CS 880: Advanced Complexity Theory	4/12/06
Lecture 26: Resolution	
Instructor: Dieter van Melkebeek	Scribe: Vinay Choudhary and Michael Kowalczyk

In the last lecture, we introduced two rules of inference in resolution refutation.

- Resolution rule: $\frac{A \vee x \quad B \vee \overline{x}}{A \vee B}$, and
- Weakening rule: $\frac{A}{A \vee \ell}$

Recall that in the method of resolution refutation, our aim is to derive the empty clause beginning with a given set of clauses, by using these two rules. We also defined the width (w) of such a proof as the maximum number of literals in any clause occurring in the proof and the length (ℓ) of as the number of lines in the proof. We note here that if a clause C can be derived from ϕ using resolution and weakening in width w and length ℓ, then there exists $C' \subseteq C$ that can be derived using resolution only in width $\leq w$ and length $\leq \ell$. This can be proved inductively, and though we do not give a proof here, note that it implies that the weakening rule cannot help to derive the empty clause in any less width or length than we could without it.

We now define the same parameters for a CNF $\phi \notin SAT$:

- $\ell(\phi) = \min\{\text{length of a resolution refutation of } \phi\}$
- $w(\phi) = \min\{\text{width of a resolution refutation of } \phi\}$

1 Towards Lower Bounds on Proofs by Resolution

In this section, we prove a theorem that will lead to exponential lower bounds on resolution refutation proofs for the Pigeonhole Principle which will be discussed in the next lecture. The theorem we want to prove is that the minimum length of a resolution refutation of ϕ is at least exponential in its minimum width.

Theorem 1. *If $\phi \notin SAT$ is a k-CNF on n variables, then*

$$w(\phi) \leq k + a\sqrt{n \log \ell(\phi)}$$

where a is a universal constant.

Corollary 1. $\ell(\phi) = 2^{\Omega\left(\frac{(w(\phi)-k)^2}{n}\right)}$

We don't need to worry about whether we use weakening or not because even if we don't use weakening, the width and length can only decrease, and the length cannot decrease in this case because it will have less power. So we will just assume that we are using both the rules in our proof system and prove lower bounds for that. Let's now turn to the proof of the theorem.

Proof. The key fact we will use is that $\phi \notin SAT$ if and only if $\phi|_{x=0} \notin SAT$ and $\phi|_{x=1} \notin SAT$, where $\phi|_{x=1}$ and $\phi|_{x=0}$ are the formulas obtained by setting the literal x to true and false, respectively.

We assume that the formula ϕ itself is unsatisfiable. Let us consider the effect of setting a literal x to true on the formula and on its resolution refutation.

- On ϕ: All clauses which have x disappear and each instance of $\overline{x}$ can be removed.
- On refutation: Similarly, every line containing x becomes 1 and is therefore useless for the refutation. Every time $\overline{x}$ appears, we can remove it. Under these changes, we claim we now have a refutation for the restricted formula. It is easily seen that any resolution step on any other variable will still hold, so we only need to consider what happens when a resolution step is performed on x. That is, we start with clauses $A \vee x$ and $B \vee \overline{x}$ and we would normally derive $A \vee B$. But now the former evaluates to 1 and the later to B so deriving $A \vee B$ amounts to a number of weakenings. Strictly speaking, we have a potential increase in length, but a more careful argument can prove that each step will either disappear or can be achieved with resolution and the previous steps.

Now we try to build a refutation for ϕ out of refutations for $\phi|_{x=0}$ and $\phi|_{x=1}$.

1. Start from derivation of $\emptyset$ from $\phi|_{x=0}$.
2. Adding x to each line of the refutation, we get a derivation of x from ϕ. To see this, first note that by this modification, we only get clauses of ϕ, since every clause in the resolution is either going to already have x or will be a step in which we removed x (and it holds in both cases). And adding a literal to both A and B in one line still gives a valid inference. The width can go up by one in this case whereas length doesn't change.
3. Note that if $\phi|_{x=1}$ contains A, then ϕ contains A or contains $A \vee \overline{x}$. Remove x from each clause of the form $A \vee \overline{x}$ by resolving with x which has been derived earlier. The width of this derivation is at most k.
4. Now derive the empty clause using the refutation for $\phi|_{x=1}$.

Therefore,

$$\begin{aligned} width(\phi) &\leq \max(w(\phi|_{x=0}) + 1, k, w(\phi|_{x=1})) &(1)\\ length(\phi) &\leq length(\phi|_{x=0}) + |\phi| + length(\phi|_{x=1}) &(2)\end{aligned}$$

Pick a threshold t appropriately. Call a clause C wide if $|C| > t$. In successive steps, we set a literal each time such that at least $\frac{t}{2n}$ fraction of the wide clauses in the refutation disappear. This is possible because every clause contains at least t literals and there are $2n$ literals in all. So the most popular literal appears in at least $\frac{t}{2n}$ fraction of the clauses.

Claim 1. *If ϕ has at most n variables and less than $\left(1 - \frac{t}{2n}\right)^{-b}$ clauses of width more than t in some refutation, then $w(\phi) \leq \max(k, t) + b$.*

Proof. The proof is by induction on $b + n$. The case $n = 1$ is trivial. When $b = 0$, there are no clauses of width more than t so $w(\phi) \leq t \leq \max(k, t) + b$.

For the induction step, let y be the most popular literal among the clauses of width greater than t in the refutation proof. Choosing y to be true in the above transformation results in a $(1 - \frac{t}{2n})$ factor reduction in the number of wide clauses because y appears in at least a $\frac{t}{2n}$ fraction of them. Since this modified resolution refutation has less than $\left(1 - \frac{t}{2n}\right)^{-(b-1)}$ wide clauses, the inductive hypothesis implies that $w(\phi|_{x=0}) \leq \max(k, t) + b - 1$ where $x = \neg y$. Also, since the number of variables is reduced by 1 in $\phi|_{x=1}$, we have $w(\phi|_{x=1}) \leq \max(k,t)+b$. Finally, we put these together using inequality (1) to get

$$\begin{aligned} w(\phi) &\leq \max(w(\phi|_{x=0})+1, k, w(\phi|_{x=1})) \\ &\leq \max(\max(k,t)+b-1+1, k, \max(k,t)+b) \\ &= \max(k, \max(k,t)+b) \\ &= \max(k,t)+b \end{aligned}$$

□

Now we choose a value of b such that the hypothesis of this claim holds. Since the number of wide clauses can be at most $\ell = \ell(\phi)$ (regardless of t), it suffices to show that $\ell \leq (1-\frac{t}{2n})^{-b}$, or equivalently, $-\frac{\ln \ell}{b} \geq \ln(1-\frac{t}{2n})$. Using the fact that $x \leq e^x - 1$, we get

$$\ln(1-\frac{t}{2n}) \leq -\frac{t}{2n} \leq -\frac{\ln \ell}{b}$$

as long as $b \geq \frac{2n \ln \ell}{t}$.

Therefore,

$$\begin{aligned} w(\phi) &\leq \max(k,t) + \frac{2n \ln \ell(\phi)}{t} \\ &\leq k + 2\sqrt{2n \ln \ell(\phi)} \qquad\qquad (\text{if } t = \sqrt{2n \ln \ell(\phi)}) \end{aligned}$$

□

2 Expansion Implies Large Width

In this section, we relate the width of resolution to expansion in certain bipartite graphs. Let ϕ be a k-CNF formula. Construct a bipartite graph G_ϕ with the clauses on the left and the variables on right. There is an edge between two vertices if the variable appears, either positively or negated, in the clause. The left degree of this graph is at most k. We show that if the bipartite graph is expanding, the width of the formula has to be large.

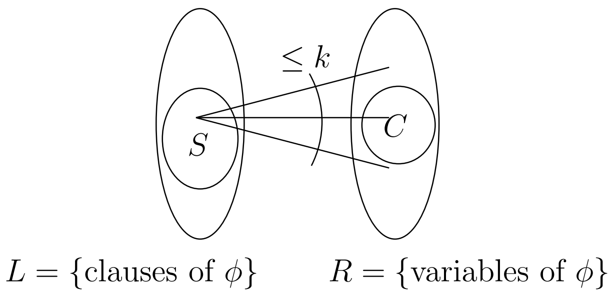

Figure 1: The underlying graph G_ϕ corresponding to formula ϕ

The intuition behind the theorem is that if G_ϕ is expanding, then no small clause C can be derived from a small set S of clauses of ϕ. Consider a small clause C which contains some variables (or their negations) from a set R' on the right. Consider its neighbors, $S = \Gamma(R')$. If C is small, then S is also small and the expansion property applies to it. That means the set of neighbors of the clauses in S will have many variables outside R' and therefore, we can possibly, set all the literals in C to false and still have enough freedom (i.e. variables remaining in $\Gamma(S)$) so as to make each clause in S true. Now we move to the proof.

Claim 2. *If G_ϕ is (K, c) expanding, then for any minimal set $S \subseteq L$ from which a clause C can be derived and such that $|S| \leq k$. Then $|C| \geq |\Gamma(S)| - |S| \geq (c-1)|S|$.*

We will prove this claim in the next lecture. Here we show how to use this to complete the proof of the theorem below.

Theorem 2. *ϕ is a k-CNF and $\phi \notin SAT$. If G_ϕ is (K, c)-expanding, then $w(\phi) \geq \frac{K}{2}(c-1)$.*

Proof. Let $\mu_\phi(C) = \min\{|S| : S \subseteq L\text{and } C \text{ can be derived from } S\}$.

We have, $\mu_\phi(\Phi) > k$. To see this, assume $|S| < k$ and apply claim 2 with $C = \Phi$. We will need at least one clause to get Φ. So, $|C| = 0$ and $|S| > 0$. This claim gives $0 \geq (c-1)|S|$ which is not possible. The only reason could be that $|S| > k$.

Now note that if C can be derived from A and B, then $\mu_\phi(C) \leq \mu_A\phi(A) + \mu_\phi(B)$.

So, in any refutation of ϕ, there is a line C such that $\frac{k}{2} < \mu_C \leq k$. The theorem now follows by the expansion property. □

2.1 Application to Random Bipartite Graphs

Now we will see an application to random k-CNFs where we see that the resolution length is $2^{\Omega(n)}$ with high probability. Recall that in Lecture 2 we showed that a random left d-regular bipartite graph is highly expanding with high probability. We state a slightly different version here, which follows from the same argument.

Theorem 3. *For any constants d, Δ and $b > 1$, there exists $\alpha(d, \Delta, b) > 0$ such that*

$$\Pr[\text{left } d\text{-regular } N \times M \text{ bipartite graph is } (\alpha \cdot N, d-b) \text{ expanding}] \geq .99$$

where $N = \Delta M$.

Now we let ϕ be a 3-CNF with N clauses, $M = n$ variables, and each literal chosen uniformly at random. The first thing we need is to see that $\phi \notin SAT$ with high probability. The probability that a fixed assignment to the M variables satisfies all of N random clauses is $(1 - (\frac{1}{2})^3)^N = (\frac{7}{8})^N$. Thus by a union bound, if $(\frac{7}{8})^N 2^M \ll 1$ then the probability that $\phi \notin SAT$ is high. The probability is at least .99 when $\frac{M}{N}$ is sufficiently larger than 2^k (i.e. $N = 40M$).

Once we can apply Theorem 3 we see that there is some constant α where G_ϕ will be $(\alpha N, 1.9)$-expanding with high probability. Therefore by a union bound we have that ϕ is not satisfiable and highly expanding with high probability. Now by Theorem 2, $w(\phi) > \frac{\alpha N}{2}(1.9 - 1) = \Omega(N)$ and by Corollary 1 the length is

$$\ell(\phi) = 2^{\Omega\left(\frac{(\Omega(N)-k)^2}{N}\right)} = 2^{\Omega(N)}.$$