A Note on the Power of Threshold Circuits

Eric Allender[1]
Department of Computer Science
Rutgers University
New Brunswick, NJ 08903

SUMMARY

Every language in AC^0 can be recognized by depth three threshold circuits of size $2^{\log^{O(1)} n}$.

1 Introduction

This paper presents a very simple proof of the fact that any language accepted by polynomial-size depth-k unbounded-fan-in circuits of AND and OR gates is accepted by depth-three threshold circuits of size $n^{O(\log^k n)}$. The proof uses much of the intuition of Toda's important result that the polynomial hierarchy is contained in $P^{\#P}$ [To-89] (making use of some known connections between the polynomial hierarchy and constant-depth circuits [FSS-84]). The proof also makes use of some observations of Razborov [Ra-87] and Smolensky [Sm-87].

For our purposes, a threshold circuit is a circuit with inputs $x_1, \ldots, x_n$ and their negations $\overline{x_1}, \ldots, \overline{x_n}$ and the constants $\{0, 1\}$, where the only gates are MAJ (majority) gates. A MAJ gate takes the value 1 iff more than half of its inputs have the value 1. Note that MAJ gates can easily simulate AND and OR gates. MAJ gates are quite powerful; Smolensky [Sm-87], building on the work of Razborov [Ra-87], has shown that the MAJ function requires exponential size to compute using constant-depth circuits with AND, OR, and (MOD p) gates, for prime p.

[1]Supported in part by National Science Foundation Research Initiation Grant number CCR-8810467. Some of this research was performed while the author was a visiting professor at Institut für Informatik, Universität Würzburg, D-8700 Würzburg, Federal Republic of Germany.

Threshold circuits have been studied by a number of authors; among others, we mention [BIS-88, Br-89, HM-87, PS-88, Re-87]. It is shown in [HM-87] that there is a language recognized by a family of polynomial-size depth three threshold circuits that cannot be recognized by polynomial-size depth two threshold circuits. On the other hand, it is not known if there is any language in NP that cannot be recognized by depth three polynomial-size threshold circuits.

The class of languages accepted by polynomial-size threshold circuits of depth $O(1)$ is denoted by TC^0. There are competing conjectures concerning the relationship between TC^0 and NC^1. Immerman and Landau have conjectured that $TC^0 = NC^1$ [IL-89], while Yao [Ya-89] discusses a conjecture that $TC^0 \neq NC^1$.

AC^0, the class of sets which can be recognized by circuits of polynomial size, with unbounded fan-in AND and OR gates, of depth $O(1)$, has also been studied by many authors in the past few years. Of particular interest has been the question of how difficult it is to simulate AC^0 circuits of depth k for fixed k. Sipser [Si-83] proved that polynomial-size depth $k-1$ AC^0 circuits are less powerful than polynomial-size depth k AC^0 circuits. Yao [Ya-85] and Håstad [Hå-86] improved this, showing that there are sets accepted by depth k AC^0 circuits, which require exponential size on AND, OR circuits of depth $k-1$.

Recently, Yao has shown that depth k AC^0 circuits require exponential size to simulate on *monotone* depth $k-1$ threshold circuits [Ya-89]. Thus a corollary of the work presented here is that (unrestricted) threshold circuits are much more powerful than monotone threshold circuits.

CH2806-8/89/0000/0580/$01.00

Results about constant-depth circuit families can often be interpreted as results about other complexity classes, and vice-versa. Furst, Saxe, and Sipser were among the first to make this connection explicit; in [FSS-84] they showed how to relate AC^0 to the polynomial-time hierarchy. The relation between PP and threshold circuits was noted in [PS-88]. Other papers which have made use of similar connections include [Bab-87, Ca-89, Hå-86, IN-88, NW-88, St-83, To-89, Ya-85]. Seen in this setting, it is clear that Toda's theorem [To-89] that the polynomial hierarchy is contained in P^{PP} has certain consequences concerning the power of threshold functions. The purpose of this paper is to give a direct and simple exposition of those consequences.

2 Main Result

This first lemma may be viewed as being a much weaker version of a result of Razborov [Ra-87] (generalized by Barrington [Bar-87] and Smolensky [Sm-87]) showing that circuits with small depth can be approximated by polynomials of small degree.

Lemma 1 For any polynomial p, there is a family of probabilistic depth-two circuits of size $n^{O(\log n)}$, computing the OR of n bits, with error less than $1/p(n)$. The first level of this circuit consists of ANDs of fan-in $O(\log n)$, and the second level consists of a PARITY gate.

Proof: In order to compute the OR of $b_1, b_2, \ldots b_n$, first consider the circuit B_n with one PARITY gate, where the inputs to the parity gate are

$$\{1\} \cup \{\mathrm{AND}(b_i, p_i) : 1 \leq i \leq n\},$$

where the p_i are probabilistic bits. It is easy to see that if $\mathrm{OR}(b_1 \ldots b_n) = 0$, then B_n outputs 1, and if $\mathrm{OR}(b_1 \ldots b_n) = 1$, then B_n outputs 0 with probability exactly 1/2.

Now take $k \log n$ separate copies of B_n (with independent probabilistic inputs for each copy of B_n) and AND these $k \log n$ circuits together. Call this new circuit C_n. It is immediate that if $\mathrm{OR}(b_1 \ldots b_n) = 0$, then C_n outputs 1, and if $\mathrm{OR}(b_1 \ldots b_n) = 1$, then C_n outputs 0 with probability $1 - 1/n^k$.

One may view C_n as a polynomial over GF(2). Using the distributive laws, one may rewrite the polynomial as

$$\Sigma\,(\Pi(b_{i_1} p_{1,i_1} b_{i_2} p_{2,i_2} \ldots b_{i_{k \log n}} p_{k \log n, i_{k \log n}})),$$

where the b_{i_j}'s range over the bits $\{b_1, \ldots, b_n\} \cup \{1\}$, and the p_{j,i_j}'s are the associated probabilistic bits. Clearly this polynomial can be implemented as a PARITY gate of $n^{O(\log n)}$ AND gates, where each AND gate has fan-in $O(\log n)$. Let D_n be the circuit that computes the negation of this polynomial (e.g., the PARITY gate has an additional 1 input). Then D_n is a circuit with the properties claimed by the lemma. ■

Corollary 2 For any polynomial p, there is a family of probabilistic depth-two circuits of size $n^{O(\log n)}$, computing the AND of n bits, with error less than $1/p(n)$. The first level of this circuit consists of ANDs of fan-in $O(\log n)$, and the second level consists of a PARITY gate.

Lemma 3 Let L be accepted by an AC^0 circuit of depth k, and let p be any polynomial. Then L is accepted by a probabilistic circuit of depth two with error less than $1/p(n)$, where the first level of the circuit consists of $n^{O(\log^k n)}$ ANDs of fan-in $O(\log^k n)$, and the second level consists of a PARITY gate.

Proof: The proof proceeds by induction on k. The basis case is proved in Lemma 1. For the induction step, let L be accepted by a family of depth k circuits of polynomially-many unbounded-fan-in AND and OR gates. Consider the circuit C_n for inputs of length n. Assume without loss of generality that the output gate of C_n is an AND gate (the proof is entirely symmetric when it is an OR gate). Thus C_n is the AND of at most n^l circuits of depth $k-1$. By the inductive hypothesis, each of these n^l circuits may be replaced by a probabilistic circuit of size $n^{O(\log^{k-1} n)}$, having error

probability at most $1/n^a$ (where a may be any constant). The resulting circuit has error probability at most n^l/n^a. Also, the top-level AND in this circuit can be replaced by a probabilistic depth-two circuit of the sort guaranteed by Corollary 2; the resulting circuit will have a PARITY gate on level 4, $n^{O(\log n)}$ AND gates on level 3, at most n^l PARITY gates on level 2, and $n^{O(\log n)}$ AND gates on level 1, and may be constructed to have error probability less than $1/p(n)$. The AND gates on level 3 have fan-in $O(\log n)$, and those on level 1 have fan-in $O(\log^{k-1} n)$.

Consider any AND gate on level 3. It has $O(\log n)$ PARITY gates as input, where each PARITY gate has $2^{O(\log^{k-1} n)}$ AND gates as input. Again using the distributive property, we can rewrite each such level-3 AND circuit as the Boolean sum of $n^{O(\log^k n)}$ AND gates, where each AND has fan-in $\log^k n$. The resulting circuit now has one PARITY gate on level 3, $n^{O(\log n)}$ PARITY gates on level 2, and $n^{O(\log^k n)}$ AND gates on level 1.

Any AND gate on level 1 that is input to an even number of PARITY gates on level 2 may be deleted without changing the output of the circuit. Thus we may assume that each AND on level 1 is input to an odd number of PARITY gates on level 2. It is now easy to verify that an odd number of the level 2 PARITY gates are on iff an odd number of the AND gates are on. Thus the circuit is in fact equivalent to a circuit consisting of a single PARITY gate connected to the $n^{O(\log^k n)}$ AND gates on level 1. ∎

Theorem 4 Every set in AC^0 is accepted by a depth three threshold circuit of size $n^{O(\log^k n)}$.

Proof: The depth two probabilistic circuit constructed in Lemma 3 can be converted into a deterministic depth three circuit using the technique of Proposition 4.2 in [HM-87]. The resulting circuit is only polynomially larger than the original probabilistic circuit, and consists of AND gates on level 1, PARITY gates on level 2, and a Majority gate as the output gate.

The AND gates are easily replaced with MAJ gates. Proposition 5 below shows how to replace the PARITY gates with MAJ gates with only a small increase in the size of the circuit. That proves the theorem. ∎

Proposition 5 If C is a depth two circuit with one MAJ gate as output and r PARITY gates on level 1, where no PARITY gate has more than $2m$ inputs, then C is equivalent to a depth two threshold circuit with at most $1 + 2rm$ MAJ gates.

Proof: We may assume without loss of generality that all of the PARITY gates have exactly $2m$ inputs. (Some of these inputs may be constant 1 or 0.)

Let G be one of the PARITY gates. For each odd integer $i, 1 \leq i \leq 2m$, build two MAJ gates, one which accepts if at least i of the $2m$ inputs are 1, and one which accepts if at most i of the $2m$ inputs are 1. It is immediate that if an even number of the $2m$ inputs are 1, then exactly m of the MAJ gates will be 1, while if an odd number of the $2m$ inputs are 1, then exactly $m + 1$ of the MAJ gates will be 1. Replace G with the $2m$ MAJ gates constructed in this way.

Now have the MAJ gate at level 2 accept iff at least $rm + r/2$ of the MAJ gates on level 1 have value 1. ∎

3 Generalizations

The proof presented in the previous section actually suffices to prove much stronger results. A few of these results are listed below.

Definitions:

- SIZE($s(n)$)DEPTH($d(n)$)GATES(S) denotes the class of languages which can be recognized by circuit families of size $s(n)$ and depth $d(n)$ where the types of gates which can be used are in the set S.
- BPSIZE($s(n)$)DEPTH($d(n)$)GATES(S) denotes the analogous class defined in terms of probabilistic circuits. (That is, the circuits have $n^{O(1)}$ probabilistic

bits as auxiliary inputs. A probabilistic circuit C with n inputs (and n^k probabilistic inputs) recognizes a set $L \subseteq \Sigma^n$ if for all $x \in L$, $\text{Prob}(C(x) = 1) > 3/4$, and for all $x \notin L$, $\text{Prob}(C(x) = 1) < 1/4$.)

Corollary 6
$\text{BPSIZE}(2^{\log^{O(1)} n})\text{DEPTH}(O(1))\text{GATES}(\{\wedge, \vee, \oplus\})$
$= \text{BPSIZE}(2^{\log^{O(1)} n})\text{DEPTH}(2)\text{GATES}(\{\wedge, \vee, \oplus\})$
$= \text{SIZE}(2^{\log^{O(1)} n})\text{DEPTH}(5)\text{GATES}(\{\wedge, \vee, \oplus\})$
$\subseteq \text{SIZE}(2^{\log^{O(1)} n})\text{DEPTH}(3)\text{GATES}(\{\text{MAJ}\})$

Proof: The first equality can be proved in exactly the same manner as Lemma 3. (The proof works for circuits of size $2^{\log^{O(1)} n}$ as well as for polynomial-size circuits. Clearly the introduction of PARITY gates causes no problems.)

The second equality follows from the results of [AB-84] which show how to simulate probabilistic constant-depth circuits.

The final inclusion follows exactly as in Theorem 4. ∎

All of the results in this paper have been stated in terms of "nonuniform" circuit complexity. That is, a language L is considered to have small circuit complexity if there exists a family of small circuits $\{C_n\}$ accepting L, regardless of whether or not the function $n \rightarrow C_n$ is computable. A family of circuits $\{C_n\}$ is "uniform" if the function $n \rightarrow C_n$ is "efficiently computable" in some sense. (See [BIS-88] for a discussion of notions of uniformity for constant-depth circuits.)

The proofs presented in this paper are not suitable for uniform circuit complexity, since the simulations of probabilistic circuits by deterministic circuits given in [HM-87] and [AB-84] are nonuniform. With some more work, however, uniform versions of most of the inclusions presented here can be proved; details will be presented in [AH-89]. The proofs of [AH-89] make use of the [NW-88] technique of simulating probabilistic circuits using only $\log^{O(1)} n$ probabilistic bits. The technique for building deterministic circuits for simulating probabilistic circuits using few probabilistic bits is quite similar to the proof that BPP is in the polynomial hierarchy [Si-83a].

4 Open Problems

Of course the obvious question is: can these results be improved? Currently it is not known if there is any set in NC^1 (or even in NP) that cannot be accepted with polynomial-size threshold circuits of depth 3. The results of this paper show that threshold circuits are quite powerful; it will be interesting to see if all sets in NC^1 have efficient small-depth threshold circuits.

5 Acknowledgments

I benefitted greatly from discussions with Ravi Boppana. I thank Lane Hemachandra for telling me about Toda's result, and I thank Osamu Watanabe for telling me about Toda's proof. Finally, I thank Klaus Wagner, Gerhard Buntrock, Ulrich Hertrampf, and Miroslaw Kowaluk for providing a stimulating environment this summer at Universität Würzburg.

References

[AB-84] M. Ajtai and M. Ben-Or, *A theorem on probabilistic constant depth computations*, Proc. 16th ACM Symposium on Theory of Computing, pp. 471–474.

[AH-89] E. Allender and U. Hertrampf, in preparation.

[Bab-87] L. Babai, *Random oracles separate PSPACE from the polynomial time hierarchy*, Information Processing Letters 26, 51–53.

[Bar-87] D. A. Barrington, *A note on the theorem of Razborov*, unpublished.

[BIS-88] D. A. Mix Barrington, N. Immerman, and H. Straubing, *On uniformity within NC^1*, Proc. 3rd IEEE Structure in Complexity Theory Conference, pp. 47–59.

[Br-89] J. Bruck, *Harmonic analysis of polynomial threshold functions*, to appear in SIAM J. Disc. Math.

[Ca-89] J. Cai, *With probability 1, a random oracle separates PSPACE from the polynomial-time hierarchy*, J. Computer and System Science 38, 68–85.

[FSS-84] M. Furst, J. Saxe, M. Sipser, *Parity, circuits, and the polynomial-time hierarchy*, Mathematical Systems Theory 17, 13–27.

[HM-87] A. Hajnal, W. Maass, P. Pudlák, M. Szegedy, G. Turán, *Threshold circuits of bounded depth*, Proc. 28th IEEE Symposium on Foundations of Computer Science, pp. 99–110.

[Hå-86] J. Håstad, *Almost optimal lower bounds for small depth circuits*, Proc. 18th ACM Symposium on Theory of Computing, pp. 6–20.

[IL-89] N. Immerman and S. Landau, *The complexity of iterated multiplication*, Proc. 4th IEEE Structure in Complexity Theory Conference, pp. 104–111.

[IN-88] R. Impagliazzo and M. Naor, *Decision trees and downward closures*, Proc. 3rd IEEE Structure in Complexity Theory Conference, pp. 29–38.

[NW-88] N. Nisan and A. Wigderson, *Hardness vs. Randomness*, Proc. 29th IEEE Symposium on Foundations of Computer Science, pp. 2–11.

[PS-88] I. Parberry and G. Schnitger, *Parallel computation with threshold functions*, J. Computer and System Science 36, 278–302.

[Ra-87] A. A. Razborov, *Lower bounds on the size of bounded depth networks over a complete basis with logical addition*, Mathematicheskie Zametki 41(4), 598–607. English translation in Mathematical Notes of the Academy of Sciences of the USSR 41:4, 333-338.

[Re-87] J. Reif, *On threshold circuits and polynomial computation*, Proc. 2nd IEEE Structure in Complexity Theory Conference, pp. 118–123.

[Si-83] M. Sipser, *Borel sets and circuit complexity*, Proc. 15th ACM Symposium on Theory of Computing, pp. 61–69.

[Si-83a] M. Sipser, *A complexity theoretic approach to randomness*, Proc. 15th ACM Symposium on Theory of Computing, pp. 330–335.

[Sm-87] R. Smolensky, *Algebraic methods in the theory of lower bounds for Boolean circuit complexity*, Proc. 19th ACM Symposium on Theory of Computing, pp. 77–82.

[St-85] L. Stockmeyer, The complexity of approximate counting (preliminary version), Proc. 15th ACM Symposium on Theory of Computing, pp. 118–126.

[To-89] S. Toda, *PP is $\leq_T^p$-hard for the polynomial-time hierarchy*, these proceedings.

[Ya-85] A. C. Yao, *Separating the polynomial-time hierarchy by oracles*, Proc. 26th IEEE Symposium on Foundations of Computer Science, pp. 1–10.

[Ya-89] A. C. Yao, *Circuits and local computation*, Proc. 21st ACM Symposium on Theory of Computing, pp. 186–196.