The quantifier complexity of polynomial-size iterated definitions in first-order logic

Samuel R. Buss[*] and **Alan S. Johnson**

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

Key words Iterated definition, recursive definition, quantifier complexity, finite quantifier
MSC (2000) 03B10, 03D80

We refine the constructions of Ferrante-Rackoff and Solovay on iterated definitions in first-order logic and their expressibility with polynomial size formulas. These constructions introduce additional quantifiers; however, we show that these extra quantifiers range over only finite sets and can be eliminated. We prove optimal upper and lower bounds on the quantifier complexity of polynomial size formulas obtained from the iterated definitions. In the quantifier-free case and in the case of purely existential or universal quantifiers, we show that $\Omega(n/\log n)$ quantifiers are necessary and sufficient. The last lower bounds are obtained with the aid of the Yao-Håstad switching lemma.

1 Introduction

Consider the situation where predicates $R_n(\vec{x})$ are defined by an iterated definition in first-order logic. Namely, assume $R_0(\vec{x})$ is an explicitly given first-order formula, and that the predicate $R_n(\vec{x})$ is defined to be a first-order formula $A(R_{n-1})$ involving R_{n-1}. Formally speaking, the notation $A(R_{n-1})$ indicates that $A(P)$ is a formula containing occurrences of a new predicate symbol P, and $A(R_{n-1})$ is obtained by replacing all occurrences of P with the predicate R_{n-1}. In particular, the different occurrences of R_{n-1} in $A(R_{n-1})$ may have different terms as arguments.

We take R_n to be the first-order formula obtained by unwinding the recursive definition. However, since $A(R_{n-1})$ may contain multiple occurrences of R_{n-1}, R_n may be exponentially large. For languages that contain $\leftrightarrow$, this exponential size was reduced by the classic results of Ferrante-Rackoff [1], who showed that the iteratively defined property $R_n(\vec{x})$ can be expressed as a first-order formula $F_n(\vec{x})$ which has size polynomially bounded by the number of iterations n and the sizes of the formulas R_0 and A. (In general, we use F_n to denote a polynomial size formula that expresses R_n). Solovay (unpublished) showed that the Ferrante-Rackoff bounds also apply to first-order logic without $\leftrightarrow$ in the language. An exposition of these results can also be found in the expository article of Pudlák [2]; Pudlák applies these polynomial size bounds to the give a polynomial upper bound on the size of proofs of partial self-consistency statements of the type given by [3, 4].

Pudlák defines predicates Sat_n which state that formulas with n logical connectives are satisfiable, and uses the results of Ferrante-Rackoff-Solovay to express Sat_n by polynomial size formulas. There are polynomial size proofs that these formulas for Sat_n satisfy the needed inductive properties: this allows Pudlák to obtain his bounds on proofs of partial self-consistency. More generally, one can define the formulas for Sat_n as Σ_n-formulas. This can be shown directly, but the goal of the present paper generalize this by characterizing the logical complexity of properties defined by arbitrary first-order recursive definitions.

The essential idea of the Ferrante-Rackoff-Solovay constructions is to show that $A(R_{n-1})$ can be converted into an equivalent formula that contains only a single occurrence of R_{n-1}. Indeed, if there is only a single occurrence of R_{n-1} in $A(R_{n-1})$, then unwinding the recursion shows R_n can be expressed by a polynomial size (in fact, linear size) formula F_n. The present paper sharpens this construction by giving a more careful analysis of the quantifier complexity of F_n. The Ferrante-Rackoff-Solovay constructions introduce additional quantifiers,

[*] Both authors were supported in part by NSF grant DMS-0700533. E-mail: sbuss@math.ucsd.edu, asj002@math.ucsd.edu

Complexity of $A(P)$	Occurrences of P	Lower Bound	Upper Bound	Theorem
$\Sigma_{2k}, k > 0$	Pos.	Σ_{2kn}	$\Sigma_{2kn}(R_0)$	5.3/4.3
$\Sigma_{2k}, k > 0$	Pos./Neg.	Σ_{2kn}	$\Delta_{2kn+1}(R_0)$	5.3/6.1
$\Sigma_{2k+1}, k > 0$	Pos.	Σ_{2kn+1}	$\Sigma_{2kn+1}(R_0)$	5.6/4.5
$\Sigma_{2k+1}, k > 0$	Pos./Neg.	$\Sigma_{(2k+1)n}$	$\Delta_{(2k+1)n+1}(R_0)$	6.3/6.1
$\Delta_{k+1}, k > 0$	Pos.	Δ_{kn+1}	$\Delta_{kn+1}(R_0)$	5.8/4.8
$\Delta_{k+1}, k > 0$	Pos./Neg.	Δ_{kn+1}	$\Delta_{kn+1}(R_0)$	5.8/6.1
Σ_1	Pos.	$\Sigma_{\frac{\varepsilon n}{\log n}} \cup \Pi_{\frac{\varepsilon n}{\log n}}$	$\Delta_{\frac{\delta n}{\log n}}(R_0)$	5.9/4.7
Σ_1	Pos./Neg.	Σ_n	$\Delta_{n+1}(R_0)$	6.3/6.1
Quantifier free	Pos.	$\Sigma_{\frac{\varepsilon n}{\log n}} \cup \Pi_{\frac{\varepsilon n}{\log n}}$	$\Delta_{\frac{\delta n}{\log n}}(R_0)$	5.9/4.7
Quantifier free	Pos./Neg.	$\Sigma_{\frac{\varepsilon n}{\log n}} \cup \Pi_{\frac{\varepsilon n}{\log n}}$	$\Delta_{\frac{\delta n}{\log n}}(R_0)$	5.9/6.2

Fig. 1 The obtained upper and lower bounds on the complexity of polynomial size formulas for R_n, where $\delta, \varepsilon > 0$ are arbitrarily small constants which depend on the degree of the polynomial growth rate.

and additional quantifier alternation. However, these additional quantifiers range over only finite sets, and it is shown that these "finite quantifiers" can be eliminated from F_n.

Let $k \geq 0$. A formula B is defined to be Σ_k (respectively, Π_k) provided its quantifier block is Σ_k (respectively, Π_k) after the application of prenex operations. It is permitted that a block of quantifiers is empty, so that Σ_k contains Π_{k-1}, and Π_k contains Σ_{k-1}. A Δ_k formula is one which is provably equivalent to a Σ_k and a Π_k formula by polynomial size proofs. We shall later define classes $\Sigma_k(\varphi)$, $\Pi_k(\varphi)$, and $\Delta_k(\varphi)$. They are defined similarly to Σ_k, Π_k, and Δ_k by counting alternations of quantifiers, but ignoring quantifiers appearing in φ. The precise definition is given in Section 4.

We always measure the *size* of a proof by the number of symbols in the proof. All the formal proofs discussed in this paper are done in the sequent calculus, a definition of which can be found in [5]. The sequent calculus is polynomially equivalent to many other common proof systems, such as Hilbert systems and natural deduction, so our results hold for these other systems as well.

Figure 1 summarizes our results. The first column lists the quantifier complexity of $A(P)$ and the second column lists whether P appears only positively or appears both positively and negatively in $A(P)$. The third and fourth columns give the upper and lower bounds on the complexity of polynomial size first-order formulas which are equivalent to the property R_n defined recursively as $A(R_{n-1})$.

The results of Figure 1 apply to validity in first-order models that have two or more elements. The case of models with only a single element can safely be ignored, since in such models, quantifiers become vacuous and first-order logic becomes just propositional logic (namely, with each predicate symbol corresponding to only a boolean variable).

The outline of the paper is as follows. Section 2 begins by defining the notion of a *finite* quantifier, i.e., a quantifier that effectively quantifies over an explicit finite set. In many cases, it is possible to exchange quantifier order, namely one can move a finite quantifier rightward (inward) past an ordinary quantifier. Section 2 gives precise bounds on how this increases the size of a formula. Section 2 also introduces notations and conventions for quantifying over a block of quantifiers, in particular for the case where there are a finite number of possible values for the block of quantified variables. Section 3 uses finite quantifiers to equivalently rewrite a formula A with multiple positive instances of a predicate P using only one occurrence of P. Section 4 proves the upper bounds in Figure 1 for positive occurrences of P. Section 5 proves the lower bounds in Figure 1 with positive occurrences. There are two techniques to prove the lower bounds. One is to use complete problems from the arithmetic hierarchy. The other is to use lower bounds on circuits calculating parity, specifically the Yao-Håstad switching lemma. Section 6 proves the bounds for formulas A in which A occurs both positively and negatively.

An example application of our results is to the predicate $T_n(x, y)$ that says there is a path of length at most 2^n from x to y in a graph (G, E). Specifically, let $T_0(x, y)$ be $E(x, y) \vee x = y$, and for $n > 0$, let $T_n(x, y)$ be $\exists z(T_{n-1}(x, z) \wedge T_{n-1}(z, y))$. As is well-known, $T_n(x, y)$ is also expressed $\exists z \forall u((u = x \vee u = y) \rightarrow T_{n-1}(u, z))$.

This construction is a simple special case of the Ferrante-Rackoff method; indeed, the quantifier "$\forall u$" is an example of what we call a finite quantifier in Section 2, since u effectively ranges over the finite set $\{x, y\}$. Unwinding the definition of T_n above gives a polynomial size Σ_{2n}-definition for T_n. Our Theorem 4.7 below improves this by showing that T_n can also be expressed by polynomial size formulas in $\Sigma_{\delta n/\log n}$, for all $\delta > 0$.

2 Finite Quantification

This section discusses quantification over fixed-length blocks of variables which effectively range over finitely many terms; such a quantifier will be termed *finite*. This section introduces notation for finite quantifiers, and then proves some basic properties. All the constructions and theorems in this section hold in pure logic, rather than systems with a pairing function, such as Peano arithmetic.

As an example of finite quantification, consider the formulas

$$\forall x((x = s_1 \vee x = s_2 \vee x = s_3) \to B(x))$$

and

$$\exists x((x = s_1 \vee x = s_2 \vee x = s_3) \wedge B(x)).$$

The variable x effectively ranges over the set $\{s_1, s_2, s_3\}$, that is, over a set of size at most three. The terms s_i may contain variables, so the range of quantification is not fixed, but can vary with the values of the variables. Note that the two formulas above contain only a single occurrence of the subformula B, but are equivalent to $B(s_1) \wedge B(s_2) \wedge B(s_3)$ and $B(s_1) \vee B(s_2) \vee B(s_3)$, respectively.

In general, finite quantification will be based on vectors (or, blocks) of variables. The following definition introduces notations for vectors of variables and for equality between vectors of terms.

Definition 2.1 Let $s_1, \ldots, s_p$ and $t_1, \ldots, t_p$ be vectors of terms. The formula $\vec{s} = \vec{t}$ abbreviates the conjunction $\bigwedge_{j=1}^{p} s_j = t_j$. A block of existential quantifiers $\exists z_1 \cdots \exists z_p$ will be denoted $\exists^p \vec{z}$. Similarly, $\forall^p \vec{z}$ denotes $\forall z_1 \cdots \forall z_p$.

The superscript p on the quantifier indicates the number of quantified variables, and may be suppressed if unimportant or if clear from the context.

Definition 2.2 Let $s_{i,j}$ be terms for $1 \leq i \leq k$ and $1 \leq j \leq p$. Let $\vec{s}_i$ be the vector $s_{i,1}, \ldots, s_{i,p}$, and define S to be the set $S = \{\vec{s}_1, \ldots, \vec{s}_k\}$. Then $\forall^p_{k,S} \vec{z}\, B(\vec{z})$ is the formula

$$\forall^p \vec{z} \left[\left(\bigvee_{i=1}^{k} \vec{z} = \vec{s}_i \right) \to B(\vec{z}) \right].$$

Dually, $\exists^p_{k,S} \vec{z}\, A(\vec{z})$ is

$$\exists^p \vec{z} \left[\left(\bigvee_{i=1}^{k} \vec{z} = \vec{s}_i \right) \wedge B(\vec{z}) \right].$$

As an example of finite quantifiers, consider the formula

$$\forall^1_{2,S} \vec{z} \forall^1_{2,T} \vec{w} B(\vec{z}, \vec{w});$$

here $\vec{z}$ and $\vec{w}$ are of length 1. Each of $\vec{z}$ and $\vec{w}$ range over two values, therefore $(\vec{z}, \vec{w})$ ranges over four values. Thus the formula can equivalently be rewritten as $\forall^2_{4,U} \vec{u}\, B(\vec{u})$, for properly chosen U. This example typifies how to combine like, finite quantifiers. The next proposition generalizes this example, and gives a size bound on the resulting formula.

Proposition 2.3 *Let φ be $\forall^p_{k,S} \vec{z}\, \forall^q_{l,T} \vec{w}\ B(\vec{z}, \vec{w})$ and ψ be $\forall^{p+q}_{kl,S\times T}(\vec{z}, \vec{w})\ B(\vec{z}, \vec{w})$, where $S \times T$ is the set $\{(\vec{s}_i, \vec{t}_j) | 1 \leq i \leq k, 1 \leq j \leq l\}$. Then $\vdash \varphi \leftrightarrow \psi$ by a proof of size polynomial in $|\varphi|$. Furthermore, the sizes of φ and ψ are related by*

$$|\psi| - |B(\vec{z}, \vec{w})| \leq \max\{k, l\}(|\varphi| - |B(\vec{z}, \vec{w})|).$$

The same result holds if the displayed universal quantifiers in φ and ψ are replaced with existential quantifiers.

P r o o f. The proof of the finite existential quantifier case is omitted because it is dual to the finite universal quantifier case. Let χ be

$$\bigwedge_{\vec{s}\in S}\bigwedge_{\vec{t}\in T} B(\vec{s},\vec{t}).$$

It is clear that $\chi \leftrightarrow \varphi$ and $\chi \leftrightarrow \psi$. Formalizing the polynomial size proofs in the sequent calculus is a straightforward exercise. To prove the size bound, write φ according to the above definitions,

$$\forall^p \vec{z}\left[\left(\bigvee_{i=1}^{k} \vec{z}=\vec{s}_i\right) \rightarrow \left[\forall^q \vec{w}\left[\left(\bigvee_{j=1}^{l} \vec{w}=\vec{t}_i\right) \rightarrow B(\vec{z},\vec{w})\right]\right]\right].$$

Similarly, ψ is the formula

$$\forall^p \vec{z}\forall^q \vec{w}\left[\left(\bigvee_{i=1}^{k}\bigvee_{j=1}^{l} \vec{z}=\vec{s}_i \wedge \vec{w}=\vec{t}_j\right) \rightarrow B(\vec{z},\vec{w})\right].$$

Count the number of logical symbols in φ and ψ while disregarding B, remembering that vector equality is an abbreviation. In φ there are $p+q$ $\forall$'s, $k+l-2$ $\vee$'s, 2 $\rightarrow$'s, and $k(p-1)+l(q-l)$ $\wedge$'s, for a total of $p(k+1)+q(l+1)$ logical connectives. On the other hand, ψ has $p+q$ $\forall$'s, $kl-1$ $\vee$'s, 1 $\rightarrow$, and $kl(p+q-1)$ $\wedge$'s, for a total of $p(kl+1)+q(kl+1)$ logical connectives. Thus the number of logical connectives in ψ is less than $\max\{k,l\}$ times the number of logical connectives in φ (not counting any logical connectives in B). Similarly, φ has $kp+lq$ $=$'s, kp z's, lq w's, k $\vec{s}_i$'s, and l $\vec{t}_j$'s, while ψ has $kl(p+q)$ $=$'s, klp z's, klq w's, kl $\vec{s}_i$'s, and kl $\vec{t}_j$'s (B is still not being counted). Thus the number of $=$'s, z's, w's, $\vec{s}_i$'s, and $\vec{t}_j$'s in ψ is at most $\max\{k,l\}$ times the number of $=$'s, z's, w's, $\vec{s}_i$'s, and $\vec{t}_j$'s in φ. □

The combination of like, finite quantifiers as in Proposition 2.3 is not a literal combination of multiple quantifiers into a single quantifier, as would be the case if pairing were used. Indeed, this paper never makes use of pairing to combine like quantifiers. Though Proposition 2.3 makes it appear that two universal quantifiers are turned into one, this is due to the suppression of quantifiers in the notation for finite quantifiers. As the proof of Proposition 2.3 shows, the number of quantifiers remains unchanged when combining like, finite quantifiers.

Proposition 2.3 is stated in a way that makes it clear that $B(\vec{z},\vec{w})$ does not participate in the size increase from φ to ψ. This fact makes it possible to combine multiple instances of adjacent like, finite quantifiers in parallel.

Proposition 2.4 *Let φ be a formula which contains occurrences of*

$$(Q_i)^{p_i}_{k_i,S_i}\vec{z}_i\ (Q_i)^{q_i}_{l_i,T_i}\vec{w}_i,$$

where, for $i=1,\ldots,n$, the quantifiers Q_i are either both $\exists$ or both $\forall$. Also suppose that the indicated quantifier blocks are all disjoint. Let ψ be the formula obtained by simultaneously combining the indicated adjacent like, finite quantifiers as in Proposition 2.3. Then $\vdash \varphi \leftrightarrow \psi$ by a proof polynomial in $|\varphi|$ and $|\psi| \le \max\{k_1,l_1,\ldots,k_n,l_n\}|\varphi|$.

P r o o f. Use induction on n and make use of Proposition 2.3. □

An example of a property that finite quantifiers enjoy over regular quantifiers is quantifier exchange for non-like quantifiers. For instance, consider the formula $\forall^1_{2,S}\vec{z}\,\exists w B(\vec{z},w)$, where $S=\{\vec{s}_1,\vec{s}_2\}$. This asserts there is a w_1 such that $B(\vec{s}_1,w_1)$ holds, and similarly there is a w_2 such that $B(\vec{s}_2,w_2)$ holds. This implies that the formula $\exists w_1\exists w_2\forall^2_{2,S^*}\vec{v}\ B(\vec{v})$ holds, where $S^*=\{(\vec{s}_1,w_1),(\vec{s}_2,w_2)\}$. This process of quantifier exchange is generalized by the next proposition.

Proposition 2.5 *Let φ be $\forall^p_{k,S}\vec{z}\,\exists^q\vec{w}\ B(\vec{z},\vec{w})$ and let ψ be*

$$\exists^q\vec{w}_1\cdots\exists^q\vec{w}_k\ \forall^{p+q}_{k,S^*}(\vec{z},\vec{u})\ B(\vec{z},\vec{u}),$$

where $S^* = \{(\vec{s}_i, \vec{w}_i) | 1 \leq i \leq k\}$. Then $\vdash \varphi \leftrightarrow \psi$ by a proof of size polynomial in $|\varphi|$. Furthermore, the sizes of φ and ψ are related by

$$|\psi| - |B(\vec{z}, \vec{w})| \leq Ck(|\varphi| - |B(\vec{z}, \vec{w})|),$$

where C is a fixed constant. A dual result holds for moving a finite existential quantifier past a universal quantifier.

P r o o f. The proof only considers the case of moving a finite universal quantifier past an existential quantifier, the other case being dual. Let χ be

$$\bigwedge_{i=1}^{k} \exists^q \vec{w}_i \; B(\vec{s}_i, \vec{w}_i).$$

There are straightforward polynomial size proofs of $\chi \leftrightarrow \varphi$ and $\chi \leftrightarrow \psi$. To prove the size bound, write φ as

$$\forall^p \vec{z} \left[\left(\bigvee_{i=1}^{k} \vec{z} = \vec{s}_i \right) \rightarrow (\exists^q \vec{w} \; B(\vec{z}, \vec{w})) \right].$$

Similarly, ψ is

$$\exists^q \vec{w}_1 \cdots \exists^q \vec{w}_k \; \forall^p \vec{z} \forall^q \vec{u} \left[\left(\bigvee_{i=1}^{k} \vec{z} = \vec{s}_i \wedge \vec{u} = \vec{w}_i \right) \rightarrow B(\vec{z}, \vec{u}) \right].$$

Disregarding the subformula $B(\vec{z}, \vec{w})$ of ψ and φ, the size of φ is $p + O(kp) + \sum_{i,j} |s_{i,j}| + q$ and the size of ψ is $kq + p + q + O(k(p+q)) + \sum_{i,j} |s_{i,j}|$, and the size bound follows. A detailed calculation similar to that in Proposition 2.3 shows that C can be taken to be 5. □

In contrast to combination of like, finite quantifiers, in the above exchange property φ and ψ do not have the same number of quantifiers. In fact, ψ has $(k+1)q + p$ quantifiers whereas φ only has $p + q$.

Similar to finite quantifier combination, the quantifier exchange of Proposition 2.5 can be done in parallel without an increase in the size.

Proposition 2.6 *Let φ be a formula which contains occurrences of*

$$(Q_i)_{k_i, S_i}^{p_i} \vec{z}_i \; (\overline{Q}_i)^{q_i} \vec{w}_i,$$

where, for $i = 1, \ldots, n$, Q_i and $\overline{Q}_i$ are existential and universal, or vice-versa. Also suppose that the indicated quantifier blocks are all disjoint. Let ψ be the formula obtained by simultaneously exchanging the indicated quantifiers as in Proposition 2.5. Then $\vdash \varphi \leftrightarrow \psi$ by a proof of size polynomial in $|\varphi|$. Moreover, there is a fixed constant C such that $|\psi| \leq C \max\{k_1, \ldots, k_n\} |\varphi|$.

P r o o f. Use induction on n and make use of Proposition 2.5. □

Note that S does not affect the size bounds in Propositions 2.3-2.6. Thus, the S may be suppressed in the notation if S is either clear from context or unnecessary to the argument. The subscript k will always be kept on a finite quantifier to distinguish it from a regular quantifier.

3 The Construction

The purpose of this section is to prove that a formula A containing only positive occurrences of a predicate P can be transformed into a logically equivalent formula containing only one (positive) occurrence of P. This result is due to Ferrante-Rackoff [1], who cite the work of Fischer-Meyer [6] and Meyer-Stockmeyer [7]. This section reproves the result of Ferrante-Rackoff with a different construction. The new feature of our proof is the explicit treatment of finite quantifiers, which will be important for our later constructions.

Before beginning the construction, notation for this section is set. In general, A can be any first-order formula, but for the present section assume that A is quantifier-free. A may contain free variables $\vec{x}$, but these will be suppressed in the notation as they do not play a role. Let P be a k-ary predicate such that P only occurs positively in A. Let n be the number of occurrences of P in A, and suppose the i^{th} occurrence of P in A occurs as $P(\vec{t}_i)$, where $\vec{t}_i$ is $t_{i,1}, \ldots, t_{i,k}$. Equality, $=$, is included in the language; indeed, the presence of $=$ plays a crucial role in the following. It is more convenient to work A in a certain normal form defined here.

Definition 3.1 A first-order formula B is in De Morgan normal form if the only negated subformulas of B are atomic.

Because A is provably equivalent to its De Morgan normal form by a proof of size polynomial in $|A|$, assume without loss of generality that A is in De Morgan normal form. Define a formula $\delta_P(A)$, which is related to the De Morgan normal form of $\neg A$, as follows.

Definition 3.2 The formula $\delta_P(A)$ is defined inductively by:

(i) $\delta_P(B)$ is B, if B is of the form $P(\vec{t})$.

(ii) $\delta_P(B)$ is $\neg B$, if B is an atomic formula not of the form $P(\vec{t})$.

(iii) $\delta_P(B \wedge C)$ is $\delta_P(B) \vee \delta_P(C)$.

(iv) $\delta_P(B \vee C)$ is $\delta_P(B) \wedge \delta_P(C)$.

By convention, delete any $\neg\neg$'s that are created due to the presence of a negated atomic formula not of the form $P(\vec{t})$. Note that $\delta_P(A)$ is the negation of A but with subformulas $P(\vec{t})$ left unnegated.

Substitution into $\delta_P(A)$ is given its own notation.

Definition 3.3 Let $\varphi_1, \ldots, \varphi_n$ be formulas. Define $\delta_P(A)(\varphi_1, \ldots, \varphi_n)$ as follows. First form $\delta_P(A)$. Then, for $i = 1, \ldots, n$, replace the i^{th} occurrence of P, namely replace the subformula $P(\vec{t}_i)$, with φ_i.

If B is a subformula of A, then $\delta_P(B/A)(\varphi_1, \ldots, \varphi_n)$ is the corresponding subformula of $\delta_P(A)(\varphi_1, \ldots, \varphi_n)$.

As an example, let A be $(B \vee P(\vec{t}_1)) \wedge (\neg C \vee D \vee P(\vec{t}_2))$, where B, C, D are atomic formulas not of the form $P(\vec{t})$, and let E be the second conjunct of A, namely $\neg C \vee D \vee P(\vec{t}_2)$. Then $\delta_P(A)(\varphi_1, \varphi_2)$ is $(\neg B \wedge \varphi_1) \vee (C \wedge \neg D \wedge \varphi_2)$ and $\delta_P(E/A)(\varphi_1, \varphi_2)$ is $C \wedge \neg D \wedge \varphi_2$.

There is a nice interplay between A and $\delta_P(A)$. Continuing with the example of the preceding paragraph, if A and $\delta_P(A)$ are both true, then $P(\vec{t}_1)$ and φ_1 are true if the first disjunct of $\delta_P(A)$ is true and $P(\vec{t}_2)$ and φ_2 are true if the second disjunct of $\delta_P(A)$ is true. Thus when A and $\delta_P(A)$ are true, they isolate the $P(\vec{t}_i)$ and φ_i that "make the formulas true." This is generalized and made precise in the following theorem.

Theorem 3.4 *Suppose, for $i = 1, \ldots, n$, φ_i^* are formulas such that $t_{i,j}$ is substitutable for u_j in φ_i^*, $j = 1, \ldots, k$. Let φ_i be $\varphi_i^*(\vec{t}_i/\vec{u})$, where, for each j, all occurrences of u_j are replaced by $t_{i,j}$. Then for any subformula B of A*

$$B \wedge \delta_P(B/A)(\vec{\varphi}) \to \exists^k \vec{u} \left[\left(\bigvee_{i=1}^{n} \varphi_i^* \right) \wedge P(\vec{u}) \right]$$

has a proof of size polynomial in $|A|$.

Proof. The proof is by induction on subformulas B of A.

(i) If B is $P(\vec{t}_j)$, where $1 \leq j \leq n$, then $\delta_P(B/A)(\vec{\varphi})$ is φ_j, and it is clear that the sequent has a short proof.

(ii) If B is an atomic subformula not of the form $P(\vec{t})$, then $\delta_P(B/A)(\vec{\varphi})$ is $\neg B$, and again it is clear that the sequent has a short proof.

(iii) Suppose B is $C \wedge D$. By induction, the sequent

$$C, \delta_P(C/A)(\vec{\varphi}) \longrightarrow \exists^k \vec{u} \left[\left(\bigvee_{i=1}^{n} \varphi_i^* \right) \wedge P(\vec{u}) \right]$$

has a polynomial size proof. By $\wedge$:left introduction derive

$$C \wedge D, \delta_P(C/A)(\vec{\varphi}) \longrightarrow \exists^k \vec{u} \left[\left(\bigvee_{i=1}^{n} \varphi_i^* \right) \wedge P(\vec{u}) \right].$$

Similarly,

$$C \wedge D, \delta_P(D/A)(\vec{\varphi}) \longrightarrow \exists^k \vec{u} \left[\left(\bigvee_{i=1}^{n} \varphi_i^* \right) \wedge P(\vec{u}) \right]$$

has a polynomial size proof. Since $\delta_P(B/A)(\vec{\varphi})$ is

$$\delta_P(C/A)(\vec{\varphi}) \vee \delta_P(D/A)(\vec{\varphi}),$$

the induction is finished by a $\vee$:left inference.

(iv) The case where B is $C \vee D$ is handled similarly to (iii).

□

Introduce new vectors of variables $\vec{r}_i$ of length k and let φ_i^* be $\vec{u}_i = \vec{r}_i$, for $1 \leq i \leq n$. Note that $\exists^k \vec{u} \, [(\bigvee_{i=1}^n \varphi_i^*) \wedge P(\vec{u})]$ is exactly $\exists^k_{n,T'} \vec{u} \, P(\vec{u})$, where $T' = \{\vec{r}_i | 1 \leq i \leq n\}$. The following corollary is immediate by $\supset$:right introduction and quantifying out the $\vec{r}_i$'s by kn $\forall$:right inferences.

Corollary 3.5 *Let* $\Psi_P^-(A)$ *be*

$$\forall^k \vec{z}_1 \cdots \forall^k \vec{z}_n \left[\delta_P(A)(\vec{t}_1 = \vec{z}_1, \ldots, \vec{t}_n = \vec{z}_n) \rightarrow \exists^k_{n,T} \vec{u} \, P(\vec{u}) \right],$$

where $T = \{\vec{z}_i | 1 \leq i \leq n\}$. *Then* $A \rightarrow \Psi_P^-(A)$ *has a proof of size polynomial in* $|A|$.

The subscript P will be suppressed if it is clear from context.

Note that $\Psi^-(A)$ only has one subformula of the form $P(\vec{t})$, so $\Psi^-(A)$ is a good candidate for the final goal of finding a formula equivalent to A that has only one occurrence of P. The question is now whether the reverse implication $\Psi^-(A) \rightarrow A$ holds. To show that as currently stated it does not, consider a model $\mathfrak{M}$ where P is constantly true. Then $\exists^k_{n,T} \vec{u} \, P(\vec{u})$ is true regardless of the $\vec{z}_i$'s, so $\Psi^-(A)$ is true. But if A happens to be false in $\mathfrak{M}$, then A is not equivalent to $\Psi^-(A)$. The next theorem shows that as long as P is false for some arguments, then the reverse implication holds.

Theorem 3.6 *There is a proof of size polynomial in* $|A|$ *of*

$$\exists \vec{x} \neg P(\vec{x}) \rightarrow (\Psi^-(A) \rightarrow A).$$

P r o o f. We argue informally as follows. Assume $\Psi^-(A)$. Let $\vec{w}_\perp$ be a vector of elements such that $\neg P(\vec{w}_\perp)$. Define the values $\vec{r}_i$ by

$$\vec{r}_i = \begin{cases} \vec{w}_\perp & \text{if } P(\vec{t}_i) \\ \vec{t}_i & \text{otherwise.} \end{cases}$$

Since $\neg P(\vec{r}_i)$, for $i = 1, \ldots, n$, $\exists^k_{n,T'} \vec{u} \, P(\vec{u})$ is false, where $T' = \{\vec{r}_i | 1 \leq i \leq n\}$. Let B be

$$\neg \delta_P(A)(\vec{t}_1 = \vec{r}_1, \ldots, \vec{t}_n = \vec{r}_n).$$

Then, by $\Psi^-(A)$, B is true. Put B in De Morgan normal form, which is exactly A with each occurrence of $P(\vec{t}_i)$ replaced with $\vec{t}_i \neq \vec{r}_i$. Finally, conclude A, since $P(\vec{t}_i) \leftrightarrow \vec{t}_i \neq \vec{r}_i$. □

After some modifications, the $\exists \vec{x} \neg P(\vec{x})$ appearing in Theorem 3.6 can be replaced with $\exists x_0 \exists x_1 (x_0 \neq x_1)$. The idea is to modify P so as to make it false on some values $\vec{w}_\perp$. Specifically, given a k-ary predicate P, form a $(k+1)$-ary predicate $\tilde{P}$, where $\tilde{P}(t_0, t_1, \ldots, t_k)$ is intended to express $t_0 = t_1 \wedge P(t_1, \ldots, t_k)$. It is clear that $\exists x_0 \exists x_1 (x_0 \neq x_1) \to \exists \vec{x} \neg \tilde{P}(\vec{x})$. It remains to equivalently formulate A using $\tilde{P}$ instead of P. To accomplish this, replace the occurrences of $P(t_{i,1}, t_{i,2}, \ldots, t_{i,k})$ in A by $\tilde{P}(t_{i,1}, t_{i,1}, t_{i,2}, \ldots, t_{i,k})$, for $i = 1, \ldots, n$, and call the resulting formula $\tilde{A}$. The main result of this section is ready to be proved.

Definition 3.7 Let φ be a first-order formula. Then $\vdash_{\geq 2} \varphi$ is an abbreviation for

$$\vdash \exists x_0 \exists x_1 (x_0 \neq x_1) \to \varphi.$$

Theorem 3.8 *Let P be a k-ary predicate and A be a quantifier-free formula. Assume that A contains n occurrences of P, all of which are positive. Then there exists a Π_2 formula $\Psi(A)$ with one (positive) occurrence of P such that*

$$\vdash_{\geq 2} \Psi(A) \leftrightarrow A$$

by a proof of size polynomial in $|A|$. Moreover, $\Psi(A)$ is of the form

$$\forall^{k+1} \vec{z}_1 \cdots \forall^{k+1} \vec{z}_n \, \exists^{k+1}_{n,T} \vec{u} \, \Psi_M(A),$$

where $\Psi_M(A)$ is quantifier-free and $T = \{\vec{z}_i | 1 \leq i \leq n\}$.

Proof. We view $\tilde{A}$ as a formula with $(k+1)$-variables, $\tilde{A} = \tilde{A}(x_0, \ldots, x_k)$, where the $x_1, \ldots, x_k$ are the free variables of A, and x_0 does not actually occur in $\tilde{A}$ at all. By Theorem 3.6, there is a polynomial size proof of

$$\exists \vec{x} \neg \tilde{P}(x_0, x_1, \ldots, x_k) \to (\Psi^-_{\tilde{P}}(\tilde{A}) \leftrightarrow \tilde{A}).$$

Replace all occurrences of $\tilde{P}(t_0, t_1, \ldots, t_k)$ in $\tilde{A}$ by $t_0 = t_1 \wedge P(t_1, \ldots, t_k)$, and call the resulting formula B. Do a similar replacement for $\Psi^-_{\tilde{P}}(\tilde{A})$, put the formula into prenex form, and call the resulting formula $\Psi(A)$. Then there is a polynomial size proof of

$$\exists \vec{x} \neg (x_0 = x_1 \wedge P(x_1, \ldots, x_k)) \to (\Psi(A) \leftrightarrow B).$$

Clearly there are polynomial size proofs for $\vdash B \leftrightarrow A$ and $\vdash_{\geq 2} \exists \vec{x} \neg (x_0 = x_1 \wedge P(x_1, \ldots, x_k))$. Thus there is also a polynomial size proof for $\vdash_{\geq 2} \Psi(A) \leftrightarrow A$. □

A dual result to Theorem 3.8 also holds.

Theorem 3.9 *Let P be a k-ary predicate and A be a quantifier-free formula. Assume that A contains n occurrences of P, all of which are positive. Then there exists a Σ_2 formula $\Theta(A)$ with one (positive) occurrence of P such that*

$$\vdash_{\geq 2} \Theta(A) \leftrightarrow A$$

by a proof of size polynomial in $|A|$. Moreover, $\Theta(A)$ is of the form

$$\exists^{k+1} \vec{z}_1 \cdots \exists^{k+1} \vec{z}_n \, \forall^{k+1}_{n,T} \vec{u} \, \Theta_M(A),$$

where $\Theta_M(A)$ is quantifier-free and $T = \{\vec{z}_i | 1 \leq i \leq n\}$.

Proof. Arguing informally, let B be $\neg A$ in De Morgan normal form. Replace subformulas of the form $\neg P(\vec{t})$ in B by $Q(\vec{t})$, where Q is a new predicate meant to express $\neg P$. Apply Theorem 3.8 to B with respect to Q to get $\Psi(B)$. Put $\neg \Psi(B)$ into De Morgan normal form, and replace the one occurrence of $\neg Q(\vec{t})$ by $P(\vec{t})$. Let $\Theta(A)$ be the resulting formula. □

Depending on the context in which it is being used, $\Psi(A)$ can be defined with slightly different properties. Note in Theorem 3.8 that the variables in the universal quantification only range over finitely many values; specifically, $\vec{z}_i$ is either $\vec{w}_\perp$ or $\vec{t}_i$. Thus it seems natural that the universal quantifiers in $\Psi(A)$ can be taken to be finite. However, finite quantification requires writing out all the terms that are being quantified over, and thus, to change the universal quantifiers in $\Psi(A)$ to finite universal quantifiers, extra assumptions must be made. Here are two possibilities.

(i) If there are terms $\vec{t}_\perp$ such that $\vdash \neg P(\vec{t}_\perp)$, then the universal quantifiers in $\Psi(A)$ can be changed to finite universal quantifiers. Specifically, $\Psi(A)$ would have the form

$$\forall^k_{2,S_1}\vec{z}_1 \cdots \forall^k_{2,S_n}\vec{z}_n\ \exists^k_{n,T}\vec{u}\ \Psi_M(A),$$

where $\Psi_M(A)$ is quantifier-free with one occurrence of P, S_i is $\{\vec{t}_i, \vec{t}_\perp\}$ and T is $\{t_i | 1 \leq i \leq n\}$. Note that it is unnecessary to introduce $\tilde{P}$ and that because $\vdash \neg P(\vec{t}_\perp)$, the proof of the equivalence between A and $\Psi(A)$ becomes unconditional.

(ii) If there are two unequal terms t_0, t_1, then the universal quantifiers in $\Psi(A)$ can be changed to finite universal quantifiers. Introduce $\tilde{P}$ as above. If $\vec{t}_\perp$ is $t_0, t_1, t_1, \ldots, t_1$, let $\Psi(A)$ be

$$\forall^{k+1}_{2,S_1}\vec{z}_1 \cdots \forall^{k+1}_{2,S_n}\vec{z}_n\ \exists^{k+1}_{n,T}\vec{u}\ \Psi_M(A),$$

where $\Psi_M(A)$ is quantifier-free with one occurrence of P, S_i is $\{\vec{t}_i, \vec{t}_\perp\}$ and T is $\{t_i | 1 \leq i \leq n\}$. Then

$$\vdash t_0 \neq t_1 \rightarrow (\Psi(A) \leftrightarrow A).$$

We do not examine the situation in (ii) in the rest of the paper. Although it seems advantageous to have both the universal and existential quantifiers in $\Psi(A)$ be finite, this does not improve the upper bounds we prove on quantifier complexity. There are two reasons for this. When proving our upper bounds, if the universal quantifiers in $\Psi(A)$ are adjacent to non-finite universal quantifiers, then the fact the universal quantifier in $\Psi(A)$ is finite becomes moot. When this situation does not arise, we prove our upper bound by the exchange property of Proposition 2.6, which does not require the outer quantifier be finite in order to move it inwards.

A dual discussion of (i) and (ii) holds for $\Theta(A)$.

4 Upper Bounds, Positive Occurrences

The general context for this section is the following. Let R_n be recursively defined as $A(R_{n-1})$, where R_{n-1} occurs only positively in $A(R_{n-1})$, and let R_n be obtained by unwinding the definition. In particular, R_n is linear size if A contains one occurrence of R_{n-1} and is exponential size if A contains multiple occurrences of R_{n-1}. This section shows how to construct polynomial size F_n equivalent to R_n, such that there are polynomial size proofs that the F_n's follow the same recursion as the R_n's. This result relies heavily on the construction of Section 3. Furthermore, F_n is equivalent to R_n, but this statement does not have a polynomial size proof (except in trivial cases) because R_n is exponentially large.

Our main goal in this section is to place upper bounds on the resulting quantifier complexity of the F_n's. The notion of quantifier complexity is made precise in the following definitions.

Definition 4.1 Let $k \geq 0$. A first-order formula is Σ_k (respectively, Π_k) if, after the application of prenex operations, its quantifier block starts with $\exists$ (respectively, $\forall$) and has at most k quantifier alternations.

A formula B is Δ_k for $k \geq 0$ if there are first order formulas α in Σ_k and β in Π_k such that there are proofs of $B \leftrightarrow \alpha$ and $B \leftrightarrow \beta$. A family of formulas $\{F_n\}$ is $\Delta_{k(n)}$ provided there are formulas $\alpha_n \in \Sigma_{k(n)}$ and $\beta_n \in \Pi_{k(n)}$ such that there are proofs of size polynomial in $|F_n|$ of $F_n \leftrightarrow \alpha_n$ and $F_n \leftrightarrow \beta_n$.

The condition of polynomial provability for Δ_k formulas in Definition 4.1 is important when R_n is defined as $A(R_{n-1})$ and A is Δ_k. In this case, A has two different expressions, and the polynomial size F_n's are defined using both expressions. Thus to prove the F_n's follow the same recursion as the R_n's via a polynomial size proof, A must be provably equivalent to both of its expressions by polynomial size proofs. Note that any Boolean combination of Σ_k and/or Π_k formulas is Δ_{k+1}.

The next definition counts quantifier alternations, modulo some formula φ.

Definition 4.2 Let φ be a first-order formula. Define the following sets of formulas $\Sigma_k(\varphi)$ and $\Pi_k(\varphi)$ for $k \geq 0$ as follows.

(i) If B is quantifier-free or a substitution instance of φ, then B is $\Sigma_0(\varphi)$ and $\Pi_0(\varphi)$.

(ii) If B and C are in $\Sigma_k(\varphi)$ (respectively, $\Pi_k(\varphi)$), then so are $B \wedge C$, $B \vee C$, and $\exists x B$ (respectively, $\forall x B$).

(iii) If B is in $\Sigma_k(\varphi)$ (respectively, $\Pi_k(\varphi)$), then $\neg B$ is in $\Pi_k(\varphi)$ (respectively, $\Sigma_k(\varphi)$).

(iv) If B is in $\Sigma_k(\varphi)$ (respectively, $\Pi_k(\varphi)$), then $\forall x B$ (respectively, $\exists x B$) is in $\Pi_{k+1}(\varphi)$ (respectively, $\Sigma_{k+1}(\varphi)$).

A formula B is $\Delta_k(\varphi)$ for $k \geq 0$ if there are first-order formulas α in $\Sigma_k(\varphi)$ and β in $\Pi_k(\varphi)$ such that there are proofs of $B \leftrightarrow \alpha$ and $B \leftrightarrow \beta$. Intuitively, B is $\Sigma_k(\varphi)$ if B is Σ_k after replacing each occurrence of φ with its own new predicate symbol.

In Theorem 4.3 we consider the case where A is Σ_{2k}, $k > 0$.

If A is a formula containing the predicate symbol Q, and P is a new predicate symbol such that P and Q have the same arity, then $A(P)$ is obtained by replacing all subformulas of the form $Q(\vec{t})$ in A by $P(\vec{t})$. Therefore, the different occurrences of P in $A(P)$ may have different terms as arguments. On the other hand, Ψ and Θ act as operators. Thus $\Psi(A)$ and $\Theta(A)$ are formulas obtained from A by the process of Section 3, and do not indicate substitution as in the notation $A(P)$.

Theorem 4.3 *Suppose R_n is recursively defined as $A(R_{n-1})$ with R_{n-1} occurring only positively in A. If A is Σ_{2k} for $k > 0$, then there are formulas F_n such that F_0 is R_0 and for $n > 0$, F_n is $\Sigma_{2kn}(R_0)$ and $\vdash_{\geq 2} F_n \leftrightarrow A(F_{n-1})$ by a proof of size polynomial in $|A|$, n, and $|R_0|$. Moreover, $\vdash_{\geq 2} F_n \leftrightarrow R_n$.*

Proof. The proof proceeds by induction on n. Let A contain occurrences of a new predicate P such that $A(R_{n-1})$ is obtained by replacing, in A, subformulas of the form $P(\vec{t})$ with $R_{n-1}(\vec{t})$. Also, suppose w.l.o.g. that A is of the form

$$\exists \vec{x}_1 \forall \vec{x}_2 \cdots \exists \vec{x}_{2k-1} \forall \vec{x}_{2k} A_M,$$

where A_M is quantifier-free. For the base case, take F_0 to be R_0. For $n > 0$, assume F_{n-1} has the stated properties, and define F_n to be

$$\exists \vec{x}_1 \forall \vec{x}_2 \cdots \exists \vec{x}_{2k-1} \forall \vec{x}_{2k} \Psi_P(A_M)(F_{n-1}).$$

F_n polynomial size because it is defined by a formula with one instance of F_{n-1}. By Theorem 3.8,

$$\vdash_{\geq 2} \Psi_P(A_M) \leftrightarrow A_M$$

by a polynomial size proof. Then $\vdash_{\geq 2} \Psi_P(A_M)(F_{n-1}) \leftrightarrow A_M(F_{n-1})$, and hence $\vdash_{\geq 2} F_n \leftrightarrow A(F_{n-1})$, by polynomial size proofs.

It remains to show that F_n is $\Sigma_{2kn}(R_0)$, for $n > 0$. By using the form of $\Psi_P(A_M)$ given in Theorem 3.8, F_n is

$$\exists \vec{x}_1 \forall \vec{x}_2 \cdots \exists \vec{x}_{2k-1} \forall \vec{x}_{2k} \forall \vec{z}\, \exists_m \vec{u}\, \Psi_M(A_M)(F_{n-1}),$$

where $\Psi_M(A_M)$ is quantifier-free, and m is the size of the finite set of terms $\vec{u}$ ranges over. If $n = 1$, for a one time increase of size, the $\exists_m \vec{u}$ in F_1 can be expanded as a disjunction of size m, so that F_1 is $\Sigma_{2k}(R_0)$. If $n > 1$, assume by induction that F_{n-1} is $\Sigma_{2k(n-1)}(R_0)$. Since $\Psi_M(A_M)(F_{n-1})$ contains only positive occurrences (one, in fact) of F_{n-1}, it is $\Sigma_{2k(n-1)}$. Thus the leftmost $\exists_m \vec{u}$ in F_n adds no complexity because the outermost quantifier of $\Psi_M(A_M)(F_{n-1})$ is also $\exists$. Thus F_n is $\Sigma_{2kn}(R_0)$. □

A dual result holds when A is Π_{2k} with $k > 0$. The argument remains the same, except that Θ is used instead of Ψ:

Theorem 4.4 *Suppose R_n is recursively defined as $A(R_{n-1})$ with R_{n-1} occurring only positively in A. If A is Π_{2k} for $k > 0$, then there are formulas F_n such that F_0 is R_0 and for $n > 0$, F_n is $\Pi_{2kn}(R_0)$ and $\vdash_{\geq 2} F_n \leftrightarrow A(F_{n-1})$ by a proof of size polynomial in $|A|$, n, and $|R_0|$. Moreover, $\vdash_{\geq 2} F_n \leftrightarrow R_n$.*

The fact that the quantifier block of A was even in length seemingly played an important role in Theorem 4.3 because the quantifiers introduced by the application of Ψ were able to be absorbed into the first and last quantifier blocks of A. Suppose A is Σ_{2k+1} for $k > 0$. Then the quantifier block of A would begin and end with the same type of quantifier. If the argument of Theorem 4.3 were applied, only one of the quantifier types of $\Psi_P(A_M)(F_{n-1})$ could be absorbed into quantifier blocks introduced of A, resulting in an extra n quantifier alternations in F_n. These extra alternations could be eliminated using the quantifier exchange property of Proposition 2.6. Instead, however, we prove the following theorem by reducing it to Theorem 4.4.

Theorem 4.5 *Suppose R_n is recursively defined as $A(R_{n-1})$ with R_{n-1} occurring only positively in A. If A is Σ_{2k+1} for $k > 0$, then there are formulas F_n such that F_0 is R_0 and for $n > 0$, F_n is $\Sigma_{2kn+1}(R_0)$ and $\vdash_{\geq 2} F_n \leftrightarrow A(F_{n-1})$ by a proof of size polynomial in $|A|$, n, and $|R_0|$. Moreover, $\vdash_{\geq 2} F_n \leftrightarrow R_n$.*

Proof. Suppose that $A(R_{n-1})$ in prenex form is

$$\exists \vec{x}_1 \forall \vec{x}_2 \exists \vec{x}_3 \cdots \forall \vec{x}_{2k} \exists \vec{x}_{2k+1} A_M(R_{n-1}),$$

where A_M is quantifier-free. For $n \geq 0$, define S_n to be

$$\forall \vec{x}_2 \exists \vec{x}_3 \cdots \forall \vec{x}_{2k} \exists \vec{x}_{2k+1} A_M(R_n),$$

then R_n is $\exists \vec{x}_1 S_{n-1}$ and S_n is

$$\forall \vec{x}_2 \exists \vec{x}_3 \cdots \forall \vec{x}_{2k} \exists \vec{x}_{2k+1} A_M(\exists \vec{z} S_{n-1}).$$

Since the occurrences of $\exists \vec{z} S_{n-1}$ are all positive in $A_M(\exists \vec{z} S_{n-1})$, S_n can be equivalently rewritten as

$$\forall \vec{x}_2 \exists \vec{x}_3 \cdots \forall \vec{x}_{2k} \exists \vec{x}_{2k+1} \exists \vec{z}_1 \cdots \exists \vec{z}_l A_M(S_{n-1}),$$

where l is the number of occurrences of R_{n-1} in A. Let $B(S_{n-1})$ be this formula. Apply Theorem 4.4 to B and S_n to get polynomial size formulas G_n in $\Pi_{2kn}(R_0)$ such that $\vdash_{\geq 2} G_n \leftrightarrow B(G_{n-1})$ by a polynomial size proof. The theorem is proved by letting F_0 be R_0 and F_n be $\exists \vec{x}_1 G_n$ for $n > 0$. □

Theorem 4.5 has a dual argument that uses Theorem 4.3 instead of Theorem 4.4:

Theorem 4.6 *Suppose R_n is recursively defined as $A(R_{n-1})$ with R_{n-1} occurring only positively in A. If A is Π_{2k+1} for $k > 0$, then there are formulas F_n such that F_0 is R_0 and for $n > 0$, F_n is $\Pi_{2kn+1}(R_0)$ and $\vdash_{\geq 2} F_n \leftrightarrow A(F_{n-1})$ by a proof of size polynomial in $|A|$, n, and $|R_0|$. Moreover, $\vdash_{\geq 2} F_n \leftrightarrow R_n$.*

Now consider the cases where A is purely universal, purely existential, or quantifier-free. The above arguments fail in these cases, since they depended on the fact that both universal and existential quantifiers were present in A so they could be combined with the quantifiers introduced by applications of Ψ and Θ. The next theorem uses the quantifier exchange property of Propositions 2.5 and 2.6 to give a better bound on the quantifier complexity of F_n.

Theorem 4.7 *Suppose R_n is recursively defined as $A(R_{n-1})$ with R_{n-1} occurring only positively in A. Fix $\delta > 0$. If A is purely existential, purely universal, or quantifier-free, then there are formulas F_n such that F_0 is R_0 and for $n > 0$, F_n is $\Delta_{\delta n/\log n}(R_0)$ and $\vdash_{\geq 2} F_n \leftrightarrow A(F_{n-1})$ by a proof of size polynomial in $|A|$, n, and $|R_0|$. Moreover, $\vdash_{\geq 2} F_n \leftrightarrow R_n$.*

Proof. First assume that A is purely existential. Let A contain occurrences of a new predicate P such that $A(R_n)$ is obtained by replacing, in A, subformulas of the form $P(\vec{t})$ with $R_n(\vec{t})$. Suppose A is of the form $\exists \vec{x}\, A_M$, where A_M is quantifier-free, and let F_0^- and F_0 both be R_0. For $n > 0$, let F_n^- be $\exists^l \vec{x}\, \Theta_P(A_M)(F_{n-1}^-)$. For $n > 0$, F_n will be constructed below from F_n^- using the polynomial size equivalences in Propositions 2.4 and 2.6. Because the size bounds in Propositions 2.4 and 2.6 do not depend on the lengths of the vectors of variables or the set of terms that the finite quantifiers range over, only the sizes of the sets a finite quantifier ranges over will be displayed. Often, the quantifying variables will be suppressed to improve readability.

By Theorem 3.9, $\vdash_{\geq 2} \Theta_P(A_M) \leftrightarrow A_M$ by a polynomial size proof, and hence $\vdash_{\geq 2} F_n^- \leftrightarrow A(F_{n-1}^-)$ by a polynomial size proof. If m is the number of occurrences of R_{n-1} in $A(R_{n-1})$, then Theorem 3.9 states that

$\Theta_P(A_M)(F^-_{n-1})$ is $\exists\vec{z}\forall_m\vec{u}\Theta_M(A_M)(F^-_{n-1})$, where $\Theta_M(A_M)$ is quantifier-free and has one, positive occurrence of P. Then F^-_n has the form $\exists\vec{w}\forall_m\vec{u}\,\Theta_M(A_M)(F^-_{n-1})$, where $\vec{w}$ is $(\vec{x},\vec{z})$. By unwinding its recursive definition, F^-_n can be written as

$$\exists\vec{w}_1\forall_m\vec{u}_1\cdots\exists\vec{w}_n\forall_m\vec{u}_n\,B,$$

where B is a polynomial sized, $\Delta_0(R_0)$ formula with one, positive occurrence of R_0. Let D_0 be this formula.

The construction proceeds in stages moving finite quantifiers rightwards past unlike quantifiers and then combining like, finite quantifiers, as in Propositions 2.3-2.6, at each stage cutting the number of quantifier alternations in half. For simplicity we assume that n is a power of two; this assumption can removed by padding with vacuous quantifiers. The construction produces formulas D_i in $\Sigma_{n/2^{i-1}}(R_0)$ with one, positive occurrence of R_0 such that D_i is of the form

$$\exists\forall_{m^{2^i}}\exists\forall_{m^{2^i}}\cdots\exists\forall_{m^{2^i}}\varphi,$$

where φ is $\Delta_0(R_0)$, and for $i>0$, $\vdash D_i\leftrightarrow D_{i-1}$ by a proof of size polynomial in $|D_i|$. Note that D_0 satisfies the base case, so it is enough to show how to construct D_{i+1} from D_i.

Suppose D_i has been constructed with the stated properties and is of the form

$$\exists\forall_{m^{2^i}}\exists\forall_{m^{2^i}}\exists\forall_{m^{2^i}}\exists\forall_{m^{2^i}}\cdots\exists\forall_{m^{2^i}}\exists\forall_{m^{2^i}}\exists\forall_{m^{2^i}}\exists\forall_{m^{2^i}}\varphi,$$

where there are $n/2^{i-1}$ quantifier alternations and φ is $\Delta_0(R_0)$. Use Proposition 2.6 to simultaneously change the first, third, fifth, etc., occurrences of $\forall_{m^{2^i}}\exists$ to be $\exists\forall_{m^{2^i}}$. This results in a formula of the form

$$\exists\exists\forall_{m^{2^i}}\forall_{m^{2^i}}\exists\exists\forall_{m^{2^i}}\forall_{m^{2^i}}\cdots\exists\exists\forall_{m^{2^i}}\forall_{m^{2^i}}\exists\exists\forall_{m^{2^i}}\forall_{m^{2^i}}\varphi.$$

The application of Proposition 2.6 multiplies the size by Cm^{2^i}, where C is a fixed constant. Simultaneously combine the like universal quantifiers using Proposition 2.4, which multiplies the size by a factor of m^{2^i}, and call the resulting formula D_{i+1}. Then D_{i+1} is clearly $\Sigma_{n/2^i}(R_0)$ and, by Propositions 2.4 and 2.6, $\vdash D_i\leftrightarrow D_{i+1}$ by a proof of size polynomial in $|D_i|$. Also, $|D_{i+1}|\le Cm^{2^i}m^{2^i}|D_i| = Cm^{2^{i+1}}|D_i|$, and hence

$$|D_j|\le|D_0|\Pi_{i=0}^{j-1}Cm^{2^{i+1}} = C^jm^{2^{j+1}-2}|D_0|.$$

Fix a constant c. Let j be $\log\log n+c+1$ and F_n be D_j. Then F_n is in $\Sigma_d(R_0)$, where $d=\frac{n}{2^c\log n}$, and since D_0 is F^-_n, F_n has polynomial size. By fixing c large enough, $d\le\delta\frac{n}{\log n}-1$. Since Σ_k is contained in Δ_{k+1} for any k, F_n is in $\Delta_{\delta\frac{n}{\log n}}(R_0)$. By combining the proofs of $D_i\leftrightarrow D_{i+1}$, $i=1,\ldots,j$, there is a polynomial size proof of $F^-_n\leftrightarrow F_n$ (since D_0 is F^-_n and D_j is F_n). Since there are polynomial size proofs for $\vdash_{\ge2}F^-_n\leftrightarrow F_n$ and $\vdash_{\ge2}F^-_n\leftrightarrow A(F^-_{n-1})$, $\vdash_{\ge2}F_n\leftrightarrow A(F_{n-1})$ by a polynomial size proof.

Taking the dual of the above argument shows the same result holds if A is purely universal. A quantifier-free formula can be taken to be purely existential (or purely universal) by adding a vacuous quantifier. □

The last case to consider is when A is Δ_k. If $k=0$, then A is quantifier-free and Theorem 4.7 applies. If $k=1$, then A can be taken to be, in particular, purely existential, and Theorem 4.7 again applies. The next theorem considers the case $k>1$. This will be useful in Section 6 when removing the assumption of only positive occurrences of R_{n-1} .

Theorem 4.8 *Suppose R_n is recursively defined as $A(R_{n-1})$ with R_{n-1} occurring only positively in A. Let $k>0$. Suppose A is Δ_{k+1}, and that the two formulas $\alpha\in\Sigma_{k+1}$ and $\beta\in\Pi_{k+1}$ equivalent to $A(P)$ contain only positive occurrences of P. Then there are formulas F_n such that F_0 is R_0 and for $n>0$, F_n is $\Delta_{kn+1}(R_0)$ and $\vdash_{\ge2}F_n\leftrightarrow A(F_{n-1})$ by a proof of size polynomial in $|A|$, n, and $|R_0|$. Moreover, $\vdash_{\ge2}F_n\leftrightarrow R_n$.*

Proof. Let A contain occurrences of a new predicate P such that $A(R_n)$ is obtained by replacing, in A, subformulas of the form $P(\vec{t})$ with $R_n(\vec{t})$. Suppose there are polynomial size proofs of $A\leftrightarrow\alpha$ and $A\leftrightarrow\beta$, where α is Σ_{k+1} and β is Π_{k+1}. Let α be

$$\exists\vec{x}_1\forall\vec{x}_2\cdots Q\vec{x}_{k+1}\alpha_M,$$

where Q is $\exists$ if k is even and is $\forall$ if k is odd, and α_M is quantifier-free. Let β be

$$\forall \vec{y}_1 \exists \vec{y}_2 \cdots \overline{Q} \vec{y}_{k+1} \beta_M,$$

where $\overline{Q}$ is $\forall$ if k is even and is $\exists$ if k is odd, and β_M is quantifier-free.

First assume that k is odd. Define two sets of formulas, F_n^- and G_n^-, that will be transformed below into the F_n's and G_n's. Let F_0^- be R_0 and F_n^- be

$$\exists \vec{x}_1 \forall \vec{x}_2 \cdots \forall \vec{x}_{k+1} \Psi_P(\alpha_M)(G_{n-1}^-)$$

for $n > 0$. Let G_0^- be R_0 and G_n^- be

$$\forall \vec{y}_1 \exists \vec{y}_2 \cdots \exists \vec{y}_{k+1} \Theta_P(\beta_M)(F_{n-1}^-)$$

for $n > 0$. Let p be the number of occurrences of R_{n-1} in $\alpha(R_{n-1})$ and q be the number of occurrences of R_{n-1} in $\beta(R_{n-1})$. By Theorem 3.8, $\Psi_P(\alpha_M)$ is of the form $\forall \vec{z}\, \exists_p \vec{u}\, \Psi_M(\alpha_M)$, where $\Psi_M(\alpha_M)$ is quantifier-free. Further,

$$\vdash_{\geq 2} \Psi_P(\alpha_M) \leftrightarrow \alpha_M,$$

and thus

$$\vdash_{\geq 2} F_n^- \leftrightarrow \alpha(G_{n-1}^-),$$

by polynomial size proofs. Similarly, by Theorem 3.9, $\Theta(\beta_M)$ is of the form $\exists \vec{z} \forall_q \vec{u}\, \Theta_M(\beta_M)$, where $\Theta_M(\beta_M)$ is quantifier-free. Further,

$$\vdash_{\geq 2} \Theta_P(\beta_M) \leftrightarrow \beta_M,$$

and thus

$$\vdash_{\geq 2} G_n^- \leftrightarrow \beta(F_{n-1}^-)$$

by polynomial size proofs. Since F_0^- and G_0^- are R_0 and $\vdash \alpha \leftrightarrow \beta$ by a polynomial size proof, induction shows that for all n, $\vdash_{\geq 2} F_n^- \leftrightarrow G_n^-$ by a polynomial size proof. Then $\vdash_{\geq 2} F_n^- \leftrightarrow A(F_{n-1}^-)$, because there are polynomial size proofs of $A \leftrightarrow \beta$ and $A \leftrightarrow \alpha$.

If $n > 0$, F_n^- has one occurrence of G_{n-1}^- and is of the form

$$\exists \vec{x}_1 \forall \vec{x}_2 \cdots \forall \vec{x}_{k+1} \forall \vec{z} \exists_m \vec{u} \Psi_M(A_M)(G_{n-1}^-).$$

Similarly if $n > 0$, G_n^- has one occurrence of F_{n-1}^- and is of the form

$$\forall \vec{y}_1 \exists \vec{y}_2 \cdots \exists \vec{y}_{k+1} \exists \vec{z} \forall_q \vec{u} \Theta_M(A_M)(F_{n-1}^-).$$

Let $n > 0$. The unwinding the recursion shows that F_n^- is a polynomial size formula with quantifier block

$$\exists \forall \exists \cdots \exists \forall \exists_p | \forall \exists \forall \cdots \forall \exists \forall_q | \exists \forall \exists \cdots \exists \forall \exists_p | \forall \exists \forall \cdots \forall \exists \forall_q | \exists \forall \cdots \overline{Q} Q_m.$$

The last quantifiers $\overline{Q}Q_m$ are either $\forall \exists_p$ or $\exists \forall_q$, depending on whether n is even or odd. Call each set of quantifiers set apart by vertical lines a cell. Each cell is created after one stage of the unwinding of F_n^-, thus there are n cells. Each cell contains $k+2$ quantifiers, corresponding to the sizes of the quantifier blocks in the recursions. Note that the rightmost quantifier in every cell is finite. Apply Proposition 2.6 to simultaneously exchange the quantifiers on either side of each of the $n-1$ vertical lines. This increases the size of the formula by a factor of $\max\{p, q\}$, and the quantifier block of the resulting formula is

$$\exists \forall \exists \cdots \exists \forall \forall | \exists_p \exists \forall \cdots \forall \exists \exists | \forall_q \forall \exists \cdots \exists \forall \forall | \exists_p \exists \forall \cdots \forall \exists \exists | \forall_q \forall \cdots \overline{Q} Q_m.$$

The innermost quantifier Q_m is still finite, so for a one time size cost expand it as a conjunction of size q (if it is universal) or disjunction of size p (if it is existential) and call the resulting formula F_n^+. Then the first cell of F_n^+'s quantifiers has $k+1$ alternations, the last cell has k alternations, and each of the $n-2$ middle cells have k alternations, so that F_n^+ is $\Sigma_{kn+1}(R_0)$. Proposition 2.6 implies that $\vdash F_n^+ \leftrightarrow F_n^-$ by a polynomial size proof. If $n=0$, let F_0^+ be R_0. Then we have a polynomial size proof for $\vdash_{\geq 2} F_n^+ \leftrightarrow A(F_{n-1}^+)$.

All that is left to show is that F_n^+ is actually Δ_{kn+1}. This is accomplished by dually defining formulas G_n^+ in Π_{kn+1} such that G_0^+ is R_0 and $\vdash G_n^+ \leftrightarrow G_n^-$ by a polynomial size proof. Since $\vdash_{\geq 2} F_n^- \leftrightarrow G_n^-$, for $n>0$, $\vdash_{\geq 2} F_n^+ \leftrightarrow G_n^+$ by a polynomial size proof. If $n=0$, $\vdash F_0^+ \leftrightarrow G_0^+$ because they are both R_0. This would show that F_n^+ is $\Delta_{kn+1}(R_0)$ if the assumption of two distinct elements were removed.

To this end, for $n>0$, let F_n be

$$F_n^+ \vee \forall x_0 \forall x_1 (x_0 = x_1),$$

let G_n be

$$G_n^+ \vee \forall x_0 \forall x_1 (x_0 = x_1),$$

and let F_0 and G_0 be R_0. Clearly, F_n is $\Sigma_{kn+1}(R_0)$, G_n is $\Pi_{kn+1}(R_0)$, and $\vdash F_n \leftrightarrow G_n$ by a polynomial size proof. Therefore, F_n is $\Delta_{kn+1}(R_0)$. It is also clear that there is a polynomial proof for $\vdash_{\geq 2} F_n \leftrightarrow A(F_{n-1})$.

If k is even, the proof is slightly easier because F_n^- can instead be defined as $\exists \vec{x}_1 \forall \vec{x}_2 \cdots \exists \vec{x}_{k+1} \Psi_P(\alpha_M)(F_{n-1}^-)$ for $n>0$. Similarly, G_n^- can be instead defined as $\forall \vec{y}_1 \exists \vec{y}_2 \cdots \forall \vec{y}_{k+1} \Theta_P(\beta_M)(G_{n-1}^-)$ for $n>0$. The rest of the proof is similar to the case when k is even, and so is omitted. □

Section 6 removes the condition that α and β contain only positive occurrences of P.

5 Lower Bounds, Positive Occurrences

The present section places lower bounds on the quantifier complexity of polynomial size formulas expressing some recursively defined predicate R_n. Once again we assume R_{n-1} occurs only positively in the definition of R_n. These lower bounds closely match the upper bounds that were proved in Section 4.

There are two ideas in proving lower bounds. One is to use complete problems from the arithmetic hierarchy, and the other is to use lower bounds of the Yao-Håstad switching lemma on circuits calculating parity. In either case, the main effort is in showing that these problems can be defined recursively by a formula with the desired quantifier complexity.

Let $T(e,x,u)$ be the Kleene T predicate which expresses the statement that u codes a computation of Turing machine with Gödel number e on input x. Sequences $x_1, \ldots, x_m$ will be coded by Gödel numbers $\langle x_1, \ldots, x_m \rangle$. Let $*$ denote concatenation of two sequences and *Accept*(u) denote the statement that u codes an accepting computation. Let $\beta(i, \langle x_1, \ldots, x_m \rangle)$ be the Gödel β function that returns the i^{th} element of the sequence. Let $\mathfrak{N}$ be the standard model of the integers in the language that contains all primitive recursive functions and predicates. Since the setting for equivalence is in models with more than one element, it is enough to give lower bounds on the quantifier complexity of a formula expressing R_n over $\mathfrak{N}$.

The following two definitions simplify notation in the following.

Definition 5.1 Let $m>0$. Let $\mu(e,x,y_1,y_2,\ldots,y_{m-1},y_m)$ be the formula

$$T(e, \langle x, y_1, y_2, \ldots, y_{m-1}, \beta(1,y_m) \rangle, \beta(2,y_m)) \rightarrow \textit{Accept}(\beta(2,y_m)).$$

Define $\nu(e,x,y_1,y_2,\ldots,y_{m-1},y_m)$ to be

$$T(e, \langle x, y_1, y_2, \ldots, y_{m-1}, \beta(1,y_m) \rangle, \beta(2,y_m)) \wedge \textit{Accept}(\beta(2,y_m)).$$

Define $\mathfrak{A}^m(e,x)$ to be

$$\exists y_1 \forall y_2 \cdots Q y_m \pi(e,x,y_1,y_2,\ldots,y_m),$$

where Q is $\forall$ and π is μ if m is even, and Q is $\exists$ and π is ν if m is odd. Define $\mathfrak{B}^m(e,x)$ by

$$\forall y_1 \exists y_2 \cdots Q y_m \pi(e,x,y_1,y_2,\ldots,y_m),$$

where Q is $\exists$ and π is ν if m is even, and Q is $\forall$ and π is μ if m is even.

Clearly, $\mathfrak{A}^m$ and $\mathfrak{B}^m$ are Σ_m- and Π_m-complete, respectively.

The first lower bound we prove is the case when is A in Σ_{2k}, $k > 0$.

Definition 5.2 Let $C_0^{2k}(e, x, u)$ be the formula

$$T(e, x, u) \to \mathit{Accept}(u).$$

For $n > 0$, let $C_n^{2k}(e, x, u)$ be the formula

$$\exists x_1 \forall x_2 \cdots \exists x_{2k-1} \forall x_{2k} C_{n-1}^{2k}(e, \langle x_1, x_2, \ldots, x_{2k-1}, \beta(1, x_{2k})\rangle * x, \beta(2, x_{2k})).$$

Note that, unless $n = 0$, the argument u is ignored. If $k, n > 0$, then $C_n^{2k}(e, x, u)$ is equivalent to $\mathfrak{A}^{2kn}(e, x)$, and thus is Σ_{2kn}-complete. This proves the following lower bound.

Theorem 5.3 *Let $k > 0$. Then there exist formulas C_n^{2k} recursively defined to be $A(C_{n-1}^{2k})$, where A is Σ_{2k}, C_0^{2k} is quantifier-free, and C_{n-1}^{2k} occurs only positively in $A(C_{n-1}^{2k})$, such that if $n > 0$ and $\mathfrak{N} \vDash \varphi \leftrightarrow C_n^{2k}$, then $\varphi \notin \Pi_{2kn}$. Hence $\varphi \notin \Sigma_l$ for $l < 2kn$.*

There is a dual construction for the case when A is Π_{2k}, $k > 0$. Let $D_n^{2k}(e, x, u)$ be $T(e, x, u) \wedge \mathit{Accept}(u)$ if $n = 0$, and

$$\forall x_1 \exists x_2 \cdots \forall x_{2k-1} \exists x_{2k} D_{n-1}^{2k}(e, \langle x_1, x_2, \ldots, x_{2k-1}, \beta(1, x_{2k})\rangle * x, \beta(2, x_{2k}))$$

if $n > 0$. Then $D_n^{2k}(e, x, u)$ is equivalent to $\mathfrak{B}^{2kn}(e, x)$ when $k, n > 0$, and hence $D_n^{2k}(e, x, u)$ is Π_{2kn}-complete. The following theorem is immediate.

Theorem 5.4 *Let $k > 0$. Then there exist formulas D_n^{2k} recursively defined to be $A(D_{n-1}^{2k})$, where A is Π_{2k}, D_0^{2k} is quantifier-free, and D_{n-1}^{2k} occurs only positively in $A(D_{n-1}^{2k})$, such that if $n > 0$ and $\mathfrak{N} \vDash \varphi \leftrightarrow D_n^{2k}$, then $\varphi \notin \Sigma_{2kn}$. Hence $\varphi \notin \Pi_l$ for $l < 2kn$.*

If A is Σ_{2k+1}, $k > 0$, a lower bound can be proved after slightly modifying C_n^{2k}. The case when A is Σ_1 is handled separately.

Definition 5.5 Let $k > 0$ be fixed, and let $C_0^{2k+1}(e, x, u)$ be the formula

$$T(e, x, u) \wedge \mathit{Accept}(u).$$

For $n > 0$ let $C_n^{2k+1}(e, x, u)$ be the formula

$$\exists x_1 \forall x_2 \cdots \forall x_{2k} \exists x_{2k+1} C_{n-1}^{2k+1}(e, \langle x_1, \ldots, x_{2k}, \beta(1, x_{2k+1})\rangle * x, \beta(2, x_{2k+1})).$$

Unwinding the definition shows that $C_n^{2k}(e, x, u)$ is equivalent to $\mathfrak{A}^{2kn+1}$, when $k, n > 0$. The same reasoning as before Theorem 5.3 proves a lower bound for the odd block length case.

Theorem 5.6 *Let $k > 0$. Then there exist formulas C_n^{2k+1} recursively defined to be $A(C_{n-1}^{2k+1})$, where A is Σ_{2k+1}, C_0^{2k+1} is quantifier-free, and C_{n-1}^{2k+1} occurs only positively in $A(C_{n-1}^{2k+1})$, such that if $n > 0$ and $\mathfrak{N} \vDash \varphi \leftrightarrow C_n^{2k+1}$, then $\varphi \notin \Pi_{2kn+1}$. Hence $\varphi \notin \Sigma_l$ for $l < 2kn + 1$.*

By constructing D_n^{2k+1} dual to C_n^{2k+1}, $D_n^{2k+1}(e, x, u)$ is $\mathfrak{B}^{2kn+1}(e, x)$ when $k, n > 0$, hence is Π_{2kn+1}-complete, and the lower bound is proved as in the previous cases:

Theorem 5.7 *Let $k > 0$. Then there exist formulas D_n^{2k+1} recursively defined to be $A(D_{n-1}^{2k+1})$, where A is Π_{2k+1}, D_0^{2k+1} is quantifier-free, and D_{n-1}^{2k+1} occurs only positively in $A(D_{n-1}^{2k+1})$, such that if $n > 0$ and $\mathfrak{N} \vDash \varphi \leftrightarrow D_n^{2k+1}$, then $\varphi \notin \Sigma_{2kn+1}$. Hence $\varphi \notin \Pi_l$ for $l < 2kn + 1$.*

The next case to consider is when A is Δ_{k+1}, $k > 0$. Here, k is required to be strictly greater than zero; the case when A is Δ_1 is open. The above constructions can be altered to produce formulas that incorporate aspects of both C_n^k and D_n^k. The construction is dependent on whether k is even or odd; only the even case is presented. Let $k > 0$. Define $E_0^{2k}(e, x, u, a)$ to be

$$(a = 0 \wedge (T(e, x, u) \to \mathit{Accept}(u))) \vee (a \neq 0 \wedge (T(e, x, u) \wedge \mathit{Accept}(u))).$$

For $n > 0$, define $E_n^{2k}(e, x, u, a)$ to be

$$(a = 0 \wedge \exists y_1 \forall y_2 \cdots \exists y_{2k-1} \forall y_{2k} E_{n-1}^{2k}(e, \langle y_1, y_2, \ldots, y_{2k-1}, \beta(1, y_{2k})\rangle * x, \beta(2, y_{2k}), 0))$$
$$\vee (a \neq 0 \wedge \forall z_1 \exists z_2 \cdots \forall z_{2k-1} \exists z_{2k} E_{n-1}^{2k}(e, \langle z_1, z_2, \ldots, z_{2k-1}, \beta(1, z_{2k})\rangle * x, \beta(2, z_{2k}), 1)).$$

The argument a in E_n^{2k} can be thought of as a flag that chooses which path of the recursion to follow. Clearly, E_n^{2k} is defined recursively to be $A(E_{n-1}^{2k})$, where A is some Δ_{2k+1} formula. By unwinding the recursion, $E_n^{2k}(e, x, u, a)$ is equivalent to $(a = 0 \wedge \mathfrak{A}^{2kn}(e, x)) \vee (a \neq 0 \wedge \mathfrak{B}^{2kn}(e, x))$. It is straightforward to rework the recursion in the odd case so that if $k > 0$, then $E_n^{2k+1}(e, x, u, a)$ is recursively defined to be a $\Delta_{(2k+1)+1}$ formula such that $E_n^{2k+1}(e, x, u, a)$ is equivalent to $(a = 0 \wedge \mathfrak{A}^{(2k+1)n}(e, x)) \vee (a \neq 0 \wedge \mathfrak{B}^{(2k+1)n}(e, x))$. The next theorem follows easily.

Theorem 5.8 *Let $k > 0$. Then there exist formulas E_n^{k+1} recursively defined to be $A(E_{n-1}^{k+1})$, where A is Δ_{k+1}, E_0^{k+1} is quantifier-free, and E_{n-1}^{k+1} occurs only positively in $A(E_{n-1}^{k+1})$, with the property that, if $n > 0$ and $\mathfrak{N} \vDash \varphi \leftrightarrow E_n^{k+1}$, then $\varphi \notin \Sigma_{kn}$ and $\varphi \notin \Pi_{kn}$.*

Proof. Let $n > 0$ and suppose $E_n^{k+1}(e, x, u, a)$ is equivalent to φ. If $a = 0$, then φ is equivalent to $\mathfrak{A}^{kn}$, and so $\varphi \notin \Pi_{kn}$. If $a = 1$, then φ is $\mathfrak{B}^{kn}$, and so $\varphi \notin \Sigma_{kn}$. □

The only lower bounds left to prove are when A is purely universal, purely existential, or quantifier-free. In all cases the lower bound arises from a recursive definition of the parity function. Without function symbols in the language the recursion is either Σ_1 or Π_1, and with function symbols in the language the recursion is quantifier-free. The lower bound rests on the Yao-Håstad Switching Lemma [8, 9], which places a lower bound on the depth of a circuit calculating parity.

For Theorem 5.9, let the language contain the predicates $=, T$, and *Mid* and the constant symbols 0 and 1. T is a unary predicate and *Mid*(i, j, k) is a ternary relation intended to express $k = \lfloor \frac{i+j}{2} \rfloor$. In the cases where A is purely universal or purely existential, function symbols are not allowed. In the quantifier-free case, function symbols are allowed, and the one function used is the binary function *Mid*(i, j), which is intended to calculate $\lfloor \frac{i+j}{2} \rfloor$.

Theorem 5.9 *There exist formulas P_N^* recursively defined to be $A(P_{N-1}^*)$, where A is purely existential, P_0^* is quantifier-free, and P_{N-1}^* occurs only positively in $A(P_{N-1}^*)$, such that the following property holds. If φ_N are formulas of size polynomial in N such that $\vDash \varphi_N \leftrightarrow P_N^*$ and φ_N is Σ_{d_N} (or Π_{d_N}), then d_N is $\Omega(N/\log N)$. More precisely, for each polynomial p, there is a constant ε such that if $|\varphi_N| < p(N)$, then $d_N > \varepsilon N/\log N$, for all large N. The same statement holds if A is purely universal or quantifier-free.*

Proof. To prove the lower bound, it is enough to work in one particular model and prove the lower bound there. The model used in this proof will be defined more precisely later, but is essentially the standard model over a finite subset of the integers.

Define $P_0(i, j)$ to be the formula $T(i)$, and, for $N > 0$, define $P_N(i, j)$ to be

$$\exists k[\textit{Mid}(i, j, k) \wedge (P_{N-1}(i, k) \oplus P_{N-1}(k, j))],$$

where $B \oplus C$ is an abbreviation for $(B \wedge \neg C) \vee (\neg B \wedge C)$. With the intended meanings of the predicates, $P_N(0, n)$ calculates the parity of $T(0), \ldots, T(n-1)$, where $n = 2^N$.

The above definition of P_N uses positive and negative occurrences of P_{N-1}. To avoid this issue, introduce a new predicate that encodes P_N and whether it occurs positively or negatively. Define $P_N^*(i, j, a)$ to be

$$(P_N(i, j) \wedge a = 0) \vee (\neg P_N(i, j) \wedge a \neq 0).$$

The recursive definition of P_N can now be restated as a recursive definition of P_N^* by breaking into cases on a. Let $R_{N-1}^*(i, j, k, a)$ be

$$a = 0 \wedge [(P_{N-1}^*(i, k, 0) \wedge P_{N-1}^*(k, j, 1)) \vee (P_{N-1}^*(i, k, 1) \wedge P_{N-1}^*(k, j, 0))],$$

let $S^*_{N-1}(i,j,k,a)$ be

$$a \neq 0 \wedge [(P^*_{N-1}(i,k,0) \wedge P^*_{N-1}(k,j,0)) \vee (P^*_{N-1}(i,k,1) \wedge P^*_{N-1}(k,j,1))],$$

and let $P^*_N(i,j,a)$ be recursively defined on P^*_{N-1} by:

$$\exists k[\mathit{Mid}(i,j,k) \wedge (R^*_{N-1}(i,j,k,a) \vee S^*_{N-1}(i,j,k,a))].$$

Note P^*_{N-1} occurs only positively in the definition of P^*_N. Before continuing, the following slight generalization of the Yao-Håstad Switching Lemma is needed.

Lemma 5.10 (Yao-Håstad Switching Lemma) *Let $\{C_n | n \in I\}$ be an infinite family of circuits where C_n has size $n^{(\log n)^{O(1)}}$. If each C_n calculates parity on n objects for each $n \in I$, then C_n has depth $\Omega(\log n / \log\log n)$.*

Lemma 5.10 follows from the slightly stronger bounds for the switching lemma proved in [10]. The bounds state that the size of any circuit calculating parity is at least $2^{\frac{1}{14}n^{1/(d-1)}}$, where d is the depth of the circuit. For fixed $r > 0$, if the size of the circuit is $n^{(\log n)^r}$, then a straightforward calculation shows the depth is $\Omega(\frac{\log n}{\log\log n})$. The constant suppressed by the Ω notation is dependent on r.

We now finish the proof of Theorem 5.9. Suppose that there are formulas φ_N such that $\models \varphi_N \leftrightarrow P^*_N$, where $|\varphi_N| = N^{O(1)}$. Also suppose that φ_N is in prenex form with a Σ_{d_N} quantifier block (the argument would carry through if the quantifier block were Π_{d_N}) and quantifier-free part ψ_N. Let $\mathfrak{M}_N$ be the model with universe $0, 1, \ldots, n$, where $n = 2^N$, $\mathit{Mid}(i,j,k)$ expresses $k = \lfloor \frac{i+j}{2} \rfloor$, and T is arbitrary.

The following translates $\varphi_N(0,n,0)$ into a circuit C_n, which will calculate the parity of n objects. First, let ψ'_N be $\psi_N(0,n,0)$ in CNF. Since $|\psi_N| = N^{O(1)}$, we have $|\psi'_N| = 2^{N^{O(1)}}$. Next, working from the outside in replace any subformula in $\varphi_N(0,n,0)$ of the form $\exists y \chi(y)$ (respectively, $\forall y \chi(y)$) by $\bigvee_{i=0}^n \chi(i)$ (respectively, $\bigwedge_{i=0}^n \chi(i)$), since they are equivalent in $\mathfrak{M}_N$. Replace instances of $\mathit{Mid}(i,j,k)$ in φ_N with $\top$ if $k = \lfloor \frac{i+j}{2} \rfloor$, and $\bot$ if $k \neq \lfloor \frac{i+j}{2} \rfloor$. Introduce new propositional variables $p_0, \ldots, p_n$ and replace occurrences of $T(i)$ in φ_N with p_i. The formula is now propositional, and naturally translates into a circuit, call it C_n. The quantifier block of φ_N translates into a circuit of depth d_N with $(n+1)^{d_N}$ inputs, where the inputs are occurrences of the circuit translation of ψ'_N. Thus depth of C_n is at most $d_N + 2$. The $(n+1)^{d_N}$ occurrences of ψ'_N in C_n contribute $(n+1)^{d_N} 2^{N^{O(1)}} = n^{(\log n)^{O(1)}}$ to its size, since $(n+1)^{d_N}$ is $O(n^{d_N})$. The initial branching of the circuit coming from the quantifier block of φ_N adds $n^{N^{O(1)}} = n^{(\log n)^{O(1)}}$ to its size, so overall C_n has size $n^{(\log n)^{O(1)}}$. Since $P^*_N(0,n,0)$ calculates the parity of $T(0), \ldots, T(n-1)$, C_n calculates the parity of $p_0, \ldots, p_{n-1}$. Then Lemma 5.10 applied to $\{C_n | n = 2^N, N \geq 0\}$ implies that d_N is $\Omega(\log n / \log\log n) = \Omega(N / \log N)$.

The proof would still hold if recursion of P^*_N on P^*_{N-1} were instead defined to be

$$\forall k[\mathit{Mid}(i,j,k) \rightarrow (R^*_{N-1}(i,j,k,a) \vee S^*_{N-1}(i,j,k,a))].$$

If the language is allowed to contain function symbols, quantifiers are not needed in defining P^*_N, as k can be replaced by $\mathit{Mid}(i,j)$, where $\mathit{Mid}(i,j)$ is a binary function symbol with intended meaning $\mathit{Mid}(i,j) = \lfloor \frac{i+j}{2} \rfloor$. The proof of Theorem 5.9 still holds in this case, because any function symbols in $P^*_N(0,n,0)$ can replaced in φ_N by their correct value. □

6 Positive and Negative Occurrences

This section removes the restriction of the previous sections that R_{n-1} occur only positively in the definition of R_n. Theorems 6.1 and 6.2 prove upper bounds and Theorem 6.3 improves the lower bound in one case.

Theorem 6.1 *Suppose R_n is recursively defined as $A(R_{n-1})$. Let $k > 0$. If A is Σ_k, Π_k, or Δ_{k+1}, then there are formulas F_n such that F_0 is R_0 and for $n > 0$, F_n is Δ_{kn+1} and $\vdash_{\geq 2} F_n \leftrightarrow A(F_{n-1})$ by a proof of size polynomial in $|A|$, n, and $|R_0|$. Moreover, $\vdash_{\geq 2} F_n \leftrightarrow R_n$.*

Proof. We first assume there are two unequal constant symbols c_0, c_1; this condition will be removed later. Introduce new $(k+1)$-ary predicates $R^*_n(\vec{x}, a)$. The intuition is that $R^*_n(\vec{x}, a)$ expresses the formula S_n:

$$(a = c_0 \wedge R_n(\vec{x})) \vee (a \neq c_0 \wedge \neg R_n(\vec{x})).$$

We need to recursively express R_n^* in terms of only positive occurrences of R_{n-1}^*. For this, form $A^+(R_{n-1}^*)$ as follows: Put A into De Morgan normal form, and replace positive occurrences of $R_{n-1}(\vec{t})$ in $A(R_{n-1})$ by $R_{n-1}^*(\vec{t}, c_0)$ and occurrences of $\neg R_{n-1}(\vec{t})$ by $R_{n-1}^*(\vec{t}, c_1)$. Similarly, form $A^-(R_{n-1}^*)$ as follows: Put $\neg A$ into De Morgan normal form, and replace positive occurrences of $R_{n-1}(\vec{t})$ in $\neg A(R_{n-1})$ by $R_{n-1}^*(\vec{t}, c_0)$ and occurrences of $\neg R_{n-1}(\vec{t})$ by $R_{n-1}^*(\vec{t}, c_1)$. Let $A^*(R_{n-1}^*)$ be

$$(a = c_0 \wedge A^+(R_{n-1}^*)) \vee (a \neq c_0 \wedge A^-(R_{n-1}^*)).$$

Note $A^*(R_{n-1}^*)$ has only positive occurrences of R_{n-1}^*, and A^* is Δ_{k+1} if A is Σ_k or Π_k, for $k > 0$. If A is Δ_{k+1}, then A^* can be made Σ_{k+1} by choosing the Σ_{k+1} representation for A in A^+ and the Π_{k+1} representation of A in A^-. Similarly, A^* can also be made Π_{k+1} by appropriately choosing the representation for A in A^+ and A^-. This shows that A^* is still Δ_{k+1}. In contrast to Section 4, the cases where A is Σ_1 or Π_1 are not treated differently. By Theorem 4.8, there exist formulas $F_n(\vec{x}, a)$ such that F_0 is S_0, and for $n > 0$,

$$\vdash_{\geq 2} F_n(\vec{x}, a) \leftrightarrow A^*(F_{n-1})$$

have polynomial size proofs. Furthermore, since S_0 is quantifier-free, F_n is $\Delta_{kn+1}(R_0)$. Also, we have polynomial size proofs of

$$c_0 \neq c_1 \rightarrow (F_n(\vec{x}, c_0) \leftrightarrow A^+(F_{n-1})) \tag{1}$$

and

$$c_0 \neq c_1 \rightarrow (F_n(\vec{x}, c_1) \leftrightarrow A^-(F_{n-1})). \tag{2}$$

By definition, there is a proof of $c_0 \neq c_1 \rightarrow (F_0(\vec{x}, c_0) \leftrightarrow \neg F_0(\vec{x}, c_1))$. Since A^+ and A^- are essentially the negations of each other in De Morgan normal form, induction on n and (1) and (2) imply that for $n \geq 0$, there is a polynomial size proof of

$$c_0 \neq c_1 \rightarrow (F_n(\vec{x}, c_0) \leftrightarrow \neg F_n(\vec{x}, c_1)). \tag{3}$$

Let $G_n(\vec{x})$ be $F_n(\vec{x}, c_0)$. Then $G_0(\vec{x})$ is $R_0(\vec{x})$, and there is a polynomial size proof of

$$c_0 \neq c_1 \rightarrow (G_n \leftrightarrow A(G_{n-1})). \tag{4}$$

We now eliminate the condition that $c_0 \neq c_1$. Write $G_n(\vec{x})$ as $G_n(\vec{x}, c_0, c_1)$ to display the dependence of G_n on c_0 and c_1. Let $G_n(\vec{x}, d_0, d_1)$ be $G_n(\vec{x}, c_0, c_1)$ with all occurrences of c_0 replaced by d_0 and all occurrences of c_1 replaced by d_1. We claim there is a polynomial size proof of

$$(c_0 \neq c_1 \wedge d_0 \neq d_1) \rightarrow (G_n(\vec{x}, c_0, c_1) \leftrightarrow G_n(\vec{x}, d_0, d_1)). \tag{5}$$

This is proved by induction. The base case follows easily from the definition of S_0, and the inductive step follows from (3) and (4). Let X_n be

$$\exists y_0 \exists y_1 (y_0 \neq y_1 \wedge G_n(\vec{x}, y_0, y_1))$$

and Y_n be

$$\forall y_0 \forall y_1 (y_0 \neq y_1 \rightarrow G_n(\vec{x}, y_0, y_1)),$$

It follows from (5) that there are polynomial size proofs of

$$c_0 \neq c_1 \rightarrow (G_n(\vec{x}, c_0, c_1) \leftrightarrow X_n)$$

and

$$c_0 \neq c_1 \rightarrow (G_n(\vec{x}, c_0, c_1) \leftrightarrow Y_n).$$

Hence

$$\vdash_{\geq 2} X_n \leftrightarrow Y_n,$$

$$\vdash_{\geq 2} X_n \leftrightarrow A(X_{n-1}),$$

and

$$\vdash_{\geq 2} Y_n \leftrightarrow A(Y_{n-1})$$

by polynomial size proofs. Since G_n is $\Delta_{kn+1}(R_0)$, it can be equivalently be written either as a $\Sigma_{kn+1}(R_0)$ or $\Pi_{kn+1}(R_0)$ formula. Then X_n is $\Sigma_{kn+1}(R_0)$ and Y_n is $\Pi_{kn+1}(R_0)$. Since $\vdash_{\geq 2} X_n \leftrightarrow Y_n$, X_n would be $\Delta_{kn+1}(R_0)$ if the condition of more than two elements could be eliminated. This is done in a similar manner to the construction at the end of the proof of Theorem 4.8. □

The remaining case is when A is quantifier-free. The above proof is essentially unchanged, and thus Theorem 4.7 extends easily to the following theorem.

Theorem 6.2 *Suppose R_n is recursively defined as $A(R_{n-1})$. Fix $\delta > 0$. If A is purely existential, purely universal, or quantifier-free, then there are formulas F_n such that F_0 is R_0 and for $n > 0$, F_n is $\Delta_{\delta n/\log n}(R_0)$ and $\vdash_{\geq 2} F_n \leftrightarrow A(F_{n-1})$ by a proof of size polynomial in $|A|$, n, and $|R_0|$. Moreover, $\vdash_{\geq 2} F_n \leftrightarrow R_n$.*

The corresponding lower bounds when $A(R_{n-1})$ is allowed to have positive and negative occurrences of R_{n-1} also follow easily from previous results. If A is Σ_{2k}, Π_{2k}, or Δ_{k+1} for $k > 0$, then Theorems 5.3, 5.4, and 5.8 still apply, respectively, and give the same lower bounds. If A is Σ_{2k+1} with $k \geq 0$, the lower bound of Theorem 5.6 can be slightly improved, and the case $k = 0$ can also be included.

Theorem 6.3 *Let $k \geq 0$. Then there exists formulas F_n^{2k+1} recursively defined to be $A(F_{n-1}^{2k+1})$, where A is Σ_{2k+1} and F_0^{2k+1} is quantifier-free, such that if $n > 0$ and $\mathfrak{N} \vDash \varphi \leftrightarrow F_n^{2k+1}$, then $\varphi \notin \Pi_{(2k+1)n}$. Hence $\varphi \notin \Sigma_l$ for $l < (2k+1)n$.*

A dual result holds for A in Π_{2k+1}. The proofs are similar to the arguments in Section 5.

References

[1] J. Ferrante and C. W. Rackoff, The Computational Complexity of Logical Theories, Lecture Notes in Mathematics #718 (Springer Verlag, Berlin, 1979).

[2] P. Pudlák, The lengths of proofs, in: Handbook of Proof Theory, edited by S. R. Buss (Elsevier North-Holland, 1998), pp. 547–637.

[3] P. Pudlák, On the lengths of proofs of finitistic consistency statements in first order theories, in: Logic Colloquium '84, (North-Holland, 1986), pp. 165–196.

[4] P. Pudlák, Improved bounds to the lengths of proofs of finitistic consistency statements, in: Logic and Combinatorics, edited by S. G. Simpson, Contemporary Mathematics Vol. 65 (American Mathematical Society, 1987), pp. 309–331.

[5] A. S. Troelstra and H. Schwichtenberg, Basic Proof Theory, 2nd edition, Tracts in Theoretical Computer Science #43 (Cambridge University Press, Cambridge, 2000).

[6] M. J. Fischer and M. O. Rabin, Super-exponential complexity of Presburger arithmetic, in: Proc. SIAM-AMS Symposium in Applied Mathematics, vol. 7, (Massachusetts Institute of Technology, 1974), pp. 27–41.

[7] L. J. Stockmeyer and A. R. Meyer, Word problems requiring exponential time (preliminary report), in: Proc. of the Fifth Annual ACM Symposium on Theory of Computing (STOC'73), (1973), pp. 1–9.

[8] J. Håstad, Computational Limitations for Small-depth Circuits (MIT Press, 1987).

[9] A. C. C. Yao, Separating the polynomial time hierarchy by oracles, in: Proceedings of the 26th Annual Symposium on Foundations of Computer Science (FOCS'85), (IEEE Computer Society, 1985), pp. 1–10.

[10] P. Beame, A switching lemma primer, Typeset manuscript, 1994.