Capturing Complexity Classes by their Reasoning Power

Stephen Cook
(joint work with Phuong Nguyen)

April 11, 2006

advertisement advertisement advertisement

Foundations of Proof Complexity: Bounded Arithmetic and Propositional Translations

Stephen Cook

Phuong Nguyen

New partial draft (300 pages) available on our web sites.

Comments and Corrections Welcome!

Complexity Classes:

$\mathbf{AC^0} \subset \mathbf{AC^0}(2) \subset \mathbf{TC^0} \subseteq \mathbf{NC^1} \subseteq \mathbf{L} \subseteq \mathbf{NL} \subseteq \mathbf{NC} \subseteq \mathbf{P}$

Complexity Questions:
$\mathbf{AC^0}(6) = \mathbf{TC^0} = \ldots = \mathbf{P} = \mathbf{NP} = \mathbf{PH}$??

Proof Complexity Questions:
(1) Given a combinatorial principle, what is the least complexity class containing enough concepts to prove the principle?
(LOW LEVEL REVERSE MATHEMATICS)

(2) Can the correctness of an algorithm solving a problem in a given complexity class be proved using concepts in that class?

Examples:
pigeonhole principle
Jordan curve theorem
matrix identities
computing the determinant
prime recognition
L = SL
SL = coSL

Example: Pigeonhole Principle

If $n+1$ pigeons are assigned to n holes, then some hole has at least two pigeons.

Theorem (Buss, 1987) $\mathbf{NC^1}$ suffices.
His proof shows $\mathbf{TC^0}$ suffices.
Theorem (Ajtai, 1988) $\mathbf{AC^0}$ does not suffice.
Open Question: Does $\mathbf{AC^0}(2)$ suffice?

Example: Matrix Identities

$$AB = I \supset BA = I$$

where A, B are $n \times n$ matrices (say over $GF(2)$).

Theorem (Soltys, 2001) **P** suffices.
Theorem (Soltys-Urquhart, 2003) $\mathbf{AC^0}$ does not suffice.
Conjecture: $\mathbf{NC^2}$ suffices.

Goal: For each complexity class, introduce a *cononical* minimal logical theory which captures reasoning using concepts in that class.

History:
C 1975 Introduced an equational theory $\mathbf{PV}$ (like Skolem's **PRA**) for reasoning about polynomial time. **PV** has function symbols for all polytime functions
$f : \mathbb{N}^k \to \mathbb{N}$ with equational axioms based on Cobham's Theorem.

Buss ca 1983 (**Independently of [C 75]**) thought up the theory $\mathbf{S}^1_2$ for polytime.
Theorem: [Buss 86] The Σ^b_1 theorems of **QPV** and $\mathbf{S}^1_2$ are essentially the same.

Question Can **PV** and $\mathbf{S}^1_2$ be modified to capture small complexity classes?

Difficulty: Their vocabulary contains $x \cdot y$, which in NOT in $\mathbf{AC^0}$ or $\mathbf{AC^0}(2)$.

Solution: Use two-sorted theories.

Two-Sorted Theories [Buss 86], [Zambella 96]

"number" variables $x, y, z...$ (range over $\mathbb{N}$)

"string" variables $X, Y, Z...$
(range over finite subsets of $\mathbb{N}$)

Language $\mathcal{L}_A^2 = [0, 1, +, \cdot, |\ |; \in, \leq, =]$

Standard model $\mathbb{N}_2 = \langle \mathbb{N}, \text{finite}(\mathbb{N}) \rangle$

$0, 1, +, \cdot, \leq, =$ usual meaning over $\mathbb{N}$

$$|X| = \begin{cases} 1 + \sup(X) & \text{if } X \neq \emptyset \\ 0 & \text{if } X = \emptyset \end{cases}$$

$x \in X$ (set membership)

number terms $s, t, u...$ defined as usual
only string terms are variables $X, Y, Z, ...$

Notation: $X(t) \equiv t \in X$, t a term

Definitions: Σ_0^B formula: All number quantifiers bounded.
No string quantifiers. (Free string variables allowed.)

Σ_1^B formula has the form

$$\exists Y_1 \leq t_1 ... \exists Y_k \leq t_k \ \phi$$

$k \geq 0$, ϕ is Σ_0^B.

$\exists X \leq t \ \phi$ stands for $\exists X(|X| \leq t \wedge \phi)$, where t does not involve X.

Σ_1^1 is the class of formulas

$$\exists \vec{Y} \phi \qquad \phi \in \Sigma_0^B$$

Σ_i^B formulas begin with at most i blocks of bounded string quantifiers $\exists\forall\exists...$ followed by a Σ_0^B formula.

Note: Σ_i^B corresponds to **strict** $\Sigma_i^{1,b}$.

Two-Sorted Complexity Classes

In general, number inputs $x, y, z...$ are presented in unary.
String inputs $X, Y, Z, ...$ are presented as bit strings.

Definition A relation $R(\vec{x}, \vec{X})$ is in $\mathbf{AC^0}$ iff some ATM (alternating Turing machine) accepts R in time $O(\log n)$ with a constant number of alternations.

Representation Theorem [BIS,I,Wrathall]
(a) The Σ_0^B formulas $\phi(\vec{x}, \vec{X})$ represent precisely the relations $R(\vec{x}, \vec{X})$ in $\mathbf{AC^0}$.

(b) The Σ_1^B formulas represent precisely the **NP** relations.

(c) The Σ_i^B formulas, $i \geq 1$, represent precisely the Σ_i^p relations.

Function Classes and Bit Graphs

Definition If **C** is a class of relations, then the function class **FC** contains

(a) All p-bounded number-valued functions $f(\vec{x}, \vec{X})$ s.t. its graph

$$G_f(y, \vec{x}, \vec{X}) \equiv (y = f(\vec{x}, \vec{X}))$$

is in **C**.

(b) All p-bounded string-valued functions $F(\vec{x}, \vec{X})$ such that its bit graph

$$B_F(i, \vec{x}, \vec{X}) \equiv F(\vec{x}, \vec{X})(i)$$

is in **C**.

p-bounded means for some polynomial $q(\vec{x}, \vec{X})$:
$f(\vec{x}, \vec{X}) \leq q(\vec{x}, |\vec{X}|)$
$|F(\vec{x}, \vec{X})| \leq q(\vec{x}, |\vec{X}|)$

All functions in $\mathbf{FAC^0}$ must have graphs (or bit graphs) representable by Σ_0^B formulas

Example: $Plus(X, Y) = X + Y$ (binary $+$)
$Plus \in \mathbf{FAC^0}$

$$Plus(X, Y)(i) \equiv X(i) \oplus Y(i) \oplus Carry(X, Y, i)$$

$$\begin{aligned} Carry(i, X, Y) \;\equiv\; & \exists j < i[X(j) \wedge Y(j) \wedge \\ & \forall k < i(j < k \supset (X(k) \vee Y(k))] \end{aligned}$$

NON-Examples:
$X \cdot Y$ (binary multiplication) NOT in $\mathbf{FAC^0}$.
$Parity(X) \equiv X$ has an odd number of ones.
$Parity \notin \mathbf{AC^0}$ (Ajtai, FSS)
$Parity(X)$ NOT representable by a Σ_0^B formula.

Hierarchy of Theories $\mathbf{V^0} \subset \mathbf{V^1} \subseteq \mathbf{V^2} \subseteq ...$
All have underlying vocabulary $\mathcal{L}_A^2$
For $i \geq 1$, $\mathbf{V}^i$ is "RSUV" isomorphic to $\mathbf{S}_2^i$.

2-BASIC Axioms for $\mathbf{V}^i, i \geq 0$ [Zam96]

B1. $x + 1 \neq 0$
B2. $x + 1 = y + 1 \supset x = y$
B3. $x + 0 = x$
B4. $x + (y + 1) = (x + y) + 1$
B5. $x \cdot 0 = 0$
B6. $x \cdot (y + 1) = (x \cdot y) + x$
B7. $(x \leq y \wedge y \leq x) \supset x = y$
B8. $x \leq x + y$
B9. $0 \leq x$
B10. $x \leq y \vee y \leq x$
B11. $x \leq y \leftrightarrow x < y + 1$
B12. $x \neq 0 \supset \exists y \leq x(y + 1 = x)$
L1. $X(y) \supset y < |X|$
L2. $y + 1 = |X| \supset X(y)$
SE. $[|X| = |Y| \wedge \forall i < |X|(X(i) \leftrightarrow Y(i))] \supset X = Y$

Also $\mathbf{V}^i$ needs $\mathbf{\Sigma^B_i}$-COMP (Comprehension)

$$\exists Z \leq y \forall j < y[Z(j) \leftrightarrow \phi(j, \vec{x}, \vec{X})]$$

where $\phi(j, \vec{x}, \vec{X})$ is a $\mathbf{\Sigma^B_i}$ formula without Z.

Theorem $\mathbf{V^0}$ proves
(because $|X| = 1+$ largest element of X...)

X-MIN

$$0 < |X| \supset \exists x < |X|(X(x) \wedge \forall y < x\ \neg X(y))$$

and X-IND

$$[X(0) \wedge \forall y < z(X(y) \supset X(y+1))] \supset X(z)$$

Therefore for $i = 0, 1, 2, ...$
$\mathbf{V}^i$ proves (using Σ_i^B-COMP)

Σ_0^B-IND: $\quad [\phi(0) \wedge \forall x(\phi(x) \supset \phi(x+1))] \supset \forall z \phi(z)$

and

Σ_0^B-MIN: $\quad \exists x \phi(x) \supset \exists x[\phi(x) \wedge \neg \exists y(y < x \wedge \phi(y))]$

where $\phi(x)$ is any Σ_i^B-formula (with parameters).

Fact: $\mathbf{V^0}$ is a conservative extension of $\mathbf{I\Delta_0}$.
Thus $\mathbf{V^0}$ proves all the usual properties of $x + y, x \cdot y, |x|, Bit(i, x)$.

Fact: $\mathbf{V}^i$ is finitely axiomatizable ($i \geq 0$).

Pairing Function: $\langle x, y\rangle$ is a **term** of $\mathcal{L}_A^2$.

$$\langle x, y\rangle =_{def} (x + y)(x + y + 1) + 2y$$

$\mathbf{V^0}$ proves $(x, y) \mapsto \langle x, y\rangle$ is one-one $\mathbb{N} \times \mathbb{N} \to \mathbb{N}$.

A **two-dimensional array** is represented by a string X. Define

$$X(i, j) = X(\langle i, j\rangle)$$

Then $X^{[i]}$ is row i of the array X. We bit-define the string function $X^{[i]}$ by

$$X^{[i]}(j) \leftrightarrow j < |X| \wedge X(i, j)$$

Example: $PHP(n, X)$ (Pigeonhole Principle)
This is a Σ_0^B formula.
Think $X(i, j)$ means pigeon $i \longrightarrow$ hole j.

$$\forall i \leq n \exists j < n X(i, j) \supset$$
$$\exists i \leq n \exists j \leq n \exists k < n(i < j \wedge X(i, k) \wedge X(j, k))$$

Theorem: (Ajtai) $\mathbf{V^0}$ does not prove $PHP(n, X)$.

$\mathbf{V^0}$ is like $\mathbf{I\Delta_0}(R)$

$\overline{\mathbf{V^0}}$: A universal conservative extension of $\mathbf{V^0}$
Vocabulary $\mathcal{L}_{\mathbf{FAC^0}}$ (In the spirit of **PV**.)

Axioms of $\mathbf{V^0}$ include **2-BASIC** except replace predecessor axiom **B12** by function $pd(x)$
$\mathbf{B12'}$ $pd(0) = 0$
$\mathbf{B12''}$ $x \neq 0 \supset pd(x) + 1 = x$

$\mathcal{L}_{\mathbf{FAC^0}}$ includes $\mathcal{L}_A^2 \cup \{pd\}$ and

$$\begin{aligned} f_{\alpha,t}(\vec{x}, \vec{X}) &= \min z < t\, \alpha(z, \vec{x}, \vec{X}) \\ &= t \text{ if } \neg\exists z < t \alpha(z, \vec{x}, \vec{X}) \end{aligned}$$

for every open formula α of $\mathcal{L}_{\mathbf{FAC^0}}$ and every term t of $\mathcal{L}_A^2$.

Open defining axioms for $f_{\alpha,t}$:

$$f_{\alpha,t}(\vec{x}, \vec{X}) \leq t(\vec{x}, \vec{X})$$
$$f_{\alpha,t}(\vec{x}, \vec{X}) < t \supset \alpha(f_{\alpha,t}(\vec{x}, \vec{X}), \vec{x}, \vec{X}))$$
$$z < f_{\alpha,t}(\vec{x}, \vec{X}) \supset \neg\alpha(z, \vec{x}, \vec{X})$$

$\mathcal{L}_{\mathbf{FAC^0}}$ also includes string funciton $F_{\alpha,t}$ with defining axiom

$$F_{\alpha,t}(\vec{x}, \vec{X})(z) \leftrightarrow z < t \wedge \alpha(z, \vec{x}, \vec{X})$$

Definition: $\overline{\mathbf{V^0}}$ has vocabulary $\mathcal{L}_{\mathbf{FAC^0}}$ and the **2-BASIC'** axioms of $\mathbf{V^0}$ together with the defining axioms for all new function symbols.

Lemma:

$$\overline{\mathbf{V^0}} \vdash \exists z < t\, \alpha(z, \vec{x}, \vec{X}) \;\leftrightarrow\; f_{\alpha,t}(\vec{x}, \vec{X}) < t$$

Lemma: $\overline{\mathbf{V^0}}$ proves equivalence of Σ_0^B-formulas of $\mathcal{L}_A^2$ and open formulas of $\mathcal{L}_{\mathbf{FAC^0}}$.

Theorem: $\overline{\mathbf{V^0}}$ is a universal theory and a conservative extension of $\mathbf{V^0}$. Its language $\mathcal{L}_{\mathbf{FAC^0}}$ has function symbols for the functions in $\mathbf{FAC^0}$.

Proof: Note that $\overline{\mathbf{V^0}}$ proves Σ_0^B-COMP
$\exists Z \forall i < y[Z(i) \leftrightarrow \phi(i, \vec{x}, \vec{X})]$
(Just bit-define $F(\vec{x}, \vec{X})(i) \leftrightarrow \phi(i, \vec{x}, \vec{X})$)

Conservativity: Every function in $\mathcal{L}_{\mathbf{FAC^0}}$ is Σ_0^B-definable in $\mathbf{V^0}$.

Witnessing (Finding Skolem functions)

Definition: Functions $\vec{F}$ *witness* $\exists \vec{Y} \phi$ in T if

$$T \vdash \phi(\vec{x}, \vec{X}, \vec{F}(\vec{x}, \vec{X}))$$

Theorem: (**Witnessing**) Suppose T is a universal theory which extends $\mathbf{V^0}$, and is defined over a language $\mathcal{L}$ and suppose that for every open formula $\alpha(i, \vec{x}, \vec{X})$ and term $t(\vec{x}, \vec{X})$ over $\mathcal{L}$ there is a function symbol F in $\mathcal{L}$ such that

$$T \vdash F(\vec{x}, \vec{X})(i) \leftrightarrow i < t \wedge \alpha(i, \vec{x}, \vec{X})$$

Then every theorem of T of the form $\exists \vec{Y} \alpha(\vec{x}, \vec{X}, \vec{Y})$, where α is open, is witnessed in T by functions in $\mathcal{L}$.

Proof: This is an easy consequence of the Herbrand Theorem.

Corollary: Every Σ^1_1 theorem of $\overline{\mathbf{V^0}}$ (and $\mathbf{V^0}$) is witnessed in $\overline{\mathbf{V^0}}$ by functions in $\mathcal{L}_{\mathbf{FAC^0}}$.

Definability of functions in theories

Definition: A string function $F(\vec{x}, \vec{X})$ is Φ-*definable* in T if it satisfies

$$Y = F(\vec{x}, \vec{X}) \leftrightarrow \phi(\vec{x}, \vec{X}, Y) \tag{1}$$

for some formula ϕ in Φ, and

$$T \vdash \forall \vec{x} \forall \vec{X} \exists ! Y \phi(\vec{x}, \vec{X}, Y) \tag{2}$$

Lemma: (Conservativity) If T is an extension of $\mathbf{V^0}$ which satisfies (2) and F is not in the language of T and T' is the result of adding F to the language and adding (1) as an axiom, then F' is a conservative extension of T.

Definition: A theory T is *p-bounded* if its axioms have all quantifiers bounded by terms of $\mathcal{L}_A^2$.

Lemma: If T is a p-bounded theory, then a function (string or number) is Σ_1^1-definable in T iff it is Σ_1^B-definable in T.

Proof: Follows from Parikh's Theorem.

Summary for $\mathbf{V^0}$ (finitely axiomatizable [CK03])

Theorem: $\overline{\mathbf{V^0}}$ is a universal theory and a conservative extension of $\mathbf{V^0}$. Its language $\mathcal{L}_{\mathbf{FAC^0}}$ has function symbols for the functions in $\mathbf{FAC^0}$.

Theorem: A function is Σ^1_1-definable in $\mathbf{V^0}$ iff it is in $\mathbf{FAC^0}$. (The same for $\overline{\mathbf{V^0}}$.)

Recall Goal: Introduce a minimal cononical theory **VC** for each complexity class **C**.

- **VC** has vocabulary $\mathcal{L}^2_A$.
- $\mathbf{VC} = \mathbf{V^0} + \{\text{one axiom}\}$ (finitely axiomatizable) [Nguyen: see Chapter 9]
- The Σ^1_1-definable functions in **VC** are those in **FC**. (NOTE: ALWAYS Σ^1_1-definable)
- **VC** has a universal convervative extension $\overline{\mathbf{VC}}$ in the style of **PV**.

class	$\mathbf{AC^0}$	$\subset$	$\mathbf{AC^0}(2)$	$\subset$	$\mathbf{TC^0}$	$\subseteq$	$\mathbf{NC^1}$
theory	$\mathbf{V^0}$	$\subset$	$\mathbf{V^0}(2)$	$\subset$	$\mathbf{VTC^0}$	$\subseteq$	$\mathbf{VNC^1}$

class	$\mathbf{L}$	$\subseteq$	$\mathbf{NL}$	$\subseteq$	$\mathbf{NC}$	$\subseteq$	$\mathbf{P}$
theory	$\mathbf{VL}$	$\subseteq$	$\mathbf{VNL}$	$\subseteq$	$\mathbf{VNC}$	$\subseteq$	$\mathbf{VP} = \mathbf{TV^0}$

Theories for polynomial time

In general, for complexity class $\mathbf{C}$,

$$\mathbf{VC} = \mathbf{V^0} + \mathbf{Axiom_C}$$

where $\mathbf{Axiom_C}$ is a Σ^B_1-formula

$$\mathbf{Axiom_C} = (\mathbf{Complete_C} \text{ has a solution})$$

$\mathbf{Complete_P} = MVCP$ (Monotone Circuit Value Problem) (Complete wrt $\mathbf{AC^0}$-reductions)

$$MVPC \equiv \exists Y\, \delta_{MCVP}(g, d, E, G, I, Y)$$

where $\delta_{MCVP}(g, d, E, G, I, Y)$ is

$$\forall x < g \forall z < d,\ (Y(0, x) \leftrightarrow I(x)) \wedge$$
$$[Y(z{+}1, x) \leftrightarrow (G(z{+}1, x) \wedge \forall u < g,\ E(z, u, x) \supset Y(z, u))$$
$$\vee\ (\neg G(z + 1, x) \wedge \exists u < g,\ E(z, u, x) \wedge Y(z, u))]$$

Robustness Theorem:
$\mathbf{VP} = \mathbf{TV}^0 = V^1$-HORN $= \mathbf{V^0} + \Sigma^B_0$-BIT-REC
$\mathbf{V}^1$ is Σ^B_1-conservative over **VP**

$\mathbf{TV}^0 = \mathbf{V^0} +$ String IND on Σ^B_0 formulas
$\mathbf{TV}^0 \simeq \mathbf{T}^0_2(MSP)$ [Jeřábec]
$\mathbf{V}^1 \simeq \mathbf{S}^1_2$

Define $\overline{\mathbf{VP}} = \mathbf{VPV}$ (two-sorted **QPV**)
$\overline{\mathbf{VP}}$ is a conservative extension of **VP**.

Theories VC for other classes C

Recall $\mathbf{VC} = \mathbf{V^0} + \mathbf{Axiom_C}$
where $\mathbf{Axiom_C} = (\mathbf{Complete_C}$ has a solution)

class	theory	$\mathbf{Complete_C}$
$\mathbf{AC^0}$	$\mathbf{V^0}$	none
$\mathbf{AC^0}(2)$	$\mathbf{V^0}(2)$	$Parity(X)$
$\mathbf{TC^0}$	$\mathbf{VTC^0}$	$numones(X)$
$\mathbf{NC^1}$	$\mathbf{VNC^1}$	tree-$MCVP$
$\mathbf{L}$	$\mathbf{VL}$	$UniConn(z, a, E)$
$\mathbf{NL}$	$\mathbf{VNL}$	$Conn(z, a, E)$
$\mathbf{P}$	$\mathbf{VP}$	$MCVP$

Robustness Theorems

$\mathbf{VTC^0} \simeq \Delta_1^B$-**CR** [JP] (proved in [Nguyen])
$\mathbf{VNC^1} \simeq \mathbf{AID}$ [Arai] (proved in [CM])
$\mathbf{VNC^1} \simeq ALV \simeq ALV'$ [Clote] (proved by [Nguyen])
$\mathbf{VL} = \Sigma_0^B$-Rec [Zam97]
$\mathbf{VNL} = \mathbf{V}$-$\mathbf{Krom}$ [Kolokolova]

Power of theories

$\mathbf{V^0} \subset \mathbf{V^0}(2) \subset \mathbf{VTC^0} \subseteq \mathbf{VNC^1} \subseteq \mathbf{VL} \subseteq \mathbf{VNL} \subseteq \mathbf{VP}$
Interesting cases: $\mathbf{VTC^0} \subseteq \mathbf{VNC^1} \subseteq \mathbf{VL}$

$\mathbf{VNL} \vdash \mathbf{NL} = co\mathbf{NL}$ [Kol] (Imm-Sze)
$\mathbf{V^0}$ proves Discrete Jordan Curve Thm (seq of edges) [Ng]
$\mathbf{V^0}(2)$ proves Discrete Jordan Curve Thm (set of edges) [Ng]

Limitations on theories

$\mathbf{V^0} \nvdash$ PHP
$\mathbf{VP} \nvdash$ Primes in P [AKS] (unless factoring in P)
$\mathbf{VL} \vdash \mathbf{SL} = \mathbf{L}$ [R] ???
$\mathbf{VP} \vdash \mathbf{SL} = \mathbf{L}$ [R] ???
V-SymKrom $\vdash$ **SL** $=$ **coSL** ??? [Kol]
LAP $\vdash$ properties of determinants ??? [Soltys]

Advantages of Two-Sorted Theories

• Can capture small complexity classes: $\mathbf{AC^0}, \mathbf{AC^0}(m)$

• More natural for *all* classes:
e.g. community missed the fact that $\mathbf{T}^0_2(MSP)$ captures polytime, until $\mathbf{TV}^0$ was defined.
Also $\mathbf{VTC}^0$ is nicer than Δ^B_1-**CR**.

• Propositional Translations are simpler: They are Paris-Wilkie style instead of PV style.

• Bounded Theorems of two-sorted theories translate naturally into polysize families of quantified propositional theorems.

Benefit of Formalizing Proofs

Razborov came up with a much better Switching Lemma proof by formalizing it in $\mathbf{V}^1$.

Open Questions

What complexity classes are needed to prove $\mathbf{SL} = \mathbf{L}$?
Properties of determinants?
Correctness of randomized algorithms? [J]

What about other classes such as **LogCFL** and search problem classes (maximum matching).