INFORMATION AND CONTROL 36, 102–117 (1978)

The Bounded Arithmetic Hierarchy

KEITH HARROW

Department of Computer and Information Science, Brooklyn College, CUNY, Brooklyn, New York 11210

The class of bounded arithmetic predicates (BA) is the smallest class containing the polynomial predicates and closed under bounded quantification $((\exists w)_{\leqslant y} R(\mathbf{x}, y, w)$ or $(\forall w)_{\leqslant y} R(\mathbf{x}, y, w))$. The bounded arithmetic predicates are a small subset of the recursively enumerable, but they include most of the standard examples from recursive function theory and form a basis for the r.e. sets. BA is closed under Boolean operations, and quantification bounded by a polynomial, but it is not closed under quantification bounded by x^y. In analogy with Kleene's arithmetic hierarchy, there is a bounded arithmetic hierarchy of predicate classes within BA, based on the number of alternations of bounded quantifiers. The closure properties of these classes are also studied. Although the existence of a strict hierarchy is not established, necessary and sufficient conditions for the hierarchy to be strict are shown. The relationship of BA to other known classes of predicates is also discussed.

1. INTRODUCTION

Most computer scientists and mathematicians are aware that there is an exact mathematical formulation of what it means for a function to be calculable in a mechanical way. By Church's thesis, the class of recursive functions is precisely the class of functions which can be calculated by actual computers using unlimited time and space. Although they are aware of this fact, most practical computer people justifiably tend to ignore it. First, considerations of time and space are paramount in the real world. Any model of computation which does not recognize this is impractical. Also, there are recursive functions which exhibit pathological properties. While all recursive functions have computations that halt on all inputs, these computations may require an arbitrarily large (e.g., not bounded by any primitive recursive function) amount of time. Manuel Blum and others have found recursive functions which have no best program, or which use rapidly growing amounts of resources, or sets which cannot be enumerated quickly (see Blum, 1967; Young, 1969).

From the point of view of actual computing practice, the recursive function (or its counterpart, the Turing machine) is an unrealistic model. Present-day computers, while fast, cannot in general handle computations involving functions that grow more rapidly than a simple exponential x^y or even 2^x. In fact, polyno-

0019-9958/78/0361-0102$02.00/0
Copyright © 1978 by Academic Press, Inc.
All rights of reproduction in any form reserved.

mial bounded computations provide a much more feasible model. This is one of the reasons why so much research is focused on the P = NP and related problems. These problems deal with computations which are at the limits of what can actually be computed, and thus provide a link between theoretical and applied areas.

We will consider classes of functions and relations (including the polynomial bounded) whose computations do grow at reasonable rates; at the same time, they include almost all of the interesting examples from the literature. In particular, we focus our attention on the class of bounded arithmetic predicates (BA). This class is small enough that is lies well within the relations of Grzegorczyk's $\mathscr{E}^3$, ensuring that there exist computations of at worst one or more exponentials in length. (A class which is contained in $\mathscr{E}^3$ is sometimes called subelementary.) On the other hand the class forms a basis for the r.e. sets and thus has an undecidable equivalence problem. Many unsolved problems in automata and formal language theory concern classes which contain BA, or are contained within BA (but it is not known if the containments are proper). A discussion of these and some other open questions will appear in the last section.

We assume that the reader is familiar with the work and notation of Grzegorczyk (1953). We use $\mathbf{x}$ as an abbreviation for the n-tuple ($n \geqslant 1$) $x_1, \ldots, x_n$. The term predicate or relation will mean a function which assumes only the values 0 (true) and 1 (false). In the well-known way, we identify sets with predicates. If $f(\mathbf{x})$ is a function, then the associated relation of f is the predicate $R(\mathbf{x}, y) \Leftrightarrow y = f(\mathbf{x})$. Given a predicate $R(\mathbf{x}, y)$, we say that $S(\mathbf{x})$ is a bounded existential quantification of R if $S(\mathbf{x}) \Leftrightarrow (\exists y)_{\leqslant x_i} R(\mathbf{x}, y)$, where x_i is one of the variables in $\mathbf{x}$. Similarly, $T(\mathbf{x})$ is a bounded universal quantification of R if $T(\mathbf{x}) \Leftrightarrow (\forall y)_{\leqslant x_i} R(\mathbf{x}, y)$.

DEFINITION. Let Poly be the class of polynomial predicates. A predicate $R(x_1, \ldots, x_n)$ is in Poly if there is a polynomial Q with integral coefficients such that $R(x_1, \ldots, x_n) \leftrightarrow Q(x_1, \ldots, x_n) = 0$.

We will sometimes rewrite the right-hand side of the equivalence as $Q_1(\mathbf{x}) = Q_2(\mathbf{x})$, where Q_1, Q_2 have only positive coefficients. For example, the addition predicate $A(x, y, z) \Leftrightarrow x + y = z$ is in Poly since $A(x_1, x_2, x_3) \Leftrightarrow x_1 + x_2 - x_3 = 0$. Similarly, the multiplication predicate $M(x, y, z) \Leftrightarrow x \cdot y = z$ is also polynomial.

We now define two hierarchies E_i, A_i of predicate classes.

DEFINITION. Let $A_0 = E_0 = \text{Poly}$.

For $n \geqslant 0$, E_{n+1} is the smallest class of predicates containing A_n and closed under bounded existential quantification; A_{n+1} is the smallest class containing E_n and closed under bounded universal quantification.

DEFINITION. The class of bounded arithmetic predicates (BA) is the smallest class containing Poly and closed under both bounded universal and bounded existential quantification.

In analogy with the arithmetic hierarchy, to determine which class a predicate R belongs to, we only count alternations of bounded quantifiers, ignoring adjacent quantifiers of the same type. Observe that $E_i \cup A_i \subseteq E_{i+1} \cap A_{i+1}$ for all $i \geqslant 0$. Each class includes both classes below it, by adding an appropriate dummy quantifier. We will call the E_i existential classes, and the A_i universal classes.

Note that a predicate R is BA if R can be expressed as a prefix of bounded quantifiers in front of a polynomial equation in the free and bound variables. We will often write a BA predicate $R(\mathbf{x})$ as $M\mathbf{y}Q(\mathbf{x}, \mathbf{y}) = 0$, where the prefix M is a sequence of bounded quantifiers, one for each y_i in $\mathbf{y}$, and Q is a polynomial in $\mathbf{x}$ and $\mathbf{y}$. It is easy to see that BA $= \bigcup_{n=0}^{\infty} A_n = \bigcup_{n=0}^{\infty} E_n$. Thus, the E_i and A_i each form a hierarchy of classes within BA.

BA is a subset, in fact quite a small subset, of the recursive predicates. But it is curious to note that until Matiyasevich's result proving that the Diophantine predicates were identical to the r.e. predicates, it was not known if the Diophantine included the bounded arithmetic. In particular, the predicate Prime(x) is BA, but no Diophantine definition for it was known. (In fact, showing that BA $\subseteq$ Diophantine would have also shown that the Diophantine predicates were precisely the recursively enumerable predicates, since $z = x^y$ is BA.) By the Davis normal form for r.e. sets (see Davis, 1958), a single unbounded existential quantifier in front of an appropriate BA predicate will give any r.e. predicate, proving that BA forms a basis for the r.e. sets.

We now summarize some results of this paper. In Section 2, we discuss elementary closure properties of BA. We prove that BA is closed under Boolean operations and quantification with a polynomial as an upper bound. Section 3 uses these results to study a hierarchy of predicate classes within BA, the bounded arithmetic hierarchy. We do not establish the existence of a strict hierarchy, but do give necessary and sufficient conditions for the hierarchy to be strict. We show that there can be no gaps in such a hierachy; either the hierarchy collapses at some stage (and all succeeding classes are identical), or it is strict. In Section 4, we discuss possible extensions and comparisons to other classes.

2. CLOSURE PROPERTIES OF THE BOUNDED ARITHMETIC PREDICATES

LEMMA 1. *Each class A_i, E_i $(i \geqslant 0)$ is closed under conjunction and disjunction. That is, if R, S are predicates in a class, then so are $R \vee S$, $R \& S$.* BA *is also closed under these operations.*

Proof. Let $R(\mathbf{x}) \Leftrightarrow M\mathbf{y}P(\mathbf{x}, \mathbf{y}) = 0$ and $S(\mathbf{x}) \Leftrightarrow N\mathbf{w}Q(\mathbf{x}, \mathbf{w}) = 0$, where

without loss of generality we can assume that the $\mathbf{w}$ and $\mathbf{y}$ variables are disjoint, and M and N are both prefixes from the same class A_i or E_i.

Now $R \vee S \Leftrightarrow M\mathbf{y}N\mathbf{w}[P(\mathbf{x}, \mathbf{y}) = 0 \vee Q(\mathbf{x}, \mathbf{w}) = 0] \Leftrightarrow M\mathbf{y}N\mathbf{w}P(\mathbf{x}, \mathbf{y}) \cdot Q(\mathbf{x}, \mathbf{w}) = 0$, since the product of two polynomials is 0 precisely when at least one is 0. $P(\mathbf{x}, \mathbf{y}) \cdot Q(\mathbf{x}, \mathbf{w})$ is a polynomial in $(\mathbf{x}, \mathbf{y}, \mathbf{w})$ with integral coefficients, and $M\mathbf{y}N\mathbf{w}$ is a bounded quantifier prefix. But no y_i ever bounds or even refers to a w_j, and vice versa. Therefore, we can interweave the prefixes so that the resulting predicate is in the same class as R, S. (This idea of combining two quantifier prefixes in "parallel" will be exploited throughout this article.)

Similarly, $R \,\&\, S \Leftrightarrow M\mathbf{y}N\mathbf{w}[P(\mathbf{x}, \mathbf{y})^2 + Q(\mathbf{x}, \mathbf{w})^2 = 0]$ since the sum of the squares is 0 precisely when both P and Q are 0. Again by interweaving, $R \,\&\, S$ will be in the same class as the original predicates.

Obviously, BA is itself closed under conjunction and disjunction.

LEMMA 2. *Each class A_i, E_i $(i \geqslant 0)$ is closed under explicit transformation.* BA *is also closed under explicit transformation.*

Proof. Clearly, each class is closed under permuting or identifying variables and adding redundant variables.

To handle substitution of a constant for a variable, observe that if $S(\mathbf{x}) \Leftrightarrow (\exists w)_{\leqslant k} R(\mathbf{x}, w)$ then $S(\mathbf{x}) \Leftrightarrow R(\mathbf{x}, 0) \vee \cdots \vee R(\mathbf{x}, k)$, while if $S(\mathbf{x}) \Leftrightarrow (\forall w)_{\leqslant k} R(\mathbf{x}, w)$ then $S(\mathbf{x}) \Leftrightarrow R(\mathbf{x}, 0) \,\&\, \cdots \,\&\, R(\mathbf{x}, k)$. By Lemma 1, S is in the same class as R. Substitution of k for w in the polynomial will of course not change the level of the predicate.

Therefore E_i, A_i, and BA are all closed under explicit transformation.

Proving that BA is closed under negation is more complex. Before starting the proof, we list some simple predicates. By Lemma 1, we are free to use $\vee$ or $\&$ in a BA definition. In the list below, note that predicates 1 to 10 are in E_1, while 11 and 12 are in A_2.

1. $x \leqslant y \Leftrightarrow (\exists w)_{\leqslant y} \quad x + w = y$,
2. $x < y \Leftrightarrow (\exists w)_{\leqslant y} \quad x + w + 1 = y$,
3. $x \neq y \Leftrightarrow x < y \vee y < x$
   $\Leftrightarrow (\exists w)_{\leqslant y}(\exists u)_{\leqslant x}(y = w + x + 1 \vee x = u + y + 1)$,
4. $z = x \dot{-} y \Leftrightarrow x = y + z \vee (z = 0 \,\&\, x < y)$,
5. $z = [x^{1/2}] \Leftrightarrow (\exists w)_{\leqslant x}(w + z^2 = x \,\&\, w < 2z + 1)$,
6. $x \mid y \Leftrightarrow (\exists z)_{\leqslant y}\, xz = y$,
7. $\sim(x \mid y) \Leftrightarrow (\exists w)_{\leqslant y}(\exists u)_{\leqslant x}(y = wx + u \,\&\, 0 < u \,\&\, u < x)$,
8. $\mathrm{Nonprime}(x) \Leftrightarrow x \leqslant 1 \vee (\exists u)_{\leqslant x}(\exists w)_{\leqslant x}(u \neq x \,\&\, w \neq x \,\&\, uw = x)$,
9. $\mathit{Relpr}(x, y) \Leftrightarrow x$ and y are relatively prime
   $\Leftrightarrow (\exists n)_{\leqslant x}(\exists m)_{\leqslant y}(xm - yn)^2 = 1$,

10. $d = Gcd(x, y) \Leftrightarrow d$ is the greatest common divisor of x and y
$\Leftrightarrow (\exists u)_{\leqslant x}(\exists w)_{\leqslant y}(du = x \,\&\, dw = y \,\&\, Relpr(u, w))$,
11. $\text{Prime}(x) \Leftrightarrow x > 1 \,\&\, (\forall w)_{\leqslant x}(w = 1 \vee w = x \vee \sim(w \mid x))$,
12. $\text{Pow2}(x) \Leftrightarrow x > 0 \,\&\, (\forall w)_{\leqslant x}(w = 1 \vee (2 \mid w) \vee \sim(w \mid x))$.

The proof that BA is closed under negation will be a simple corollary of the proof that the class is closed under a more general type of quantification: bounded by a polynomial with positive coefficients, instead of a single variable. First, we show that only free variables need be used as bounds in a BA definition.

LEMMA 3. *If $R(\mathbf{x}, y)$ is a* BA *predicate, then only y and the x_i need serve as bounds.*

Proof. Assume that $R(\mathbf{x}, y) \Leftrightarrow (Mw)_{\leqslant y} \cdots (Nv)_{\leqslant w} Q(\mathbf{x}, y, w, v)$, where Q is bounded arithmetic and the quantifiers M and N can each be either a bounded universal or a bounded existential. An equivalent definition for R is

$$R(\mathbf{x}, y) \Leftrightarrow (Mw)_{\leqslant y} \cdots (\exists u)_{\leqslant y}\, w \geqslant u \,\&\, Q(\mathbf{x}, y, w, u) \text{ if } N \text{ is existential;}$$
$$R(\mathbf{x}, y) \Leftrightarrow (Mw)_{\leqslant y} \cdots (\forall u)_{\leqslant y}\, w < u \vee Q(\mathbf{x}, y, w, u) \text{ if } N \text{ is universal.}$$

Note that $w \geqslant u$ is expressible as $(\exists v)_{\leqslant y}\, w = u + v$, while $w < u$ is expressible as $(\exists v)_{\leqslant y}\, u = w + v + 1$. In either case, the resulting predicate is bounded arithmetic. Repeat this procedure for each bound variable which itself occurs as a bound.

LEMMA 4. *If P, Q are polynomials with positive coefficients such that any* BA *predicate quantified with P or Q as an upper bound is still BA, then the same holds true for $P + Q$ and $P \cdot Q$.*

Proof. $(Nw)_{\leqslant P+Q} R(\mathbf{x}, w) \Leftrightarrow (Nw_1)_{\leqslant P}(Nw_2)_{\leqslant Q} R(\mathbf{x}, w_1 + w_2)$, where $R(\mathbf{x}, w_1 + w_2)$ simply means "substitute $w_1 + w_2$ for w in the polynomial part of R" (using the same idea as in Lemma 3, first ensure that w is not used as a bounding variable). Clearly, $R(\mathbf{x}, w_1 + w_2)$ is BA. By hypothesis, quantifying it up to Q is also BA; if we quantify once more up to P, the result is still BA. Thus, $P + Q$ is a valid bound.

For $P \cdot Q$, simply use $w_2 P + w_1$ as a replacement for w, with $w_2 < Q$ and $w_1 \leqslant P$. This will represent precisely those w less than or equal to $P \cdot Q$, so we have neither gained nor lost possible values.

$$(Nw)_{\leqslant P \cdot Q} R(\mathbf{x}, w) \Leftrightarrow (Nw_1)_{\leqslant P}(Nw_2)_{\leqslant Q} R(\mathbf{x}, w_2 P + w_1) \,\&\, w_2 < Q.$$

The predicate $w_2 < Q$ is easily seen to be bounded arithmetic (e.g., $(\exists v)_{\leqslant Q}\, w_2 + v + 1 = Q$).

PROPOSITION 1. *Let $P(\mathbf{x})$ be a polynomial with positive coefficients. Let $R(\mathbf{x}, w)$ be a* BA *predicate. Then so are $(\exists w)_{\leqslant P(\mathbf{x})} R(\mathbf{x}, w)$ and $(\forall w)_{\leqslant P(\mathbf{x})} R(\mathbf{x}, w)$. That is, class* BA *is closed under quantification bounded by a polynomial.*

Proof. Whenever necessary, apply Lemma 3 so that only free variables appear as bounds; the proof then follows from Lemma 4 by building up the polynomial $P(\mathbf{x})$ by repeated additions and multiplications.

THEOREM 1. *If $R(\mathbf{x})$ is a* BA *predicate, then so is $\sim R(\mathbf{x})$.*

That is, class BA is closed under negation.

Proof. Assume that $R(\mathbf{x}) \Leftrightarrow M\mathbf{y}P(\mathbf{x}, \mathbf{y}) = 0$. Then $\sim R(\mathbf{x}) \Leftrightarrow M'\mathbf{y}P(\mathbf{x}, \mathbf{y}) \neq 0$ where M' is obtained from M by changing existentials to universals, and vice versa. M' is a valid quantifier prefix, so we need only express $P \neq 0$ in a BA format.

$$P \neq 0 \Leftrightarrow P_1 \neq P_2 \qquad \text{(where } P_1, P_2 \text{ have only positive coefficients)}$$
$$\Leftrightarrow P_1 < P_2 \vee P_2 < P_1 .$$

We show that $P_1 < P_2$ is bounded arithmetic; the proof for $P_2 < P_1$ is of course analogous.

$$P_1 < P_2 \Leftrightarrow (\exists w)_{\leqslant P_2} P_1 + w + 1 = P_2 .$$

This completes the proof that BA is closed under negation, and thus closed under all Boolean operations.

Once we know that BA is closed under Boolean operations, we can show that BA is identical to two classes of predicates used by Smullyan (1961): the constructive arithmetic predicates (CA) and the rudimentary predicates (RUD). Briefly, CA is the smallest class containing the addition and multiplication predicates, and closed under Boolean operations, bounded quantification, and explicit transformation. RUD is the smallest class containing the concatenation predicate ($C(x, y, z) \Leftrightarrow$ the string x followed by the string y is identical to the string z), and closed under the same operations as CA.

COROLLARY 1. BA = CA = RUD.

Proof. Bennett (1962) showed that CA = RUD. Clearly, BA contains the initial predicates of CA, and is closed under the CA operations. Thus, CA $\subseteq$ BA. But BA $\subseteq$ CA since any polynomial predicate can be built up in CA by composition of addition and multiplication predicates.

Therefore, any predicate known to be in either of these classes is also BA. Bennett proved that $z = x^y$ is rudimentary. Using a similar technique, Finkelstein (1977) showed that the associated relations for the nth Grzegorczyk

function ($z = f_n(x, y)$), the combinatorial coefficient ($z = \binom{n}{k}$), and the number of divisor functions are all rudimentary. By Corollary 1, they are also BA (although no direct BA definitions are known). One of the very few common number theoretic predicates not known to be bounded arithmetic is $z = \Pr(n)$, z is the nth prime number.

We remark that one can prove many other closure properties of the class of bounded arithmetic predicates. For example, the class is closed under substitution of a polynomial or a polynomially bounded function whose associated relation is bounded arithmetic (e.g., $[x^{1/2}]$). A more detailed discussion of these questions can be found in Harrow (1973).

3. The Bounded Arithmetic Hierarchy

We now apply the results of the previous section to the predicate hierarchies. First, we note the following.

Lemma 5. $A_1 = Poly$.

Proof. A bounded universal quantification of a polynomial predicate is still a polynomial predicate. See Davis (1958, p. 104) for a proof.

Corollary 2. $A_{2m+1} = A_{2m}$, $E_{2m+2} = E_{2m+1}$ *for all* $m \geqslant 0$.

Since the rightmost quantifier in a BA predicate must be a bounded existential, there is really just one set of classes to consider, rather than two as in Kleene's arithmetic hierarchy.

Definition. The bounded arithmetic hierarchy consists of the classes Poly, E_1, A_2, E_3, ... Q_i ($i \geqslant 0$) will denote the ith level of this hierarchy.

We note a few closure results that can be proved about the levels of the bounded arithmetic hierarchy.

Lemma 6. *If* $z = f(\mathbf{x}, y) \in Q_m$, $y = g(\mathbf{w}) \in Q_n$, *and* $f(\mathbf{x}, y) \geqslant y$ *for all* y, *then* $z = f(\mathbf{x}, g(\mathbf{w})) \in Q_k$ *where* $k = \max(m, n) + 1$.

Proof. $z = f(\mathbf{x}, g(\mathbf{w})) \Leftrightarrow (\exists y)_{\leqslant z}\, z = f(\mathbf{x}, y) \,\&\, y = g(\mathbf{w})$. By the assumption on f, z can serve as a bound for y. By interweaving quantifiers, the prefixes for $z = f(\mathbf{x}, y)$ and $y = g(\mathbf{w})$ can be combined in parallel. The resulting predicate will be at the maximum of the levels of the two original predicates, plus 1 if the larger did not already have a leftmost existential.

Corollary 3. *Each* Q_m ($m > 0$) *is closed under quantification* (*of the appropriate type*) *bounded by a polynomial.*

Proof. By Proposition 1, BA is closed under quantification bounded by a

polynomial. A check of Lemmas 3 and 4 shows that the only new quantifiers introduced are existential, which can be moved to the right (since we can interweave with the original predicate). But then they can be absorbed by the original rightmost level of quantification, which had to be existential by Lemma 5.

BA is not closed under a more general type of quantification, e.g.,a bound of the form $(\exists w)_{\leqslant x^y}$. Let f_n be the nth Grzegorczyk function ($f_0(x, y) = x + 1$; $f_1(x, y) = x + y$; $f_2(x, y) = (x + 1) \cdot (y + 1)$; f_3 is of roughly exponential growth, and so on). Let $\mathscr{E}^n$ be the nth Grzegorczyk class, and let $(\mathscr{E}^n)_*$ be the 0–1 functions of $\mathscr{E}^n$. See Grzegorczyk (1953) for the explicit definitions.

DEFINITION. For each $n \geqslant 0$, $\mathrm{BA}(f_n)_*$ is the smallest class containing BA and closed under quantification with f_n as an upper bound.

Restating the previous results, we have:

COROLLARY 4. $\mathrm{BA} = \mathrm{BA}(f_0)_* = \mathrm{BA}(f_1)_* = \mathrm{BA}(f_2)_*$.

But this does not extend past f_2.

LEMMA 7. $\mathrm{BA} \subseteq (\mathscr{E}^0)_*$.

Proof. $(\mathscr{E}^0)_*$ contains the addition and multiplication predicates, and is closed under the operations used to define *BA*. Thus, $BA \subseteq (\mathscr{E}^0)_*$. It is not known if this inclusion is strict.

PROPOSITION 2. $\mathrm{BA} \subsetneq \mathrm{BA}(f_3)_*$.

Proof. For $n \geqslant 3$, Harrow (1973) showed that $\mathrm{BA}(f_n)_* = (\mathscr{E}^n)_*$. Grzegorczyk proved that $(\mathscr{E}^0)_* \subsetneq (\mathscr{E}^3)_*$. Thus,

$$\mathrm{BA} \subseteq (\mathscr{E}^0)_* \subsetneq (\mathscr{E}^3)_* = \mathrm{BA}(f_3)_*.$$

We now study the question of whether the BA hierarchy is indeed a strict hierarchy, i.e., if for each $m \geqslant 0$, there is a predicate in $Q_{m+1} \backslash Q_m$.

COROLLARY 5. *If* $R(\mathbf{x}) \in Q_m$ $(m \geqslant 0)$, *then* $\sim R(\mathbf{x}) \in Q_{m+1}$.

Proof. This is a corollary of the proof that BA is closed under negation (Theorem 1). The only new quantifiers introduced are existential, which will appear on the right; every other quantifier flips (universal to existential and vice versa). If the original prefix is E_m, then the negated prefix will become A_m, and then A_{m+1} because of the rightmost existential. If the original is A_m, the negated prefix will be E_m, and then E_{m+1}.

The predicate obtained by the negation process outlined above need not be

the most efficient representation of the complement. For example, $x \mid y \Leftrightarrow (\exists w)_{\leqslant y}\, y = wx$.

$$\sim(x \mid y) \Leftrightarrow (\forall w)_{\leqslant y}(\exists q)_{\leqslant wx}(\exists u)_{\leqslant y}(y = wx + u + 1) \vee (wx = y + q + 1).$$

Directly,

$$\sim(x \mid y) \Leftrightarrow (\exists w)_{\leqslant y}(\exists u)_{\leqslant x}(y = wx + u \,\&\, 0 < u \,\&\, u < x),$$

which shows that the predicate and its complement are both E_1. Thus, the negation technique provides only an upper bound on the complexity of the complement. No predicate lies more than one level of quantification from its complement, but they can in fact be at the same level of complexity.

PROPOSITION 3. *If for some* $m \geqslant 0$ $Q_m = Q_{m+1}$, *then* Q_m *and all higher levels of the hierarchy are equal to* BA.

Proof. Assume without loss of generality that Q_m is an existential class. Then Q_m is closed under bounded existential quantification, while Q_{m+1} is closed under bounded universal quantification. By hypothesis, Q_m is also closed under bounded universal quantification. Obviously, Q_m includes Poly. But then $\mathrm{BA} \subseteq Q_m$, by the definition of BA as the smallest class containing Poly and closed under both bounded existential and bounded universal quantification. Therefore, $Q_m = Q_{m+1} = Q_{m+2} = \cdots = \mathrm{BA}$, and the hierarchy collapses.

DEFINITION. For $i \geqslant 0$, let $D_i = \{R \mid R, \sim R \in Q_i\}$. D_i consists of those predicates in Q_i which are as difficult to express as their complements. Note that D_i is closed under negation. By Corollary 5, $Q_i \subseteq D_{i+1} \subseteq Q_{i+1}$ for all $i \geqslant 0$.

PROPOSITION 4. *If* $D_{i+1} = Q_{i+1}$ *or if* $D_{i+1} = Q_i$ *for any* $i \geqslant 0$, *then the hierarchy collapses.*

Proof. By hypothesis, either Q_{i+1} or Q_i is closed under negation. But if any Q_j is closed under negation, then it is closed under both bounded universal and bounded existential quantification, since $(\exists w)_{\leqslant w_i} R(\mathbf{x}, w) \Leftrightarrow \sim(\forall w)_{\leqslant x}(\sim R(\mathbf{x}, w))$ and vice versa. This implies that $Q_j = \mathrm{BA}$, and thus the hierarchy collapses.

Using the previous results, we can give necessary and sufficient conditions for the hierarchy to be strict.

THEOREM 2. *The following are all equivalent*:

(1) *The bounded arithmetic hierarchy is strict.*

(2) *No* Q_j $(j \geqslant 0)$ *is closed under negation.*

(3) *No* Q_j $(j \geqslant 0)$ *is closed under both bounded universal and bounded existential quantification.*

(4) *No* D_j $(j \geqslant 1)$ *is closed under either bounded universal or bounded existential quantification.*

(5) *For each* $j \geqslant 0$, $Q_j \subsetneq D_{j+1} \subsetneq Q_{j+1}$.

Proof. We show that (1) and (2) are equivalent; it is easy to see that (3), (4), and (5) are equivalent to (2).

If no Q_j is closed under negation, then for any $j \geqslant 0$, there is a predicate R in Q_j such that $\sim R$ is not in Q_j. By Corollary 5, $\sim R$ is in Q_{j+1}. Thus, $\sim R \in Q_{j+1} \backslash Q_j$, which means that the hierarchy is strict.

If some Q_j is closed under negation, then the proof of Proposition 4 shows that the hierarchy collapses.

We note one last lifting lemma.

LEMMA 8. *If* $D_i \subsetneq Q_i$, *then* $Q_i \subsetneq D_{i+1}$.

Proof. If $D_i \subsetneq Q_i$, then Q_i is not closed under negation (since D_i is the part of Q_i which is closed under negation), and therefore cannot be equal to D_{i+1}.

4. Discussion and Open Problems

So far, we have not shown that the BA hierarchy is strict, or in fact that $Q_m \subsetneq Q_{m+1}$ for any m. We have only partial results in this direction.

Poly $(=Q_0)$ is a trivial class of predicates. The unary polynomial predicates are precisely the finite sets and the entire set of integers, since every polynomial in one variable has either a finite number of roots or else vanishes identically. For $n \geqslant 2$, n-ary polynomial predicates are also essentially trivial. For example, it is easy to see that $x \leqslant y$ is in E_1 but not in Poly. The only n-ary predicates in D_0 (i.e., those $R(\mathbf{x})$ such that both R and $\sim R$ are polynomial predicates) will be the always-false predicate (e.g., $R(\mathbf{x}) \Leftrightarrow x_i - x_i + 1 = 0$) and the always-true predicate $(R(\mathbf{x}) \Leftrightarrow x_i - x_i = 0)$.

However, E_1 does contain several fairly complex predicates (see predicates 1 to 10 in the list in Section 2). We still do not know if $E_1 \subsetneq A_2$ (which by the results of Section 3 is equivalent to asking if $E_1 \subsetneq$ BA). One possible line of research is to study the growth rates of solutions of Diophantine and "almost"-Diophantine equations (i.e., a Diophantine equation in which one of the variables bounds the others). $Pow2(x)$ and $Prime(x)$ seem to be likely candidates for this approach. For n-ary predicates and larger classes Q_i, the problems become more complex, and soon begin to face gaps in our knowledge about the solutions of algebraic equations.

An interesting point: If there is an n-ary $(n \geqslant 2)$ predicate $R(\mathbf{x})$ in some

class Q_{i+1} but not in Q_i, then we can show that there is also a unary predicate $S(z)$ in $Q_{i+1}\backslash Q_i$.

We use the standard pairing functions, $J(x, y)$, $K(z)$, $L(z)$ of Davis (1958):

$$z = J(x, y) \Leftrightarrow 2z = (x + y) \cdot (x + y + 1) + 2x,$$
$$x = K(z) \Leftrightarrow (\exists y)_{\leqslant z} \quad z = J(x, y), \qquad y = L(z) \Leftrightarrow (\exists x)_{\leqslant z} \quad z = J(x, y).$$

Thus, $z = J(x, y)$ is a polynomial predicate, while the other two are in E_1.

LEMMA 9. (a) *If* $S(z) \in Q_n$ $(n > 0)$, *then the predicate* $R(x, y)$ *defined by* $R(x, y) \Leftrightarrow S(J(x, y))$ *is also in* Q_n.

(b) *If* $R(x, y) \in Q_n$ $(n > 0)$, *then the predicate* $S(z)$ *defined by* $S(z) \Leftrightarrow R(K(z), L(z))$ *is also in* Q_n.

Proof. (a) Assume that Q_n is an existential class. Then $R(x, y) \Leftrightarrow (\exists z)_{\leqslant P(x,y)} z = J(x, y)$ & $S(z)$, where P is any polynomial that dominates J. Clearly, R is in the same class as S. If Q_n is a universal class, then $R(x, y) \Leftrightarrow (\forall z)_{\leqslant P(x,y)} z \neq J(x, y) \vee S(z)$. Since $z \neq J(x, y)$ is in E_1, R is in Q_n regardless.

(b) A similar idea works here, namely either:

$$S(z) \Leftrightarrow (\exists x)_{\leqslant z} (\exists y)_{\leqslant z} \quad z = J(x, y) \,\&\, R(x, y)$$

or

$$S(z) \Leftrightarrow (\forall x)_{\leqslant z} (\forall y)_{\leqslant z} \quad z \neq J(x, y) \vee R(x, y).$$

PROPOSITION 5. *If there is a predicate* $R(x, y)$ *in* Q_{m+1} *but not in* Q_m $(m > 0)$, *then there is a unary predicate* $S(z)$ *in* $Q_{m+1}\backslash Q_m$.

Proof. Define $S(z) \Leftrightarrow R(K(z), L(z))$. Then by Lemma 9, part (b), since $R \in Q_{m+1}$, we also know that $S \in Q_{m+1}$. But S is not in Q_m; otherwise $R(x, y)$ would be in Q_m by Lemma 9, part (a). Therefore, $S(z) \in Q_{m+1}\backslash Q_m$.

Trivially, if there is a unary predicate $S(z)$ in $Q_{m+1}\backslash Q_m$, then there is also a binary predicate $R(x, y)$ satisfying these conditions. Just define $R(x, y) \Leftrightarrow S(x)$. Lemma 9 and Proposition 5 can be extended from binary to n-ary $(n > 2)$ predicates.

Wrathall (1975) and Finkelstein (1977) have studied formal language theoretic properties of RUD, e.g., closure under AFL operations. Since RUD is identical to BA, these are of interest to a discussion of the bounded arithmetic hierarchy. Let us look more closely at the relationship of BA and RUD to some well-known classes from formal language theory.

Jones (1968, 1969) has extensively studied these properties of the class of rudimentary predicates. Among other things, he showed that the class of context-free languages (CFL) is strictly contained in RUD. By looking at

elementary closure properties of RUD, Wrathall proved that the class of quasi-realtime languages (Q) discussed by Book and Greibach (1970) is a subset of RUD, although it is not known if this containment is proper. Myhill (1960), in his paper defining what we now call a deterministic linear bounded automaton, showed that RUD is a subset of the class of sets (DLBA) accepted by these machines. Ritchie (1963) proved that DLBA $= (\mathscr{E}^2)_*$. Clearly, DLBA is a subset of the class of sets accepted by nondeterministic linear bounded automata, which is identical to the class of context-sensitive languages (CSL) (see Hopcroft and Ullman (1969) for a proof). The famous LBA conjecture asks if this containment is proper.

Harrow (1973) showed that CSL is strictly contained in $\mathrm{BA}(f_3)_* = (\mathscr{E}^3)_*$; Harrow (1975) showed that BA is also identical to the 0–1 functions of several Grzegorczyk-like classes defined from f_2 by limited minimum rather than limited recursion. To summarize, it is known that

$$\begin{array}{ccccccc} \mathrm{BA} & \subseteq & (\mathscr{E}^0)_* & \subseteq & (\mathscr{E}^1)_* & \subseteq & (\mathscr{E}^2)_* \subsetneq (\mathscr{E}^3)_* \\ = & & & & & & \\ \mathrm{CA} & & & = & & & \subsetneq \\ = & & & & & & \\ \mathrm{CFL} \subsetneq Q \subseteq \mathrm{RUD} & \subseteq & & & \mathrm{DLBA} & \subseteq & \mathrm{CSL} \end{array}$$

It is not known if any of the inclusions not shown to be strict are in fact strict. For example, the question of whether nondeterministic linear time (Q) is strictly weaker than nondeterministic linear space (CSL) is still open.

Note that BA and RUD could play pivotal roles in solving many open questions. Proving that BA $= (\mathscr{E}^2)_*$ would show that the three smallest Grzegorczyk classes of relations are identical. Similarly, if RUD $\subsetneq$ CSL, then $Q \subsetneq$ CSL also. Thus, in studying BA and the bounded arithmetic hierarchy we can hope to shed light on many unsolved problems, and also to open up new areas of research in formal language theory.

The bounded arithmetic hierarchy is intended as an analog to Kleene's arithmetic hierarchy. Let us see how closely this new hierarchy mirrors the old. (See Rogers (1967) for properties of the arithmetic hierarchy.) First we look at some differences. For each class Σ_n or Π_n in the arithmetic hierarchy, there is a predicate in that class but not in the corresponding class in the other part of the hierarchy. In addition, for $n > 0$ we have $\Sigma_n \cup \Pi_n \subsetneq \Delta_{n+1} = \Sigma_{n+1} \cap \Pi_{n+1}$ (for $n = 0$, there is equality). In the bounded arithmetic hierarchy, there is only one set of classes Q_i, and we do not know if $Q_n \subsetneq Q_{n+1}$ for any $n > 0$. D_n (the bounded arithmetic analog of Δ_n) is not defined to be $E_n \cap A_n$, since that would degenerate to either E_{n-1} or A_{n-1}, depending upon which class collapsed to the one below (Corollary 2). But there are also similarities between the bounded arithmetic and the arithmetic hierarchies. Each class (Q_i, Σ_i,

or Π_i) is closed under conjunction and disjunction; if R is a member of any class, then $\sim R$ belongs to the appropriate next class. The reader should be able to note other similarities and differences.

There have been previous generalizations of the arithmetic hierarchy to subrecursive predicate classes. We mention two that are most relevant to the present study.

Meyer and Stockmeyer (1972) and Stockmeyer (1975) define a polynomial-time hierarchy. The class of predicates recognizable in polynomial time on a deterministic Turing machine takes the place of the class of recursive predicates in Kleene's arithmetic hierarchy (or Poly in the bounded arithmetic hierarchy). The levels of the polynomial-time hierarchy are defined in terms of sets recognizable using oracle Turing machines (the arithmetic hierarchy can be defined in an analogous way). Meyer and Stockmeyer were unable to prove the existence of a strict hierarchy, but they did use these classes to discuss the P = NP problem. Stockmeyer also gives a syntactic characterization of the polynomial-time hierarchy in terms of classes defined by a series of alternating polynomial bounded quantifiers, similar to our definition of the bounded arithmetic hierarchy. See also Adleman and Manders (1976).

Wrathall (1975) defines a linear hierarchy (starting from just the empty set) within RUD, again using oracle Turing machines, operating within linear time. She also gives a characterization of the levels of the hierarchy based on alternations of bounded quantifiers. Wrathall does not prove that the linear hierarchy is strict, but she does show that if the hierarchy is strict, then RUD is a proper subset of DLBA (we prove a similar result for the bounded arithmetic hierarchy below). She also relates the linear and polynomial-time hierarchies, showing that sets in the polynomial-time hierarchy have "padded" representations in the linear hierarchy (the padding enables them to be recognized in linear rather than polynomial time).

Both Wrathall and Stockmeyer discuss the important concept of reducibility between sets and the question of "complete" sets (see Aho *et al.* (1974) for a definition of these terms). Recently, Jones (1975) and others have used RUD in classifying reducibilities among combinatorial problems, including many known to be complete for NP. We are attempting to extend these notions to the bounded arithmetic predicates. Here is an example of a result of this type, a theorem analogous to Wrathall's on the relationship between a strict hierarchy and proper inclusion in DLBA:

THEOREM 3. *If* BA = DLBA, *then the* BA *hierarchy collapses.*

Proof. Let $M_0, M_1, \ldots$ be a suitable numbering of all deterministic Turing machines on some fixed alphabet, e.g., {0, 1, 2}. (The particular details of such a numbering can be found in Wrathall (1975) but they are unimportant for our purposes.)

Define the following 3-ary predicate R: $R(x, i, c) \Leftrightarrow (\exists z)_{\leqslant x}\ x = z^c$ and "M_i accepts input z while using an amount of space bounded by the length of x."

Observe these two facts about R:

(1) R is in DLBA (see, for example, Wrathall, 1975).

(2) if S is a unary predicate in DLBA, then there exist fixed i, c such that $S(z) \Leftrightarrow R(z^c, i, c)$.

These two conditions imply that R is a "complete" predicate for DLBA.

Assume now that BA = DLBA. Then R is in BA and thus R is in Q_m for some m.

CLAIM. *$Q_m = Q_{m+1}$ and the hierarchy collapses.*

Proof of claim. Pick any unary $S(z)$ in Q_{m+1}. S is in DLBA and therefore by condition (2), $S(z) \Leftrightarrow R(z^c, i, c)$ for some fixed i, c.

But this means that S is obtained from R by substituting constants or fixed powers of z (such as z^3 or z^7 depending upon c). By Lemma 2 and Corollary 3, S is in the same class of the hierarchy as R, implying that S is in Q_m. Thus the unary predicates in Q_{m+1} are in Q_m as well.

By Proposition 5 and its extension, there cannot exist an n-ary ($n \geqslant 2$) predicate in $Q_{m+1} \backslash Q_m$. Therefore, $Q_m = Q_{m+1}$ and the hierarchy collapses.

It follows immediately that if the BA hierarchy is strict, then BA $\subsetneq$ DLBA. Of course, it is also possible that BA is strictly contained in DLBA and the hierarchy still collapses.

There are some other interesting results concerning the class of rudimentary predicates. Finkelstein has shown that for each $m > 0$, there is a rudimentary predicate expressible using $m + 1$ quantifiers but not m. Note that he is counting the number of quantifiers, not the number of alternations; Finkelstein is also starting from the concatenation predicate rather than a polynomial predicate, so his results are not directly applicable to the bounded arithmetic hierarchy. Independently, Nepomnyashchii (1970a, b) and Finkelstein have shown that any set recognizable on a nondeterministic Turing machine in space $n^{1-\beta}$ and polynomial time (where $\beta > 0$, and n is the length of the input) is rudimentary. Can we extend this to a complete machine characterization of BA and RUD? For example, can we find necessary and sufficient conditions on the time or tape used by a machine to ensure that it recognizes only rudimentary predicates? Is there some machine class, possibly with an appropriate definition of an oracle, corresponding to the levels of the bounded arithmetic hierarchy?

We conclude by listing four of the major avenues for further research in this field.

1. The question of a strict bounded arithmetic (or linear or polynomial-time) hierarchy.

2. A comparison of the levels of the various hierarchies discussed (assuming that the hierarchies are strict).

3. The relation of BA and RUD to the well-known formal language theory classes.

4. A machine characterization of RUD, or the levels of the bounded arithmetic hierarchy, including a definition of BA reducibility and BA complete set.

Acknowledgments

The author would like to thank Martin Davis for originally suggesting the question of a bounded arithmetic hierarchy, Sheldon Finkelstein for his careful reading of the manuscript and simplification of many of the proofs, and the referee for suggesting Theorem 3 and several other improvements.

Received: April 16, 1976; revised: February 4, 1977

References

Adleman, L., and Manders, K. (1976), Number theoretic aspects of computational complexity, unpublished manuscript.

Aho, A., Hopcroft, J., and Ullman, J. (1974), "The Design and Analysis of Computer Algorithms," Addison–Wesley, Reading, Mass.

Bennett, J. H. (1962), "On Spectra," Ph.D. Dissertation, Princeton University.

Blum, M. (1967), A machine-independent theory of the complexity of recursive functions, *J. Assoc. Comput. Mach.* **14**, 322–336.

Book, R., and Greibach, S. (1970), Quasi-realtime languages, *Math. Systems Theory* **4**, 97–111.

Davis, M. (1958), "Computability and Unsolvability," McGraw–Hill, New York.

Finkelstein, S. (1977), On rudimentary languages, unpublished manuscript.

Grzegorczyk, A. (1953), Some classes of recursive functions, *Rozprawy Mat.* **4**.

Harrow, K. (1973), "Sub-elementary Classes of Functions and Relations," Ph.D. Dissertation, New York University.

Harrow, K. (1975), Small Grzegorczyk classes and limited minimum, *Z. Math. Logik Grundlagen Math.* **11**, 417–426.

Hopcroft, J., and Ullman, J. (1969), "Formal Languages and Their Relation to Automata," Addison–Wesley, Reading, Mass.

Jones, N. (1968), Classes of automata and transitive closure, *Inform. Contr.* **13**, 207–229.

Jones, N. (1969), Context-free languages and rudimentary attributes, *Math. Systems Theory* **3**, 102–109.

Jones, N. (1975), Space-bounded reducibility among combinatorial problems, *J. Comput. System Sci.* **11**, 68–85.

Meyer, A., and Stockmeyer, L. (1972), The equivalence problem for regular expressions with squaring requires exponential space, *in* "Proceedings of the IEEE Thirteenth Annual Symposium on Switching and Automata Theory," pp. 125–129.

Myhill, J. (1960), "Linear Bounded Automata," WADD Technical Note 60–165.

NEPOMNYASCHII, V. (1970*a*), Rudimentary interpretation of two-tape Turing computation, *Kibernetika* **2**, 43–50 (English transl.).

NEPOMNYASCHII, V. (1970*b*), Rudimentary predicates and Turing computations, *Soviet Math. Dokl.* **11**, 1462–1465.

RITCHIE, R. (1963), Classes of predictably computable functions, *Trans. Amer. Math. Soc.* **106**, 139–173.

ROGERS, H. (1967), "Theory of Recursive Functions and Effective Computability," McGraw–Hill, New York.

SMULLYAN, R. (1961), "Theory of Formal Systems," Princeton Univ. Press, Princeton, N.J.

STOCKMEYER, L. (1975), "The Polynomial-Time Hierarchy," IBM Research Report RC 5379.

YOUNG, P. (1969), Towards a theory of enumerations, *J. Assoc. Comput. Mach.* **16**, 328–348.

WRATHALL, C. (1975), "Subrecursive Predicates and Automata," Yale University Research Report No. 56.