on all positive examples), or outputs true on at least half of the negative examples.*

Proof. If the crude circuit is not identically false, then it accepts at least those graphs that have a clique on some set X of nodes, with $|X| \leq \ell$. But from lemma 2 we know that at least half of the colorings assign different colors to the nodes of $X \Rightarrow$ half of the negative examples have a clique at X and are accepted.

□

Now we will make the last step of proving the Razborov's theorem(here we use Stirling's formula) :

- let's defune $p = \sqrt[8]{n} \log n$, $\ell = \sqrt[8]{n} \Rightarrow$

  $$M = (p-1)^\ell \ell! < n^{\frac{1}{3}\sqrt[8]{n}}$$

  for large enough n .

- if the final crude circuit is identically false $\Rightarrow$ all possitive examples were introduced as false negatives at some step.

- $\Rightarrow$ the original monotone circuit for $CLIQUE_{n,k}$ had (lemma 3) $\leq$

  $$\frac{\binom{n}{k}}{M^2\binom{n-\ell-1}{k-\ell-1}}$$

  $$\geq \frac{1}{M^2(\frac{n-\ell}{k})^\ell} \geq n^{c\sqrt[8]{n}},$$

  with $c = \frac{1}{12}$

- lemma 4 states that there are $\geq \frac{1}{2}(k-1)^n$ false positives and each approximation step introduces $\leq M^2 2^{-p}(k-1)^n$ (lemma 2) of them.

- $\Rightarrow$ the original monotone circuit had at least $2^{p-1}M^{-2} > n^{c\sqrt[8]{n}}$, with $c = \frac{1}{3}$.

Thus the proof of the theorem is completed.

4 Resolution

Definition 4.0.9. The propositional resolution proof system is the one which uses elementary disjunctions i. e., disjunctions of literals, as formulas, and the cut rule as the only one rule

$$\frac{\Gamma \vee p, \Delta \vee \neg p}{\Gamma \vee \Delta}$$

Where Γ, Δ are elementary disjunctions.

Definition 4.0.10. The ternary connective *sel* (selector) is defined by $sel(0, x, y) = x$ and $sel(1, x, y) = y$

Theorem 2. *(Effective interpolation for Resolution)*

Let P be a resolution proof of the empty clause from clauses $A_i(\bar{p}, \bar{q}), i \in I, B_j(\bar{p}, \bar{r}), j \in J$ where $\bar{p}, \bar{q}, \bar{r}$ are disjoint sets of propositional variables. Then there exists a circuit $C(\bar{p})$ such that for every $0-1$ assignment $\bar{a}$ for $\bar{p}$

$$C(\bar{a}) = 0 \Rightarrow A_i(\bar{p}, \bar{q}), i \in I$$

are unsatisfiable, and

$$C(\bar{a}) = 1 \Rightarrow B_j(\bar{p}, \bar{r}), j \in J$$

are unsatisfiable;

the circuit C is in basis $\{0, 1, \vee, \wedge\}$ and its underlying graph is the graph of the proof P.

Moreover, we can construct in polynomial time a resolution proof of the empty clause from clauses $A_i(\bar{p}, \bar{q}), i \in I$ if $C(\bar{a}) = 0$, respectively $B_j(\bar{p}, \bar{r}), j \in J$ if $C(\bar{a}) = 1$; the length of this proof is less than or equal to the length of P.

Proof. Let's consider the transformation of the proof for a given assignment $\bar{p} \to \bar{a}$

- 1. We replace each clause of P by a subclause so that each clause in the proof is either q-clause or r-clause. We start with initial clause, which are left unchanged and continue along the derivation P.

- Case 1.
$$\frac{\Gamma \vee p_k, \Delta \vee \neg p_k}{\Gamma \vee \Delta}$$
and we have replaced $\Gamma \vee p_k$ by Γ' and $\Delta \vee \neg p_k$ by Δ'. Then we replace $\Gamma \vee \Delta$ by Γ' if $p_k \to 0$ and by Δ' if $p_k \to 1$

- Case 2.
$$\frac{\Gamma \vee q_k, \Delta \vee \neg q_k}{\Gamma \vee \Delta}$$
and we have replaced $\Gamma \vee q_k$ by Γ' and $\Delta \vee \neg q_k$ by Δ'. If one of Γ', Δ' is an r-clause $\to$ replace $\Gamma \vee \Delta$ by this clause. If both Γ' and Δ' are q-clauses $\to$ resolve along q_k, or take one without q_k.

- Case 3.
$$\frac{\Gamma \vee r_k, \Delta \vee \neg r_k}{\Gamma \vee \Delta}$$
This is the dual case to case 2.

- 2. Delete the clauses which contain a $\bar{p}$ literal with value 1, and remove all $\bar{p}$ literals from the remaining clauses.

- We got a valid derivation of the final empty clause from the reduced initial clauses. If this final clause is a q-clause, the proof contains a subproof using only the reduced clauses $A_i, i \in I$; if an r-clause $\Rightarrow$ $B_j, j \in J$

Construction of C: The value computed at a gate corresponding to a clause Γ will determine if it is transformed into a q(r)-clause. We assign 0 to q-clauses and 1 to r-clauses. We put constant 0 gates on clauses $A_i, i \in I$ and constant 1 gates on clauses $B_j, j \in J$.

Now consider 3 cases as above.

- Case 1. If the gate on $\Gamma \vee p_k$ gets value x and the gate on $\Delta \vee \neg p_k$ gets value y, then the gate on $\Gamma \vee \Delta$ should get the value $z = sel(p_k, x, y)$. We place the *sel* gate on $\Gamma \vee \Delta$.

- Case 2. If the gate on $\Gamma \vee q_k$ gets value x and the gate on $\Delta \vee \neg q_k$ gets value y, then the gate on $\Gamma \vee \Delta$ should get the value $z = x \vee y$). We place the $\vee$ gate on $\Gamma \vee \Delta$.

- Case 3. This is dual to case 2.

□

Theorem 3. *Suppose moreover that either all variables $\bar{p}$ occur in $A_i(\bar{p},\bar{q}), i \in I$ only positively or all variables $\bar{p}$ occur in $\bar{p}$ occur in $B_j(\bar{p},\bar{r}), j \in J$ only negatively, then one can replace the selector connective sel by a monotone ternary connective.*

Proof. W. l. o. g. assume that all $\bar{p}$'s are positive in clauses $A_i, i \in I$. Hence in case 1, if Δ' is a q-clause, it cannot contain $\neg p_k$, hence we can take it for $\Gamma \vee \Delta$, even if $p_k \to 0$. Thus we can replace $sel(p_k, x, y)$ by $(p_k \vee x) \wedge y$ which is monotone and differs from selector exactly on one input $(p_k = 0, x = 1, y = 0)$ which corresponds to the above situation.

□

There is a construction of a formula which states that the graph is n-colorable and that it has a n-clique. Obviously this formula is false for every assignment. It can be written in the manner which was described in the statement of the theorem. So there is an interpolation circuit for it and the Razborov's theorem tells as that the size of this circuit is exponential. Thus we can make the conclution that the lengh of the resolution proof for it was also exponential.

5 Cutting planes:

Definition 5.0.11. We use propositional variables $\bar{p}$ with the interpretation $0 = \mathit{false}, 1 = \mathit{true}$.

- A proof line is an inequality

$$\sum_k c_k p_k \geq C$$

- Axiom: $0 \leq p_k \leq 1$

The rules:

- *Addition*: $\sum_k c_k p_k \geq C$ and $\sum_k d_k p_k \geq D \longrightarrow \sum_k (c_k + d_k) p_k \geq C + D$
- *Division*: $d > 0, d \in Z, d | c_k$ and $\sum_k c_k p_k \geq C \longrightarrow \sum_k \frac{c_k}{d} p_k \geq \lceil \frac{C}{d} \rceil$

- *Multiplication*: $d > 0, d \in Z$ and $\sum_k c_k p_k \geq C \longrightarrow \sum_k dc_k p_k \geq dC$

Theorem 4. *Let P be a cutting plane proof of the contradiction $0 \geq 1$ from inequalities*

$$\sum_k c_{i,k} p_k + \sum_l b_{i,l} q_l \geq A_i, i \in I$$

,

$$\sum_k c'_{j,k} p_k + \sum_m d_{j,m} q_m \geq B_j, j \in J$$

where $\bar{p}, \bar{q}, \bar{r}$ are disjoint sets of propositional variables. Then there exists a circuit $C(\bar{p})$ such that for every $0-1$ assignment $\bar{a}$ for $\bar{p}$

$$C(\bar{a}) = 0 \Rightarrow \sum_k c_{i,k} p_k + \sum_l b_{i,l} q_l \geq A_i, i \in I$$

are unsatisfiable, and

$$C(\bar{a}) = 1 \Rightarrow \sum_k c'_{j,k} p_k + \sum_m d_{j,m} q_m \geq B_j, j \in J$$

are unsatisfiable.

The size of the circuit is polynomial in the binary length of the numbers $A_i, i \in I, B_j, j \in J$ and the number of inequalities in P.

Moreover, we can construct in polynomial time a cutting plane proof of the contradiction $0 \geq 1$ from inequalities $\sum_k c_{i,k} p_k + \sum_l b_{i,l} q_l \geq A_i, i \in I$ if $C(\bar{a}) = 0$, respectively $\sum_k c'_{j,k} p_k + \sum_m d_{j,m} q_m \geq B_j, j \in J$ if $C(\bar{a}) = 1$; the length of this proof is less than or equal to the length of P.

Proof. Let P and assignment $\bar{p} \to \bar{a}$ be given. We will gradually replace each inequality in P

$$\sum_k e_k p_k + \sum_l f_l q_l + \sum_l f_l q_l \geq D$$

by a pair of inequalities

$$\sum_l f_l q_l \geq D_0$$

and

$$\sum_m g_m r_m \geq D_1$$

where D_0, D_1 are some integers. The have to ensure that the pair is at least as strong as the original one for the assignment $\bar{a}$, which means that

$$D_0 + D_1 \geq D - \sum_k e_k p_k$$

.

An initial inequality

$$\sum_k c_{i,k} p_k + \sum_l b_{i,l} q_l \geq A_i$$

will be replaced by the pair

$$\sum_l b_{i,l} q_l \geq A_i - \sum_k c_{i,k} p_k, \qquad 0 \geq 0.$$

Dually, an initial inequality

$$\sum_k c'_{j,k} p_k + \sum_m d_{j,m} r_m \geq B_j$$

will be replaced by the pair

$$\sum_m d_{j,m} r_m \geq B_j - \sum_k c'_{j,k} p_k, \qquad 0 \geq 0.$$

The rules are performed in parallel on the 2 inequalities in the pair. The pair corresponding to the last inequality $0 \geq 1$ is $0 \geq D_0$, $0 \geq D_1$ where $D_0 + D_1 \geq 1 \Rightarrow D_0 \geq 1 \vee D_1 \geq 1$ Thus we have a proof of contradiction either from $\sum_k c_{i,k} p_k + \sum_l b_{i,l} q_l \geq A_i, i \in I$ or from $\sum_k c'_{j,k} p_k + \sum_m d_{j,m} q_m \geq B_j, j \in J$. Also we have to use the fact that each proof P can be transformed in proof P' wich is at most polynomially longer and all the coefficients have polynomially bounded binary length (Clote and Buss). So we can consider that all D_i have polynomially bounded binary length $\Rightarrow$ the above procedure can be done in polynomial time in the binary length of $A_i, i \in I, B_j, j \in J$ and the number of inequalities. After it we use the transformation of polynomial time algorithms into sequences of polynomial size circuits.

□

Now to prove the lower bounds for cutting planes we have to use the generalisation of Razborov's thereme for so-called real monotone circuits. In fact the idea of the proof of it is the same. With the help of it we can make the same reasoning as for the resolution proof system.

References

[1] Christos M. Papadimitriou, Computational complexity (1994), 343-349

[2] Pudlák, P. , Lower bounds for resolution and cutting plane proofs and monotone computations, Journal of Symbolic Logic (3), (1997), 981–998