Relating the bounded arithmetic and polynomial time hierarchies

Samuel R. Buss**

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

Received 22 May 1993; revised 30 November 1994; communicated by S. Artemov

Abstract

The bounded arithmetic theory $S_2$ is finitely axiomatized if and only if the polynomial hierarchy provably collapses. If $T^i_2$ equals $S^{i+1}_2$ then $T^i_2$ is equal to $S_2$ and proves that the polynomial time hierarchy collapses to $\Sigma^p_{i+3}$, and, in fact, to the Boolean hierarchy over $\Sigma^p_{i+2}$ and to $\Sigma^p_{i+1/poly}$.

1. Introduction

Theories of bounded arithmetic are theories of arithmetic obtained by putting restrictions on induction axioms; namely, allowing induction only for certain classes, $\Sigma^b_i$, of bounded formulas, and using polynomial, or length, induction (PIND or LIND) in place of successor induction (IND). The most important subtheories of bounded arithmetic are the theories $S^i_2$, axiomatized with $\Sigma^b_i$-PIND (or equivalently, $\Sigma^b_i$-LIND, if $i \geq 1$), and the theories $T^i_2$, axiomatized by $\Sigma^b_i$-IND. The following inclusions are known for these theories:

$$S^0_2 \subseteq T^0_2 \subseteq S^1_2 \subseteq T^1_2 \subseteq S^2_2 \subseteq T^2_2 \subseteq \cdots$$

and their union is the theory $S_2 = T_2$ [2]. However, with the exception of $S^2_2 \neq T^0_2$ (see [13]), it is not known whether the rest of the theories of bounded arithmetic are distinct. It is a well-known fact that $S^i_2$ and $T^i_2$ are finitely axiomatized for $i > 0$, and thus it is immediate that this hierarchy of theories collapses if and only if $S_2$ is finitely axiomatized. This latter condition is equivalent to $IA_0 + \Omega_1$ being finitely axiomatized (see [11, 14] for this alternate, and original, approach to bounded arithmetic).

* Email: sbuss@ucsd.edu.

** Supported in part by NSF grant DMS-9205181.

0168-0072/95/$09.50 © 1995 - Elsevier Science B.V. All rights reserved
SSDI 0168-0072(94)00057-3
There are close connections between theories of bounded arithmetic and the polynomial hierarchy. First, the class of predicates definable by $\Sigma^b_i$ (or $\Pi^b_i$) formulas is precisely the class of predicates in the $i$th level $\Sigma^p_i$ (or $\Pi^p_i$, respectively) of the polynomial hierarchy. For instance, $S^2_1$ and $T^1_2$ are axiomatized with their induction axioms restricted to NP-predicates (since NP = $\Sigma^p_1$ is the class of predicates definable by $\Sigma^b_1$-formulas). Second, it is known that the $\Sigma^b_i$-definable functions of $S^2_1$ are precisely the $\Pi^p_i$-functions, which are the functions which are polynomial time computable with an oracle for $\Sigma^p_{i-1}$. For instance, the $\Sigma^b_i$-definable functions of $S^2_1$ are precisely the polynomial time computable functions.

Since it is open whether the polynomial time hierarchy collapses, it is natural to ask whether there is any connection between the possible collapses of the hierarchy of bounded arithmetic theories and the polynomial hierarchy. This question has already been partially answered by the work of Krajíček et al. [10] who showed that if $T^1_i = S^1_{i+1}$ for any $i \geq 1$, then the polynomial hierarchy collapses with $\Sigma^p_{i+2} = \Pi^p_{i+2}$ (in fact, they show that in this case, $\Sigma^p_{i+1} \subset \Delta^p_i/poly$).

The main results of this paper strengthen the results of Krajíček et al. by proving that if $T^1_2 = S^1_{2+1}$ holds, then the following conditions must hold: (1) $T^1_2 = S_2$, so that the hierarchy of bounded arithmetic theories collapses, and (2) $T^1_2$ can prove that the polynomial time hierarchy collapses to $\mathcal{B}(\Sigma^p_{i+2})$ and to $\Sigma^{\Delta^p_{i+1}}$, where $\mathcal{B}(\Sigma^p_{i+2})$ is the class of Boolean combinations of $\Sigma^b_{i+2}$-predicates. Our proofs are easier, in a combinatorial sense, than the proofs of [10] and this makes it possible to formalize them in $T^1_2$.

We believe that the results of this paper are nearly the strongest that are obtainable relating the possibility that $T^1_2 = S^1_{2+1}$ to the possible collapse of the polynomial time hierarchy – at least with current techniques. To support this belief, consider the three conditions:

(a) the polynomial hierarchy collapses,

(β) $S_2$ proves that the polynomial hierarchy collapses,

(γ) $S_2$ is finitely axiomatized.

Our results show that (β) and (γ) are equivalent; however, we do not expect to show that (α) is equivalent to (γ) using current techniques. The reason for this is that (α) is a $\Sigma^0_2$-condition whereas, since (β) is a $\Sigma^0_2$-condition, the results of the current paper show that (γ) is a $\Sigma^0_2$-condition; and, based on the history of attempts to solve the P versus NP problem, it seems to be difficult even to establish that the collapse of the polynomial time hierarchy is equivalent to a natural $\Sigma^0_2$-condition like (γ).

It is known that $S^1_{2+1}$ is conservative over $T^1_2$ with respect to $\forall \Sigma^b_{i+1}$-sentences [3]. On the other hand, the axioms of $S^1_{2+1}$ can be expressed as $\forall \Pi^b_{i+2}$-sentences (in this formulation, an induction axiom of $S^1_{2+1}$ will become a $\forall \Pi^b_{i+2}$-formula with a sharply bounded existential quantifier in its outermost block of bounded universal quantifiers). Thus saying $S^1_{2+1}$ is $\Pi^b_{i+2}$-conservative over $T^1_2$ is equivalent to saying that $S^1_{2+1} = T^1_2$.

An open problem is to try to relate the condition $S^1_2 = T^1_2$ to the possible collapse of the polynomial hierarchy. Krajíček [9] shows that if $S^1_2 = T^1_2$, then the set $\leq^b_\mu(\Sigma^p_i)$ of
predicates logspace, Turing reducible to \( \Sigma^p_i \) is equal to the set \( \leq^p_T(\Sigma^p_i) \) of predicates polynomial time, Turing reducible to \( \Sigma^p_i \). However, it is open whether this last condition implies the polynomial hierarchy collapses. See [4–6] for more on this connection.

The prerequisites for reading this paper are a basic knowledge of bounded arithmetic theories as contained in [2]. The reader would also benefit from knowledge of [10, 3]. In the next section we will review the necessary background material needed from [10].

After preparing the first draft of this paper, we learned that D. Zambella has independently discovered the main results of this paper [15].

2. The KPT witnessing theorem for \( T^i_2 \)

There are two important witnessing theorems for \( T^i_2 \). The first follows from the ‘Main Theorem’ for \( S^i+1_2 \) and the fact that \( S^i+1_2 \) is \( \Sigma^p_{i+1} \)-conservative over \( T^i_2 \): this witnessing theorem states that the \( \Sigma^p_{i+1} \)-definable functions of \( T^i_2 \) are precisely the functions which can be computed in polynomial time with a \( \Sigma^p_i \)-oracle (i.e., the \( \Pi^p_{i+1} \)-functions). The second witnessing theorem puts a necessary condition on the \( \Pi^p_i \) and \( \Sigma^p_{i+1} \)-definable functions of \( T^i_2 \); we call this the ‘KPT witnessing theorem’. It is this latter witnessing theorem that we need for our proofs:

**Theorem 1.** Let \( i \geq 1 \). Suppose \( T^i_2 \models (\forall x)(\exists y)(\forall z)B(x, y, z) \), where \( B \) is a \( \exists \Pi^1_i \)-formula, with only \( x, y, z \) as free variables. There exists \( k > 0 \) and functions \( f_1, \ldots, f_k \) such that each \( f_m \) is \( m \)-ary and is \( \Sigma^p_{i+1} \)-definable by \( T^i_2 \) and such that

\[
T^i_2 \models B(a, f_1(a), b_1) \lor B(a, f_2(a, b_1), b_2) \lor B(a, f_3(a, b_1, b_2), b_3) \lor \cdots \lor B(a, f_k(a, b_1, b_2, \ldots, b_{k-1}), b_k).
\]

For \( i = 0 \), the same result holds for \( PV_1 \) in place of \( T^0_2 \). As usual, \( PV_1 \) denotes the conservative extension of \( PV \) to first-order logic, or equivalently, \( PV_1 \) is \( S^0_2 \) or \( T^0_2 \) enlarged to have function symbols and their defining equations for all polynomial time functions.

Note that since the functions \( f_m \) are \( \Sigma^p_{i+1} \)-definable by \( T^i_2 \), they must be \( \Pi^p_{i+1} \)-functions.

Theorem 1 is due to [10]; some later, related results can be found in [8, 12, 1]. We do not include a proof here.

We next use Theorem 1 to establish a consequence of the condition \( T^i_2 = S^i+1_2 \). We assume that \( i \geq 0 \) and work with the theory \( T^i_2 \); when \( i = 0 \) our results are intended to hold for \( PV_1 \) in place of \( T^0_2 \).

**Definition.** A quantified Boolean formula is a formula constructed from Boolean connectives (say, \( \land, \lor \) and \( \neg \)) and quantifiers ranging over Boolean values. A quantifier \( (\forall p) \) or \( (\exists p) \) indicates quantification allowing \( p \) to range over the values \( True \) and \( False \).
Given a truth assignment to the free variables of a quantified Boolean formula, it is obvious how the truth value of the formula should be defined. A quantified Boolean formula is \textit{satisfiable} if there is some truth assignment to its free variables which gives it value \textit{True}. A \( \Pi^P_i \)-formula is a quantified Boolean formula which is in prenex form with \( i \) blocks of like quantifiers starting with a universal block. It is well-known that the set of satisfiable \( \Pi^P_i \)-formulas is \( \Sigma^P_{i+1} \)-complete.

\textbf{Definition.} Let \( i \geq 0 \). \( TRU^i \) and \( SAT^i \) are bounded arithmetic formulas which express:

\[
TRU^i(\varphi, w) \iff \varphi \text{ codes a } \Pi^P_i \text{-formula and } w \text{ codes a satisfying assignment of } \varphi,
\]

\[
SAT^i(\varphi) \iff (\exists w \leq \varphi) TRU^i(\varphi, w).
\]

In the definition of \( TRU^i \) and \( SAT^i \) we presume that quantified Boolean formulas and truth assignments are coded in some natural and efficient way by integers; we use Greek letters \( \varphi, \ldots \) as variables that range over integers which are intended to code quantified Boolean formulas. Since the code of a truth assignment can w.l.o.g. always be less than the code of a formula, \( SAT^i(\varphi) \) expresses the condition that \( \varphi \) is satisfiable. Standard bootstrapping techniques allow \( TRU^i \) to be a \( \Delta^P_{i+1} \)-formula with respect to the theory \( T_2^i \); in fact, for \( i \geq 1 \), \( TRU^i \) is a \( \Pi^P_i \)-formula. Hence \( SAT^i \) is a \( \Sigma^P_{i+1} \)-formula. Also, \( T_2^i \) can prove basic properties of the \( TRU^i \) and \( SAT^i \) predicates. Most importantly, \( T_2^i \) can prove that \( SAT^i \) is many-one complete for \( \Sigma^P_{i+1} \)-formulas; i.e., for any \( \Sigma^P_{i+1} \)-formula \( A(\vec{b}) \), there is a polynomial time function \( f \) so that \( A(\vec{b}) \) is \( T_2^i \)-provably equivalent to \( SAT^i(f(\vec{b})) \).

As an application of Theorem 1, consider the formula

\[
\forall \langle \varphi_0, \ldots, \varphi_n \rangle (\exists l \leq n)(\exists \langle w_0, \ldots, w_l \rangle)
\]

\[
[(\forall j \leq l) TRU^i(\varphi_j, w_j) \land (l < n \rightarrow \neg (\exists w_{l+1}) TRU^i(\varphi_{l+1}, w_{l+1}))].
\]

(1)

The meaning of formula (1) requires some explanation. First, a notation like \( \langle \forall \varphi_0, \ldots, \varphi_n \rangle B(\vec{\phi}, n) \) means the same as ‘there is an integer \( \varphi^* \) which codes a sequence of \( \Pi^P_i \)-formulas \( \varphi_0, \ldots, \varphi_n \) so that \( B(\vec{\phi}, n) \) holds’. The quantifier \( \exists l \leq n \) is a sharply bounded quantifier since \( l \) can be bounded by the length of the code for \( \langle \vec{\phi} \rangle \), and the quantifiers \( \exists w_j \) and \( \exists w_{l+1} \) are bounded quantifiers since each \( w_j \) may be bounded by \( \varphi_j \). By using prenex operations and using the fact that \( l \) can be computed from \( \langle w_0, \ldots, w_l \rangle \), formula (1) is equivalent to the formula

\[
(\forall \langle \varphi_0, \ldots, \varphi_n \rangle) (\exists \langle w_0, \ldots, w_l \rangle) (\forall w_{l+1})
\]

\[
[(\forall j \leq l) TRU^i(\varphi_j, w_j) \land (l < n \rightarrow \neg TRU^i(\varphi_{l+1}, w_{l+1}))].
\]

(2)

which is a \( \forall \exists \leq \forall \leq \Delta^P_{i+1} \)-formula.

The intuitive meaning of formula (1) or (2) is, of course, that every sequence \( \varphi_0, \ldots, \varphi_n \) of \( \Pi^P_i \)-formulas has an initial sequence of maximal length \( l \) of satisfiable
formulas. Furthermore, the formula (1) is a theorem of $S_{2}^{i+1}$. This is because $S_{2}^{i+1}$ can use length induction on the $\Sigma_{i+1}^{b}$-formula $S(\langle \phi \rangle, l)$ expressing the condition that the first $l$ formulas of the sequence are satisfiable. (An equivalent way to see this is to note that $S_{2}^{i+1}$ can prove the $\Sigma_{i+1}^{b}$-length-maximization principle.)

Now suppose $T_{2}^{i}$ is equal to $S_{2}^{i+1}$; in particular, $T_{2}^{i}$ proves the formula (2). By Theorem 1, this means that there is an integer $k \geq 0$ and there are $\Sigma_{i+1}^{b}$-defined functions $f_{0}, \ldots, f_{k}$ so that, letting $\mathcal{A}(\langle \phi \rangle, \langle \bar{w} \rangle, w_{i+1})$ be the subformula of (2) enclosed in square brackets, we have that

$$
T_{2}^{i} \models (\forall \langle \phi \rangle)[A(\langle \phi \rangle, f_{0}(\langle \phi \rangle), b_{0}) \vee A(\langle \phi \rangle, f_{1}(\langle \phi \rangle, b_{0}), b_{1}) \vee \cdots \vee A(\langle \phi \rangle, f_{k}(\langle \phi \rangle, b_{0}, b_{1}, \ldots, b_{k-1}), b_{k})]
$$

(3)

We henceforth shall use (3) restricted to the case where $n = k$, so that the sequence $\bar{\phi}$ is $\varphi_0, \ldots, \varphi_k$.

Without loss of generality, each $f_{j}$ satisfies the following property (provably in $T_{2}^{i}$): whenever $\text{TRU}^{i}(\varphi_{r}, h_{r})$ holds for $r = 0, \ldots, j - 1$, then the value $f_{j}(\langle \phi \rangle, b_{0}, \ldots, h_{j-1})$ is the Gödel number of a sequence $\langle v_{0}, \ldots, v_{j-1} \rangle$ of length $l \geq j$ so that $\text{TRU}^{i}(\varphi_{r}, v_{r})$ holds for all $r = 0, \ldots, l - 1$.

Recall that $\beta$ represents the Gödel $\beta$ function so that $\beta(i, w)$ is equal to the $i$th integer in the sequence coded by $w$. Define

$$
g_{j}(\varphi_{0}, \ldots, \varphi_{k}, w_{0}, \ldots, w_{j-1}) = \beta(j + 1, f_{j}(\langle \varphi_{0}, \ldots, \varphi_{k}, w_{0}, \ldots, w_{j-1} \rangle)).
$$

Suppose that $\varphi_{0}, \ldots, \varphi_{k}$ are codes for satisfiable $\Pi^{b}_{k}$-Boolean formulas and let $w_{0}, \ldots, w_{k}$ be satisfying assignments. Define $b_{0}, b_{1}, \ldots$ inductively as follows: if $f_{j}(\langle \phi \rangle, b_{0}, \ldots, b_{j-1})$ is a sequence of length $l + 1 \leq k$, then let $b_{j}$ equal $w_{l+1}$. It is obvious that whenever $f_{j}(\langle \phi \rangle, b_{0}, \ldots, b_{j-1})$ has length $l + 1 \leq k$ then $b_{j}$ gives a "counterexample" so that $A(\langle \phi \rangle, f_{j}(\langle \phi \rangle, b_{0}, \ldots, b_{j-1}), b_{j})$ is false. Now, by (2), there is some $j \leq k$ for which $f_{j}(\langle \phi \rangle, b_{0}, \ldots, b_{j-1})$ has length $k + 1$. Let $j_{0}$ be the least value such that $f_{j_{0}}(\langle \phi \rangle, b_{0}, \ldots, b_{j_{0} - 1})$ has length $\geq j_{0} + 1$. It must be that $\text{TRU}^{i}(\varphi_{j_{0}}, g_{j_{0}}(\bar{\phi}, w_{0}, \ldots, w_{j_{0} - 1}))$ holds. This argument formalizes in $T_{2}^{i}$ and thus we have proven:

**Lemma 2.** Suppose $T_{2}^{i} = S_{2}^{i+1}$. Then there is $k \geq 0$ and there are $\Sigma_{i+1}^{b}$-definable functions $g_{0}, \ldots, g_{k}$ of $T_{2}^{i}$ so that

$$
T_{2}^{i} \models (\forall \varphi_{0}, \ldots, \varphi_{k})(\forall w_{0}, \ldots, w_{k})\left[ \bigwedge_{j=0}^{k} \text{TRU}^{i}(\varphi_{j}, w_{j}) \right.

\rightarrow \text{TRU}^{i}(\varphi_{0}, g_{0}(\bar{\phi}))

\vee \text{TRU}^{i}(\varphi_{1}, g_{1}(\bar{\phi}, w_{0}))

\vee \text{TRU}^{i}(\varphi_{2}, g_{2}(\bar{\phi}, w_{0}, w_{1}))

\vee \cdots \vee \text{TRU}^{i}(\varphi_{k}, g_{k}(\bar{\phi}, w_{0}, \ldots, w_{k-1}))

\left. \right]
$$

71
3. Collapsing bounded arithmetic

In this and the next section, we examine consequences of the condition $T^i_1 = S^i_{2+1}$. In this section we show that this implies that $S_2$ collapses to $T^i_1$.

Our point of departure is Lemma 2; we henceforth fix $k$ and $g_0, \ldots, g_k$. This lemma states that at least one of the functions $g_j$ can find a satisfying assignment for $\varphi_j$ using only the vector $\vec{f}$ and arbitrary satisfying assignments $w_0, \ldots, w_{j-1}$. However, it need not always be the same $g_j$ that succeeds in this way; different vectors of formulas $\vec{f}$ and even different witnesses $\vec{w}$ may cause different $g_j$'s to succeed. We define $\text{SucceedBy}(l, \vec{f}, \vec{w})$ to be the following formula which states that one of the first $l + 1$ $g$'s succeeds in this way; namely, it is defined as

$$\text{SucceedBy}(l, \vec{f}, \vec{w}) \iff \bigwedge_{j=0}^{k} \left[j < l \land \text{TRU}^i(\varphi_j, g_j(\vec{f}, w_0, \ldots, w_{j-1})) \right].$$

Our first goal is to show that $\Sigma^p_{l+1} = \Pi^p_{l+1}/\text{poly}$ where the 'poly' means that polynomial amount of advice is needed. As a preliminary to defining what constitutes advice, we define 'preadvice' by letting $\text{PreAdvice}^i(a, \langle \varphi_{l+1}, \ldots, \varphi_k \rangle)$ be the formula

$$\bigwedge_{j=0}^{k} (l < j \rightarrow \varphi_j < 2^{|a|}) \land (\forall \langle \varphi_0, \ldots, \varphi_l \rangle)(\forall w_0, \ldots, w_l) \text{TRU}^i(\varphi_j, w_j).$$

Several points to note are: firstly, in defining $\text{PreAdvice}$ we are continuing our practice of letting variables $\varphi_j$ represent integers that must code $\Pi^p_2$ formulas; secondly, the value of $l$ is determined by the second argument to $\text{PreAdvice}$ ($k$ is fixed and $l$ varies, namely, $l$ equals $k + 1$ minus the length of the sequence coded by the second argument of $\text{PreAdvice}^i$); thirdly, the quantifiers are bounded quantifiers since the $\varphi_j$'s and $w_j$'s are bounded by $2^{2^{|a|}}$. The reason for bounding everything by $2^{2^{|a|}}$ is that we need only define 'advice' that works for $\varphi$'s with $|\varphi| < a$ for $a$ an arbitrary integer. Also note that $\text{PreAdvice}^i$ is a $\Pi^b_{l+1}$ formula.

We can now define 'advice' for formulas of length $\leq |a|$ by

$$\text{Advice}^i(a, \langle \varphi_{l+1}, \ldots, \varphi_k \rangle)$$

$$\iff \text{PreAdvice}^i(a, \langle \varphi_{l+1}, \ldots, \varphi_k \rangle) \land \neg (\exists \varphi_l) \text{PreAdvice}^i(a, \langle \varphi_1, \ldots, \varphi_k \rangle).$$

Note that $\varphi_l$ is bounded by $2^{2^{|a|}}$; thus $\text{Advice}^i$ is a $\Pi^p_{l+2}$ formula. The next lemma shows that $T^i_2$ can prove that there always does exist advice:

**Lemma 3.** Suppose $T^i_2 = S^i_{2+1}$. Then

$$T^i_2 \vdash (\forall a)(\exists \langle \varphi \rangle) \text{Advice}^i(a, \langle \varphi \rangle).$$
Proof. First, note that Lemma 2 implies that $T_1^2$ proves that $PreAdvice^i(a, \langle \phi \rangle)$ holds. Since $k$ is a constant, it follows (without using induction) that there is a least $l$ such that $(\exists \langle \phi_{l+1}, \ldots, \phi_k \rangle) PreAdvice^i(a, \langle \phi \rangle)$ holds. For this $l$, any 'preadvice' is actually advice. □

Next we give the key lemma that shows how 'advice' can be used to make $\Sigma^b_{l+1}$-IND hold and the polynomial time hierarchy collapse, probably in $T_1^2$.

Lemma 4. Suppose $T_1^2 = S_2^{l+1}$. Then $T_1^2$ proves

$$Advice^i(a, \langle \phi_{l+1}, \ldots, \phi_k \rangle) \land \phi_l < 2^{u_i}$$

$$\implies \left[ -SAT^i(\phi_l) \leftrightarrow (\exists \langle \phi_0, \ldots, \phi_{l-1} \rangle)(\exists \langle w_0, \ldots, w_{l-1} \rangle)$$

$$\{ VTRU^i(\langle \phi \rangle, \langle w_0, \ldots, w_{l-1} \rangle, a)$$

$$\land \neg \text{SucceedBy}(l-1, \phi, \bar{w})$$

$$\land \neg TRU^i(\phi_l, g_l(\phi_0, \ldots, \phi_k, w_0, \ldots, w_{l-1})) \} \right].$$

Proof. Let $RHS(\phi_l, \langle \phi_{l+1}, \ldots, \phi_k \rangle, a)$ denote the formula on the right-hand side of the $\leftrightarrow$ connective in the formula above; we often suppress the variables $\phi_{l+1}, \ldots, \phi_k$ and $a$ that occur freely in $RHS$ and write just $RHS(\phi_l)$.

We shall argue informally in $T_1^2$ to prove the lemma. Suppose $\phi_l, \ldots, \phi_k \leq 2^{u_i}$ are formulas and that $Advice^i(a, \langle \phi_{l+1}, \ldots, \phi_k \rangle)$ holds. The latter condition obviously implies that $-PreAdvice^i(a, \langle \phi_{l+1}, \ldots, \phi_k \rangle)$. By the definition of $PreAdvice$, there must exist $\Pi^b_1$-formulas $\phi_0, \ldots, \phi_{l-1}$ satisfied by witnesses $w_0, \ldots, w_{l-1}$ such that $\text{SucceedBy}(l-1, \phi, \bar{w})$ is forced to be false. First suppose that $\phi_l$ is not satisfiable. Then clearly $TRU^i(\phi_l, g_l(\phi, \bar{w}))$ must be false. Thus $RHS(\phi_l)$ follows from $-SAT^i(\phi_l)$. Second, suppose that $\phi_l$ is satisfiable. By $PreAdvice^i(a, \langle \phi_{l+1}, \ldots, \phi_k \rangle)$, it must be that $\text{SucceedBy}(l, \phi, \bar{w})$ holds. On the other hand, $\text{SucceedBy}(l-1, \phi, \bar{w})$ is false. Thus $TRU^i(\phi_l, g_l(\phi, \bar{w}))$ is forced to be false and we have shown that $SAT^i(\phi_l)$ implies $-RHS(\phi_l)$. □

In the subformula $RHS$, the leading existential quantifiers are actually bounded existential quantifiers since the formulas $\phi_j$ and their witnesses $w_j$ are bounded by $2^{u_i}$. This means that $RHS(\phi_l)$ is a $\Sigma^b_{l+1}$-formula.

Lemma 5. Suppose $T_1^2 = S_2^{l+1}$. Then $T_1^2 \vdash \Sigma^b_{l+1}$-IND and $T_1^2 = T_1^{l+1}$.

Proof. The proof is based on the fact that $SAT^i(\cdots)$ is complete for $\Sigma^b_{l+1}$-formulas and is also equivalent on bounded ranges to the $\Pi^b_{l+1}$-formula $-RHS(\cdots)$ (under the assumption that $T_1^2 = S_2^{l+1}$, as always). Indeed, for any $\Sigma^b_{l+1}$-formula $B(c, \bar{d})$, there is a polynomial time and $\Sigma^b_1$-computable function $f(c, \bar{d})$ so that $B(c, \bar{d})$ is $T_1^2$-provably equivalent to $SAT^i(f(c, \bar{d}))$. The induction axiom for the formula $B(c, \bar{d})$ can be
expressed as

\[ B(0, \bar{d}) \land (\forall x)(B(x, \bar{d}) \rightarrow B(x + 1, \bar{d})) \rightarrow B(c, \bar{d}). \]

Let us prove this by reasoning informally in \( T_2^i \) which is presumed to equal \( S_2^{i+1} \).

Considering particular values for \( c \) and \( \bar{d} \), there is a value \( a \) so that \( f(x, \bar{d}) < 2^a \) for all \( x \leq c \). Let \( \phi_{i_1}, \ldots, \phi_k \) be formulas such that \( Advice'(a, \langle \phi_{i_1}, \ldots, \phi_k \rangle) \) holds. Then, with these parameters, by Lemma 4, we have that the \( \Sigma_{i+1}^b \)-formula \( B(x, \bar{d}) \) is equivalent to the \( \Pi_{i+1}^b \)-formula \( \neg RHS(f(x, \bar{d})) \) for all \( x \leq c \). Now, it is known that \( S_2^{i+1} \) proves \( \Delta_{i+1}^b \)-IND and the usual proof (see [2, Theorem 2.22]) shows that \( T_2^i = S_2^{i+1} \) proves induction for \( B \), since \( B \) is \( \Delta_{i+1}^b \) with parameters" on the range \( 0 \leq x \leq c \).

Iterating the method of this proof, we obtain:

**Theorem 6.** If \( T_2^i = S_2^{i+1} \), then \( T_2^i = S_2 \). Thus, if \( T_2^i = S_2^{i+1} \), then \( S_2 \) is finitely axiomatized.

Also, if \( PV_1 = S_2^i(PV) \), then \( PV_1 = S_2(PV) \).

**Proof.** Analogous to the method of proof of Lemma 5, we must show that any bounded formula is equivalent to a \( \Sigma_{i+2}^b \)-formula with parameters, where the parameters vary with the range of the induction variable. From this, using Lemma 5, it will follow that \( T_2^i \) proves induction for any bounded formula.

We do the case of \( B(c, \bar{d}) \in \Sigma_{i+2}^b \) in some detail. We may suppose that \( B(x, \bar{d}) \) is of the form \( (\exists y \leq t(x, \bar{d}))C(x, y, \bar{d}) \) for some \( \Pi_{i+1}^b \)-formula \( C \). We argue informally in \( T_2^i \). By the method of Lemma 5, there is an \( a \), given by a polynomial time function \( a = a(c, \bar{d}) \) of \( c \) and \( \bar{d} \), and there is a polynomial time function \( f \), so that for all \( x \leq c \), and \( y \leq t(x, \bar{d}) \) and for advice \( \langle \phi \rangle \) satisfying \( Advice'(a \langle \phi \rangle) \), the \( \Pi_{i+1}^b \)-formula \( C(x, y, \bar{d}) \) is equivalent to the \( \Sigma_{i+1}^b \)-formula \( RHS(f(x, y, \bar{d}), \langle \phi \rangle, a(c, \bar{d})) \). Thus, for \( 0 \leq x \leq c \), \( B(x, \bar{d}) \) is equivalent to a \( \Sigma_{i+2}^b \)-formula, and full induction holds for \( B \) up to \( c \) by Lemma 5. Hence \( T_2^i = T_{2+i}^i \).

A slight modification of the construction of the last paragraph shows that if \( A(\bar{x}) \) is a \( \Sigma_{i+2}^b \)-formula (respectively, a \( \Pi_{i+2}^b \)-formula), then there is a polynomial growth rate function \( a(c) \) and a \( \Sigma_{i+1}^b \)-formula (respectively, \( \Pi_{i+1}^b \)-formula) \( A^*(\bar{x}, \langle \phi \rangle, a(c)) \) such that for all \( \bar{x} \) such that \( max \{ \bar{x} \} \leq c \) and all \( \langle \phi \rangle \) such that \( Advice'(a(c), \langle \phi \rangle) \), \( A(\bar{x}) \) is equivalent to \( A^*(\bar{x}, \langle \phi \rangle, a(c)) \), provably in \( T_2^i \). This further implies that if \( A(\bar{x}) \) is a \( \Sigma_{i+3}^b \)-formula, then \( A^* \) may be taken to be a \( \Sigma_{i+3}^b \)-formula, because, if \( A(\bar{x}) \) is \( (\exists y \leq t(\bar{x}))B(\bar{x}, y) \), then there is a \( \Sigma_{i+2}^b \)-formula \( B^* \) so that \( A(\bar{x}) \) will be equivalent to \( (\exists y \leq t(\bar{x}))B^*(\bar{x}, \langle \phi \rangle, a) \) for \( a \) given by a polynomial growth rate function of \( c \geq max \bar{x} \) and for \( \langle \phi \rangle \) such that \( Advice'(a, \langle \phi \rangle) \). This fact is sufficient to imply that \( T_2^i = T_{2+i}^i \).

By iterating the above method of proof, one can show that \( T_2^i \) is equal to all of \( S_2 \).

We shall leave the details of this to the reader, and remark instead that an alternative proof is given by Theorem 7 below where it is shown that \( T_2^i \) can prove that every bounded formula is equivalent to a Boolean combination of \( \Sigma_{i+2}^b \)-formulas without
any additional parameters or advice. Then since $T^i_2 = T^i_2 = S^{i^2}_2$ and $S^{i^2}_2$ proves induction for Boolean combinations of $\Sigma^b_{i^2 + 2}$-formulas [3], it follows that $T^i_1 = S_2$.

4. Collapsing the polynomial hierarchy

All the work of this section is predicated on the condition that $T^i_2 = S^{i^2 + 1}_2$. We have shown above that if $T^i_2 = S^{i^2 + 1}_2$, then $T^i_2$ proves that the $\Sigma^p_{i^2 + 1}$-predicates are contained in $\Sigma^p_1/poly$. From this, the methods of Karp–Lipton [7] imply that the entire polynomial time hierarchy is contained in $\Sigma^p_{i + 1}/poly$ and in $\Pi^p_{i + 1}/poly$; furthermore, the proof of this containment can be formalized in $T^i_2$. The methods of Karp–Lipton also imply immediately that the polynomial hierarchy collapses to $\Sigma^p_{i + 3} = \Pi^p_{i + 3}$. However, we shall prove a somewhat stronger result; namely, if $T^i_2 = S^{i^2 + 1}_2$, then every polynomial hierarchy predicate (i.e., bounded formula) is $T^i_2$-provably equivalent to a Boolean combination of $\Sigma^b_{i^2 + 2}$-formulas.

To prove this, it will suffice to prove that every $\Sigma^b_{i + 3}$-formula is equivalent to a Boolean combination of $\Sigma^b_{i^2 + 2}$-formulas. Let $A(b)$ be an arbitrary $\Sigma^b_{i + 3}$-formula. From the previous section, we know that $T^i_2$ proves that $A(b)$ is equivalent to

$$
(\exists \langle \phi \rangle) [\text{Advice}^i(a(b), \langle \phi \rangle) \land A^*(b, \langle \phi \rangle)]
$$

and to

$$
(\forall \langle \phi \rangle) [\text{Advice}^i(a(b), \langle \phi \rangle) \rightarrow A^*(b, \langle \phi \rangle)],
$$

where $A^*$ is a $\Sigma^b_{i + 2}$-formula and $a = a(b)$ is function of sufficiently large polynomial growth rate. Unfortunately, $\text{Advice}^i$ is a $\Pi^p_{i + 2}$-formula and the quantifier complexity of these equivalent formulations of $A(b)$ is higher than we desire; namely, formula (4) is a $\Sigma^b_{i + 3}$-formula and formula (5) is a $\Pi^p_{i + 3}$-formula. This implies that every bounded formula is $A^i_{i + 3}$ with respect to $T^i_2$, but we wish to prove a yet stronger result.

To reduce the complexity of these formulas we would like to use $\text{PreAdvice}^i$ in place of $\text{Advice}^i$. However, this can not be done directly since if $\langle \phi \rangle$ satisfies $\text{PreAdvice}^i$, then it is not necessarily true that $A^*(b, \langle \phi \rangle)$ is equivalent to $A(b)$. Instead, we look for a longest vector $\langle \phi \rangle$ which satisfies $\text{PreAdvice}^i$; namely, consider the formula $A'(b)$ defined as:

$$
(\exists \langle \phi_1, \ldots, \phi_k \rangle) [\text{PreAdvice}^i(a(b), \langle \phi_1, \ldots, \phi_k \rangle) \land A^*(b, \langle \phi_1, \ldots, \phi_k \rangle)]
$$

$$
\lor \left( \bigwedge_{l=2}^k \neg (\exists \langle \phi_{l-1}, \ldots, \phi_k \rangle) \text{PreAdvice}^i(a(b), \langle \phi_{l-1}, \ldots, \phi_k \rangle) \right)
$$

$$
\land (\exists \langle \phi_1, \ldots, \phi_k \rangle) [\text{PreAdvice}^i(a(b), \langle \phi_1, \ldots, \phi_k \rangle) \land A^*(b, \langle \phi_1, \ldots, \phi_k \rangle)].
$$

We claim that $A'(b)$ is equivalent to $A(b)$. The proof of this now quite easy. First, there must exist a least $l \geq 1$ such that there exists $\langle \phi_1, \ldots, \phi_k \rangle$ which satisfies $\text{PreAdvice}^i$. Second, if $\text{PreAdvice}^i(\langle \phi_1, \ldots, \phi_k \rangle)$ holds and if there is no $\langle \phi_{l-1}, \ldots, \phi_k \rangle$ which
satisfies \( \text{PreAdvice}^i \), then clearly \( \langle \varphi_1, \ldots, \varphi_k \rangle \) satisfies \( \text{Advice}^i \). And for this advice, \( A^*(b, \langle \varphi \rangle) \) is equivalent to \( A(b) \).

Since \( \text{PreAdvice}^i \) is a \( \Pi^b_{i+1} \)-formula and \( A^* \) is a \( \Sigma^b_{i+2} \)-formula, \( A' \) is a Boolean combination of \( \Sigma^b_{i+2} \)-formulas. This establishes:

**Theorem 7.** If \( T_2 = S_2^{i+1} \), then every bounded formula is \( T_2 \)-provably equivalent to a Boolean combination of \( \Sigma^b_{i+2} \) formulas. In other words, if \( T_2 = S_2^{i+1} \), then the polynomial hierarchy \( T_2 \)-provably collapses to (a finite level of) the Boolean hierarchy over \( \Sigma^b_{i+2} \). Also, in this case, \( T_2 \) proves that the polynomial time hierarchy collapses to \( \Sigma^p_{i+1} / \text{poly} \).

If \( PV_1 = S_1^1(PV) \), then every bounded formula is \( PV_1 \)-provably equivalent to a Boolean combination of \( \Sigma^b_2 \)-formulas, so the polynomial time hierarchy provably collapses to the Boolean hierarchy over \( \Sigma_2 \). Also, in this case, \( PV_1 \) proves that the polynomial time hierarchy collapses to \( \text{NP}/\text{poly} \).

It should be noted again that [10] have shown that if \( T_2^{i+1} = S_2^{i+2} \) then the polynomial hierarchy collapses to \( \Sigma^p_{i+2} = \Pi^p_{i+2} \) and to \( \Delta^p_{i+1} / \text{poly} \): we do not know how to prove that this stronger collapse would be \( T_2 \)-provable.

**References**


