Bounded Arithmetic, Proof Complexity and Two Papers of Parikh

Samuel R. Buss [1]

Department of Mathematics, University of California, San Diego, La Jolla, CA 92093-0112, USA

Abstract

This article surveys R. Parikh's work on feasibility, bounded arithmetic and the complexity of proofs. We discuss in depth two of Parikh's papers on these subjects and some of the subsequent progress in the areas of feasible arithmetic and lengths of proofs.

1 Introduction

This article discusses two papers of Rohit Parikh on feasibility and bounded arithmetic and on the complexity of proofs: the first is the 1971 paper "Existence and Feasibility in Arithmetic" [30] and the second is the 1973 paper "Some Results on Length of Proofs" [31]. Both papers were seminal and influential and led to large research areas which are still active and fruitful 25 years later. The first paper addressed the intuitive concept of feasibility, discussed the infeasibility of exponentiation, and presented the original definition of bounded arithmetic ($I\Delta_0$). The second paper solved a special case of a conjecture of Kreisel's and additional problems in proof speed-up, and introduced important tools for the analysis of the complexity of proofs in first-order logic and other formal systems.

We will discuss first the "feasibility" paper, in section 2. Section 3 takes up the "length of proof" paper, and section 4 concludes with a discussion of the connections between these topics.

[1] Supported in part by NSF grant DMS-9503247 and by a cooperative research grant INT-9600919/ME-103 from the NSF (USA) and the MŠMT (Czech republic)

2 Existence and Feasibility in Arithmetic

The most important aspect of the 1971 "feasibility" paper was arguably the introduction of the first-order theory of bounded arithmetic, now referred to as $I\Delta_0$. This paper starts by considering the issue of whether a fast-growing function such as exponentiation gives rise to numbers which are intuitively infeasible or non-constructible. Of course, from the point of view of classical proof theory, exponentiation is a rather slow-growing function — it is, of course, a primitive recursive function and finitists and intuitionists certainly accept all primitive recursive functions. On the other hand, there are a number of reasons to doubt the feasibility of exponentiation, and a number of logicians and philosophers have doubted the concrete existence of large numbers such as $67^{257^{729}}$ formed with exponential terms. (See Yessenin-Volpin [45] and the later works building on Parikh's bounded arithmetic of E. Nelson [26] and Sazanov [38,39]). And for computer scientists, the computational infeasibility of exponentiation was already well-recognized. Parikh endorsed this infeasibility of exponentiation, saying

> ... there is a large element of phantasy in conventional mathematics which one may accept if one finds it pleasant, but which one could equally sensibly (perhaps more sensibly) reject.

and the principle themes of the paper were to give justifications for viewing exponentiation as infeasible in formal theories of arithmetic and to give an alternative formalization of arithmetic (i.e., bounded arithmetic) which would be closer to feasible.

There are four aspects of the "feasibility" paper that we shall discuss: (i) the suggestion that bounded formulas and linear space computations are feasible, whereas exponentiation is not, (ii) the definition of bounded arithmetic, (iii) the "Parikh theorem" for bounded arithmetic, and (iv) the extent to which exponentiation is needed for the arithmetization of metamathematics.

(i) In section 3 of the "feasibility" paper, Parikh presents a model-theoretic result illustrating the gap between exponentiation and the feasible operations of addition and multiplication in models of arithmetic. Specifically, consider the axioms

$$f(x,0) = 1 \qquad\qquad f(x, y+z) = f(x,y)\cdot f(x,z)$$

$$f(x,S(y)) = x\cdot f(x,y) \qquad f(x, y\cdot z) = f(f(x,y),z)$$

which uniquely characterize $f(x,y)$ as the exponentiation function x^y in the standard integers. Parikh proved, however, that there is a non-standard model M of $Th(\mathbb{N})$ and two distinct functions f_1 and f_2 on M both of which satisfy the above four axioms for all values $x, y, z \in M$.

In the next section, Parikh proposes an "anthropomorphic" system based on classes of predicates and functions which are more feasible than exponentiation. His first class of predicates is the predicates which can be defined by *bounded formulas* of the form

$$(Q_1 x_1 < t_1)(Q_2 x_2 < t_2) \cdots (Q_n x_n < t_n) B(x_1, \ldots, x_n, y_1, \ldots, y_m)$$

where the terms t_i involve only the variables $y_1, \ldots, y_m$ and where B is a quantifier-free formula. Nowadays, the notation Δ_0 is commonly used to denote either the set of bounded formulas or the set of predicates definable by bounded formulas; it is also now known that this set of predicates is precisely the set of predicates recognized by constant alternation, linear time Turing machines, which are called the *linear-time hierarchy* predicates.

The linear time hierarchy was first introduced by Smullyan [41] in the guise of "rudimentary predicates". Shortly thereafter, Bennett [2] proved that Smullyan's rudimentary predicates are precisely the predicates that can be defined with a bounded formula over the integers (using a m-adic representation of integers). During the later 1960's, a number of people investigated the relationship between the rudimentary predicates and computational complexity classes (see, e.g., Jones [19]). In later work, subsequent to Parikh's "feasibility" paper, the rudimentary predicates were studied extensively by Wrathall [44], Harrow [17], Nepomnjaščiĭ [27], and Wilkie [42]; both Wrathall and Wilkie essentially proved that this class was equal to the linear time hierarchy, but Lipton [23] was the first to explicitly prove the fact that the set of Δ_0 predicates is equal to the linear time hierarchy.

Parikh was definitely interested in connections between computational complexity and the Δ_0-formulas, but instead of discussing the linear-time hierarchy (since, in any event, the linear-time hierarchy was not yet defined, much less known to be related to the Δ_0-hierarchy), he turned to the class of predicates which are recognized by deterministic linear bounded automata (dlba's), or in modern day terminology, to the class of predicates which can be computed in linear space by a deterministic Turing machine. Parikh called the linear space predicates "concrete": he noted the theorem of Myhill that every Δ_0-formula defines a concrete predicate and also noted that the converse inclusion was open (and this is still open today!).

From a present-day computer science viewpoint, the suggestion that concrete (linear-space) predicates are "feasible" seems odd, since it is commonly conjectured that some of these predicates require exponential time to compute. Likewise, the class of Δ_0-predicates, which equals the linear-time hierarchy, is conjectured to contain predicates which require exponential time to compute; for instance, the NP-complete predicate SAT is in the linear time hierarchy and is commonly conjectured to be infeasible. Parikh does not offer any strong reasons in the "feasibility" paper for why the concrete or the Δ_0-predicates should be considered feasible; however, there are at least two reasons that might support taking them to be feasible: firstly, they do not involve the use of any exponentially large numbers, and secondly, the Δ_0-predicates at least do not seem to be vulnerable to the kind of model-theoretic separation from addition and multiplication that was shown to hold for exponentiation.

(ii) The most important contribution of the "feasibility" paper was probably the definition of the theory of bounded arithmetic, denoted PB in that paper, but now usually denoted $I\Delta_0$. (We'll use the modern notation in this paper.)

Definition 1 (Parikh [30]) $I\Delta_0$ *(or PB) is the first-order theory with non-logical symbols* 0*,* S*,* $+$ *and* $\cdot$ *and containing the axioms*

(1)	$0 \neq S(x)$	*(5)*	$x + S(y) = S(x + y)$
(2)	$S(x) = S(y) \rightarrow x = y$	*(6)*	$x \cdot 0 = 0$
(3)	$x = 0 \vee \exists y(x = S(y))$	*(7)*	$x \cdot S(y) = x \cdot y + x$
(4)	$x + 0 = x$		
(8_n)	$A(0) \wedge \forall x(A(x) \rightarrow A(S(x))) \rightarrow \forall x A(x)$		

where A may be any bounded formula, possibly involving free variables other than x.

The syntactic details of defining bounded quantifiers in the language of $I\Delta_0$, which does not contain a $<$ relation symbol, were not discussed in the "feasibility" paper, but presumably "$x < y$" was intended to abbreviate the formula "$\exists z(x + S(z) = y)$."

(iii) The primary evidence presented in the "feasibility" paper for the feasibility of bounded arithmetic is the nondefinability of functions of superlinear growth rate. Parikh proves this in two parts. First, he shows that there is no formula, bounded or otherwise, which defines exponentiation as a provably total function of $I\Delta_0$; more precisely,

Theorem 2 (Parikh [30]) *There is no formula $A(x, y, z)$ such that $I\Delta_0$ proves*

(1) $\forall x \forall y \exists ! z A(x, y, z)$.
(2) $\forall x \forall y \forall z (A(x, 0, 1) \wedge (A(x, y, z) \rightarrow A(x, S(y), z \cdot x)))$.
(3) $\forall x \forall y \forall z (A(x, y, z) \rightarrow z \neq 0)$. [2]

The theorem was proven by a proof-by-contradiction using a model-theoretic argument; namely, let N be a nonstandard model of Peano arithmetic and α be an infinite integer in N. Define M to the initial segment of N containing the integers of N which are less than α^n for some standard $n \in \mathbb{N}$. Then, by the absoluteness of Δ_0-formulas, $M \vDash I\Delta_0$. Consider the value b such that $M \vDash A(\alpha, \alpha, b)$. From the construction of M, $b < \alpha^k$ for some $k \in \mathbb{N}$. Let c be the value such that $M \vDash A(\alpha, \alpha - k, c)$. Then, in M, $c \neq 0$ and $c \cdot \alpha^k = b < \alpha^k$, which is a contradiction. q.e.d.

The second result concerning the nondefinability of superlinear growth rate functions is the theorem commonly called "Parikh's theorem" in bounded arithmetic. It introduced the very important idea of Δ_0-definable functions:

Theorem 3 (Parikh [30]) *If $A(x, y)$ is a bounded formula with no additional free variables, and if $I\Delta_0$ can prove $\forall x \exists y A(x, y)$, then for some $k, \ell > 0$,*

$$I\Delta_0 \vdash \forall x \exists y (y < x^k + \ell \wedge A(x, y))$$

Dimitracopoulos later noted that the model-theoretic proof of the previous theorem can readily be modified to yield a proof of this theorem; however, Parikh presented a proof-theoretic proof based on Herbrand's theorem. He first noted that by adding Skolem functions

$$f_{A,t}(x, \vec{y}) \quad := \quad \mu x (A(x, \vec{y}) \vee x = t)$$

for all Δ_0-formulas A and all terms $t = t(\vec{y})$, one obtains a conservative extension PB' of $I\Delta_0$ such that PB' is axiomatizable by universal formulas and such that any Δ_0-formula is PB'-provably equivalent to a quantifier-free formula. By Herbrand's theorem and by closure of the set of PB' function

[2] Parikh omitted the condition (3) from the statement of the theorem and unfortunately the theorem is false without this assumption, since one may take $A(x, y, z)$ to be the formula

$$exp(x, y, z) \vee (z = 0 \wedge \neg \exists u (exp(x, y, u))).$$

where $exp(x, y, z)$ is a formula defining the graph of exponentiation.

symbols under definition by cases, it follows the $I\Delta_0$-provability of $\forall x \exists y A(x,y)$ implies that there is a PB'-term t such that

$$PB' \vdash \forall x A(x, t(x)).$$

Since any PB'-term is provably bounded by a polynomial, and since PB' is a conservative extension of $I\Delta_0$, the theorem follows.

(iv) Parikh discusses as an open question the issue of whether the exponentiation function is required for the arithmetization of metamathematics necessary for the Gödel incompleteness theorems. Rephrasing his arguments slightly, he notes that if a formula $A(x)$ has m symbols and a term t has n symbols, then the formula $A(t)$ may have number of symbols proportional to $m \cdot n$. Hence if an efficient Gödel numbering is used, A will have Gödel number $\ulcorner A \urcorner \approx 2^{O(m)}$ and t will have Gödel number $\ulcorner t \urcorner \approx 2^{O(n)}$, and therefore $A(t)$ will have Gödel number

$$\ulcorner A(t) \urcorner \approx 2^{O(n \cdot m)} \approx (\ulcorner A \urcorner)^{O(n)} \approx (\ulcorner A \urcorner)^{O(\log(\ulcorner t \urcorner))}.$$

The value of $\ulcorner A(t) \urcorner$ cannot be bounded by a polynomial of $\ulcorner A \urcorner$ and $\ulcorner t \urcorner$ and, as corollary to the previous two theorems, $I\Delta_0$ cannot prove that $\ulcorner A(t) \urcorner$ exists from only the assumption that $\ulcorner A \urcorner$ and $\ulcorner t \urcorner$ exist. Since the arithmetization of substitution is an important component of the arithmetization of metamathematics, this gives an indication that the usual "intensional" Gödel incompleteness theorems cannot be carried out in the theory $I\Delta_0$.

Subsequent developments showed however that full exponentiation is not needed for the arithmetization of metamathematics. Already, Parikh's size analysis shows that only the function $(x,y) \mapsto x^{\log y}$ is needed. Wilkie and Paris [43] first noted this and considered the theory $I\Delta_0$ augmented with an axiom Ω_1 stating $\forall x, y \exists z (z = x^{\lfloor \log y \rfloor})$: they showed that all the usual arithmetization of metamathematics can be carried out in $I\Delta_0 + \Omega_1$. Later, Nelson [26] and Buss [3] used a function $\#$ of similar growth rate: $x\#y := 2^{|x| \cdot |y|}$ where $|x| \approx \log_2 x$. The growth rates of Ω_1 and the $\#$ function are generally felt to be much closer to multiplication than to exponentiation, and indeed, $x^{\log y}$ and $x\#y$ are generally viewed as feasible.

Even without the axiom Ω_1 or the function $\#$, the arithmetization of metamathematics and the intensional treatment of Gödel's incompleteness theorems can be (mostly) carried out in $I\Delta_0$. Namely, Solovay, in an unpublished 1976 letter to Hájek, showed that $I\Delta_0$ can prove the existence of a cut or initial segment of the integers which is closed under $\#$. $I\Delta_0$ can thus carry out the arithmetization of metamathematics relativized to this initial segment, and this is sufficient for most applications of Gödel incompleteness theorems. (See

Pudlak [35] and Nelson [26] for the development of this.)

However, it is still open whether $I\Delta_0$ can formalize the metamathematics needed for the Gödel incompleteness theorems without the use of relativization to initial segments. It turns out that the substitution function is not the essential problem, but rather the problem is to formalize the provability predicate $Thm_{I\Delta_0}$ so that $I\Delta_0$ can prove that if a formula is provable, then it is provable that it is provable (in other words, so that the third Hilbert-Bernays-Löb derivability condition holds).

Since Parikh's original definition, bounded arithmetic has grown into a large and actively studied area. In the late 1970's to mid 1980's, Paris and Wilkie and other authors developed a large body of results on bounded arithmetic and especially its connections to computational complexity. Of particular note is the influential paper of Paris and Dimitracopoulos [33] relating model-theoretic results in bounded arithmetic, Peano arithmetic and true arithmetic to open questions in computational complexity concerning P, NP, the polynomial time hierarchy, and polynomial space. In the mid-1980's, the present author introduced a new formalization S_2 of bounded arithmetic which is essentially equivalent to $I\Delta_0 + \Omega_1$. The theory S_2 contains natural subtheories S_2^i and T_2^i which can be related directly to complexity classes in the polynomial time hierarchy such as P and NP. Since the mid-1980's, there has been extensive research on S_2 and related theories; however, it is beyond the scope of the present paper to survey this work.

For more complete treatments of bounded arithmetic, the reader can refer to Wilkie-Paris [43], Buss [3], Hájek-Pudlák [16] and Krajíček [21].

In addition to the aspects of Parikh's "feasibility" paper discussed above, two additional topics were covered in the paper. The first topic was the description of a formula A which is provable in PA and therefore the formula $P(A) := Thm_{\mathrm{PA}}(\ulcorner A \urcorner)$ is provable in PA, such that the shortest PA-proof of A is significantly longer than the shortest PA-proof of $P(A)$. "Significantly longer" means that any primitive recursive gap between the proof lengths is obtainable. Parikh then generalized this to formulas $P(P(\cdots P(A)\cdots))$ based on iterated use of the provability predicate. Subsequent research on this topic includes [10,8,6,40].

The second additional topic was "almost consistent theories" which are inconsistent theories extending PA in which the shortest proof of an inconsistency is infeasibly long. Let PA^+ denote Peano arithmetic formalized in the usual fashion, but in a language containing function symbols for all primitive recursive functions and the defining axioms for these function symbols. Let PA_F^+ denote PA^+ extended with the inclusion of a new unary predicate symbol F, where $F(x)$ has the intuitive meaning "x is a feasible integer", plus a finite list of axioms including (at least)

$$F(0), \qquad (\forall x)(F(x) \to F(S(x))),$$

and $\neg F(\theta)$ for some *closed* term θ. The new function symbol F is not allowed to appear in induction formulas. The theory PA_F^+ is obviously inconsistent, since θ is closed term and can hence be proved to equal $S^i(0)$ for some $i \geq 0$. However the term θ may have an extremely large value, and since F cannot be used in induction formulas, one might expect that any proof of an inconsistency in PA_F^+ must be extremely long. This is, in fact, exactly what Parikh proves in a strengthened form. Namely, he proves theorems stating that if $PA_F^+ \vdash B$ for some PA^+-formula B and if the PA_F^+-proof is short enough relative to the value of the closed term θ, then already PA^+ can prove B. We will omit a detailed description of these theorems, but instead only remark that the value of θ is superexponentially larger than the size of the PA_F^+-proof of F. Further work on these almost consistent theories has been done by Dragalin [11] and Carbone [7].

3 Lengths of Proofs

In the "length of proofs" paper [31], Parikh proved a remarkable theorem about lengths of proofs, of a type first conjectured by Kreisel. Let PA^* be a formalization of Peano arithmetic in a first-order language with a constant symbol 0, a unary symbol S, and two 3-ary relation symbols $A(\cdots)$ and $M(\cdots)$ defining the graphs of addition and multiplication. PA^* has axioms describing the properties of 0, S, A and M, plus induction for all formulas. Further assume PA^* is axiomatized in a "schematic" way with a finite number of axiom schemes and inference rule schemes (we'll discuss schematic systems in more detail below).

Proof length will be measured in terms of the number of steps, i.e., the number of lines or formulas, which appear in the proof. The notation "$PA^* \overset{k}{\vdash} A$" means that A has a PA^*-proof of size at most k. Let $\underline{n}$ denote the term $S^n(0) = S(S(\cdots S(0)\cdots))$ with value equal to $n \in \mathbb{N}$.

Theorem 4 (Parikh [31]) *Let $A(x)$ be a formula. Suppose there is a fixed $k > 0$ such that $PA^* \overset{k}{\vdash} A(\underline{n})$ for all $n \in \mathbb{N}$. Then $PA^* \vdash \forall x A(x)$.*

A quick observation is that the converse of this theorem is trivially true, since $\forall x A(x) \to A(\underline{n})$ is provably in a constant number of steps, independent of A and n (and since any schematic system can simulate modus ponens in a constant number of steps). As a second observation, note that since we are measuring proof lengths in terms of the number steps in the proof instead of in terms of the number of symbols in the proof, there are infinitely many proofs of size less than or equal to k. This is because a k-step proof may contain arbitrarily complex formulas.

Now it should be noted that Theorem 4 cannot possibly be true of all possible formalizations of PA^* — indeed, since PA^*'s consequences form a recursively enumerable set, a method of Craig's shows that it is possible to give a recursive axiomatization of PA^* in which every theorem of PA^* has a PA^*-proof in a constant number of lines. This is done by letting PA^* be axiomatized by the formulas $A \wedge (S^k0 = S^k0)$ where k is the Gödel number of a PA^*-proof of A. Clearly this axiomatization is recursive. Also, A can be proved from this axiom in a constant number of steps via modus ponens with the formula $(A \wedge (S^k0 = S^k0)) \to A$.

This recursive formalization of PA^* is of course rather pathological and to avoid pathological axiomatizations, Parikh required that PA^* be formalized as a *schematic system*. A schematic system is a formal system with a finite set of *schematic rules* modified by *admissible restrictions*. Rather than reproduce Parikh's definitions, we present four examples of schematic rules:

1. Modus Ponens. The following inference rule is a binary schematic rule:

$$\frac{P \quad P \to Q}{Q}$$

   where any formulas may be substituted for P and Q. It has no associated restrictions.
2. $P \to (Q \to P)$ is a nullary schematic rule (i.e., an axiom scheme). Any formulas may be substituted for P and Q and it has no associated restrictions.
3. $\forall x P(x) \to P(t)$ is another nullary schematic rule. Any formula may be substituted for P and any variable for x and any term for t, provided t is free for x in P. The condition "t is free for x in P" is the admissible restriction modifying this schematic inference rule.
4. $P(0) \wedge \forall x(P(x) \to P(S(x))) \to \forall x P(x)$ is another example of a schematic inference rule. Again, any formula may be substituted for P and any variable for x. It has the associated admissible restriction that "x does not appear bound in P".

Other types of admissible restrictions are possible other than the above example. They include, for instance, "x does not occur in P" and "x does not occur free in P".

To properly define schematic systems, Parikh needed to augment the language of first-order logic with special "formula variables", "term variables" and "meta-variables". A schematic rule and its associated admissible restrictions are expressed in the augmented language; a *substitution* is a mapping from the formula variables, term variables and metavariables to formulas, terms and variables (respectively). Then, a schematic rule indicates that any substitution instance of the rule is a valid inference provided the substitution satisfies the conditions of the associated admissible restrictions.

It is an important property of schematic systems that they are specified by only a finite set of schematic rules. For instance, this excludes the pathological axiomatization of PA^* discussed above. Examples of schematic proof systems include the usual Hilbert-style proof systems. (The Gentzen sequent calculus is not strictly speaking a schematic system; however, by slightly extending the definition of schematic systems, it can also be viewed as a schematic system.)

We'll now sketch the main ideas of the proof of Theorem 4. For this proof, it is necessary to circumvent the problem of having infinitely many proofs of k steps. Toward this end, Parikh introduced the important notion of a proof analysis (subsequently called a "proof skeleton" or "proof scheme" by other researchers).

A proof analysis describes a (possible) proof by giving a precise statement about which schematic rule is used to derive each line in the proof, including a specification of which earlier lines in the proof (if any) were used as hypotheses to the rule. However, a proof analysis does not list the precise formulas and terms appearing in the proof. It is an easy observation that there are only finitely many proof analyses for proof of size at most k.

Of course, every proof has a proof analysis, but not every proof analysis corresponds to an actual proof. However, Parikh shows that for PA^* it is decidable whether a given proof analysis corresponds to an actual proof of a given formula. For this, he first establishes the following:

Lemma 5 (Parikh [31, Lemma A]) *Given a k-step proof analysis $\mathfrak{A}$, one can effectively produce a sequence of formulas $F_1, \ldots, F_m, G_1, \ldots, G_m, H$ and a finite set K of admissible restrictions so that any formula A has a proof with analysis $\mathfrak{A}$ if and only if there is a substitution $\mathcal{S}$ satisfying $\mathcal{S}(F_i) = \mathcal{S}(G_i)$ for all i and $\mathcal{S}(H) = A$ and satisfying the restrictions in K. Furthermore, for any such substitution, $\mathcal{S}(F_1), \ldots, \mathcal{S}(F_k), A$ is a valid proof of A and, conversely, any proof of A with proof analysis $\mathfrak{A}$ can be obtained in this way.*

Lemma 5 holds for any schematic proof system, not just for PA^*. It is proved by induction on k, and we shall omit the proof here. The number m turns out to be $\leq c \cdot k$ where c is the maximum number of hypotheses in a schematic rule.

The problem of determining whether there is a substitution that makes formulas F_i equal to G_i is a kind of unification problem with the admissible restrictions placing extra conditions on the solution of the unification problem.

Thus Lemma 5 states that the question of whether a proof analysis corresponds to an actual proof of A is equivalent to a unification problem.

Lemma 6 (Parikh [31, Lemma B]) *Given a proof analysis $\mathfrak{A}$ for the theory PA^* and a formula $A(x)$, one can effectively find a formula $B_{\mathfrak{A}}(x)$ in the language of Presburger arithmetic such that (PA^* proves that) for all $n \in \mathbb{N}$, $B_{\mathfrak{A}}(n)$ is true if and only if $A(\underline{n})$ has a PA^*-proof with proof analysis $\mathfrak{A}$.*

Lemma 6 is proved by showing that the unification problem of Lemma 5 can be rephrased as a formula of Presburger arithmetic. It is useful to split this unification into two parts: the first part is an ordinary unification of the type due to Herbrand [18] and Robinson [37], which is sometimes called "first-order" unification. This part of the unification problem can be solved to get the *logical* and *relational* structure of all the formulas in a (simplest possible) proof with analysis $\mathfrak{A}$. That is to say, we can solve a first-order unification problem to determine the logical symbols (propositional connectives and quantifiers but not the identity of the quantified variables) and all the relation symbols ($=$, A and M; undetermined relations may be set to "$=$").

The remaining part of the unification problem is a special case of "second-order unification" and it is the problem of determining which *terms* must appear in a proof with analysis $\mathfrak{A}$. The second-order unification problem is find terms $t_1, \ldots, t_n$ satisfying conditions such as

(i) $t_i = t_j$
(ii) $t_i = S^r(t_j)$
(iii) $t_i(x/t_j) = t_k$
(iv) $t_i = S^r(0)$.

where, since we are working with PA^*, each t_i is of the form $t_i = S^{\ell_i}0$ or $t_i = S^{\ell_i}(x_{v_i})$ where $\ell_i, v_i \geq 0$ and x_{v_i} is a variable. There are, up to renaming of variables, only finitely many choices for the innermost symbols 0 or x_{v_i} of t_i. Fixing one such choice, finding a solution of the unification problem reduces to finding an integer solution to a set of simultaneous equations of the forms

(i′) $\ell_i = \ell_j$ (provided $v_i = v_j$)
(ii′) $\ell_i = r + \ell_j$ (provided $v_i = v_j$)
(iii′) $\ell_i + \ell_j = \ell_k$ (provided $x = x_{v_i}$ and $v_j = v_k$)
(iv′) $\ell_i = r$ (provided 0 was chosen instead of a variable x_{v_i})

Since Presburger arithmetic includes 0, S and $+$, the property of being able to satisfy equations of the forms (i′)-(iv′) is expressible by a formula of Presburger arithmetic (in fact, by an existential formula). We get at most one such Presburger arithmetic formula for each set of choices of innermost symbols of t_i; and, for each such choice it is easy to check whether the admissible restrictions will be satisfied. Thus the formula $B_{\mathfrak{A}}(n)$ is defined to be the disjunction of these Presburger arithmetic formulas for choices of the $0/x_{v_i}$ values for which a substitution could give a valid proof.

The above argument concludes our outline of the proof of Lemma 6. From this lemma, Theorem 4 follows readily by observing the following facts: (1) There are only finitely many proof analyses $\mathfrak{A}$ for proofs of size less than or equal to k. Thus, taking the disjunction of finitely many formulas $B_{\mathfrak{A}}$ gives a Presburger arithmetic formula $B_{A,k}(n)$ expressing the condition $A(\underline{n})$ has a proof of size $\leq k$. (2) Presburger arithmetic is complete. Therefore if $A(\underline{n})$ has a PA^*-proof of $\leq k$ steps for all n, the Presburger formula $B_{A,k}(n)$ is true for all $n \in \mathbb{N}$ and by completeness, $\forall x B_{A,k}(x)$ is provable in Presburger arithmetic and hence in PA^*. (3) From the solution of the first-order portion of the unification problem, there is a uniform upper bound on the quantifier complexity needed in PA^*-proofs of A of size $\leq k$. It follows that PA^* proves that every instance of $A(\underline{n})$ has a PA^*-proof of bounded logical complexity. Finally, Theorem 4 follows by the reflexivity of PA^*; that is to say, by the fact that Peano arithmetic proves the reflection principle for formulas proved by proofs with a constant upper bound on their logical complexity.

As a corollary to the method of proof of Theorem 4 above, we get

Theorem 7 (Parikh [31]) *Let $A(x)$ be a PA^*-formula and $k \in \mathbb{N}$. Then the set $\{n \mid PA^* \overset{k}{\vdash} A(\underline{n})\}$ consists of a finite set of integers plus a finite union of arithmetic progressions.*

The "length of proofs" paper concluded with a proof of a special case of a theorem stated by Gödel in his lengths of proofs paper [14]. Let PA_k^* denote k-th order arithmetic in the language 0, S, A, M. Thus $PA_1^* = PA^*$ and PA_2^* is a system of analysis.

Theorem 8 (Parikh [31]) *PA_2^* has unbounded proof speedup over PA_1^*.*

The theorem is proved by noting that PA_2^* proves $Con(PA^*)$, but PA^* does not. Let $Con(PA^*, n)$ be a formula stating that there is no PA^*-proof of $0 = 1$ of size less than or equal to n. Then there is a uniform upper bound on the size of the shortest PA_2^*-proofs of $Con(PA^*, \underline{n})$, with the same upper bound applying for all values of n. But, by Theorem 4, there is no uniform upper bound on the size of the shortest PA^*-proofs of $Con(PA^*, \underline{n})$ since otherwise PA^* would prove its own consistency.

Gödel actually stated that $(k + 1)$-st order arithmetic has unbounded speedup over k-th order arithmetic; however, he gave no proof of this. Parikh extended the unbounded speedup in [32], showing that PA_{k+1}^* has unbounded speedup over PA_k^*. Buss [5] later established the same results for Peano arithmetic formulated in the usual style with addition and multiplication as function symbols.

Since the "length of proofs" paper [31] there has been a large number of papers dealing with the lengths of first-order proofs. The results in this area are too numerous to list them all, so we mention only a few of the papers most closely related to Parikh's work. Richardson [36], Yukami [46,47] and Miyatake [25,24] gave extensions of Parikh's work on proof lengths. Krajíček [20] sharpened the bounds on the logical complexity of formulas in proofs of bounded size. Goldfarb [15] proved the undecidability of general second-order unification when arbitrary function symbols are allowed in place of the single unary functions symbol S. Farmer [12,13] proved the decidability of second-order monadic unification where all function symbols are unary and gave applications to lengths of proofs. Orevkov [28,29] and Krajíček-Pudlák [22] established the undecidability of the problem of determining whether a proof analysis corresponds to an actual proof. Buss [4] proved the undecidability of the k-step provability problem for a particular version of the Gentzen sequent calculus. (It is still open whether the k-step provability problem is undecidable for all schematic formulations of first-order logic or Peano arithmetic with a binary function symbol in the language.) Finally, Baaz and Pudlák [1] proved Kreisel's conjecture for the theory $L\Sigma_1$ with the least number principle for Σ_1-formulas.

4 Feasible Arithmetic and Proof Lengths

At first glance, the two topics of feasibility in theories of arithmetic and of length of first-order proofs seem to have little in common. However, subsequent to the publication of the two papers of Parikh's, surprisingly close connections have been discovered between bounded arithmetic and proof length. Firstly, Paris and Wilkie [34] give a direct translation between provability in $I\Delta_0$ and provability in constant depth propositional theories. Secondly, Cook [9] introduced an equational theory PV of polynomial time computable functions and showed that PV-proofs can be translated into schematic propositional proofs; thus, by the conservativity of S^1_2 over PV [3], the same holds for proofs in the fragment S^1_2 of bounded arithmetic. In this way, there are two ways of obtaining direct connections between provability in theories of bounded arithmetic and lengths of proofs in propositional logic: this has led to renewed interest in the lengths of proofs in both propositional and first-order logic, and a large number of significant results in this area have been obtained, largely inspired by the connections to computational complexity. It is beyond the scope of this paper to survey this research, but the the book length treatment by Krajíček [21] contains a comprehensive and fairly up-to-date treatment of recent progress in this area.

References

[1] M. Baaz and P. Pudlák, *Kreisel's conjecture for $L\exists_1$*, in Arithmetic, Proof Theory, and Computational Complexity, Clarendon Press, Oxford, 1993, pp. 30–60. Includes a postscript by G. Kreisel.

[2] J. Bennett, *On Spectra*, PhD thesis, Princeton University, 1962.

[3] S. R. Buss, *Bounded Arithmetic*, Bibliopolis, 1986. Revision of 1985 Princeton University Ph.D. thesis.

[4] ——, *The undecidability of k-provability*, Annals of Pure and Applied Logic, 53 (1991), pp. 75–102.

[5] ——, *On Gödel's theorems on lengths of proofs I: Number of lines and speedup for arithmetics*, Journal of Symbolic Logic, 59 (1994), pp. 737–756.

[6] A. Carbone, *Provable fixed points in $I\Delta_0 + \Omega_1$*, Notre Dame Journal of Formal Logic, 32 (1991), pp. 562–572.

[7] ——, *Cycling in proofs, feasibility and no speed-up for nonstandard arithmetic.* IHES preprint, 1996.

[8] A. Carbone and F. Montagna, *Much shorter proofs: a bimodal investigation*, Zeitschrift für Mathematische Logik und Grundlagen der Mathematik, 36 (1990), pp. 47–66.

[9] S. A. Cook, *Feasibly constructive proofs and the propositional calculus*, in Proceedings of the Seventh Annual ACM Symposium on Theory of Computing, 1975, pp. 83–97.

[10] D. H. J. de Jongh and F. Montagna, *Provable fixed points*, Zeitschrift für Mathematische Logik und Grundlagen der Mathematik, 34 (1988), pp. 229–250.

[11] A. G. Dragalin, *Correctness of inconsistent theories with notions of feasibility*, in Computation Theory: Fifth Symposium, Springer-Verlag, 1985, pp. 58–79.

[12] W. M. Farmer, *A unification algorithm for second order monadic terms*, Annals of Pure and Applied Logic, 39 (1988), pp. 131–174.

[13] ——, *A unification-theoretic method for investigating the k-provability problem*, Annals of Pure and Applied Logic, 51 (1991), pp. 173–214.

[14] K. Gödel, *Über die Länge von Beweisen*, Ergebnisse eines Mathematischen Kolloquiums, (1936), pp. 23–24. English translation in *Kurt Gödel: Collected Works, Volume 1, pages 396-399, Oxford University Press, 1986.*

[15] W. D. Goldfarb, *The undecidability of the second-order unification problem*, Theoretical Comput. Sci., 13 (1981), pp. 225–230.

[16] P. Hájek and P. Pudlák, *Metamathematics of First-order Arithmetic*, Springer-Verlag, Berlin, 1993.

[17] K. Harrow, *The bounded arithmetic hierarchy*, Information and Control, 36 (1978), pp. 102–117.

[18] J. Herbrand, *Recherches sur la théorie de la démonstration*, PhD thesis, University of Paris, 1930.

[19] N. D. Jones, *Context-free languages and rudimentary attributes*, Mathematical Systems Theory, 3 (1969), pp. 102–109.

[20] J. Krajíček, *On the number of steps in proofs*, Annals of Pure and Applied Logic, 41 (1989), pp. 153–178.

[21] ——, *Bounded Arithmetic, Propositional Calculus and Complexity Theory*, Cambridge University Press, 1995.

[22] J. Krajíček and P. Pudlák, *The number of proof lines and the size of proofs in first-order logic*, Archive for Mathematical Logic, 27 (1988), pp. 69–84.

[23] R. J. Lipton, *Model theoretic aspects of computational complexity*, in Proceedings of the 19th Annual Symposium on Foundations of Computer Science, IEEE Computer Society, 1978, pp. 193–200.

[24] T. Miyatake, *On the length of proofs in a formal system of recursive arithmetic*, in Logic Symposia, Hakone, Lecture Notes in Mathematics #891, Springer-Verlag, 1980, pp. 81–108.

[25] ——, *On the length of proofs in formal systems*, Tsukuba Journal of Mathematics, 4 (1980), pp. 115–125.

[26] E. Nelson, *Predicative Arithmetic*, Princeton University Press, 1986.

[27] V. A. Nepomnjaščiĭ, *Rudimentary predicates and Turing computations*, Dokl. Akad. Nauk SSSR, 195 (1970), pp. 282–284. English translation in *Soviet Math. Dokl.* 11 (1970) 1462–1465.

[28] V. P. Orevkov, *Reconstruction of a proof from its scheme*, Soviet Mathematics Doklady, 35 (1987), pp. 326–329. Original Russian version in Dokl. Akad. Nauk. **293** (1987) 313-316.

[29] ——, *Complexity of Proofs and Their Transformations in Axiomatic Theories*, American Mathematical Society, 1991.

[30] R. J. Parikh, *Existence and feasibility in arithmetic*, Journal of Symbolic Logic, 36 (1971), pp. 494–508.

[31] ——, *Some results on the lengths of proofs*, Transactions of the American Mathematical Society, 177 (1973), pp. 29–36.

[32] ——, *Introductory note to 1936(a)*, in Kurt Gödel, Collected Works, Volume 1, Oxford University Press, 1986, pp. 394–397.

[33] J. B. Paris and C. Dimitracopoulos, *Truth definitions for Δ_0 formulae*, in Logic and Algorithmic, L'Enseignement Mathematique Monographie no 30, 1982, pp. 317–329.

[34] J. B. Paris and A. J. Wilkie, *Counting problems in bounded arithmetic*, in Methods in Mathematical Logic, Lecture Notes in Mathematics #1130, Springer-Verlag, 1985, pp. 317–340.

[35] P. Pudlák, *Cuts, consistency statements and interpretation*, Journal of Symbolic Logic, 50 (1985), pp. 423–441.

[36] D. Richardson, *Sets of theorems with short proofs*, Journal of Symbolic Logic, 39 (1974), pp. 235–242.

[37] J. A. Robinson, *A machine-oriented logic based on the resolution principle*, J. Assoc. Comput. Mach., 12 (1965), pp. 23–41.

[38] V. Y. Sazanov, *A logical approach to the problem "P=NP?"*, in Mathematics Foundations of Computer Science, Lecture Notes in Computer Science #88, Springer-Verlag, 1980, pp. 562–575. There is an unfixed problem with the proof of the main theorem in this article; see [39] for a correction.

[39] ——, *On existence of complete predicate calculus in matemathematics without exponentiation*, in Mathematics Foundations of Computer Science, Lecture Notes in Computer Science #118, Springer-Verlag, 1981, pp. 383–390.

[40] V. Y. Shavrukov, *Subalgebras of diagonalizable algebras of theories containing arithmetic*, Dissertationes mathematicae (Rozprawy matematycne), 323 (1993). Instytut Matematyczny, Polska Akademia Nauk, Warsaw.

[41] R. Smullyan, *Theory of Formal Systems*, Princeton University Press, 1961.

[42] A. J. Wilkie, *Applications of complexity theory to Σ_0-definability problems in arithmetic*, in Model Theory of Algebra and Arithmetic, Lecture Notes in Mathematics #834, Springer-Verlag, 1979, pp. 363–369.

[43] A. J. Wilkie and J. B. Paris, *On the scheme of induction for bounded arithmetic formulas*, Annals of Pure and Applied Logic, 35 (1987), pp. 261–302.

[44] C. Wrathall, *Rudimentary predicates and relative computation*, SIAM Journal on Computing, 7 (1978), pp. 194–209.

[45] A. S. Yessenin-Volpin, *The ultra-intuitionistic criticism and the antitraditional program for foundations of mathematics*, in Intuitionism and Proof Theory, A. Kino, J. Myhill, and R. E. Vesley, eds., North-Holland, 1970, pp. 1–45.

[46] T. Yukami, *A theorem on the formalized arithmetic with function symbols $\prime$ and* $+$, Tsukuba Journal of Mathematics, 1 (1977), pp. 195–211.

[47] ——, *A note on a formalized arithmetic with function symbols $\prime$ and* $+$, Tsukuba Journal of Mathematics, 2 (1978), pp. 69–73.