Notre Dame Journal of Formal Logic
Volume 34, Number 4, Fall 1993

A Parity-Based Frege Proof for the Symmetric Pigeonhole Principle

STEVE FIREBAUGH

Abstract Sam Buss produced the first polynomial size Frege proof of the pigeonhole principle. We introduce a variation of that problem and produce a simpler proof based on parity. The proof appearing here has an upper bound that is quadratic in the size of the input formula.

1 Introduction The proof complexity of the pigeonhole principle is a well studied problem. In 1985, Haken [6] proved that any resolution proof for it must have exponential size. In 1987, Buss [4] produced a polynomial size Frege proof. Ajtai [1] first showed that any polynomial Frege proof must have greater than constant depth. Most recently, Beame, Impagliazzo, Krajíček, Pitassi, Pudlák, and Woods [3] showed an $\Omega(\log\log n)$ lower bound on the depth of any polynomial Frege proof.

Buss's Frege proof proceeds by building up predicates to count the number of atoms set to true, i.e., a propositional multiplexer. With the binary results, the proof then introduces predicates to compute "less than." Starting with the constraints that each pigeon must be placed in some hole, the proof derives that n is less than the number of holes filled by the first $n+1$ pigeons. Using the constraints that each hole may contain at most one pigeon, the proof derives that the number of holes filled by the first $n+1$ pigeons is less than or equal to n. In this way, Buss produces a contradiction. He estimates the size of the proof to be $O(n^{20})$.

The main result of this paper is a proof similar in structure to Buss's proof. We introduce a variation of the pigeonhole principle which is referred to as the symmetric pigeonhole principle. With this variation it is possible to get a simpler Frege proof that is based on parity rather than less than.

The symmetric pigeonhole principle is formalized as follows. For both problems, begin with the same variable space P_{ij}, $1 \le i \le (n+1)$, $1 \le j \le n$, where P_{ij} is true iff pigeon i is placed in hole j. The traditional pigeonhole principle, abbreviated PHP_n, is the conjunction of the following two constraints.

Received January 18, 1993; revised August 16, 1993

For every pigeon, there exists some hole it is placed in.

$$\bigwedge_{1 \le i \le (n+1)} \left(\bigvee_{1 \le j \le n} P_{ij} \right).$$

For every hole, for every pair of two pigeons, one of those pigeons is not placed in the hole.

$$\bigwedge_{1 \le j \le n} \; \bigwedge_{1 \le i < z \le (n+1)} (\neg P_{ij} \vee \neg P_{zj}).$$

In order to get the symmetric pigeonhole principle, abbreviated $SPHP_n$, we add the following additional two constraints.

For every hole, there exists some pigeon that is placed in it.

$$\bigwedge_{1 \le j \le n} \left(\bigvee_{1 \le i \le (n+1)} P_{ij} \right).$$

For every pigeon, for every pair of two holes, one of those holes does not contain the pigeon.

$$\bigwedge_{1 \le i \le (n+1)} \; \bigwedge_{1 \le j < z \le n} (\neg P_{ij} \vee \neg P_{iz}).$$

The PHP_n formula represents the truth value of the existence of a 1-1 relation from $(n + 1)$ to n. By adding the first set of new clauses, the relation is forced to be onto. By adding the second set, it is forced to be a function. Thus the main proof in this paper that $\neg SPHP_n$ shows that there is no 1-1 onto function from $(n + 1)$ to n.

The variables P_{ij} may be thought of as the edges in a complete bipartite graph with the pigeon vertices on one side and the hole vertices on the other. This new problem is termed "symmetric" because the constraint clauses are now the same for both sets of vertices. Ajtai [2] defines a closely related problem, PAR_n, which is based, in a similar fashion, on a completely connected graph. PAR_n represents the truth value of the existence of a perfect matching on a set with an odd number of elements. Ajtai shows that this formulation of the parity principle is, in some sense, stronger than the pigeonhole principle. The proof which follows of $\neg SPHP_n$ works equally as well to show $\neg PAR_n$.

2 Lemmas In 1979, Cook and Reckhow [5] showed that any Frege system is polynomial equivalent to any other. With this in mind, the Frege system chosen for this proof provides a rich set of connectives, namely $\neg$ (not), $\wedge$ (and), $\vee$ (or), $\rightarrow$ (implication), and $\oplus$ (parity). The total number of symbols appearing in the proof provides the measure of the size.

For the sake of readability, the proof is presented here in extended Frege form. We allow the introduction of new variables to represent formulas over existing variables. However, the substitution is allowed only a constant number of levels deep, namely two. The size analysis for the proof is done with the introduced variables replaced by the formulas they represent.

The definition of $SPHP_n$ is derived by representing each edge in the bipartite graph with a variable. The proof proceeds by considering the edges of the graph one at a time. For this reason, it is helpful to have an ordering on the edges. The pairing function $\langle i,j\rangle$ is defined to be $(i-1)*n+j$, and the convention is that the result is represented by l. Projection of the two constituents out of the result is represented by subscripting, i.e., $\langle i,j\rangle = l$ and $l_1 = i$, $l_2 = j$. L is defined to be the maximum value for l, $L \equiv (n+1)*n$, and X is defined to be the maximum number of vertices, $X \equiv 2n+1$. The abbreviation m, $m \equiv (n+1)$, is introduced because the holes are typically labeled starting at 1; however, they are labeled starting at $m+1$ when they are used as vertices in the bipartite graph since the pigeon vertices occupy the labels $1 \ldots (n+1)$.

The first new variables introduced match one for one the vertices from the graph. The truth value of the introduced variable indicates whether that vertex has been "included" after considering a certain number of edges. (A pigeon is *included if it is placed in a hole; a hole is included if it has a pigeon placed in* it.) For $1 \le x \le X$ and $0 \le l \le L$,

$$V_l^x \equiv \bigvee_{((x=i)\,or\,(x=j+m))\,and\,(\langle i,j\rangle \le l)} P_{ij}.$$

This may be thought of inductively as

$$V_l^x = \begin{cases} \textbf{FALSE} & l = 0 \\ V_{l-1}^x & l \neq 0,\ l_1 \neq x,\ l_2 + m \neq x \\ V_{l-1}^x \vee P_{l_1 l_2} & l \neq 0,\ (l_1 = x)\ or\ (l_2 + m = x). \end{cases}$$

As this definition indicates, most of the V_l^x's are unchanged as l is increased by one. Only two of the variables are different: the two which represent vertices on the ends of edge l.

The second set of new variables is introduced to represent the parity of the vertices that are included. For $0 \le l \le L$,

$$P_l \equiv \bigoplus_{1\le x\le X} V_l^x.$$

P_l is the parity of all of the vertices which have been included after having considered the first l edges.

Lemma 1 $\neg P_0$.

Proof: By definition, $P_0 \equiv V_0^1 \oplus V_0^2 \oplus \cdots \oplus V_0^X$. Also by definition, $V_0^x \equiv$ **FALSE** for all x. Therefore, if the Frege rule **(FALSE** $\oplus$ **FALSE** $\oplus A) \mid A$ is applied $\lfloor \frac{X}{2} \rfloor$ times, the string from the formal system representing P_0 is reduced to **FALSE** and therefore $\neg P_0$.

Lemma 2 $SPHP_n \rightarrow P_L$.

Proof: Assume $SPHP_n$. It is easy to derive $(\vee_i P_{ij})$ for any j or $(\vee_j P_{ij})$ for any i from $SPHP_n$ by using Frege rules for commutativity and the Frege rule $(A \wedge B)|A$. Depending on the value of x, the formula $(\vee_i P_{ij})$ or the formula $(\vee_j P_{ij})$ is exactly V_L^x. Thus $SPHP_n \rightarrow V_L^x$, $1 \le x \le X$. Now, the formulas which have been derived are combined by using the Frege rule $(A,B,C) \mid A \oplus B \oplus C$. This is repeated $\lfloor \frac{X}{2} \rfloor$ times until the formula $V_L^1 \oplus V_L^2 \oplus V_L^3 \oplus \cdots \oplus V_L^X$ is obtained. This is exactly the definition of P_L. Therefore $SPHP_n \rightarrow P_L$.

Lemma 3 *For any l, $1 \le l \le L$, $SPHP_n \to (\neg P_{l_1 l_2} \vee \neg V_{l-1}^{l_1})$ and $SPHP_n \to (\neg P_{l_1 l_2} \vee \neg V_{l-1}^{l_2+m})$.*

Proof: Assume $SPHP_n$. It is easy to derive

$$(\neg P_{l_1 l_2} \vee \neg P_{l_1 1}) \wedge (\neg P_{l_1 l_2} \vee \neg P_{l_1 2}) \wedge \ldots \wedge (\neg P_{l_1 l_2} \vee \neg P_{l_1 (l_2-1)})$$

from $SPHP_n$. Restructuring this by using distributive Frege rules produces

$$(\neg P_{l_1 l_2} \vee (\neg P_{l_1 1} \wedge \neg P_{l_1 2} \wedge \ldots \wedge \neg P_{l_1 (l_2-1)})).$$

Using Frege rules for DeMorgan's laws produces

$$(\neg P_{l_1 l_2} \vee \neg (P_{l_1 1} \vee P_{l_1 2} \vee \ldots \vee P_{l_1 (l_2-1)})).$$

The second half of this disjunction inside the negation is exactly $V_{l-1}^{l_1}$, thus

$$(\neg P_{l_1 l_2} \vee \neg V_{l-1}^{l_1}).$$

The proof of $(\neg P_{l_1 l_2} \vee \neg V_{l-1}^{l_2+m})$ is exactly the same.

3 *Theorem and Analysis*

Theorem 1 $\neg SPHP_n$.

Proof: Assume $SPHP_n$. By Lemma 2, P_L is true. P_L is by definition

$$V_L^1 \oplus V_L^2 \oplus \cdots \oplus V_L^X.$$

By the inductive definition of V_L^x, this is

$$V_{L-1}^1 \oplus V_{L-1}^2 \oplus \cdots \oplus (V_{L-1}^{L_1} \vee P_{L_1 L_2}) \oplus \cdots \oplus (V_{L-1}^X \vee P_{L_1 L_2}).$$

In this case, $L_2 + m = X$; however, the steps that follow will work for any l. By using commutative Frege rules for $\oplus$, the two complicated terms may be moved to the right. By Lemma 3 and distributive Frege rules for $\oplus$ and $\wedge$, $(\neg P_{L_1 L_2} \vee \neg V_{L-1}^{L_1})$ can be derived for one of these terms and $(\neg P_{L_1 L_2} \vee \neg V_{L-1}^{L_2+m})$ for the other.

$$V_{L-1}^1 \oplus V_{L-1}^2 \oplus \cdots \oplus ((V_{L-1}^{L_1} \vee P_{L_1 L_2}) \wedge (\neg P_{L_1 L_2} \vee \neg V_{L-1}^{L_1})) \oplus ((V_{L-1}^{L_2+m} \vee P_{L_1 L_2}) \wedge (\neg P_{L_1 L_2} \vee \neg V_{L-1}^{L_2+m})).$$

By Frege rule $((A \vee B) \wedge (\neg A \vee \neg B)) \mid A \oplus B$,

$$V_{L-1}^1 \oplus V_{L-1}^2 \oplus \cdots \oplus (V_{L-1}^{L_1} \oplus P_{L_1 L_2}) \oplus (V_{L-1}^{L_2+m} \oplus P_{L_1 L_2}).$$

By Frege rule $(A \oplus B \oplus B \oplus C) \mid A \oplus C$ and by reusing the commutative Frege rules,

$$V_{L-1}^1 \oplus V_{L-1}^2 \oplus \cdots \oplus V_{L-1}^{L_1} \oplus \cdots \oplus V_{L-1}^X$$

which is the definition of P_{L-1}. These steps are repeated L more times to produce $P_{L-2}, P_{L-3}, \ldots$, and eventually P_0. However, by Lemma 1, P_0 is false.

$SPHP_n$ is $O(n^3)$. Writing out P_L has $2n+1$ V_L^x's, each of which has n or $n+1$ variables. The widest line includes $P_L, SPHP_n$ written twice, $(\neg P_{L_1 L_2} \vee \neg V_{L-1}^{L_1})$, and $(\neg P_{L_1 L_2} \vee \neg V_{L-1}^{L_2+m})$. Thus the widest line is only $O(n^3)$.

The application of Lemma 3 in each stage of the main proof required $O(n)$ lines. This is repeated for each of the edges, or $O(n^2)$ times. Thus the proof has $O(n^3)$ lines. The final size is $O(n^6)$. Since the representation of the problem, $SPHP_n$, is $O(n^3)$, the proof is actually quadratic as a function of the size of the input formula.

4 Conclusions By adding clauses to the pigeonhole principle, we produce a formula which expresses the fact that there is no 1-1 onto function from $(n + 1)$ to n. Because there are more constraints, this new formula is theoretically easier to prove. However, it is interesting that we are able to do away with all of the complications incurred by using "less than" in a propositional proof. Instead, we use "parity" to get a simpler proof which is of size $O(n^6)$, which is quadratic in the size of the input.

Unlike the pigeonhole principle and Ajtai's parity principle [2], the symmetric pigeonhole principle is not "minimally inconsistent." It is possible to prove $\neg SPHP_n$ without using every original clause (use Buss's proof in [4]). However, for the proof presented here, every original clause is required. In order to adapt this proof to the standard pigeonhole principle, it would be necessary to first derive the additional two sets of constraint clauses. Although this is possible, it is not obviously easier than simply proving $\neg PHP_n$. The lower bound result on the depth of any polynomial Frege proof of the pigeonhole principle found in [3] will work for the symmetric pigeonhole principle.

REFERENCES

[1] Ajtai, M., "The Complexity of the Pigeonhole Principle," *Proceedings of the 29th Annual Symposium on Foundations of Computer Science* (1988), pp. 346–358.

[2] Ajtai, M., "Parity and the Pigeonhole Principle," published in *Feasible Mathematics*, ed. Buss and Scott, Birkhauser (1990), pp. 1–23.

[3] Beame, P., R. Impagliazzo, J. Krajíček, T. Pitassi, P. Pudlák, A. Woods, "Exponential Lower Bounds for the Pigeonhole Principle," *Proceedings of the 24th Annual Symposium on the Theory of Computing* (1992).

[4] Buss, S. R., "Polynomial Size Proofs of the Propositional Pigeonhole Principle," *Journal of Symbolic Logic*, 52(4) (1987), pp. 916–927.

[5] Cook, S. A. and R. A. Reckhow, "The Relative Efficiency of Propositional Proof Systems," *Journal of Symbolic Logic*, 44(1) (1979), pp. 36–50.

[6] Haken, A., "The Intractability of Resolution," *Theoretical Computer Science*, vol. 39 (1985), pp. 297–308.

6515 16th Ave NE
Seattle, WA 98115