Logic and Proof Hilary 2026

Lower Bounds for Resolution

James Worrell

No polynomial-time algorithm is known for the SAT problem. However, if a propositional formula is satisfiable, then there is a short and easily checkable certificate of this fact—a satisfying valuation. An important question is whether there likewise exist short certificates of unsatisfiability. One candidate for such a certificate would be a resolution refutation, which leads to the question of whether there exists a sub-exponential upper bound on the length of the shortest refutation of an unsatisfiable formula. In this lecture we give a negative answer to this question by exhibiting a family of unsatisfiable formulas whose refutations have length $2^{\Omega(n)}$, where n is the number of variables.

Lower Bounds via Bottleneck Counting

We will give a lower bound on the length of resolution proofs of the pigeonhole principle:

If n pigeons are placed in $n-1$ boxes, then some box contains at least two pigeons.

Fix $n \in \mathbb{N}$. For $i \in \{1,\dots,n\}$ and $j \in \{1,\dots,n-1\}$, let propositional variable $x_{i,j}$ denote that pigeon i is in box j. We consider the following formulas:

$$P_i := \bigvee_{j=1}^{n-1} x_{i,j} \quad \textit{"pigeon } i \textit{ is in some box"}$$

$$\mathrm{CRIT}_n := \bigwedge_{j=1}^{n-1} \bigvee_{i=1}^{n} x_{i,j} \quad \textit{"every box contains some pigeon"}$$

$$\wedge \bigwedge_{j=1}^{n-1} \bigwedge_{1\le i<i'\le n} (\neg x_{i,j} \vee \neg x_{i',j}) \quad \textit{"no box contains two different pigeons"}$$

$$\wedge \bigwedge_{i=1}^{n} \bigwedge_{1\le j<j'\le n-1} (\neg x_{i,j} \vee \neg x_{ij'}) \quad \textit{"no pigeon is in two different boxes"}$$

A valuation that satisfies CRIT_n is said to be *critical*. Such a valuation corresponds to a bijective assignment of $n-1$ out of n pigeons to $n-1$ boxes, with one pigeon left unassigned. We formalise the pigeonhole principle for n pigeons to be the statement that $\mathrm{PHP}_n := \mathrm{CRIT}_n \wedge \bigwedge_{i=1}^{n} P_i$ is unsatisfiable. The rest of the lecture is devoted to a proof of the following result.

Theorem 1. Every resolution refutation of PHP_n has length at least $2^{\frac{n}{21}}$.

We lay the groundwork for the proof by introducing some key concepts. We say that a sequence of monotone clauses (i.e., clauses with only positive literals) $C_1,\dots,C_m$ is a *pseudo refutation* of PHP_n if C_m is the empty clause and for all $1 \le i \le m$ either:

PR1 $\mathrm{CRIT}_n \wedge P_j \models C_i$ for some $1 \le j \le n$, or

PR2 $\mathrm{CRIT}_n \wedge C_j \wedge C_k \models C_i$ for some $j,k < i$.

We say that $W \subseteq \{1, \ldots, n\}$ is a *witness* of a clause C if $\text{CRIT}_n \wedge \bigwedge_{i \in W} P_i \models C$ (i.e., every critical assignment that houses all pigeons in W satisfies C). Every clause in a pseudo refutation has a witness: a clause that follows by rule PR1 has a singleton witness, while a clause that follows from PR2 has as witness the union of the witnesses of its two antecedents under PR2. We may thus define $\text{weight}(C)$ to be the minimum cardinality of any witness of C.

For a clause C, let C^* be the clause in which each negative literal $\neg x_{i,j}$ is replaced by $\bigvee_{i' \neq i} x_{i',j}$. Observe that $\text{CRIT}_n \models C \leftrightarrow C^*$—that is, for all critical assignments, pigeon i is not in box j iff some other pigeon is in box j. It follows that if $C_1, \ldots, C_m$ is a resolution refutation of PHP_n, then $C_1^*, \ldots, C_m^*$ is a pseudo refutation of PHP_n. It thus suffices to prove a lower bound on the length of pseudo refutations of PHP_n.

Proposition 2. Every pseudo refutation of PHP_n contains a clause with at least $\frac{2n^2}{9}$ variables.

Proof. Consider a pseudo refutation $\rho := C_1, \ldots, C_m$ of PHP_n. Since $C_m = \square$, we have $\text{weight}(C) = n$. Thus, there exists a first clause C in ρ with $\text{weight}(C) \geq \frac{n}{3}$. Clause C must be derived by rule PR2, hence $\frac{n}{3} \leq \text{weight}(C) \leq \frac{2n}{3}$, since its two antecedents have weight at most $\frac{n}{3}$.

We now argue that the clause C, identified above, contains at least $\frac{2n^2}{9}$ variables. Let W be a minimal witness for C, with $\frac{n}{3} \leq |W| \leq \frac{2n}{3}$. For each $i_1 \in W$ we exhibit $n - |W|$ different variables in C of the form $x_{i_1,j}$. We conclude that C contains at least $|W|(n - |W|) \geq \frac{2n^2}{9}$ variables.

Fix $i_1 \in W$. By the minimality of W as a witness, there exists a critical assignment v that leaves out pigeon i_1 and does not satisfy C. Now let $i_2 \notin W$ and suppose that v assigns i_2 to box j_2. Define an assignment v' by $\mathsf{v}'[\![x_{i_1,j_2}]\!] = 1$, $\mathsf{v}'[\![x_{i_2,j_2}]\!] = 0$, and otherwise v' agrees with v. (That is, v' assigns pigeon i_1 to box j_2 and makes i_2 the unassigned pigeon.) Then v' satisfies $\text{CRIT}_n \wedge \bigwedge_{i \in W} P_i$ and hence v' satisfies C. But the fact that v' satisfies C while v does not satisfy C entails that x_{i_1,j_2} is mentioned in C. This completes the proof. □

Proof of Theorem 1. Let $\rho := C_1, \ldots, C_m$ be a pseudo refutation of PHP_n. Say that a clause is *long* if it contains at least $\frac{n^2}{8}$ variables. Suppose that ρ has ℓ long clauses. By double counting (i.e., using the fact that the sum of the number of variables in each long clause equals the sum of the number of long clauses that each variable belongs to) we see that some variable is mentioned in at least $\frac{\ell}{8}$ long clauses. By renaming variables if necessary, we can assume that the aforementioned variable is $x_{n,n-1}$. Now we transform ρ by *"assigning pigeon n to box $n-1$"*—formally we delete any clause containing $x_{n,n-1}$ and then delete from the remaining clauses every variable $x_{i,j}$ with either $i = n$ or $j = n - 1$. Then the resulting sequence $C_1', \ldots, C_{m'}'$ is a pseudo refutation of PHP_{n-1} with at most $\frac{7\ell}{8}$ long clauses (see Exercise 3 as the end of this proof.) Repeating this process $\frac{n}{4}$ times, we arrive at pseudo refutation of $\text{PHP}_{\frac{3n}{4}}$ with at most $\left(\frac{7}{8}\right)^{n/4} \ell$ long clauses. But by Proposition 2, every pseudo refutation of $\text{PHP}_{\frac{3n}{4}}$ contains a clause with $\frac{2}{9}(\frac{3n}{4})^2 = \frac{n^2}{8}$ variables, i.e., a long clause. We deduce that $\left(\frac{7}{8}\right)^{\frac{n}{4}} \ell \geq 1$ and hence $\ell \geq \left(\frac{8}{7}\right)^{\frac{n}{4}} \geq 2^{\frac{n}{21}}$. But this means that there are more than $2^{\frac{n}{21}}$ clauses in ρ. □

Exercise 3. Explain why $\text{CRIT}_n[\mathbf{true}/x_{n,n-1}] \equiv \text{CRIT}_{n-1}$ and why, for arbitrary formulas F and G, if $F \models G$, then $F[\mathbf{true}/x_{n,n-1}] \models G[\mathbf{true}/x_{n,n-1}]$. Deduce that if $C_1, \ldots, C_m$ is a pseudo refutation of PHP_n, then $C_1', \ldots, C_m'$ is a pseudo refutation of PHP_{n-1}, where $C_i' := C[\mathbf{true}/x_{n,n-1}]$ for $i \in \{1, \ldots, m\}$.

Lower Bounds via Interpolation

In this section we describe another technique to derive exponential lower bounds on the length of resolution refutations, this time using the construction of interpolants via monotone circuits and

known lower bounds on the size of monotone circuits. We will rely on the following result, whose proof lies beyond the scope of this course.[1]

Theorem 4. There is a constant $\varepsilon > 0$ such that for all $n \in \mathbb{N}$ and $m = \lfloor n^{1/4} \rfloor$, there is no monotone straight-line program of size at most $2^{\varepsilon n^{1/8}}$ that computes a function $f : \{0,1\}^{\binom{n}{2}} \to \{0,1\}$ such that for the adjacency matrix A of an n-vertex graph G, if $f(A) = 0$, then G does not have clique of size m and if $f(A) = 1$, then G does not have a proper $(m-1)$-colouring.

For all $n \in \mathbb{N}$ and $m = \lfloor n^{1/4} \rfloor$, for appropriate sets of propositional variables $\boldsymbol{p}, \boldsymbol{q}, \boldsymbol{r}$, let the formula $\varphi_{n,1}(\boldsymbol{p}, \boldsymbol{r})$ express that $\boldsymbol{p}$ represents a clique of size m in the graph represented by $\boldsymbol{r}$. Furthermore let the formula $\varphi_{n,2}(\boldsymbol{q}, \boldsymbol{r})$ express that $\boldsymbol{q}$ represents an $(m-1)$-colouring of the graph represented by $\boldsymbol{r}$. The existence of such formulas, of size bounded polynomially in n, is shown in Exercise Sheet 1. There we also show that the variables $\boldsymbol{r}$ can be assumed to occur positively in $\varphi_{n,1}$. Clearly $\varphi_{n,1} \wedge \varphi_{n,2}$ is unsatisfiable: an m-clique admits no proper $(m-1)$-colouring.

Theorem 5. There is a constant $\varepsilon > 0$ such that for all $n \in \mathbb{N}$ every resolution refutation of $\varphi_{n,1} \wedge \varphi_{n,2}$ has length at least $2^{\varepsilon n^{1/8}}$.

Proof. Recall from the previous lecture that from a resolution refutation of $\varphi_{n,1} \wedge \varphi_{n,2}$ we can construct a monotone straight-line program that computes an interpolant of $\varphi_{n,1}$ and $\varphi_{n,2}$. By Theorem 4, the above straight-line program must have size at least $2^{\varepsilon n^{1/8}}$. But then the refutation of $\varphi_{n,1} \wedge \varphi_{n,1}$ must also have length at least $2^{\varepsilon n^{1/8}}$. □

[1] See Chapter 14.3 of *Computational Complexity* by Arora and Barak for details.