Bounded Arithmetic

and

Propositional Proofs

Part I: Bounded Arithmetic

Samuel R. Buss
Department of Mathematics
University of California, San Diego
La Jolla, CA 92093-0112, USA

Feasibly Constructive Proof Systems

A **constructive** proof system is one in which proofs of existence contain, or imply the existence of, algorithms for finding the object which is proved to exist. For a feasibly constructive system, the algorithm will be feasible, not merely effective.

I.e., if $\forall x \exists y A(x, y)$ is provable then there should be an algorithm to find y as a function of x.

Effective versus **Feasible**:

"Effective" means "recursive" - Church's thesis.

"Feasible" means "Computable with a reasonable amount of time or space resources".

The usual mathematical model for *feasible* is "polynomial-time computable".

PRELIMINARIES:
COMPUTATIONAL COMPLEXITY

Def'n: P is the set of polynomial time recognizable functions. FP is the set of polynomial time computable *functions*.
Functions and predicates are *arithmetic*: polynomial time means in terms of the length $|x|$ of the input x. (Instead of having functions and predicates operate on strings of characters.)

$|x| = \lceil \log_2(x+1) \rceil$ = the length of the binary representation of x

$|\vec{x}| = |x_1|, |x_2|, \ldots, |x_k|$

Cobham (1964) defined FP as the closure of some base functions under compostion and *limited iteration on notation*.

Base Functions: 0, S (successor), $\lfloor \frac{1}{2}x \rfloor$, $2 \cdot x$,

$$x \leq y = \begin{cases} 1 & \text{if } x \leq y \\ 0 & \text{otherwise} \end{cases}$$

$$Choice(x, y, z) = \begin{cases} y & \text{if } x > 0 \\ z & \text{otherwise} \end{cases}$$

Def'n: Let q be a polynomial. f is defined from g and h by **limited iteration on notation** with space bound q iff

$$\begin{aligned} f(\vec{x}, 0) &= g(\vec{x}) \\ f(\vec{x}, y) &= h(\vec{x}, y, f(\vec{x}, \lfloor \tfrac{1}{2}y \rfloor)) \end{aligned}$$

provided $|f(\vec{x}, y)| \leq q(|\vec{x}|, |y|)$ for all $\vec{x}, y$.

Def'n: NP is the set of non-deterministic polynomial time computable predicates. *Co-NP* is the set of complements of NP predicates.

Def'n: If Ψ is a set of predicates, $PB\exists(\Psi)$ is the set of predicates A expressible as

$$\vec{x} \in A \Leftrightarrow (\exists y \leq 2^{p(|\vec{x}|)})B(\vec{x}, y)$$

for some polynomial p and some $B \in \Psi$.
$PB\forall(\Psi)$ is defined similarly with universal polynomially bounded quantification.

Now, $NP = PB\exists(P)$ and *co-NP* $= PB\forall(P)$.

Def'n: If Ψ is a set of predicates, P^{Ψ} (resp., FP^{Ψ})is the set of predicates (resp., functions) polynomial time recognizable with oracles for a finite number of predicates in Ψ.

Polynomial Time "Hierarchy" :

$$
\begin{aligned}
\Box_1^p &= FP \\
\Delta_1^p &= P \\
\Sigma_k^p &= PB\exists(\Delta_k^p) \\
\Pi_k^p &= PB\forall(\Delta_k^p) \\
\Delta_{k+1}^p &= P^{\Sigma_k^p} = P^{\Pi_k^p} \\
\Box_{k+1}^p &= FP^{\Sigma_k^p} = FP^{\Pi_k^p}
\end{aligned}
$$

$$
\begin{array}{ccc}
\vdots & & \vdots \\
\Delta_3^p & & \Box_3^p \\
\Pi_2^p & \quad & \Sigma_2^p \\
\Delta_2^p & & \Box_2^p \\
co\text{-}NP = \Pi_1^p & & \Sigma_1^p = NP \\
P = \Delta_1^p & & \Box_1^p = FP
\end{array}
$$

Open Question: Is this hierarchy proper?

The Equational Theory PV

(Cook, 1975) PV - "Polynomially Verifiable"

- an equational theory analogous to PRA except for polynomial time computability.
- based on dyadic representation of integers as strings of 0's and 1's.
- base functions include $\leq$, $Cond$, concatenation and iterated concatenation and two successor functions s_1 and s_2:

$$s_{i+1}(x) = 2x + i$$

- additional function symbols definable by limited iteration; this gives all FP functions.

- a length induction rule (for equations A):

$$\frac{A(0) \quad A(1) \quad (\forall x)(A(x) \supset A(s_1(x)) \wedge A(s_2(x)))}{(\forall x)A(x)}$$

- Statman showed that the full induction (for equations) holds:

$$A(0) \wedge A(1) \wedge (\forall x)(A(x) \supset A(x+1)) \supset (\forall x)A(x)$$

- Bounded Arithmetic will present more appropriate induction axioms for polynomial time/hierarchy complexity

- PV can define precisely the polynomial time functions; equations express precisely the polynomial time predicates.

- Connections between PV and polynomial size extended Frege proofs

Language of Bounded Arithmetic

First-order language for $\mathbf{N}$ with function symbols 0, S, $+$, $\cdot$, $\lfloor \frac{1}{2}x \rfloor$, $|x|$, $\#$ and relation symbol $\leq$, where

$$x\#y = 2^{|x|\cdot|y|}$$

The $\#$ (pronounced "smash") function allows us to express $2^{q(|\vec{a}|)}$ for q any polynomial with positive integer coefficients.

Def'n: A *bounded quantifier* is a quantifier of the form $(Qx \leq t)$ with t a term. A *sharply bounded quantifier* is one of the form $(Qx \leq |t|)$. $(\forall x)$ and $(\exists x)$ are unbounded quantifiers. A *bounded* formula is one with no unbounded quantifiers.

A hierarchy of classes Σ_k^b, Π_k^b of bounded formulas is defined by counting alternations of bounded quantifiers, ignoring sharply bounded quantifiers. (Analogous to defining the arithmetic hierarchy by counting unbounded quantifiers, ignoring bounded quantifiers.)

$\Sigma_0^b = \Pi_0^b$ is the set of formulas with only sharply bounded quantifiers.
If $A \in \Sigma_k^b$ then $(\forall x \leq |t|)A$ and $(\exists x \leq t)A$ are in Σ_k^b and $(\forall x \leq t)A$ is in Π_{k+1}^b.
Dually, if $A \in \Pi_k^b$ then $(\exists x \leq |t|)A$ and $(\forall x \leq t)A$ are in Π_k^b and $(\exists x \leq t)A$ is in Σ_{k+1}^b.
Connectives $\wedge$, $\vee$, $\neg$, $\supset$ are treated in the usual manner.

Thm: Fix $k \geq 1$. A predicate Q is in Σ_k^p iff there is a Σ_k^b formula which defines it.

Pf: (Stockmeyer-Wrathall, 1976),
(Kent-Hodgson, 1982)

Reasons the # function and sharply bounded quantifiers are natural choices:

- # has the right growth rate for polynomial time computation.
- the above theorem defines the Σ, Π classes of the polynomial hierarchy syntactically (without use of computation),
- Quantifier Exchange Principle:

$$(\forall x \leq |a|)(\exists y \leq b)A(x,y) \leftrightarrow$$
$$\leftrightarrow (\exists y \leq (2a+1)\#(4(2b+1)^2))(\forall x \leq |a|)$$
$$[A(x,\beta(x+1,y)) \wedge \beta(x+1,y) \leq b]$$

- Ed Nelson introduced # for defining substitution syntactically. Wilkie-Paris have used Ω_1 ("$x^{|\log x}$ is total") similarly.

Induction Axioms for Bounded Arithmetic

The *IND* axioms are the usual induction axioms. The *PIND* and *LIND* axioms are "polynomial" and "length" induction axioms that are intended to be feasibly effective forms of induction.

Σ^b_k-IND: For $A \in \Sigma^b_k$,

$$A(0) \wedge (\forall x)(A(x) \supset A(x+1)) \supset (\forall x)A(x)$$

Σ^b_k-PIND: For $A \in \Sigma^b_k$,

$$A(0) \wedge (\forall x)(A(\lfloor \tfrac{1}{2}x \rfloor) \supset A(x)) \supset (\forall x)A(x)$$

Σ^b_k-LIND: For $A \in \Sigma^b_k$,

$$A(0) \wedge (\forall x)(A(x) \supset A(x+1)) \supset (\forall x)A(|x|)$$

Σ^b_k-LIND and Σ^b_k-PIND typically are equivalent and are (strictly?) weaker than Σ^b_k-IND. Exponentiation is not provably total in Bounded Arithmetic.

Theories of Bounded Arithmetic

Def'n: T_2^i is the first-order theory with language $0, S, +, \cdot, \lfloor \frac{1}{2}x \rfloor, |x|, \#$ and $\leq$ and axioms:

(1) A finite set, BASIC, of (universal closures of) open axioms defining simple properties of the function and relation symbols. BASIC properly contains Robinson's Q since it has to be used with weaker induction axioms.

(2) The Σ_i^b-IND axioms.

T_2^{-1} has no induction axioms.
T_2 is the union of the T_2^i's.

T_2 is equivalent to $I\Delta_0 + \Omega_1$ (Parikh, Wilkie-Paris) except for differences in the nonlogical language.

Def'n: S_2^i is the first-order theory with language $0, S, +, \cdot, \lfloor \frac{1}{2}x \rfloor, |x|, \#$ and $\leq$ and axioms:

(1) The BASIC axioms, and

(2) The Σ_i^b-PIND axioms.

$S_2^{-1} = T_2^{-1}$ has no induction axioms.
S_2 is the union of the S_2^i's.

Thm: (Buss, 1985). Let $i \geq 1$.

$$T_2^i \vdash S_2^i$$

and

$$S_2^i \vdash T_2^{i-1}.$$

So $S_2 \equiv T_2$.

Open: Are the inclusions proper?

$$S_2^1 \subseteq T_2^1 \subseteq S_2^2 \subseteq T_2^2 \subseteq \cdots$$

Def'n: Let $f: \mathbf{N}^k \mapsto \mathbf{N}$. f is Σ_i^b-*definable* by a theory R iff there is a formula $A(\vec{x}, y) \in \Sigma_i^b$ and a term t so that

(1) For all $\vec{n} \in \mathbf{N}^k$, $A(\vec{n}, f(\vec{n}))$ is true.

(2) $R \vdash (\forall \vec{x})(\exists y \leq t) A(\vec{x}, y)$

(3) $R \vdash (\forall \vec{x}, y, z)(A(\vec{x}, y) \wedge A(\vec{x}, z) \supset y = z)$

Def'n: Let $Q \subseteq \mathbf{N}$. Q is Δ_i^b-definable by a theory R iff there is a Σ_i^b-formula A and Π_i^b-formula B that define Q so that A and B are provably equivalent in R. A formula is Δ_i^b w.r.t R iff it is provably equivalent to a Σ_i^b- and to a Π_i^b-formula.

Bootstrapping Thm: Every polynomial time function is Σ_1^b-definable by S_2^1 and every polynomial time predicate is Δ_1^b-definable by S_2^1.

Thm: Any Σ_1^b-definable function or Δ_1^b-definable predicate of S_2^i may be introduced into the non-logical language and used freely in induction axioms.

Pf: If f is Σ_1^b-defined by R:

$$R \vdash (\forall x)(\exists! y \leq r(\vec{x}))A_f(\vec{x}, y).$$

An atomic formula $\varphi(f(\vec{s}))$ is equivalent to both

$$(\exists y \leq r(\vec{s}))(A_f(\vec{s}, y) \wedge \varphi(y))$$

and

$$(\forall y \leq r(\vec{s}))(A_f(\vec{s}, y) \supset \varphi(y))$$

Thus any Σ_i^b-formula involving f is equivalent to one not involving f by tranforming atomic subformulas as above (and by removing f from the quantifier bounds). $\square$

Main Theorems for S_2^i

Theorem: (Buss, 1985) Let $i \geq 1$. Let A be a Σ_i^b-formula. Suppose $S_2^i \vdash (\forall \vec{x})(\exists y) A(\vec{x}, y)$. Then there is a Σ_i^b-formula B and a function $f \in \Box_i^p$ and a term t so that

(1) $S_2^i \vdash (\forall \vec{x}, y)(B(\vec{x}, y) \supset A(\vec{x}, y))$.

(2) $S_2^i \vdash (\forall \vec{x})(\exists! y) B(\vec{x}, y)$.

(3) $S_2^i \vdash (\forall \vec{x})(\exists y \leq t) B(\vec{x}, y)$. [Parikh]

(4) For all $\vec{n}$, $\mathbf{N} \models B(\vec{n}, f(\vec{n}))$.

Conversely: If $f \in \Box_i^p$ then there is a formula $B \in \Sigma_i^b$ and a term t so that (2), (3) and (4) hold.

Corollary: ($i \geq 1$) The Σ_i^b-definable functions of S_2^i are precisely the functions in $\Box_i^p$.

To restate in terms of predicates:

Theorem: ($i \geq 1$). Suppose $A(\vec{x}) \in \Sigma_i^b$ and $B(\vec{x}) \in \Pi_i^b$ and $S_2^i \vdash A \leftrightarrow B$. Then there is a predicate $Q \in \Delta_i^p$ so that, for all $\vec{n}$,

$$Q(\vec{n}) \Leftrightarrow \mathbf{N} \models A(\vec{n}) \Leftrightarrow \mathbf{N} \models B(\vec{n})$$

Conversely, if $Q \in \Delta_i^p$ then there are A and B so that the above holds.

So, the Δ_i^b-definable predicates of S_2^i are precisely the Δ_i^p-predicates.

Special case when $i = 1$: If A is a formula which is S_2^1-provably in *NP*∩*co-NP* then A defines a polynomial time predicate (provably in S_2^1). Being provably in *NP* ∩ *co-NP* means provably equivalent to a Σ_1^b- and to a Π_1^b-formula.

The Sequent Calculus

To prove the Main Theorem, we shall formalize S_2^i in Gentzen's sequent calculus.

$\wedge, \vee, \neg, \supset, \forall, \exists$ are the logical symbols.

$\longrightarrow$ is the sequent connective.

Def'n: A *sequent* is of the form

$$A_1, A_2, \ldots, A_n \longrightarrow B_1, B_2, \ldots, B_k$$

where the A_i's and B_i's are formulas. Its intended meaning is

$$(A_1 \wedge A_2 \wedge \cdots \wedge A_n) \supset (B_1 \vee \cdots \vee B_k)$$

Greek letters $\Gamma, \Delta, \ldots$ are used to denote finite sequences of formulas separated by commas ("cedents").

An *LK-proof* is a tree of sequents: the leaves or *initial sequents* must be of the form $A \longrightarrow A$; the root, or *endsequent*, is what is proved; and the valid inferences are:

$$\frac{\Gamma \longrightarrow \Delta, A}{\neg A, \Gamma \longrightarrow \Delta} \qquad \frac{A, \Gamma \longrightarrow \Delta}{\Gamma \longrightarrow \Delta, \neg A}$$

$$\frac{\Gamma \longrightarrow \Delta, A \qquad \Gamma \longrightarrow \Delta, B}{\Gamma \longrightarrow \Delta, A \wedge B}$$

$$\frac{A, \Gamma \longrightarrow \Delta}{A \wedge B, \Gamma \longrightarrow \Delta} \qquad \frac{B, \Gamma \longrightarrow \Delta}{A \wedge B, \Gamma \longrightarrow \Delta}$$

$$\frac{A, \Gamma \longrightarrow \Delta \qquad B, \Gamma \longrightarrow \Delta}{A \vee B, \Gamma \longrightarrow \Delta}$$

$$\frac{\Gamma \longrightarrow \Delta, A}{\Gamma \longrightarrow \Delta, A \vee B} \qquad \frac{\Gamma \longrightarrow \Delta, B}{\Gamma \longrightarrow \Delta, A \vee B}$$

$$\frac{\Gamma \longrightarrow \Delta, A \qquad B, \Gamma \longrightarrow \Delta}{A \supset B, \Gamma \longrightarrow \Delta}$$

$$\frac{A, \Gamma \longrightarrow \Delta, B}{\Gamma \longrightarrow \Delta, A \supset B}$$

$$\frac{A(b), \Gamma \longrightarrow \Delta}{(\exists x)A(x), \Gamma \longrightarrow \Delta} \qquad \frac{\Gamma \longrightarrow \Delta, A(t)}{\Gamma \longrightarrow \Delta, (\exists x)A(x)}$$

$$\frac{A(t), \Gamma \longrightarrow \Delta}{(\forall x)A(x), \Gamma \longrightarrow \Delta} \qquad \frac{\Gamma \longrightarrow \Delta, A(b)}{\Gamma \longrightarrow \Delta, (\forall x)A(x)}$$

In the quantifier inferences the free variable b is called the *eigenvariable* and must not appear in the lower sequent.

$$\frac{\Gamma \longrightarrow \Delta}{A, \Gamma \longrightarrow \Delta} \qquad \frac{\Gamma \longrightarrow \Delta}{\Gamma \longrightarrow \Delta, A}$$

$$\frac{\Gamma, A, B, \Pi \longrightarrow \Delta}{\Gamma, B, A, \Pi \longrightarrow \Delta} \qquad \frac{\Gamma \longrightarrow \Delta, A, B, \Lambda}{\Gamma \longrightarrow \Delta, B, A, \Lambda}$$

$$\frac{A, A, \Gamma \longrightarrow \Delta}{A, \Gamma \longrightarrow \Delta} \qquad \frac{\Gamma \longrightarrow \Delta, A, A}{\Gamma \longrightarrow \Delta, A}$$

Cut: $$\frac{\Gamma \longrightarrow \Delta, A \qquad A, \Pi \longrightarrow \Lambda}{\Gamma, \Pi \longrightarrow \Delta, \Lambda}$$

Theorem: (Gentzen)

- LK is complete.
- LK without the Cut inference is complete.

So if P is an LK-proof of $\Gamma \longrightarrow \Delta$ then there is a cut-free proof P^* of $\Gamma \longrightarrow \Delta$. There is an effective (but not feasible) procedure to obtain P^* ¿from P.

Because there are no cuts in P^*, every formula appearing in P^* will be a subformula (in a wide sense) of a formula in $\Gamma \longrightarrow \Delta$. This gives a bound on the logical complexity of formulas needed to prove $\Gamma \longrightarrow \Delta$.

To formulate sequent calculus systems of Bounded Arithmetic, we enlarge LK as follows:

(1) Allow equality axioms and BASIC axioms as initial sequents. An initial sequent will contain only atomic formulas.

(2) Add inferences for bounded quantifiers (the variable b occurs only as indicated):

$$\frac{b \leq s, A(b), \Gamma \longrightarrow \Delta}{(\exists x \leq s)A(x), \Gamma \longrightarrow \Delta}$$

$$\frac{\Gamma \longrightarrow \Delta, A(t)}{t \leq s, \Gamma \longrightarrow \Delta, (\exists x \leq s)A(x)}$$

$$\frac{A(t), \Gamma \longrightarrow \Delta}{t \leq s, (\forall x \leq s)A(x), \Gamma \longrightarrow \Delta}$$

$$\frac{b \leq s, \Gamma \longrightarrow \Delta, A(b)}{\Gamma \longrightarrow \Delta, (\forall x \leq s)A(x)}$$

(3) Allow induction inferences: (for $A \in \Sigma_i^b$)

$$\Sigma_i^b\text{-IND} \quad \frac{A(b), \Gamma \longrightarrow \Delta, A(b+1)}{A(0), \Gamma \longrightarrow \Delta, A(t)}$$

$$\Sigma_i^b\text{-PIND} \quad \frac{A(\lfloor \frac{1}{2} b \rfloor), \Gamma \longrightarrow \Delta, A(b)}{A(0), \Gamma \longrightarrow \Delta, A(t)}$$

S_2^i and T_2^i may be formulated as sequent calculus systems with BASIC axioms as initial sequents and with Σ_i^b-PIND and Σ_i^b-IND inference rules, respectively. With side formulas, the induction inferences are equivalent to the induction axioms.

Def'n: A cut inference

$$\frac{\Gamma \longrightarrow \Delta, A \qquad A, \Pi \longrightarrow \Lambda}{\Gamma, \Pi \longrightarrow \Delta, \Lambda}$$

is *free* unless A is the direct descendant either of a formula in an initial sequent or of a principal formula of an induction inference.

Free-Cut Elimination Theorem
(essentially due to Gentzen and Takeuti):
If P is an S^i_2-proof (or T^i_2-proof) then there is a proof P^* in the same theory with the same end-sequent which contains no free cuts.

In a free-cut free proof, every formula will be a subformula (in a wide sense) of an induction formula, of a formula in an axiom or of a formula in the conclusion. (The wide sense allows terms to change.)
In S^i_2 and T^i_2, cut formulas may be restricted to be Σ^b_i-formulas.

To prove the Main Theorem

Step 1: Start with an S_2^i-proof P of

$$\longrightarrow (\exists y) A(\vec{c}, y).$$

By free-cut elimination, there is an S_2^i proof P^* of $\longrightarrow (\exists y \leq t) A(\vec{c}, y)$ such that every formula in P^* is a Σ_i^b-formula.

Step 2: Given the proof P^* we will extract an algorithm to compute a function $f(\vec{c})$ such that $A(\vec{n}, f(\vec{n}))$ is true for all n. The function f will be in $\Box_i^p$ and will be Σ_i^b-defined by S_2^i. Furthermore, S_2^i will prove $(\forall \vec{x}) A(\vec{x}, f(\vec{x}))$.

P^* can be thought of as a program plus a proof that it is correct.

Proof of Step 2 follows...

The Witness Predicate

Def'n: Let $B(\vec{a})$ be a Σ_i^b-formula with all free variables indicated. Then $\mathit{Witness}_B^{i,\vec{a}}(w,a)$ is a formula defined inductively:

(1) If $B \in \Sigma_{i-1}^b \cup \Pi_{i-1}^b$ then

$$\mathit{Witness}_B^{i,\vec{a}}(w,\vec{a}) \Leftrightarrow B(\vec{a})$$

(2) If $B = C \vee D$ then

$$\mathit{Witness}_B^{i,\vec{a}}(w,\vec{a}) \Leftrightarrow \mathit{Witness}_C^{i,\vec{a}}(\beta(1,w),\vec{a}) \vee \mathit{Witness}_D^{i,\vec{a}}(\beta(2,w),\vec{a})$$

(3) If $B = C \wedge D$ then

$$\mathit{Witness}_B^{i,\vec{a}}(w,\vec{a}) \Leftrightarrow \mathit{Witness}_C^{i,\vec{a}}(\beta(1,w),\vec{a}) \wedge \mathit{Witness}_D^{i,\vec{a}}(\beta(2,w),\vec{a})$$

(4) If $B = (\exists x \leq t)C(\vec{a},x)$ then

$$\mathit{Witness}_B^{i,\vec{a}}(w,\vec{a}) \Leftrightarrow \beta(1,w) \leq t \wedge \mathit{Witness}_{C(\vec{a},b)}^{i,\vec{a},b}(\beta(2,w),\vec{a},\beta(1,w))$$

(5) If $B = (\forall x \leq |t|)C(\vec{a}, x)$ then

$$Witness_B^{i,\vec{a}}(w, \vec{a}) \Leftrightarrow$$

$$(\forall x \leq |t|)Witness_{C(\vec{a},b)}^{i,\vec{a},b}(\beta(x+1, w), \vec{a}, x)$$

(6) If $B = \neg C$ use prenex operations to push the negation sign inside.

Lemma: Let $B \in \Sigma_i^b$.

(1) For some term t_B, S_2^i proves

$$B(\vec{a}) \leftrightarrow (\exists w \leq t_B)Witness_B^{i,\vec{a}}(w, \vec{a})$$

(2) $Witness_B^{i,\vec{a}} \in \Delta_i^p$ $(= P$ if $i = 1)$

(3) $Witness_B^{i,\vec{a}}$ is Δ_i^b with respect to S_2^i.

Pf: Induction on complexity of B.

The Main Lemma

Lemma: Suppose $S_2^i \vdash \Gamma \longrightarrow \Delta$ where Γ and Δ contain only Σ_i^b-formulas. Let $\vec{c}$ be the free variables in Γ and Δ. Then, there is a function f such that

(1) f is Σ_i^b-defined by S_2^i

(2) S_2^i proves

$$Witness^{i,\vec{c}}_{\bigwedge\Gamma}(w,\vec{c}) \supset Witness^{i,\vec{c}}_{\bigvee\Delta}(f(w,\vec{c}),\vec{c})$$

(3) $f \in \Pi_i^p$ $(= FP$ if $i = 1)$

Proof is by induction on the number of inferences in a free-cut free S_2^i-proof of $\Gamma \longrightarrow \Delta$.

As an example of one case of the proof of the main lemma, suppose that P is a free-cut free proof and ends with the inference

$$\frac{B(\lfloor \frac{1}{2}a \rfloor) \longrightarrow B(a)}{B(0) \longrightarrow B(t)}$$

By the induction hypothesis, there is a function g so that

(1) g is Σ_i^b-defined by S_2^i

(2) g is in $\Box_i^p$ $(= FP$ if $i = 1)$

(3) $S_2^i \vdash \mathit{Witness}^{i,a,\vec{c}}_{B(\lfloor \frac{1}{2}a \rfloor)}(w, a, \vec{c}) \supset$

$$\supset \mathit{Witness}^{i,a,\vec{c}}_{B(a)}(g(w, a, \vec{c}), a, \vec{c}).$$

(4) $S_2^i \vdash (\forall a, \vec{c})[g(w, a, \vec{c}) \leq t_B(a, \vec{c})]$

Now define f by limited iteration as

$$\begin{aligned} f(w,0,\vec{c}) &= g(w,0,\vec{c}) \\ f(w,a,\vec{c}) &= g(f(w,\lfloor \tfrac{1}{2}a \rfloor,\vec{c}),a,\vec{c}) \end{aligned}$$

so $f(w,a,\vec{c}) \leq t_B(a,\vec{c})$ and the following hold:

(1) $f \in \Pi_i^p$ $(= FP$ if $i = 1)$

Pf: Since f is defined by limited iteration from g

(2) S_2^i can Σ_i^b-define f and prove that f satisfies the above conditions

(3) $S_2^i \vdash \mathit{Witness}^{i,a,\vec{c}}_{B(0,\vec{c})}(w,a,\vec{c}) \supset$

$$\supset \mathit{Witness}^{i,a,\vec{c}}_{B(a,\vec{c})}(f(w,a,\vec{c}),a,\vec{c}).$$

Pf: Since $\mathit{Witness}^{i,b,\vec{c}}_{B(a,\vec{c})}$ is a Σ_i^b-formula, S_2^i can prove this by Σ_i^b-PIND directly from the induction hypothesis.

Q.E.D. Main Lemma

Proof of Main Theorem from Main Lemma

Suppose $S_2^i \vdash (\forall \vec{x})(\exists y) A(\vec{x}, y)$. By a theorem of Parikh, there is a term t so that S_2^i proves $\longrightarrow (\exists y \leq t) A(\vec{c}, y)$. By the Main Lemma,

$$S_2^i \vdash \mathit{Witness}^{i,\vec{c}}_{(\exists y \leq t) A}(g(\vec{c}), \vec{c})$$

for some Σ_i^b-defined function g. Define

$$B(\vec{c}, y) \text{ to be } y = \beta(1, g(\vec{c}))$$

Since g is Σ_i^b-defined by S_2^i, B is a Σ_i^b-formula and by the properties of *Witness*,

$$S_2^i \vdash (\forall \vec{x}, y)(B(\vec{x}, y) \supset A(\vec{x}, y))$$

Finally, define $f(\vec{c}) = \beta(1, g(\vec{c}))$.
Q.E.D. Main Theorem

Other Axioms for Bounded Arithmetic

Let Ψ be a set of formulas. The axioms below are schemes where $A \in \Psi$:

Ψ-MIN: (Minimization)

$$(\exists x)A(x) \supset (\exists x)[A(x) \wedge (\forall y < x)(\neg A(y))]$$

Ψ-LMIN: (Length minimization)

$$(\exists x)A(x) \supset A(0) \vee (\exists x)[A(x) \wedge (\forall y \leq \lfloor \tfrac{1}{2}x \rfloor)(\neg A(y))]$$

Ψ-replacement:

$$\begin{aligned}(\forall x \leq |t|)(\exists y \leq s)A(x,y) \leftrightarrow \\ \leftrightarrow (\exists w \leq SqBd(t,s))(\forall x \leq |t|) \\ (A(x, \beta(Sx,w)) \wedge \beta(Sx,w) \leq s)\end{aligned}$$

strong Ψ-replacement:

$$\begin{aligned}(\exists w \leq SqBd(t,s))(\forall x \leq |t|) \\ [(\exists y \leq s)A(x,y) \leftrightarrow \\ \leftrightarrow A(x, \beta(Sx,w) \wedge \beta(Sx,w) \leq s]\end{aligned}$$

For $i \geq 1$, relative to S_2^1:

Σ_i^b-IND $\Leftrightarrow$ Π_i^b-IND $\Leftrightarrow$ Σ_i^b-MIN $\Leftrightarrow$ Δ_{i+1}^b-IND

$\Downarrow$

Σ_i^b-PIND $\Leftrightarrow$ Π_i^b-PIND $\Leftrightarrow$ Σ_i^b-LIND $\Leftrightarrow$ Π_i^b-LIND

$\Updownarrow$

Σ_i^b-LMIN $\Leftrightarrow$ strong Σ_i^b-replacement

$\Downarrow$ $\Updownarrow$

Σ_{i-1}^b-IND $(\Sigma_{i+1}^b \cap \Pi_{i+1}^b)$-PIND

Σ_{i+1}^b-MIN $\Leftrightarrow$ Π_i^b-MIN

Σ_{i+1}^b-replacement $\Rightarrow$ Σ_i^b-PIND $\Rightarrow$ Σ_i^b-replacement

$S_2^{i+1} \underset{\Sigma_{i+1}^b}{\succ} T_2^i$

$S_2^{i+1} \underset{\mathcal{B}(\Sigma_{i+1}^b)}{\succ} T_2^i + \Sigma_{i+1}^b$-replacement

Thm: (Buss, 1985)

$S_2^1 + \Sigma_i^b\text{-}PIND \vdash \Delta_i^b$-IND.

Hence $S_2^i \supset T_2^{i-1}$.

Pf: Suppose A is Δ_i^b w.r.t. S_2^i. Assume $(\forall x)(A(x) \supset A(x+1))$ and argue inside S_2^i. Let $B(x,z)$ be the formula

$$(\forall w \leq x)(\forall y \leq z+1)(A(w \dot{-} y) \supset A(w)).$$

So B is equivalent to a Π_i^b-formula. Now by definition of B, $(\forall x, z)(B(x, \lfloor \frac{1}{2} z \rfloor) \supset B(x,z))$ and hence by Π_i^b-PIND on $B(x,z)$ w.r.t z,

$$(\forall x)(B(x,0) \supset B(x,x)).$$

By the assumption, $(\forall x)B(x,0)$; hence $(\forall x)B(x,x)$, ¿from whence

$$(\forall x)(A(0) \supset A(x))$$

□

Conservation Results

Thm: (Buss, 1987) Let $i \geq 1$.

S_2^{i+1} is conservative over T_2^i with respect to Σ^b_{i+1}-formulas, and hence with respect to $\forall\exists\Sigma^b_{i+1}$-sentences.

Pf: (Idea). Fix $i \geq 1$ and let Z be PV or T_2^{i-1} as appropriate. First show that every Π^p_i-function is definable in Z in an appropriate sense. For $i = 1$, there is a function symbol for every polynomial time function; for $i \geq 1$, we show that every Π^p_i-function can be "Q_i-defined" — this is stronger than "Σ^b_i-defined". Second, prove a stronger version of the Main Lemma above; in essence, we partially formalize the Main Lemma in Z and prove that the witnessing function f is defined appropriately in Z. Namely, we prove:

Lemma: If $S_2^i \vdash A$ with $A \in \Sigma^b_i$ then $Z \vdash A$.

Witnessing Theorem for T_2^1

Defn: [Papadimitriou] A *Polynomial Local Search (PLS)*, problem is specified by polynomial time functions F, N, c:

(1) $c(s, x)$ is a *cost function*,

(2) $N(s, x)$ is a neighborhood function, such that for all s s.t. $F(x, s)$

$$c(N(s, x), x) \leq c(s, x)$$

(3) $\{s : F(s, x)\}$ is the solution space for input x, and $F(0, x)$ always holds, and such that, if $F(s, x)$, then $|s| < p(|x|)$ for p a polynomial.

A solution to the PLS problem is a (multivalued) function f, s.t., for all x,

$$c(N(f(x), s), s) = c(f(x), x).$$

Thm [Buss-Krajíček'94] Suppose T_2^1 proves $(\forall x)(\exists y)A(x,y)$ where $A \in \Sigma_1^b$. Then there is a PLS function $f(x) = y$ and a polynomial time function π such that

$$T_2^1 \vdash (\forall x)A(x, \pi \circ f(x)).$$

Furthermore, every PLS function (and every function $\pi \circ f$) is Σ_1^b-definable by T_2^1.

Corollary The same holds for S_2^2 by conservativity of S_2^2 over T_2^1.

Proof-idea: A free-cut free T_2^1-proof can be transformed into a PLS problem. □

The KPT Witnessing Theorem

Thm [Krajíček-Pudlák-Takeuti,91]
Suppose $A \in \Sigma^b_{i+2}$ and $T^i_2 \vdash (\forall x)(\exists y)(\forall z)A(x,y,z)$.
Then there $k > 0$ and functions $f_i(x, z_1, ..., z_{i-1})$
so that

(1) Each f_i is Σ^b_{i+1}-defined by T^i_2.

(2) T^i_2 proves

$$(\forall x)[(\forall z_1)[A(x, f_1(x), z_1) \vee$$
$$(\forall z_2)[A(x, f_2(x, z_1), z_2) \vee$$
$$(\forall z_3)[A(x, f_3(x, z_1, z_2), z_3) \vee \cdots$$
$$(\forall z_k)[A(x, f_k(x, z_1, \ldots, z_{k-1}), z_k)] \cdots]]].$$

This is called a "nocounterexample interpretation"; and is a special form of a generalized Herbrand's theorem (see [Buss'95]).

Thm [KPT'91; Buss'9?,Zambella'9?] If $T_2^i = S_2^{i+1}$, then the polynomial time hierarchy collapses, provably in T_2^i. In fact, in this case, T_2^i proves that every Σ_3^p predicate is (a) equivalent to a Boolean combination of Σ_2^p-predicates and (b) is in $\Sigma_1^p/poly$.

Proof-idea For simplicity, assume $i = 0$. Suppose $T_2^0(PV) = S_2^1$. Let $\overline{\varphi}$ represent a vector of Boolean formula $\overline{\varphi} = \langle \varphi_1, \ldots, \varphi_n \rangle$. Then $T_2^0(PV)$ proves

$$\begin{aligned} &\forall \overline{\varphi} (\exists \ell \leq n)(\exists \langle w_1, \ldots, w_\ell \rangle) \\ &\qquad [(\forall j \leq \ell)(w_j \text{ satisfies } \varphi_j) \\ &\qquad \wedge \text{“}\ell = n \text{ or } \varphi_{\ell+1} \text{ is unsatisfiable”}] \end{aligned}$$

The formula in $[\cdots]$ is in Π_1^b, so the KPT witnessing theorem can be applied to get $k > 0$ and polynomial time functions $f_1, \ldots, f_k$ so that $T_2^0(PV)$ proves (setting $n = k$) that given $\varphi_1, \ldots, \varphi_k$ satisfied by $w_1, \ldots w_k$, that one of $f_j(\overline{\varphi}, w_1, \ldots, w_{j-1})$ produces a witness to φ_j. [Note that f_j has all φ_i's as input.]

Let $PreAdvice(a, \langle \varphi_{\ell+1}, \ldots, \varphi_k \rangle)$ mean that for all $\varphi_1, \ldots, \varphi_\ell < a$ (not nec. satisfiable), that $f_j(\overline{\varphi}, w_1, \ldots, w_{j-1})$ satisfies φ_j for some $j \leq \ell$.

Let $Advice(a, \langle \varphi_{\ell+1}, \ldots, \varphi_k \rangle)$ mean that $PreAdvice$ holds, and that ℓ is the minimum possible value for which there is such *PreAdvice*.

Claim: $T_2^0(PV)$ proves, that if $\varphi_\ell < a$ and if $Advice(a, \langle \varphi_{\ell+1}, \ldots, \varphi_k \rangle)$, then φ_ℓ is satisfiable if and only if for all $\varphi_1, \ldots, \varphi_{\ell-1}$, satisfied by $w_1, \ldots, w_{\ell-1}$, there is $j \leq \ell$ such that $f_j(\overline{\varphi}, w_1, \ldots, w_{j-1})$ satisfies φ_j.

Pf: If the latter condition is true, then the only way for $\langle \varphi_\ell, \ldots, \varphi_k \rangle$ to not be "preadvice", (which it isn't, by def'n of "advice") is for φ_ℓ to be satisfied by $f_\ell(\overline{\varphi}, \vec{w})$ for some $\varphi_1, \ldots, \varphi_{\ell-1}$, $w_1, \ldots, w_{\ell-1}$. □

Note that this means that the NP complete property of satisfiability is in coNP relative to the polynomial size advice, $\langle \varphi_{\ell-1}, \ldots, \varphi_k \rangle$.

The above shows that $T_2^0(PV)$ would prove $NP \subseteq coNP/poly$. ¿From this, Karp-Lipton style methods can show that $T_2^0(PV)$ proves the polynomial time hierarchy collapses.

In fact it can be shown that $T_2^0(PV)$ proves that every polynomial time hierarchy predicate is equivalent to Boolean combination of Σ_2^p predicates. The proof idea is that the property $PreAdvice$ is in $coNP$ and therefore, property

$$PAlen(\ell) \equiv \exists \langle \varphi_{\ell+1}, \ldots, \varphi_k \rangle PreAdvice(a, \langle \vec{\varphi} \rangle)$$

is a Σ_2^p-property. Q.E.D.

Similar methods work for $i \geq 1$.