III. Bounded Arithmetic, Paris-Wilkie Translations, and Witnessing in P and PLS

Sam Buss, UCSD
sbuss@math.ucsd.edu

Prague, September 2009

Constant depth propositional LK proofs

Syntax: Tait-style calculus. Variables: p. Literals: p, $\overline{p}$.

Unbounded fanin OR's and AND's: $\bigvee$ and $\bigwedge$.

Cedent Γ is set of formulas; intended meaning is disjunction, $\bigvee \Gamma$.

Axioms: *Neg:* $p, \overline{p}$ *Taut:* Γ , where Γ is a tautology.

Rules of inference:

$\bigvee$: $\dfrac{\Gamma, \varphi_{i_0}}{\Gamma, \bigvee_{i \in \mathcal{I}} \varphi_i}$, where $i_0 \in \mathcal{I}$. $\bigwedge$: $\dfrac{\Gamma, \varphi_i \quad \text{for all } i \in \mathcal{I}}{\Gamma, \bigwedge_{i \in \mathcal{I}} \varphi_i}$

Weakening: $\dfrac{\Gamma}{\Gamma, \Delta}$ *Cut:* $\dfrac{\Gamma, \varphi \qquad \Gamma, \overline{\varphi}}{\Gamma}$

In the Cut, we can assume w.l.o.g. that outermost connective of ϕ is not an $\bigwedge$.

Depth and Σ'-depth of LK formulas and proofs

The *depth* of a formula is the maximum nesting depth of blocks of $\land$'s and $\lor$'s. Literals have depth 0.

For the Paris-Wilkie translation from bounded arithmetic formulas to propositional logic, a better notion is Σ'-depth which allows small fanin at the bottom for free:

Definition

Let S be a proof size parameter (size upper bound). The formulas that have Σ'-depth d *with respect to* S are inductively defined as follows:

a. If φ has size $\le \log S$, then φ has Σ'-depth 0.
b. If φ has Σ'-depth d, then it has Σ'-depth d' for all $d' > d$.
c. If each φ_i has Σ'-depth d, then $\bigvee_{i\in\mathcal{I}} \varphi_i$ and $\bigwedge_{i\in\mathcal{I}} \varphi_i$ have Σ'-depth $(d+1)$.

Σ'-depth d is often called "*depth* $d+\frac{1}{2}$".

Definition

Let S be a size parameter. An LK-proof P is a Σ'-depth d proof of size S provided:

a. P has $\le S$ symbols,
b. Every formula in P has Σ'-depth d,
c. Every *Taut* axiom has size at most $\log S$. That is, only small tautologies are allowed.

Σ'-depth d proofs are particularly useful for translating $s\Sigma^b_d$ and $s\Pi^b_d$ formulas to propositional logic. The inner, sharply bounded quantifiers correspond to the bottom level of small fanin gates.

Definitions similar to Σ' depth given by: [K'94] of Σ-depth; [BB'03] of Θ-depth.

Sharply Strict Bounded Arithemetic

A formula is *form restricted* Σ_i^b, or *sharply strict* Σ_i^b, denoted $ss\Sigma_i^b$ if it is of the form

$$(\exists y_1 \le t_1)(\forall y_2 \le t_2)\cdots(Qy_i \le t_i)(\overline{Q}z \le |r|)B,$$

where B is quantifier-free.

Every Σ_i^b-formula is equivalent to a sharply strict one: this fact can be proved in S_2^i using induction on only $ss\Sigma_i^b$-formulas (with $\dot{-}$ and MSP in the base language).

Therefore, by free-cut elimination, bounded arithmetic may be equivalently formulated with induction only for $ss\Sigma_i^b$-formulas.

These notions are similar to Takueti's "pure i-form", and, later, "strictly i-normal proof".

Def'n: Let P be a proof. The free variables in the endsequent, $\vec{c}$, are called *parameter variables*.
A quantifier $(Qx \leq t)$ is *restricted by parameter variables* iff t uses only parameter variables.

A proof is *restricted by parameter variables* iff (a) every quantifier is restricted by parameter variables and (b) every sequent which contains a non-parameter b contains a formula $b \leq t(\vec{c})$ in its antecedent.

Theorem

Let R be S^i_2 or T^i_2, $i \geq 1$. If $\mathfrak{S} := \Gamma \longrightarrow \Delta$ contains only $ss\Sigma^b_i$-formulas and $R \vdash \mathfrak{S}$, then it has an R-proof which is restricted by parameter variables and in which every formula is $ss\Sigma^b_i$.

Such proofs are called *restricted-Σ^b_i*. These proofs are conveniently formed for translation into propositional logic.

From S_2^i, T_2^i to LK: First Paris-Wilkie Translation

Let $d \geq 1$ and R be one of S_2^d or T_2^d. Suppose $A(x)$ is $ss\Sigma_d^b$ and $R \vdash A$. We describe how to transform a restricted proof of A into a Σ'-depth d LK proof. W.l.o.g., x is the only parameter variable.

Fix $n \in \mathbb{N}$. The translation $[\![A]\!]_n$ is a propositional formula stating that $A(x)$ is true for all x such that $|x| \leq n$. The free variables of $[\![A]\!]_n$ are variables $p_{x,i}$ representing the i-th bit of the binary representation of x.

Base case of defn: For quantifier-free formulas φ, the formula $[\![\varphi]\!]$ is any polynomial size formula that expresses the value of ϕ. Since the function and relations are computable with polynomial size formulas, $[\![\phi]\!]$ has size $m^{O(1)}$ if the free variables of φ are integers of length $\leq m$. Because we have the *Taut* axioms, the choice of translation formula $[\![\phi]\!]$ is unimportant. (In any event, elementary properties of $[\![\phi]\!]$ should have polynomial size proofs.)

More generally, $[\![\phi]\!]$ respects Boolean connectives.

Quantifier case of defn. Consider $(\forall y \le |s|)B$ or $(\exists y \le |s|)B$. Because the term s contains only parameter variables as variables, and since the parameter variables have at most n bits, we can find a bound $n_y = n^{O(1)}$ such that $|s| \le n_y$. Then,

$$[\![(\forall y \le |s|)B]\!] = \bigwedge_{i=0}^{n_y} [\![y \le |s| \to B]\!]/(y \mapsto i).$$

The notation "$\psi/(y \mapsto i)$" means replace each $p_{y,j}$ by the (constant) jth bit of the integer i.

$[\![(\forall y \le |s|)B]\!]$ has size only $n^{O(1)}$. Thus, it has Σ'-depth 0 for suitable $S(n) = 2^{n^{O(1)}}$.

General bounded quantifiers translated by exactly the same construction, but have bigger size: $2^{n^{O(1)}}$.

A Σ^b_d-formula is translated to a Σ'-depth d formula with size parameter $S(n) = 2^{n^{O(1)}}$.

To translate a sequent $\mathfrak{S}$ in a restricted R-proof, view it as a Tait-style cedent by moving all formulas to right of the sequent (negated). All non-parameter variable $y_1, \ldots, y_k$ are restricted by parameter variables. So $|y_j| \leq n_j$ for some $n_j = n^{O(1)}$.

$\mathfrak{S}$ is translated into a set of cedents, one cedent for each choice of $i_1, \ldots, i_k$ with each $|i_j| < n_j$. The cedents are just

$$[\![\mathfrak{S}]\!]/(y_1 \mapsto i_1, \ldots, y_k \mapsto i_k),$$

where the translation is applied individually to each formula.

Note: the only variables left are $p_{x,i}$.

As the next theorem states, the translated cedents Γ can be pieced together into a valid proof.

Theorem

Let $i \geq 1$. Suppose $A(x) \in ss\Sigma_i^b$. Let $[\![A]\!]_n$ denote the propositional translation of A; $[\![A]\!]_n$ has free variables $p_{x,i}$, for $i < n$.

a. *Suppose $S_2^i \vdash A$. Then there is a function $S(n) = 2^{n^{O(1)}}$ such that, for all n, $[\![A]\!]_n$ has a Σ'-depth i proof of size $S(n)$. This proof*
   i. *has height $O(\log\log S(n))$, and*
   ii. *contains only $O(1)$ many formulas in each cedent.*

b. *Suppose $T_2^i \vdash A$. Then there is a function $S(n) = 2^{n^{O(1)}}$ such that, for all n, $[\![A]\!]_n$ has a Σ'-depth i proof of size $S(n)$. This proof*
   i. *has height $O(\log S(n))$, and*
   ii. *contains only $O(1)$ many formulas per cedent.*

Defn. The *height* of a proof is the maximum length of any branch in the proof.

The same theorem applies to $S_2^i(\alpha)$ and $T_2^i(\alpha)$ under the 2nd Paris-Wilkie translation, (defined later).

Case (1): translation of $\land$:right inference

An $\land$*:right* inference

$$\frac{\Gamma, \varphi \qquad \Gamma, \psi}{\Gamma, \varphi \land \psi}$$

translates to

$$\dfrac{[\![\Gamma]\!], [\![\phi]\!] \qquad \dfrac{[\![\Gamma]\!], [\![\psi]\!] \qquad \dfrac{[\![\psi \land \phi]\!], \overline{[\![\phi]\!]}, \overline{[\![\psi]\!]}}{[\![\Gamma]\!], [\![\psi \land \phi]\!], \overline{[\![\phi]\!]}, \overline{[\![\psi]\!]}}\ \textit{Weakening}}{[\![\Gamma]\!], [\![\psi \land \phi]\!], \overline{[\![\phi]\!]}}\ \textit{Cut}}{[\![\Gamma]\!], [\![\psi \land \phi]\!]}\ \textit{Cut}$$

Note that the upper right sequent is a *Taut* axiom.

Case 2: $\forall\le$:-right inference. The inference

$$\frac{c \le t(a), \Gamma \longrightarrow \Delta, B(c)}{\Gamma \longrightarrow \Delta, (\forall y \le t(a))B(y)}$$

translates into

$$\frac{\dfrac{[\![\neg c \le t(a)]\!]/(c \mapsto i), [\![\neg\Gamma]\!], [\![\Delta]\!], [\![A(c)]\!]/(c \mapsto i)}{[\![\neg\Gamma]\!], [\![\Delta]\!], [\![c \le t(a) \to A(c)]\!]/(c \mapsto i)}}{[\![\neg\Gamma]\!], [\![\Delta]\!], [\![(\forall y \le t(a))B(y)]\!]}$$

Here the top two lines are repeated for all values of $i \le t(a)$. That is, the last inference is a $\bigwedge$ inference, with many hypotheses.

Note the added height is constant (two), independent of n.

Case (3): Consider an induction inference in P. This translates into m *Cut* inferences in the LK proof, where m is the "length" of the induction. By balancing the tree of cuts, the added height is only $O(\log m)$.
The induction bound t involves only parameter variables, so m can be bounded in terms of parameter variables.)

If R is S^i_2, the induction inference translates into $m = |t| = n^{O(1)}$ many cuts, so the added height is $O(\log n)$.

If R is T^i_2, the induction inference translates into $m = t = 2^{n^{O(1)}}$ many cuts, so the added height is $O(n^{O(1)})$. □

Important fact: The LK-proofs given by Theorem 4 are polynomial time uniform. Given a path from the root of the proof, one can determine that part of the proof in polynomial time.

The Paris-Wilkie translation is more usually defined with a predicate α adjoined to the language. In this case, there are additional propositional variables q_i that encode the truth of $\alpha(i)$. In this setting, it is usual for there to be no free (parameter) variable x, so the variables $p_{x,i}$ are not used. To keep the framework above, we just assign $x = 2^n - 1$ so that $p_{x,i}$'s are all true.

Then $[\![\alpha(t)]\!]$ is q_i where i is the value of the closed term t.

It is also possible to combine the use of the x with α.
Then $[\![\alpha(t)]\!]$ can be expressed as both a large disjunction or a large conjunction.

Main Theorem for S^1_2

Theorem (B'85)

Suppose $A(x, y) \in \Sigma^b_1$ and that S^1_2 proves $(\forall x)(\exists y)A(x, y)$. Then there is a polynomial time function $f(x) = y$ such that for all $x \in \mathbb{N}$, $A(x, f(x))$ holds.

Proof. By Parikh, $S^1_2 \vdash (\exists y \leq s(x))A(x, y)$. x is the parameter variable. Applying Theorem (a) yields a Σ'-depth 1 proof; adding a *Cut* to the end of this proof turns the proof into a refutation R of

$$[\![(\forall y \leq s(x))\neg A(x, y)]\!]. \tag{1}$$

We give a polynomial time procedure that is has as input a particular value for x, and traverses the refutation R until it arrives at a false initial cedent. Of necessity, this false initial cedent is the cedent (1), and when it is reached, the procedure will know a value y that falsifies the cedent. This value y will be $f(x)$.

The polynomial time procedure acts as follows: it starts at the root of the proof and traverses the proof upward, backtracking as needed as described below. At each stage, the procedure is at some cedent Γ in the proof that it believes (or, hopes or assumes) to be false. In particular, every Σ'-depth 0 formula in Γ is *False*. (Recall that the variables $p_{x,i}$ are the only variables in R, and the procedure has values for these.) Furthermore, for any formula in Γ which is a conjunction of Σ'-depth 0 formulas, a particular conjunct is known to be false. For the formulas which are a disjunction of Σ'-depth 1 formulas, the procedure does not know for sure that they are false, it merely tentatively assumes they are false.

At the beginning, the procedure is at the endsequent of R, which is the empty cedent.

We next describe how the procedure handles *Cut*, $\bigwedge$, and $\bigvee$ inferences.

If the procedure is at the lower cedent of a cut inference

$$\frac{\Gamma, \varphi \qquad \Gamma, \overline{\varphi}}{\Gamma}$$

If φ is Σ'-depth 0, then it can be evaluated as being either *True* or *False*. If it is true, the procedure proceeds to the right upper cedent, otherwise, it proceeds to the left upper cedent. Otherwise, φ is w.l.o.g. a disjunction, and the algorithm proceeds to the left upper cedent.

If the procedure is at the lower cedent of a $\bigwedge$-inference:

$$\frac{\Gamma, \psi_i \qquad \text{, for } i \in \mathcal{I}}{\Gamma, \bigwedge_{i \in \mathcal{I}} \psi_i}$$

the algorithm acts as follows. By assumption, the procedure knows a value i_0 such that the conjunct ψ_{i_0} is false. The algorithm proceeds to the upper cedent Γ, ψ_{i_0} where $i = i_0$.

If the procedure is at the lower cedent of a $\bigvee$-inference:

$$\frac{\Gamma, \psi_{i_0}}{\Gamma, \bigvee_{i \in \mathcal{I}} \psi_i}$$

the algorithm acts as follows. If ψ_{i_0} is false, it proceeds to the upper cedent. However, if it is true, the algorithm has discovered a disjunct of $\varphi = \bigvee_{i \in \mathcal{I}} \psi_i$ which is true, contradicting the tentative assumption that φ was false. The procedure then backtracks down the path towards the root until it finds the *Cut* inference where the formula φ was added to the cedent. It then proceeds to the other (right) upper cedent of the *Cut*, and saves the information about which conjunct of $\overline{\varphi}$ is false.

Run-time analysis: The assumption on how *Cut* hypotheses are ordered implies that if the procedure backtracks, it moves from the left sub-proof above a *Cut* to the right subproof above the *Cut*. Therefore, the procedure is always following a left-to-right-ordered depth-first traversal in the proof.

The run time therefore $O(n^{O(1)})$, because there are only this many *Cut*'s and since this is an upper bound on the height of the proof.

This upper bound of $O(n^{O(1)})$ on the size of the subproof visited during the traversal applies *even though* the proof is exponentially big! (It is big but shallow, due to large fan-in of $\bigwedge$-inferences.

The procedure can terminate only at the cedent (1), since that is the only false leaf cedent. When it reaches this, it knows a value for y that falsifies it.

This value of y satisfies $A(x, y)$. □

The Main Theorem for T^1_2

Theorem (BK'94)

Suppose $A(x,y) \in \Sigma^b_1$ and that T^1_2 proves $(\forall x)(\exists y)A(x,y)$. Then there is a Polynomial Local Search (PLS) function $f(x) = y$ such that for all $x \in \mathbb{N}$, $A(x, f(x))$ holds.

The proof is identical to before, based on exactly the same procedure. Now the procedure may need $2^{n^{O(1)}}$ steps, instead of $n^{O(1)}$. Use the position in the proof to define a decreasing cost function, based on the procedure following a left-to-right depth first traversal. □

The theorems both hold if all true Π^b_1-formulas are added as axioms (no change to proof needed).

The generalize to S^i_2 and T^i_2 for $i > 1$ by the same proof. (Improved T^i_2 results will be discussed in the next talk.)

Transforming constant depth proofs.

Theorem (K'94, R'94, see BB'03)

Let $d \in \mathbb{N}$, and $\{\mathcal{A}_n\}_n$ be a family of sets of cedents. Then the following conditions (1) and (2) are equivalent:

(1) $\mathcal{A}_n$ *has a* Σ'*-depth* d LK *refutation of sequence-size quasi-polynomial in* n*, for all* n*.*

(2) $\mathcal{A}_n$ *has a* Σ'*-depth* $(d+1)$ LK *refutation of tree-size quasi-polynomial in* n*, for all* n*.*

Furthermore, the following conditions (3) and (4) are equivalent:

(3) $\mathcal{A}_n$ *has* Σ'*-depth* d LK *refutation of tree-size quasi-polynomial in* n*, for all* n*.*

(4) $\mathcal{A}_n$ *has a* Σ'*-depth* $(d+1)$ LK *refutation which simultaneously has tree-size quasi-polynomial in* n *and height poly-logarithmic in* n*, for all* n*.*

Corollary

Let $d \geq 2$. Suppose A is a $ss\Sigma_d^b$-formula and that $T_2^d \vdash A$. Without loss of much generality, A has the form

$$(\exists y \leq t(x))(\forall z \leq r(x))C(x,y,z).$$

Let $n_t = n^{O(1)}$ bound $|t(x)|$ for all $x < 2^n$, and $n_r = n^{O(1)}$ bound $|r(x)|$ for all $x < 2^n$ Then the set $\mathcal{A}_n$ of cedents

$$\{[\![y \leq t \rightarrow (z \leq r \wedge \neg C(x,y,z)]\!]_n/(y \mapsto i, z \mapsto j) \ : \ j < 2^{n_r}\},$$

for $i < 2^{n_t}$, has a Σ'-depth $(d-2)$ LK-refutation of size $2^{n^{O(1)}}$.

Explanation: In effect, $[\![A]\!]$ has a Σ'-depth $(d-2)$ proof.

This is a depth $(d - 1\frac{1}{2})$ refutation of the clauses expressing $\neg A$.

Some selected references

- S. Buss, Bounded Arithmetic and Constant Depth Frege Proofs, Quaderni di Matematica, 2004. (This paper has the main constructions of the talk.)
- A. Beckmann, S. Buss, Separation results for the size of constant-depth propositional proof systems, APAL 136 (2005) 30-55.
- S. Buss, *Bounded Arithmetic*, Ph.D. thesis, 1985. Bibliopolis, 1986. Also available online.
- S. Buss, J. Krajíček, An application of Boolean complexity to separation problems in bounded arithmetic. Proc. LMS 69 (1994) 1-21.
- J. Krajíček, Lower bounds to the size of constant-depth Frege proofs. JSL, 59 (1994) 73-86.
- A. Razborov, On provable disjoint NP pairs, BRICS & ECCC, 1994.