Resolution and the Weak Pigeonhole Principle

Sam Buss*[1] and Toniann Pitassi**[2]

[1] Departments of Mathematics and Computer Science
University of California, San Diego, La Jolla, CA 92093-0112.
[2] Department of Computer Science, University of Arizona, Tucson, AZ 85721.

Abstract. We give new upper bounds for resolution proofs of the weak pigeonhole principle. We also give lower bounds for tree-like resolution proofs. We present a normal form for resolution proofs of pigeonhole principles based on a new monotone resolution rule.

1 Introduction

Tautologies expressing versions of the pigeonhole principle have been important test cases for obtaining bounds on the lengths of propositional proofs and for comparing the proof theoretic strengths of various propositional proof systems. The seminal paper of Cook and Reckhow [5] showed that pigeonhole principles have polynomial length extended Frege systems; later, [3] showed that they also have polynomial length Frege proofs. On the other hand, the first superpolynomial lower bound on the length of resolution refutations was Haken's proof [6], that resolution proofs of the propositional pigeonhole principle require exponential length. A significant strengthening of Haken's lower bound was obtained by Ajtai [1] who proved that constant-depth Frege proofs of the pigeonhole principle require superpolynomial size; this was later strengthened by [2, 10, 7] to show that constant-depth Frege proofs of the pigeonhole principle require exponential size.

In some cases, finer distinctions can be made using generalized forms of the pigeonhole principle. One such principle is the "m into n" generalization which states that there is no one-to-one mapping of m objects into n objects, where $m > n$. Tautologies, defined below, expressing this principle are denoted PHP^m_n. It is known that the tautologies PHP^{2n}_n have quasipolynomial size constant Frege proofs [8, 9]. In addition, Haken's lower bound for resolution proofs of the pigeonhole principle was generalized by [4] who proved superpolynomial bounds on resolution proofs of PHP^m_n for $m = o(n^2/\log n)$.

These prior upper and lower bounds on the the size resolution proofs leave open the question of the size of resolution proofs of PHP^m_n when $n^2/\log n \leq m \leq 2^n$. It has been a folklore conjecture that the shortest resolution proofs of PHP^m_n have the same length as resolution proofs of PHP^{n+1}_n; in other words,

* Supported in part by NSF grant DMS-9503247 and US-Czech Science and Technology grant 93-025.
** Research supported by NSF grant CCR-9457782.

that when $m > n + 1$, optimal length resolution proofs can be obtained by ignoring all but one of the domain elements. We show in this paper, however, that this conjecture is false for $m = 2^{\sqrt{n \log n}}$.

The results of this paper are as follows. In section 3 we present a normal form for resolution proofs of pigeonhole principle tautologies. Normal form resolution proofs contain only positive occurrences of variables; the usual resolution rule is replaced by a new monotone resolution rule. The sizes of monotone resolution proofs are polynomially related to the sizes of resolution proofs. As a corollary, we prove that resolution proofs of the onto version of the pigeonhole principle are not significantly shorter than resolution proofs of the non-onto pigeonhole principle. In section 4, we give a polynomial upper bound on the size of resolution proofs of PHP_n^m for $m = 2^{\sqrt{n \log n}}$. This improves on the upper bound $n^2 2^n$ for proofs of PHP_n^{n+1}; which shows that having additional domain elements can make the pigeonhole principle easier to prove. In section 5, we prove an exponential lower bound on the size of tree-like resolution proofs of PHP_n^m.

2 Definitions

This paper deals exclusively with propositional logic. A *literal* is either a propositional variable or a negated propositional variable. A clause is defined to be a set of literals and is identified with the disjunction of its member literals. We assume that a clause never contains both a variable and the negation of that variable. We use capital letters, usually with subscripts, e.g., $P_{i,j}$, to denote variables; lowercase letters such as x denote literals; and clauses are denoted by letters $A, B, C, \ldots$.

A *resolution inference* infers $A \vee B$ from two clauses $A \vee x$ and $B \vee \overline{x}$. A conjunctive normal form (CNF) formula ϕ is identified with the set of clauses which appear as conjuncts of ϕ. A resolution refutation of ϕ consists of a sequence $C_1, \ldots, C_s$ of clauses, where each C_i is either a conjunct of ϕ or is inferred from earlier clauses in the refutation by a resolution inference. The *size* of the refutation is equal to the number, s, of clauses in the refutation.

A refutation proof of a disjunctive normal form (DNF) formula is defined to be a resolution refutation of the negation of the formula. It is well-known that resolution is refutationally sound and complete, so a DNF formula has a resolution proof if and only if it is a tautology.

Resolution refutations are usually viewed as sequences or directed acyclic graphs. However, they can also be restricted to be *tree-like* with each clause in the refutation being used as a hypothesis of an inference at most once. Note that the same clause may appear multiple times in the tree-like proof; the size of a tree-like refutation equals the number of occurrences of clauses in the refutation.

Definition 1. *Let $m > n$. The tautology PHP_n^m expresses the pigeonhole principle that there is no one-to-one mapping from a domain of m objects (called "pigeons") into a range of n objects (called "holes"). This is easily defined* *by a DNF formula, but since it is more relevant for resolution, we describe instead the set of clauses which are the conjuncts of the CNF formula $\neg PHP_n^m$. The propositional variables are $P_{i,j}$, $i \le m$, $j \le n$, with $P_{i,j}$ having the intuitive meaning that pigeon i is mapped to hole j. The clauses of $\neg PHP_n^m$ are:*

(1) $P_{i,1} \vee P_{i,2} \vee \cdots \vee P_{i,n}$, *for each* $i \le m$; *and*
(2) $\neg P_{i,k} \vee \neg P_{j,k}$, *for each* $i, j \le m$, $k \le n$, $i \ne j$.

Note that the number of clauses in $\neg PHP_n^m$ is $m + \binom{m}{2} n < m^2 n < m^3$.

As mentioned in the introduction, Haken proved that resolution proofs of PHP_n^{n+1} require exponential size. The first author and Turán [4] showed that any resolution refutation of PHP_n^m requires size $\frac{1}{2}\left(\frac{3}{2}\right)^{\frac{1}{50}\frac{n^2}{m}}$. However, when $m \ge n^2/\log n$, this lower bound is only polynomial, and in fact there are no nontrivial lower bounds known in this case. To the best of the authors knowledge, prior to the present paper, the best upper bounds known for the sizes of resolution proofs of PHP_n^m was the bound $n^3 2^n$ of Lemma 1 below.

3 A Normal Form Theorem

In this section we define a variation of the resolution proof system, called the monotone resolution system, which is tailored for proofs of pigeonhole principles. We prove that this system is complete for proofs of pigeonhole principle tautologies, and that the sizes of monotone resolution proofs and the sizes of resolution proofs are polynomially related. The motivation for introducing the monotone resolution proof system is the hope that it will provide a better framework for obtaining lower bounds on the sizes of resolution refutations of pigeonhole principles.

We define a monotone resolution proof for PHP_n^m as follows. A *monotone* clause is a clause which contains only positive variables. We let the mn variables $P_{i,j}$ correspond to entries in an an n-by-m array, with rows labeled by the n holes, and columns labeled by the m pigeons. Thus the variable $P_{i,j}$ corresponds to the entry in the j-th row and the i-th column. A monotone clause is visualized as an n-by-m array, with $+$'s in each entry corresponding to the occurrences of variables in the clause and with array entries corresponding to variables not occurring in the clause left blank.

For $R \subseteq \{1, \ldots, m\}$ we let $P_{R,j}$ be the disjunction of the variables $P_{i,j}$ for all $i \in R$. Let $C_1 = A \vee P_{R,j} \vee P_{S,j}$ and $C_2 = B \vee P_{R,j} \vee P_{T,j}$, where R, S and T are disjoint and where A and B are both disjunctions of positive variables not in row j. Then the *monotone resolution inference rule* allows us to derive $C_3 = A \vee B \vee P_{R,j}$ from C_1 and C_2. In other words, we can infer the clause C_3 from C_1 and C_2 by the monotone resolution rule with respect to hole j, provided C_3 consists of the union of all variables in $C_1 \cup C_2$, minus all variables $P_{i,j}$, which occur in exactly one of C_1 and C_2. Implicit in the monotone resolution rule is one-to-oneness: if pigeon i is mapped to hole j, then no other pigeon i' can be mapped to hole j.

A *monotone resolution proof* is a sequence of monotone clauses, where the final clause is the empty clause; and where every clause is either an *initial clause* of the form $\vee_{j=1}^{n} P_{i,j}$, or follows from two previous clauses by the monotone resolution rule.

Strictly speaking, monotone resolution is not a proof system, since it is not complete for arbitrary sets of clauses; however, it follows from the next theorem that monotone resolution is sufficient to prove pigeonhole principle tautologies. Only such tautologies are considered in this paper.

Two proof systems are said to be *polynomially equivalent* for a class Φ of formulas if and only if there is a polynomial $q(x)$ such that if $\phi \in \Phi$ has a proof of size s in one of the systems, then it has a proof of size $\leq q(s)$ in the other system.

Theorem 1. *The resolution proof system and the monotone resolution proof system are polynomially equivalent for the pigeonhole tautologies PHP_n^m.*

Proof. Let us first show that if we have a monotone refutation, then we also have a resolution refutation of the usual kind.. For this it suffices to simulate a monotone resolution inference by only polynomially many ordinary resolution inferences. Suppose that C_3 is obtained from C_1 and C_2 by the monotone resolution rule, where $C_1 = A \vee P_{R,j} \vee P_{S,j}$ and $C_2 = B \vee P_{R,j} \vee P_{T,j}$ and $C_3 = A \vee B \vee P_{R,j}$, and where R, S and T are disjoint and A and B are sets of variables not involving hole j. We shall show how to obtain C_3 from C_1, C_2 and the initial clauses with only polynomially many resolution steps. First, generate the clauses $C_1^t = A \vee P_{R,j} \vee \neg P_{t,j}$, for all $t \in T$. Each clause C_1^t is obtained by $|S|$ many resolution inferences from C_1 and the initial clauses $(\neg P_{t,j} \vee \neg P_{s,j})$ for all $s \in S$. Then from the clauses C_1^t, where $t \in T$, and from C_2, generate $C_3 = A \vee B \vee P_{R,j}$ in $|T|$ additional inferences.

Since $|S|, |T| \leq m$, the above construction shows that a monotone resolution inference can be simulated with $\leq m^2$ usual resolution inferences.

In the other direction we want to show that if P is a resolution refutation of $\neg PHP_n^m$, then there exists a monotone resolution refutation P' of $\neg PHP_n^m$ of size polynomial in the size of P. As a first step, we will transform every clause in P into a totally monotone clause as follows: if $C = A \vee B$ is a clause in P, where A is the disjunction of positive variables, and B is the disjunction of negative variables, then $C^m = A \vee B^m$, where B^m is obtained by replacing every negative literal $\neg P_{i,k}$ in B by the (disjunction of the) set of literals $\{P_{\ell,k} \mid \ell \neq i\}$. Note that the initial clauses of the form $\vee_{k=1}^{n} P_{j,k}$ will remain unchanged, and the initial clauses of the form $(\neg P_{i,k} \vee \neg P_{j,k})$ will become $\vee_{\ell=1}^{m} P_{\ell,k}$. Note that in the latter case, the clause is not a valid initial clause for a monotone resolution refutation.

Now suppose that C_3 is inferred from the clauses C_1 and C_2 in the original resolution refutation. We want to show how to derive C_3^m from C_1^m and C_2^m. Suppose that C_1 contains $P_{i,k}$ and C_2 contains $\neg P_{i,k}$, where $P_{i,k}$ is the variable resolved upon to obtain C_3. We must show how to derive a subclause of C_3^m from C_1^m and C_2^m. (It suffices to derive a subclause of C_3^m, since it is obvious that the subsumption principle applies to monotone resolution.) There are two cases to consider. Firstly, suppose C_2 is an initial clause of the form $(\neg P_{i,k} \vee \neg P_{j,k})$. In this case, it is easy to check that C_m^1 is a subclause of C_m^3, so this resolution refutation does not need to be simulated by any monotone resolution steps. More generally, if the array representation of C_2^m has a $+$ in the position corresponding to $P_{i,k}$, then it has $+$'s in every position in row k and hence C_1^m is a subclause of C_3^m and no additional monotone resolution inference is needed. Secondly, suppose C_2^m does not have $+$ in the position for $P_{i,k}$. Let C_3^* be the clause obtained from C_1^m and C_2^m by using the monotone resolution inference with respect to row k. We shall show that C_3^* is a subclause of C_3^m. In this case, we can write $C_1^m = A \vee P_{R,k} \vee P_{i,k}$ where $i \notin R$, and can write $C_2^m = B \vee P_{R,k} \vee P_{T,k}$ where T is the complement of $R \cup \{i\}$. Thus C_m^* is the clause $A \vee B \vee P_{R,k}$. Each member $P_{j,k}$ of $P_{R,k}$ is present in C_1^m because it is already in C_1 or because $\neg P_{j',k}$ is in C_1 for some $j' \neq j$. The same literal also appears in C_3 and therefore $P_{j,k}$ is also in C_3^m. This shows that C_3^* is a subclause of C_3^m.

Therefore, we have shown that a resolution inference can be simulated by (at most) a single monotone resolution inference. □

The "onto" version of the pigeonhole principle is obtained by taking the clauses $P_{1,k} \vee P_{2,k} \vee \cdots \vee P_{n,k}$ as additional initial clauses. However, these clauses are just the monotone translation of the initial clauses $\neg P_{i,k} \vee \neg P_{j,k}$ of the usual pigeonhole principle. Examination of the above proof shows that we have proved that any (ordinary) resolution refutation of the onto pigeonhole principle of size n inferences, can be translated into a monotone resolution of size $\leq n$. From this, the following theorem is an immediate corollary of Theorem 1.

Theorem 2. *The shortest resolution proofs of PHP_n^m have size polynomially bounded by the size of resolution proofs of the onto pigeonhole principle with m pigeons and n holes.*

4 An Upper Bound

Theorem 3. *There is a $d > 0$ such that when $m = 2^{\sqrt{n \log n}}$, then PHP_n^m has a resolution proof with m^d steps. Thus, for $m \geq 2^{\sqrt{n \log n}}$, PHP_n^m has a resolution proof of size polynomially bounded by the number of variables.*

Since Haken [6] proved a size lower bound of $2^{\epsilon n}$ for proofs of PHP_n^{n+1}, where ϵ is a constant, Theorem 3 implies that the size of resolution proofs of PHP_n^{n+1} must be superpolynomially longer than the shortest resolution proof of PHP_n^m where $m = 2^{\sqrt{n \log n}}$.

By Theorem 1, it will suffice to prove Theorem 3 for monotone resolution proofs instead of ordinary resolution proofs; indeed, since there are m pigeons, the length of the shortest ordinary resolution refutation is no more than m^2

times the length of a monotone resolution refutation. First, we need the following lemma:

Lemma 1. *PHP_n^{n+1} has a monotone resolution refutation of size $O(n2^n)$.*

Note that the lemma and the proof of Theorem 1 imply that PHP_n^{n+1} has an ordinary resolution proof of size $O(n^3 2^n)$.

Proof. Let $P_{S,T}$ denote the disjunction of the variables $P_{i,j}$, where $i \in S$, $j \in T$. Also, $P_{i,T}$ denotes $P_{\{i\},T}$. Let $[i,j]$ denote the set $\{i, i+1, \ldots, j\}$.

The initial clauses of the monotone resolution refutation are $P_{i,[1,n]}$ for all $i \in [1, n+1]$. The monotone refutation first derives the clauses $P_{S^{(2)},[2,n]}$, for all sets $S^{(2)} \subset [1, n+1]$ of size 2. Each is obtained by one monotone resolution step from the initial clauses. Next, we generate the clauses $P_{S^{(3)},[3,n]}$ for all $S^{(3)} \subset [1, n+1]$ of size 3. Each is obtained by two monotone resolution steps from clauses derived in the previous stage. Continuing in this fashion, we eventually derive $P_{S^{(n)},n}$ for all $S^{(n)} \subset [1, n+1]$ of size n. Finally, we derive the empty clause from this last set of clauses. The total number of monotone resolution inferences derived is bounded by

$$n\left(\binom{n+1}{2} + \binom{n+1}{3} + \ldots + \binom{n+1}{n}\right) \leq n2^{n+1} = O(n2^n).$$

Proof. We will now prove Theorem 3 by induction on n. Let $a = b\sqrt{n \log n}$ for a fixed $b > 1$.

The base case, $n = 2$, is trivial for d sufficiently large. The induction step is argued as follows: The monotone resolution refutation we construct has two stages. The first stage splits the m pigeons into disjoint blocks of $a+1$ pigeons. For each block, we run a resolution refutation of PHP_a^{a+1}, so as to remove a 1's (range elements) from the columns. That is to say, for each block S of $a+1$ columns, we derive the clause $C_{S,T}$ where T is $[1, n-a]$. The size analysis for this part is equal to the number of blocks times the complexity of proving PHP_a^{a+1}; i.e.,

$$(m/(a+1))O((a+1)2^a) = O(2^{(b+1)\sqrt{n \log n}}).$$

In the second stage, we use the induction hypothesis applied to $n-a$ holes; we do this by keeping the disjoint blocks of $a+1$ columns (pigeons) grouped together; in essence, we have divided the number of columns by $a+1$. Therefore, we are proving an instance of $PHP_{n-a}^{m/(a+1)}$: the induction hypothesis tells us that this can be proved with the number of monotone resolution inferences bounded by

$$\begin{aligned} 2^{d\sqrt{(n-a)\log(n-a)}} &< 2^{d(\sqrt{n \log n} - a(1+\log n)/(2\sqrt{n \log n}))} \\ &= 2^{d\sqrt{n \log n} - 0.5db(1+\log n)} \\ &= o(2^{d\sqrt{n \log n}}) \end{aligned}$$

(The first inequality is obtained by letting $f(x) = \sqrt{x \log x}$ and using the fact that $f(n-a) < f(n) - af'(n)$ since f is concave down.)

Adding the size bounds from the two stages of the monotone resolution refutation gives the desired upper bound of $2^{d\sqrt{n \log n}}$, provided d is sufficiently large.

It is still left to verify that the use of the inductive hypothesis was valid, i.e., that $m/(a+1) > 2^{\sqrt{(n-a)\log(n-a)}}$. The lefthand side is equal to

$$2^{\sqrt{n \log n}}/(b\sqrt{n \log n} + 1).$$

By the calculation above, the righthand side is less than or equal to $2^{\sqrt{n \log n}}/n^{b/2}$, since $d > 1$. Thus the desired inequality holds since $b > 1$. □

5 A Lower Bound

Theorem 4. *For any $m > n$, any tree-like resolution refutation of PHP^m_n requires 2^n steps.*

Proof. We'll prove the stronger statement that any tree-like monotone resolution refutation P of PHP^m_n has at least 2^n inferences.

The proof is by induction on n. For $n = 1$, the statement is easy to verify. Now suppose $n > 1$. Let the last inference of P infer the empty clause from two clauses C_1 and C_2 by a monotone resolution inference. We have $C_1 = P_{S,k}$ and $C_2 = P_{T,k}$ for disjoint nonempty subsets S and T of $[1, n]$. Let $p_{i,k} \in S$ and $p_{j,k} \in T$. Let P_1 and P_2 be the subproofs of P which derive C_1 and C_2 respectively. We form a new refutation from P_1 by restricting $p_{j,k}$ to be true: this involves (1) removing from P_1 every clause which contains $p_{j,k}$ and (2) erasing from the clauses of P_1 every occurrence of the variables $p_{j',k}$ with $j' \neq j$. The result is (easily modified to be) a valid resolution proof of PHP^{m-1}_{n-1}. By the induction hypothesis, this proof and hence P_1 must have at least 2^{n-1} inferences. Similar reasoning shows that P_2 must have at least 2^{n-1} inferences. Therefore, P has at least $2^{n-1} + 2^{n-1} + 1$ inferences. □

6 Further Research

Subsequently to the present paper, Razborov, Widgerson and Yao [11] have investigated relationships between restricted resolution refutations of the pigeonhole principle and restricted read-once braching programs. They identified several restricted versions of resolution, including a *rectangular resolution calculus*, and they generalized the upper bound of Theorem 3 to the rectangular calculus and proved a nearly matching lower bound on the size of rectangular refutations for the weak pigeonhole principle.

For the (unrestricted) resolution calculus, the problem of proving exponential lower bounds for the weak pigeonhole principle, $\neg PHP^m_n$, where the number of pigeons, m, is polynomially large (e.g., $m = n^2$) remains open.

References

1. M. Ajtai, *The complexity of the pigeonhole principle*, in Proceedings of the 29-th Annual IEEE Symposium on Foundations of Computer Science, 1988, pp. 346–355.
2. P. Beame, R. Impagliazzo, J. Krajíček, T. Pitassi, P. Pudlák, and A. Woods, *Exponential lower bounds for the pigeonhole principle*, in Proceedings of the 24th Annual ACM Symposium on Theory of Computing, 1992, pp. 200–220.
3. S. R. Buss, *Polynomial size proofs of the propositional pigeonhole principle*, Journal of Symbolic Logic, 52 (1987), pp. 916–927.
4. S. R. Buss and György Turán, *Resolution proofs of generalized pigeonhole principles*, Theoretical Computer Science, 62 (1988), pp. 311–317.
5. S. A. Cook and R. A. Reckhow, *The relative efficiency of propositional proof systems*, Journal of Symbolic Logic, 44 (1979), pp. 36–50.
6. A. Haken, *The intractability of resolution*, Theoretical Computer Science, 39 (1985), pp. 297–308.
7. J. Krajíček, P. Pudlák, and A. Woods, *Exponential lower bound to the size of bounded depth Frege proofs of the pigeonhole principle*, Random Structures and Algorithms, 7 (1995), pp. 15–39.
8. J. B. Paris and A. J. Wilkie, Δ_0 *sets and induction*, in Open Days in Model Theory and Set Theory, W. Guzicki, W. Marek, A. Pelc, and C. Rauszer, eds., 1981, pp. 237–248.
9. J. B. Paris, A. J. Wilkie, and A. R. Woods, *Provability of the pigeonhole principle and the existence of infinitely many primes*, Journal of Symbolic Logic, 53 (1988), pp. 1235–1244.
10. T. Pitassi, P. Beame, and R. Impagliazzo, *Exponential lower bounds for the pigeonhole principle*, Computational Complexity, 3 (1993), pp. 97–140.
11. A. Razborov, A. Widgerson and A. Yao, *Read-once branching programs, rectangular proofs of the pigeon-hole principle and the transversal calculus*, in Proceedings of the 29th Annual ACM Symposium on Theory of Computing, 1997, pp. 739–748.