

LA, Permutations, and the Hajós Calculus

Michael Soltys

McMaster University, Canada

ICALP 2004

The Logical Theory **LA**

- **LA**: three-sorted:
 - indices,
 - field elements,
 - matrices
- Designed for reasoning about matrices^a.
- Bounded index quantifiers.
- Proves all ring properties of matrices.
- Field independent; here we focus on $\text{GF}(2)$.

^aM. Soltys and S. Cook, *Proof Complexity of Linear Algebra*, to appear in the Annals of Pure and Applied Logic.

Axioms:

- usual Number Theory: $i + 0 = i$, $i + sj = s(i + j)$, \dots ,
- usual Field/Commutative-ring axioms:
 $a + b = b + a$, $a(bc) = (ab)c$, \dots ,
- Matrices: properties of ΣA .
- λ -calculus for constructing matrices from matrix variables
 A, B, C, \dots

$$A + B := \lambda ij \langle n, n, e(A, i, j) + e(B, i, j) \rangle$$

$$A * B := \lambda ij \langle n, n, \Sigma \lambda kl \langle 1, n, e(A, i, l) * e(B, l, j) \rangle \rangle$$

Rules:

- Gentzen's System **LK**,
- Equality:

$$\frac{r(A) = r(B), c(A) = c(B), e(A, i, j) = e(B, i, j)}{A = B}$$

i, j index variables that do not occur free on the right-hand.

- Induction:

$$\frac{\alpha(i) \rightarrow \alpha(i + 1)}{\alpha(0) \rightarrow \alpha(n)}$$

i an index variable which does not occur free on the right-hand side, α contains no field or matrix quantifiers, and only bounded index quantifiers.

Translations:

- over $\text{GF}(2)$ into $\mathbf{AC}^0[2]$ -Frege,

$$\|(a*(b+c)) = ((a*b)+(a*c))\| \mapsto (a \wedge (b \oplus c)) \leftrightarrow ((a \wedge b) \oplus (a \wedge c))$$

$$\|A = A\|_{\sigma} \mapsto \left\{ \bigwedge_{1 \leq i \leq \sigma(r(A)), 1 \leq j \leq \sigma(c(A))} (A_{ij} \leftrightarrow A_{ij}) \right\}_{\sigma}$$

- over $\text{GF}(p)$ into $\mathbf{AC}^0[p]$ -Frege,
- over \mathbb{Q} into \mathbf{TC}^0 -Frege.

Hard Matrix Identities:

- $AB = I \supset BA = I^a$
- Other matrix identities equivalent to $AB = I \supset BA = I$ in **LA**.
For example:
 - $AB = I \supset (AC = 0 \supset C = 0)$
 - $(AB_1 = I \wedge AB_2 = I) \supset B_1 = B_2$
 - $AB = I \supset A^t B^t = I$
 - “Top row of B is 0” $\rightarrow AB \neq I$

^aProposed by Cook as a candidate for separating Frege and extended Frege. See M. Bonnet, S. Buss, T. Pitassi, *Are there hard examples for Frege Systems?*, Feasible Mathematics, II:30–56, 1994.

Quantification over Permutation Matrices

- Example:

$$\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

- **LA**-formula $\text{Perm}(P)$:

$$(\forall i \leq r(P))(\exists! j \leq c(P))e(P, i, j) = 1 \wedge PP^t = I$$

- Permutation Quantifiers:

$$(\exists P \leq n)\alpha \quad \text{is} \quad (\exists P)[(r(P) \leq n \wedge c(P) \leq n \wedge \text{Perm}(P)) \wedge \alpha]$$

$$(\forall P \leq n)\alpha \quad \text{is} \quad (\forall P)[(r(P) \leq n \wedge c(P) \leq n \wedge \text{Perm}(P)) \supset \alpha]$$

- $\exists\text{PLA}$ and $\forall\text{PLA}$: **LA** with $((\exists/\forall)P \leq n)$ -induction.
Both prove the soundness of the Hajós Calculus.

- **LAP** is **LA** with **P**, and defining axioms:

$$\mathbf{P}(0, A) = I \quad \& \quad \mathbf{P}(n + 1, A) = \mathbf{P}(n, A) * A$$

Can formalize computing the char poly:

- Berkowitz's Algorithm
- Newton's Symmetric Polynomials (over fields of char 0)

Proves their correctness?

- $\exists\text{PLAP}$ and $\forall\text{PLAP}$

Prove Cayley-Hamilton Thm., mult. of det, etc. In particular, the correctness of the above two algorithms.

**Feasible ($\exists PLAP$) Proof of the
Cayley-Hamilton Theorem**

Let p_A be the characteristic polynomial of A , computed by Berkowitz's Algorithm:

$$A = \left(\begin{array}{c|c} a & R \\ \hline S & M \end{array} \right) \quad p_A = \underbrace{\begin{bmatrix} -1 \\ -a \\ -RS \\ -RMS \\ \vdots \\ -RM^{n-2}S \end{bmatrix}}_{\text{Toeplitz}} \cdot p_M$$

- $p_A(A) = (p_A)_n A^n + (p_A)_{n-1} A^{n-1} + \cdots + (p_A)_0 I$,
an **LAP**-term.
- The Cayley-Hamilton Theorem states that $p_A(A) = 0$.

- **Theorem:**

$$\exists\text{PLAP} \vdash p_A(A) = 0$$

- **Corollary:** $\exists\text{PLAP}$ proves:

- multiplicativity of the determinant,
- hard matrix identities,
- other universal theorems of matrix algebra
(cofactor expansion of det, axiomatic definition of det).

Proof of Theorem:

- $A[n]$ is the n -th principal submatrix of A .

$$A[n] =_{\text{def}} \lambda kl \langle r(A) - n, c(A) - n, e(A, n + k, n + l) \rangle.$$

- $A[1]$ is A with the first row and column removed.

$A[r(A) - 1]$ is the 1×1 matrix (a_{nn}) .

$A[0] = A$.

- $CH(A, n)$ is an **LAP**-formula stating that the CHT holds for

$$A[n], A[n + 1], \dots, A[r(A) - 1].$$

Formally, $CH(A, n)$ is

$$(\forall n \leq i < r(A)) p_{A[i]}(A[i]) = 0$$

Lemma: $\exists\text{PLAP}$ proves:

$$\neg\text{CH}(A, n) \supset (\exists P \leq r(A)) \neg\text{CH}(PAP^t, n + 1).$$

Proof: If $\neg\text{CH}(A, n)$, then $\exists k, n \leq k \leq r(A) - 1$, such that

$$p_{A[k]}(A[k]) \neq 0.$$

We choose the *largest* such k , and consider two cases.

Case 1: If $k \neq n \Rightarrow k \geq n + 1 \Rightarrow P = I \Rightarrow \neg CH(PAP^t, n + 1)$.

Case 2: If $k = n$, then by definition of k ,

$$p_{A[n+1]}(A[n + 1]) = \cdots = p_{A[r(A)-1]}(A[r(A) - 1]) = 0$$

Find the *first* non-zero column of $p_{A[n]}(A[n])$, and call it j .

Note: $j \neq 1$

Why $j \neq 1$?

- **Long Technical Result:**

$\mathbf{LAP} \vdash p_{C[1]}(C[1]) = 0 \rightarrow$ “first column of $p_C(C)$ is zero”

- Since $k = n$, $p_{A[n+1]}(A[n+1]) = 0$
- Therefore: first column of $p_{A[n]}(A[n])$ is zero.
- Therefore: $j > 1$.

Find P and $0 \leq i < j$ such that

$$p_{(PAP^t)[n+j-i]}((PAP^t)[n+j-i]) \neq 0$$

Let I_k be I with rows k and $k+1$ transposed.

```

P ← I
i ← 0
while i < j
  if  $p_{(PAP^t)[n+j-i]}((PAP^t)[n+j-i]) = 0$  then
    P ←  $I_{n+j-i-1} \cdot P$ 
    i ← i + 1
  else
    output P
    break

```

- Before i reaches $(j - 1)$, the program finds a P such that

$$p_{(PAP^t)[n+j-i]}((PAP^t)[n+j-i]) \neq 0$$

- Otherwise,

$$p_{(PAP^t)[n+1]}((PAP^t)[n+1]) = 0 \text{ with } P = I_n I_{n+1} \cdots I_{n+j-1}.$$

- Not possible: Long Technical Result & $C = (PAP^t)[n+1]$.

Theorem: $\exists\text{PLAP} \vdash p_A(A) = 0$

Proof:

- Suppose $p_A(A) \neq 0$.
- Then $(\exists P \leq r(A)) \neg \text{CH}(PAP^t, 0)$: **Basis Case.**
- Lemma we just proved: **Induction Step.**
- Thus, by induction: $(\exists P \leq r(A)) \neg \text{CH}(PAP^t, r(A) - 1)$.
- Therefore: CHT fails for a 1×1 matrix $(PAP^t)[r(A) - 1]$.
Contradiction.

NP and co-NP Graph-Theoretic Problems

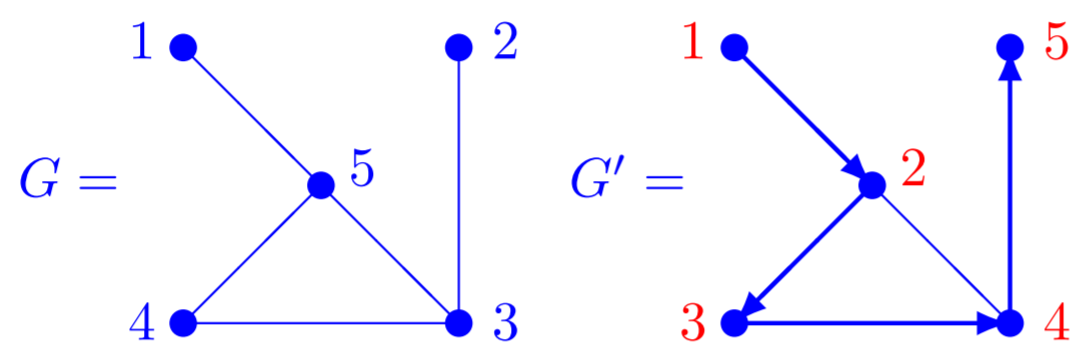
Graph Isomorphism: $(\exists P \leq r(A))[A = PBP^t]$

Path:

$(\exists P \leq r(A))[(\forall 0 < i < k)e(PAP^t, i, i+1) = 1 \wedge Ps = e_1 \wedge Pt = e_k]$

Hamiltonian Path:

$(\exists P \leq r(A))(\forall 0 < i < r(A))[e(PAP^t, i, i+1) = 1]$



k -Colorability:

Let 0_k denote the $k \times k$ matrix of zeros.

A is k -colorable, for any fixed k , can be stated as:

$$(\exists P \leq r(A))(\exists \mathbf{i} \leq r(A)) [PAP^t = \left[\begin{array}{c|c|c|c} 0_{i_1} & * & \cdots & * \\ \hline * & 0_{i_2} & \cdots & * \\ \hline \vdots & \vdots & \ddots & \vdots \\ \hline * & * & * & 0_{i_k} \end{array} \right]]$$

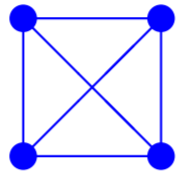
where $\mathbf{i} = i_1, i_2, \dots, i_k$.

For $k = 3$, let **Non-3-Col**(A) be the negation of the above formula, stating that A is *not* 3 colorable.

Non-3-Col(A) is a formula in the language of $\forall\text{PLA}$.

The Hajós Calculus and $\forall P\mathbf{LA}$

- HC is a non-deterministic procedure for building non-3-colorable graphs.
- Can be used as a propositional refutation system, and as such it is p -equivalent to extended Frege^a.
- $\forall P\mathbf{LA}$ proves the soundness of the HC.
- Let K_4 denote the 4-clique,

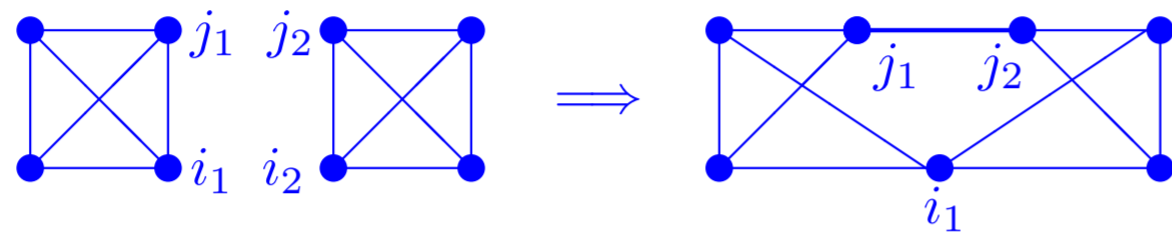


The K_4 graph is the only axiom of the HC.

- $\forall P\mathbf{LA} \vdash \mathbf{Non-3-Col}(A_{K_4})$.

^aSee T. Pitassi and A. Urquhart, *The complexity of the Hajós Calculus*, SIAM J. Disc. Math., (3):464-483, 1995.

1. **Addition Rule:** Add any number of vertices and/or edges.
2. **Join Rule:** Let G_1 and G_2 be two graphs with disjoint sets of vertices. Let (i_1, j_1) and (i_2, j_2) be edges in G_1 and G_2 , respectively. Construct G_3 as follows: remove edges (i_1, j_1) and (i_2, j_2) , and add the edge (j_1, j_2) , and contract vertices i_1 and i_2 into the single vertex i_1 .



3. **Contraction Rule:** Contract two nonadjacent vertices into a single vertex, and remove the resulting duplicated edges. The new vertex can be either of the two original vertices.

- A **derivation** is a sequence $\{G_1, G_2, \dots, G_n\}$.
Each G_i is either K_4 , or follows from previous G_j 's by a rule.
- (The HC is **sound** and **complete**.)
- **Lemma:** $\forall P\mathbf{LA}$ proves the soundness of the rules.

Proof:

Addition Rule:

- Let G' be G with new vertices/edges:

$$(\forall i, j \leq r(A_G))[e(i, j, A_G) = 1 \supset e(i, j, A_{G'}) = 1]$$

- So, $A_{G'}$ contains A_G in its upper-left corner, with, possibly, certain 0s replaced by 1s, and:

$$\forall P\mathbf{LA} \vdash \mathbf{Non-3-Col}(A_G) \rightarrow \mathbf{Non-3-Col}(A_{G'})$$

Join rule:

- Suppose that $e(A_{G_1}, i_1, j_1) = e(A_{G_2}, i_2, j_2) = 1$.
- Then A_G is given by a matrix with $r(A_{G_1}) + r(A_{G_2}) - 1$ rows (and columns), and of the form:

$$\left[\begin{array}{cc|c} A_{G_1}[i_1|i_1] & & D_1 \\ & A_{G_2}[i_2|i_2] & D_2 \\ \hline D_1^t & D_2^t & 0 \end{array} \right]$$

where: $(D_1)_{1j} = 1 \iff e(A_{G_1}, i_1, j) = 1$ and
 $(D_2)_{1j} = 1 \iff e(A_{G_2}, i_2, j) = 1$.

- $\forall P\mathbf{LA}$ proves

$$\mathbf{Non-3-Col}(A_{G_1}), \mathbf{Non-3-Col}(A_{G_2}) \rightarrow \mathbf{Non-3-Col}(A_G)$$

Same for contraction rule.

- Let $Y = [X_1 X_2 \dots X_n]$ encode a HC refutation, and let $\mathbf{HC}(Y)$ be an \mathbf{LA} -formula stating that.
- $\mathbf{HC}(Y)$ can be defined with bounded index quantifiers: for all $i \leq n$, the i -th block of Y is either A_{K_4} , or follows from previous 1 or 2 blocks by one of the three rules.
- The completeness of the HC can be stated as:

$$\mathbf{Non-3-Col}(X) \rightarrow \exists Y(\mathbf{HC}(Y) \wedge X_n = X)$$

Theorem: $\forall P\mathbf{LA}$ proves the soundness of the HC.

Proof:

- Soundness can be stated with:

$$\mathbf{HC}([X_1 X_2 \dots X_n]) \supset \mathbf{Non-3-Col}(X_n)$$

- Show by induction on k that:

$$(\forall i \leq k) [\mathbf{HC}([X_1 X_2 \dots X_n]) \supset \mathbf{Non-3-Col}(X_i)]$$

- * Since X_1 must encode K_4 , it follows that $\mathbf{Non-3-Col}(X_1)$, and hence we have the Basis Case.
- * The Induction Step follows from the lemma showing the soundness of the rules.

PK with Quantification over Permutations

- Permutation Frege: $\frac{\alpha}{\alpha^\pi}$
- Extended Frege p -simulates Permutation Frege.
- Are they p -equivalent ?
- Consider two fragments of PK with QP:
 - $\exists\sigma$ -PK: $\exists\sigma_S\alpha$
 - $\forall\sigma$ -PK: $\forall\sigma_S\alpha$ α has no quantifiers.
- Example:

$$\exists\sigma_{\{a,b,c\}}(\neg a \vee (b \wedge c))$$

Semantics:

- t^{σ_S} is the truth value assignment t where $t^{\sigma_S}(x) = t(\sigma_S(x))$.

- $t \models \exists \sigma_S \alpha$ iff there exists a σ_S such that $t^{\sigma_S} \models \alpha$.

This could also be defined as follows:

$t \models \exists \sigma_S \alpha$ iff there exists a σ_S such that $t \models \alpha^{\sigma_S}$.

- $t \models \forall \sigma_S \alpha$ iff for all σ_S , $t^{\sigma_S} \models \alpha$.

- Different from G^a where $t \models \exists x \alpha(x)$ iff $t \models \alpha(0) \vee \alpha(1)$.

^aJan Krajíček, *Bounded Arithmetic, Propositional Logic, and Complexity Theory*, Cambridge University Press.

Rules:

$$\frac{\Gamma \rightarrow \Delta, \alpha}{\Gamma \rightarrow \Delta, \exists \sigma_S \alpha^{\pi_S}} \quad \boxed{\frac{\alpha, \Gamma \rightarrow \Delta}{\exists \sigma_S \alpha^{\pi_S}, \Gamma \rightarrow \Delta}}$$

$$\boxed{\frac{\Gamma \rightarrow \Delta, \alpha}{\Gamma \rightarrow \Delta, \forall \sigma_S \alpha^{\pi_S}}} \quad \frac{\alpha, \Gamma \rightarrow \Delta}{\forall \sigma_S \alpha^{\pi_S}, \Gamma \rightarrow \Delta}$$

Restrictions:

- α has *no* quantifiers,
- boxed rules: variables in S are *not free* in the bottom sequent.

The variables in a finite set S are **not free** in a formula β :

- they do not occur in β at all, or
- $\beta =_{\text{def}} (\exists/\forall) \sigma_Q \gamma$, with γ having no quantifiers, and $S \subseteq Q$.

Complete? Equivalent to G ?

Theorem: $\exists\sigma$ -PK* and $\forall\sigma$ -PK* are p -equivalent to extended Frege.

Proof: $\exists\sigma$ -PK* and $\forall\sigma$ -PK* are p -equivalent.

- G_1^* : is tree-like quantified Frege with Σ_1 quantification.

- Extended Frege and G_1^* are p -equivalent.

- G_1^* p -simulates $\exists\sigma$ -PK* as follows:

Take a $\exists\sigma$ -PK* derivation, and replace $\exists\sigma_S$ introductions with $|S|$ -many \exists -introductions, one for each variable in S .

The restriction for $\exists\sigma_S$ -introductions left ensure that the restriction for \exists -introductions left is met.

Example: $\exists\sigma_{\{x,y\}}(\neg x \wedge y)$ replaced by $\exists x\exists y(\neg x \wedge y)$.

\therefore extended Frege p -simulates $\exists\sigma$ -PK*.

- Show $\forall\sigma$ -PK* p -simulates extended Frege using:
 1. extended Frege and the HC are p -equivalent, and
 2. Theorem: $\forall P$ LA proves the soundness of the HC.
- By 1., to show that $\forall\sigma$ -PK* p -simulates extended Frege, it is enough to show that $\forall\sigma$ -PK* p -simulates the HC.
- By 2., to show that $\forall\sigma$ -PK* p -simulates the HC, it is enough to show that the proof of the theorem can be formalized in $\forall\sigma$ -PK*.
- The theorems of LA translate into $\mathbf{AC}^0[2]$ -Frege, and universal permutation quantifiers occur in the form $(\forall P \leq n)\alpha(PAP^t)$ (in the formula **Non-3-Col**), and so they translate into $\forall\sigma_A \|\alpha(A)\|_{\sigma'}$.

Open Problems

1. Is there an **LAP** proof of the CHT?
2. Can we prove hard matrix identities in **LAP**?
3. Hard matrix identities have been proposed as candidates for separating Frege and extended Frege—do they, or can they be proven in Frege, or somewhere in between (eg., Permutation Frege, if indeed it is strictly “in between”)?
4. Can we show that hard matrix identities are independent of **LA** (i.e., can we show that they don’t follow feasibly from basic ring properties of matrices?)
5. What would be a natural definition of QPF (that ensures soundness and completeness)?
6. Is (a good definition of) QPF p -equivalent to G ?