How to Associate Propositional Proof Systems and Theories with Complexity Classes

Part I:

Second Order Bounded Theories

Stephen Cook

LC2002 Tutorial
August, 2002

Lectures based on author's course notes
"Proof Complexity and Bounded Arithmetic"

`www.cs.toronto.edu/~sacook/csc2429h/`

General Reference:
Jan Krajíček : Bounded Arithmetic, Propositional Logic, and Complexity Theory. Cambridge, 1995.

Examples

Complexity Classes:

$$\mathbf{AC^0} \subset \mathbf{NC^1} \subseteq \mathbf{L} \subseteq \mathbf{P} \subseteq \mathbf{PLS} \subseteq \mathbf{PSPACE}$$

First Order Theories:

$$\mathbf{AID} \subset \mathbf{S^1_2} \subseteq \mathbf{T^1_2}$$

Second Order Theories

$$\mathbf{V^0} \subset \mathbf{V^1} \subset \mathbf{U^1_2}$$

Proof Systems:

$$\mathbf{AC^0}\text{Frege} \leq_p \text{Frege} \leq_p \text{EFrege} \leq \mathbf{G_1} \leq \mathbf{G}$$

Example:

Complexity Class: **P** (Polynomial time)
Theory: $\mathbf{S_2^1}$ (Buss's first order theory)
Proof System: EFrege (Extended Frege)

- The functions definable in $\mathbf{S_2^1}$ are precisely the polynomial time functions

- $\forall$-theorems of $\mathbf{S_2^1}$ translate into tautology families with polysize EFrege proofs.

- $\mathbf{S_2^1}$ proves the consistency of EFrege. If $\mathbf{S_2^1}$ proves the consistency of system S, then EFrege p-simulates S.

- A line in an EFrege proof represents a Boolean circuit. A problem in nonuniform **P** is specified by a polysize family of Boolean circuits.

- EFrege is a nonuniform version of the $\forall$-fragment of $\mathbf{S_2^1}$.

Second Order Theories

We will use the "second order" theory $\mathbf{V}^1$ for polytime reasoning.
$\mathbf{V}^1$ is equivalent to Buss's first order theory $\mathbf{S}^1_2$.

Second order theories have two sorts:
• natural numbers • binary strings.
(a "string" is really a finite set of numbers)

A function on strings is definable in $\mathbf{V}^1$ iff it is polytime computable.

History

• Buss 1986 introduced second order theories $\mathbf{V}^i_j$ and $\mathbf{U}^i_j$ for PSPACE, EXPTIME,...

• Razborov 1993 pointed out usefulness of $\mathbf{V}^1_1$ for polytime reasoning. (Raz, Tak: $\mathbf{V}^1_1 \sim \mathbf{S}^1_2$)

• Zambella (JSL, 1996) introduce clean syntax.

• Our $\mathbf{V}^0$ and $\mathbf{V}^1$ based on Zambella's presentation

Descriptive Complexity

Motivation for second order theories

Reference: Neil Immerman, *Descriptive Complexity*

Consider the first order language

$$\mathcal{L} = [X; PLUS, TIMES, =, 0, max]$$

Given bit string $X \in \{0,1\}^n, n \geq 1$
$\mathcal{M}[X]$ is a finite $\mathcal{L}$-structure with
universe $M = \{0, 1, ..., n-1\}$
$X^{\mathcal{M}}(i) \leftrightarrow X_i = 1$
$max = n-1, PLUS(i,j,k) \leftrightarrow i + j = k$
$TIMES(i,j,k) \leftrightarrow i \times j = k$

Let A be an $\mathcal{L}$-sentence.

$$L(A) = \{X \in \{0,1\}^+ \mid \mathcal{M}[X] \models A\}$$

$$\mathbf{FO} = \{L(A) \mid \mathcal{L} \text{ sentence } A\}$$

FO Characterization Theorem: [BIS, I]

$$\mathrm{FO} = \mathbf{AC^0} = \mathbf{LH}$$

$\mathbf{AC^0}$: uniform bounded depth $\wedge, \vee$ polysize circuit family.

LH: Log time hierarchy for alternating Turing machines.

Example PAL = set of binary palindromes. $PAL = L(A)$, where A is

$$\forall x \forall y (x + y = max \supset (X(x) \leftrightarrow X(y))$$

Theorem: [Ajtai,FSS 83] $PARITY \notin \mathbf{AC^0}$

($PARITY$ = set of strings with an odd number of 1's.)

Thus $\langle U, V, U \cdot V \rangle \notin \mathbf{AC^0}$.

But $\langle U, V, U + V \rangle \in \mathbf{AC^0}$

Second Order Theories $\mathbf{V}^0, \mathbf{V}^1$

(Two-sorted first order theories)

"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

$X = Y$ means $(|X| = |Y| \wedge \forall z < |X|(X(z) \leftrightarrow Y(z)))$

Sets to strings:

If $\emptyset \neq X \subset \mathbb{N}$, $|X| = \sup(X) + 1 = n \geq 1$

$$w(X) = X(0)X(1)...X(n-2)$$

Note $n - 1 \in X$. w is a bijection from finite nonempty subsets of $\mathbb{N}$ onto $\{0,1\}^*$.

Definitions: $\mathbf{\Sigma_0^B}$ formula: All number quantifiers bounded.
No string quantifiers. (Free string variables allowed.)

$\mathbf{\Sigma_1^B}$ formula has the form

$$\exists Y_1 \leq t_1 ... \exists Y_k \leq t_k \; B$$

$k \geq 0$, B is $\mathbf{\Sigma_0^B}$.

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

Every formula $A(X)$ represents a language $L \subseteq \{0,1\}^*$ over the standard model $\mathbb{N}_2$

Theorem: Let L be a subset of $\{0,1\}^*$.

a) L is in $\mathbf{AC^0}$ iff some $\mathbf{\Sigma_0^B}$ formula $A(X)$ represents L.

b) L is in **NP** iff some $\mathbf{\Sigma_1^B}$ formula $A(X)$ represents L.

Proof: Note that if $A(X)$ is $\mathbf{\Sigma_0^B}$, then all number quantifiers are bounded by $|X|^d$ for fixed d depending on A.

a) From **FO** Characterization Theorem.

b) From definition of **NP**.

Axioms for $\mathbf{V^0}$ and $\mathbf{V^1}$ [Zam96]

The axioms are the universal closures of B1,...,B14, L1,L2, and Comprehension.

B1...B13 are first order quantifier-free formulas about $+, \cdot, \leq$ (essentially Robinson's $\mathbf{Q}$).

B14: $x \neq 0 \supset \exists y \leq x(y + 1 = x)$

L1: $X(y) \supset y < |X|$
L2: $y + 1 = |X| \supset X(y)$

$\mathbf{\Sigma_i^B}$-Comprehension (for $i = 0, 1$).

$$\exists X \leq y \forall z < y[X(z) \leftrightarrow A(z, \vec{w}, \vec{Y})]$$

where $A(z)$ is a $\mathbf{\Sigma_i^B}$ formula not containing X.

$\mathbf{V^0}$ has $\mathbf{\Sigma_0^B}$-Comprehension
$\mathbf{V^1}$ has $\mathbf{\Sigma_1^B}$-Comprehension

Theorem: $\mathbf{V^i}$ proves the $\mathbf{\Sigma_i^B}$-IND scheme

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

Proof: Axioms L1 and L2 show each nonempty X has a largest element. This and $\mathbf{\Sigma_i^B}$-COMP show the LNP for $\mathbf{\Sigma_i^B}$ formulas. $\Box$

Recall: $\mathbf{I\Delta_0}$ is **PA** with induction restricted to bounded formulas.

Theorem: $\mathbf{V^0}$ is a conservative extension of $\mathbf{I\Delta_0}$.

Proof: Conservativity follows because any model M for $\mathbf{I\Delta_0}$ can be extended to a model (M, M_2) for $\mathbf{V^0}$:

Let $M_2 =$ set of all bounded $\mathbf{\Sigma_0^B}$-definable subsets S of M.

$|S| = \sup(S)+1$ exists in M because $\mathbf{I\Delta_0}$ proves bounded induction.

FAC^0 and **FP** are the function versions of **AC^0** and **FP**.

FAC^0 and **FP** each have both number (valued) and string (valued) functions.

Think of numerical inputs and outputs expressed in *unary* notation.

• Number functions in **FAC^0** (resp **FP**) have the form

$$f(\vec{x}, \vec{Y}) = \min z < p(\vec{x}, |\vec{Y}|)\ R(z, \vec{x}, \vec{Y})$$

where p is a polynomial and R is an **AC^0** (resp **P**) relation.

• String functions in **FAC^0** (resp **FP**) satisfy

$$F(\vec{x}, \vec{Y})(i) \leftrightarrow i < p(\vec{x}, |\vec{Y}|) \wedge B_F(i, \vec{x}, \vec{Y})$$

where B_F is an **AC^0** (resp **P**) relation.

$B_F(i, \vec{x}, \vec{Y})$ is the *bit graph* of F.

Definition: A $\mathbf{\Sigma_1^1}$ formula is one of the form

$$\exists Y_1 ... \exists Y_k \ B$$

where $k \geq 0$ and B is $\mathbf{\Sigma_0^B}$.

A string function $F(\vec{x}, \vec{Y})$ is $\mathbf{\Sigma_1^1}$ *definable* in **T** iff there is a $\mathbf{\Sigma_1^1}$ formula A such that
(a) F has defining axiom

$$Z = F(\vec{x}, \vec{Y}) \leftrightarrow A(\vec{x}, \vec{Y}, Z)$$

(b) **T** $\vdash \forall \vec{x} \forall \vec{Y} \exists ! Z \ A(\vec{x}, \vec{Y}, Z)$

Similarly for number functions.

Main Theorem: A function (string or number) is $\mathbf{\Sigma_1^1}$ definable in $\mathbf{V^0}$ (resp $\mathbf{V^1}$) iff it is in $\mathbf{FAC^0}$ (resp **FP**).

Example in $\mathbf{FAC^0}$: Binary Addition:

$$PLUS_{bin}(X, Y) = X +_{bin} Y$$

The bit graph B_{PLUS} satisfies

$$B_{PLUS}(i, X, Y) \equiv X(i) \oplus Y(i) \oplus Carry(X, Y, i)$$

where $Carry(X, Y, i) \equiv$
$\exists j < i[X(j) \wedge Y(j) \wedge \forall k < i(j < k \supset (X(k) \vee Y(k))]$

Proof of Main Theorem $\Leftarrow$

Straightforward for $\mathbf{V^0}$

$$\begin{aligned} A(\vec{x}, \vec{Y}, Z) &\equiv Z = F(\vec{x}, \vec{Y}) \\ &\equiv \forall i < t(Z(i) \leftrightarrow B_F(i, \vec{x}, \vec{Y})) \end{aligned}$$

$\exists Z$ proved by $\mathbf{\Sigma_0^B}$-COMP.

$\Leftarrow$ for $\mathbf{V^1}$ case: Suppose Turing machine M computes $F(\vec{x}, \vec{Y})$ within time $p(\vec{x}, |\vec{Y}|)$. Define

$A(\vec{x}, \vec{Y}, Z) \equiv \quad \exists W < t(\vec{x}, |\vec{Y}|)$
$[W$ codes the computation of M on input $\vec{x}, \vec{Y}]$.

$[W, \vec{x}, \vec{Y}]$ expressed by a $\mathbf{\Sigma_0^B}$ formula.

$\exists Z$ proved by $\mathbf{\Sigma_1^B}$-COMP.

$\mathbf{V^0}, \mathbf{V^1}$ Witnessing Theorems [Buss86]
If B is a $\mathbf{\Sigma_0^B}$ formula and $\mathbf{V^0} \vdash \exists Z\ B(\vec{x}, \vec{Y}, Z)$ then there is a function $F \in \mathbf{FAC^0}$ such that

$$AX(F), \mathbf{V^0} \vdash B(\vec{x}, \vec{Y}, F(\vec{x}, \vec{Y}))$$

Similarly for $\mathbf{V^1}$, except F is in **FP**.

Proof: Based on free cut free (anchored) proofs. **(Tomorrow)**

Corollary 1: $\Rightarrow$ direction of Main Theorem.

Corollary 2: Parikh's Theorem for $\mathbf{V^0}, \mathbf{V^1}$:
If A is $\mathbf{\Sigma_0^B}$ and

$$\mathbf{V}^i \vdash \exists Z\ A(\vec{x}, \vec{Y}, Z)$$

then for some term $t(\vec{x}, \vec{Y})$

$$\mathbf{V}^i \vdash \exists Z \le t(\vec{x}, \vec{Y})\ A(\vec{x}, \vec{Y}, Z)$$

Corollary 3: Suppose A, B are $\mathbf{\Sigma_0^B}$ formulas, and

$$R(X) \equiv \exists Z < t \ A(X,Z)$$
$$S(X) \equiv \neg \exists Z < t \ B(X,Z)$$

If $\mathbf{V^1} \vdash R(X) \leftrightarrow S(X)$ then R, S represent poly-time relations.

i.e. if $\mathbf{V^1}$ proves $R \in \mathbf{NP} \cap \mathbf{coNP}$ then $R \in \mathbf{P}$.

Proof: $\mathbf{V^1} \vdash S \supset R$, so

$$\mathbf{V^1} \vdash \exists Z < t(B(X,Z) \vee A(X,Z))$$

Apply Witnessing Theorem.

(This holds even without bounding term t.) $\square$

Example: $\{Primes\} \in \mathbf{NP} \cap \mathbf{coNP}$.

RSUV Isomorphism [Razb, Tak]

$\mathbf{V}^1$ isomorphic to $\mathbf{S}_2^1$.

Buss's first order system of bounded arithmetic:

$$\mathcal{L} = [0, S, +, \cdot, \#, |x|, \lfloor \frac{1}{2} x \rfloor; =, \leq]$$

$$|x| = \log_2(x+1), \qquad x \# y = 2^{|x| \cdot |y|}$$

PIND axiom scheme:

$$[A(0) \wedge \forall x (A(\frac{1}{2} x) \supset A(x))] \supset \forall x A(x)$$

where A is Σ_1^b.

Models of $\mathbf{V}^1$ equivalent to models of $\mathbf{S}_2^1$

A model M of $\mathbf{S}_2^1$ converts to a model $\langle M_1, M_2 \rangle$ of $\mathbf{V}^1$, where
$M_1 = \{|m| : m \in M\}$
$M_2 = \{$binary notations for elements of $M\}$

A model $\langle M_1, M_2 \rangle$ of $\mathbf{V}^1$ converts to a model M of $\mathbf{S}_2^1$, where
M is the set of elements whose binary notations are strings in M_2.

$\mathbf{V}^1$ proves theorems for which polytime concepts suffice.

Graph Theory: Use strings to represent the adjacent matrix of graphs. $\mathbf{V}^1$ proves Halls theorem, Menger's Theorem, etc., but NOT Ramsey's Theorem (by $\mathbf{V}^1$ Witnessing Theorem).

Linear Algebra over nice fields.

Number Theory: (Binary Notation) $\mathbf{V}^1$ proves GCD's exist, and FLT for n=3,4:

$$x, y > 0 \supset x^3 + y^3 \neq z^3$$

FLT in general??
Complexity assumptions imply cannot prove Fermat Little Theorem, or prime factorization.

$$0 < x < y \wedge (x^{y-1} \bmod y) \neq 1 \supset \exists d (1 < d < y \wedge (d|y))$$

Boolean circuit complexity: Razborov argues most lower bounds can be formalized, provided Boolean functions represented by their graphs. However, probably not Shannon's lower bound.

Tomorrow

Prove “cut elimination” for Gentzen systems with induction, using semantic methods.

Use this to prove witnessing theorems.