How to Associate Propositional Proof Systems and Theories with Complexity Classes

Part II:

Free-Cut Free Proofs and Witnessing Theorems

Stephen Cook

LC2002 Tutorial

August, 2002

Gentzen's System LK

A *sequent* has the form

$$\Gamma \to \Delta$$

where Γ, Δ are finite sequences of formulas. The meaning is

$$\bigwedge \Gamma \supset \bigvee \Delta$$

An *axiom* (in our version) has the form

$$A \to A$$

A any formula.

Propositional Rules (System PK)

exchange rules

left $\dfrac{\Gamma_1, A, B, \Gamma_2 \to \Delta}{\Gamma_1, B, A, \Gamma_2 \to \Delta}$ **right** $\dfrac{\Gamma \to \Delta_1, A, B, \Delta_2}{\Gamma \to \Delta_1, B, A, \Delta_2}$

contraction rules

left $\dfrac{\Gamma, A, A \to \Delta}{\Gamma, A \to \Delta}$ **right** $\dfrac{\Gamma \to \Delta, A, A}{\Gamma \to \Delta, A}$

weakening rules

$$\textbf{left}\ \frac{\Gamma \to \Delta}{A, \Gamma \to \Delta} \qquad \textbf{right}\ \frac{\Gamma \to \Delta}{\Gamma \to \Delta, A}$$

$\neg$ introduction rules

$$\textbf{left}\ \frac{\Gamma \to \Delta, A}{\neg A, \Gamma \to \Delta} \qquad \textbf{right}\ \frac{A, \Gamma \to \Delta}{\Gamma \to \Delta, \neg A}$$

$\wedge$ introduction rules

$$\textbf{left}\ \frac{A, B, \Gamma \to \Delta}{(A \wedge B), \Gamma \to \Delta} \qquad \textbf{right}\ \frac{\Gamma \to \Delta, A \quad \Gamma \to \Delta, B}{\Gamma \to \Delta, (A \wedge B)}$$

$\vee$ introduction rules

$$\textbf{left}\ \frac{A, \Gamma \to \Delta \quad B, \Gamma \to \Delta}{(A \vee B), \Gamma \to \Delta} \qquad \textbf{right}\ \frac{\Gamma \to \Delta, A, B}{\Gamma \to \Delta, (A \vee B)}$$

cut rule

$$\frac{\Gamma \to \Delta, A \quad A, \Gamma \to \Delta}{\Gamma \to \Delta}$$

The formula A in the cut rule is called the *cut* formula.

Soundness of PK: All **PK** theorems are valid.

Proof: The axioms are valid, and the rules preserve validity.

Completeness of PK: All valid sequents have cut-free **PK** proofs.

Inversion Principle: For each **PK** rule except weakening, if the bottom sequent is valid then all top sequents are valid.

Proof of Completeness: Induction on the total number of connectives in the sequent. □

$$\dfrac{\dfrac{\dfrac{\dfrac{P \to P}{P \to P, \neg Q}\,(weakening)}{\to P, \neg P, \neg Q}\,(\neg\ right) \qquad \dfrac{\dfrac{Q \to Q}{Q \to Q, \neg P}\,(weakening}{\to Q, \neg P, \neg Q}\,(\neg\ right)}{\to P \wedge Q, \neg P, \neg Q}\,(\wedge\ right)}{\to P \wedge Q, \neg P \vee \neg Q}\,(\vee\ right)}{\neg(P \wedge Q) \to \neg P \vee \neg Q}\,(\neg\ left)$$

Subformula property of cut-free proofs:

If π is a **PK** proof of a sequent $\Gamma \to \Delta$, and π does not use the cut rule, then every formula in every sequent of π is a subformula of some formula in Γ or Δ.

(This assumes that every sequent of π other than the endsequent is used as a hypothesis for a later sequent.)

System LK for Predicate Calculus

Two sytactic classes of variables:
Free variables $a, b, c, ...$
Bound variables $x, y, z, ...$

$\forall$ introduction rules

left $\dfrac{A(t), \Gamma \to \Delta}{\forall x A(x), \Gamma \to \Delta}$ **right** $\dfrac{\Gamma \to \Delta, A(b)}{\Gamma \to \Delta, \forall x A(x)}$

$\exists$ introduction rules

left $\dfrac{A(b), \Gamma \to \Delta}{\exists x A(x), \Gamma \to \Delta}$ **right** $\dfrac{\Gamma \to \Delta, A(t)}{\Gamma \to \Delta, \exists x A(x)}$

Restriction: The free variable b must not occur in the conclusion in $\forall$ **right**, $\exists$ **left**.

Term t in $\forall$ **left** and $\exists$ **right** may not have bound variables x, y, z. Hence t is freely substitutable for x in $A(x)$.

Soundness of LK: All **LK** theorems are valid. (Quantifier rules preserve validity.)

Completeness of LK: All valid sequents have cut-free **LK** proofs. (No $=$).

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{Pb \to Pb}{Pb \to Pb, Qb}\,(weakening)}{Pb \to Pb \vee Qb}\,(\vee\ right)}{\forall x Px \to Pb \vee Qb}\,(\forall\ left)}{\forall x Px \to \forall x(Px \vee Qx)}\,(\forall\ right) \qquad \dfrac{\dfrac{\dfrac{\dfrac{Qb \to Qb}{Qb \to Pb, Qb}\,(wea}{Qb \to Pb \vee Qb}\,(\vee}{\forall x Qx \to Pb \vee Qb}\,(\forall}{\forall x Qx \to \forall x(Px \vee Qx)}}{\forall x Px \vee \forall x Qx \to \forall x(Px \vee Qx)}$$

Proof of Completeness:
Let $A_1, A_2, ...,\ t_1, t_2, ...$ be enumerations of formulas and terms, with each occurring infinitely often.
Let $\langle A_i, t_j \rangle$ be an enumerations of all pairs.

Try to build a tree proof of $\Gamma \to \Delta$ by working backwards in stages. A sequent at a leaf is *active* unless it contains some formula A on both left and right.

Next stage: Let $\langle A_i, t_j \rangle$ be next pair. If A_i is atomic, do nothing. Otherwise for each active leaf sequent with A_i, apply appropriate rule to A_i in reverse.

In the case of $\exists$ **right** or $\forall$ **left**, use term t_j and keep a copy of A_i, using contraction.

$$\frac{\Gamma' \to \Delta', \exists x B(x), B(t_j), \Delta''}{\Gamma' \to \Delta', \exists x B(x), \Delta''}$$

If the algorithm halts with an **LK** proof of $\Gamma \to \Delta$, then done. Otherwise it forms an infinite tree with (by Konig) an infinite branch β.

Construct a *term model* $\mathcal{M}$ and variable assignment σ which falsifies every sequent on β, including $\Gamma \to \Delta$.

NOTE: No atomic formula occurs both on the left and right in β.

Definition of $\mathcal{M}$ and σ:

Universe M is the set of all terms t, using only free variables $a, b, c....$

$\sigma(a) = a$ for all variables a.

$f^{\mathcal{M}}(t_1, ..., t_k) = ft_1...t_k$ [so $t^{\mathcal{M}[\sigma]} = t$ for all t]

$P^{\mathcal{M}}(t_1, ..., t_k)$ holds iff $Pt_1...t_k$ occurs on the left side of some formula in branch β.

Lemma: $\mathcal{M}$ and σ satisfy every A which occurs on the left in β, and falsifies every A which occurs on the right.

Proof: Structural induction on A.

Q.E.D. (cut free completeness of **LK**)

Anchored LK Proofs

$\forall A$ denotes the universal closure of A.
Let Φ be a set of formulas.
$\forall\Phi = \{\forall A \mid A \in \Phi\}$

An $\mathbf{LK} - \Phi$ proof is an **LK** proof which allows sequents $\rightarrow A$, for $A \in \Phi$, as (nonlogical) axioms.

Definition: An $\mathbf{LK} - \Phi$ proof is *anchored* if all cut formulas are in Φ.

Anchored LK Completeness Theorem: If Φ is a set of formulas closed under substitution of terms for free variables, and $\Gamma \rightarrow \Delta$ is a logical consequence of $\forall\Phi$, then there is an anchored $LK - \Phi$ proof of $\Gamma \rightarrow \Delta$.

Application: Φ is a set of quantifier-free axioms for a theory. An anchored $\mathbf{LK} - \Phi$ proof contains no universal quantifiers, unless they occur in the endsequent.

Proof of Anchored Completeness:

At each stage, if $A_i \in \Phi$, for each active leaf $\Gamma' \to \Delta'$, add the derivation

$$\frac{A_i, \Gamma' \to \Delta' \qquad \Gamma' \to \Delta', A_i}{\Gamma' \to \Delta'} \, (cut)$$

The branch β goes through

$$A_i, \Gamma' \to \Delta'$$

Hence for term model $\mathcal{M}$ and assignment σ

$$\mathcal{M}[\sigma] \models A_i$$

If $A_i(a_1, ..., a_n) \in \Phi$, then $A_i(t_1, ..., t_n) \in \Phi$ for all $t_1, ..., t_n$, so

$$\mathcal{M} \models \forall x_1 ... \forall x_n A_i(x_1, ..., x_n)$$

so

$$\mathcal{M} \models \forall \Phi$$

System $\mathbf{LK}^2$

Second order **LK**, needed for $\mathbf{V}^0, \mathbf{V}^1$.
Add to **LK**

- Free string variables $\alpha, \beta, \gamma, ...$
- Bound string variables $X, Y, Z, ...$
- String quantifier rules String $\forall$ introduction rules

left $\dfrac{A(\alpha), \Gamma \to \Delta}{\forall X A(X), \Gamma \to \Delta}$ **right** $\dfrac{\Gamma \to \Delta, A(\beta)}{\Gamma \to \Delta, \forall X A(X)}$

String $\exists$ introduction rules

left $\dfrac{A(\beta), \Gamma \to \Delta}{\exists X A(X), \Gamma \to \Delta}$ **right** $\dfrac{\Gamma \to \Delta, A(\alpha)}{\Gamma \to \Delta, \exists X A(X)}$

Restriction: The free variable β must not occur in the conclusion of $\forall$-**right** and $\exists$-**left**.

Anchored completeness theorem holds for $\mathbf{LK}^2$

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

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

$\mathbf{\Sigma^B_i}$-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 $\mathbf{\Sigma^B_i}$ formula not containing X.

$\mathbf{V^0}$ has $\mathbf{\Sigma^B_0}$-Comprehension
$\mathbf{V^1}$ has $\mathbf{\Sigma^B_1}$-Comprehension

Witnessing Theorem for $\mathbf{V^1}$:
If $\mathbf{V^1} \vdash \exists Z\ B(\vec{x}, \vec{Y}, Z)$, where B is a $\mathbf{\Sigma^1_1}$ formula, then there is a string function $F \in \mathbf{FP}$ which is $\mathbf{\Sigma^B_1}$-definable in $\mathbf{V^1}$ such that

$$\mathbf{V^1}, AX(F) \vdash B(\vec{x}, \vec{Y}, F(\vec{x}, \vec{Y}))$$

where $AX(F)$ is the $\mathbf{\Sigma^B_1}$ defining axiom for F.

Proof Ideal: We may assume B is $\mathbf{\Sigma^B_0}$ (by combining $\exists Z_1 ... \exists Z_k$ to $\exists Z$).

Thus $B(\vec{x}, \vec{Y}, Z)$ has no string quantifiers, and all number quantifiers are bounded.

Definition: A is *strict* $\mathbf{\Sigma^1_1}$ if A has the form

$$\exists Z\ C$$

where C is $\mathbf{\Sigma^B_0}$.

Want an anchored $\mathbf{V^1}$-**LK** proof of $\exists Z\ B(\vec{x}, \vec{Y}, Z)$ in which all formulas are strict $\mathbf{\Sigma^1_1}$.

Difficulty: $\mathbf{\Sigma^B_1}$-COMP is NOT $\mathbf{\Sigma^1_1}$:

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

Theorem: $\mathbf{V^1}$ remains the same if $\mathbf{\Sigma^B_1}$-COMP is replaced by $\mathbf{\Sigma^B_0}$-COMP and strict $\mathbf{\Sigma^B_1}$-IND

$$[A(0) \wedge \forall x(A(x) \supset A(x+1))] \supset A(y) \qquad (1)$$

Proof: We've already shown $\mathbf{V^1}$ proves (1). Buss proved the other direction by introducing the function $Numones(X)$. □

Scheme (1) is not $\mathbf{\Sigma^B_1}$, but we can relpace it by the rule

$$\textbf{strict } \mathbf{\Sigma^B_1}\textbf{-IND } \frac{\Gamma, A(b) \to A(b+1), \Delta}{\Gamma, A(0) \to A(t), \Delta}$$

Restriction: b does not occur in the bottom sequent.

$A(b)$ any formula of the form $\exists Y \leq tB$, where B is $\mathbf{\Sigma^B_0}$. (or $A(b)$ is $\mathbf{\Sigma^B_0}$)

An *anchored* $\mathbf{V^1}$ *proof* is an anchored $\mathbf{LK^2} - \Phi$ proof which allows this rule, and allows $A(0), A(t)$ as cut formulas, where Φ is all term substitution instances of $\{B1, ..., B14, L1, L2, \mathbf{\Sigma^B_0}\text{-COMP}\}$.

Theorem: Every $\mathbf{V^1}$ theorem has an anchored $\mathbf{V^1}$ proof.

Proof of $\mathbf{V^1}$ Witnessing Theorem

Let π be an anchored $\mathbf{V^1}$ proof of

$$\rightarrow \exists Z \ B(\vec{x}, \vec{Y}, Z) \tag{2}$$

where B is $\mathbf{\Sigma^B_0}$. We want a polytime function F so

$$AX(F), \mathbf{V^1} \vdash B(\vec{x}, \vec{Y}, F(\vec{x}, \vec{Y}))$$

Claim: Every formula in every sequent of π is $\mathbf{\Sigma^B_0}$ or strict $\mathbf{\Sigma^1_1}$ (like (2)).

Thus each sequent in π has the form

$$S = \ ..., \exists X_i \phi_i(X_i), ..., \Gamma \rightarrow \Delta, ..., \exists Y_j \psi_j(Y_j), ...$$

where all ϕ_i, ψ_j and all formulas in Γ, Δ are $\mathbf{\Sigma^B_0}$.

We show there are polytime witnessing functions $..., F_j, ..$ such that there is a $\mathbf{LK}^2 - \mathbf{V^1} + AX(F_i)$ proof of

$$S' = \ ..., \phi_i(\beta_i), ..., \Gamma \rightarrow \Delta, ..., \psi_j(F_j), ...$$

where $..., \beta_i, ...$ are distinct new free variables and F_j stands for $F_j(\vec{a}, \vec{\alpha}, \vec{\beta})$, and $\vec{a}, \vec{\alpha}$ are exactly the free variables in S.

Case I: S is one of the axioms B1,...,B14,L1,L2,$\mathbf{\Sigma_0^B}$-COMP. Nothing to witness except for $\mathbf{\Sigma_0^B}$-COMP:

$$S =\!\!\rightarrow \exists X \leq b \forall z < b(X(z) \leftrightarrow \phi(z, b, \vec{a}, \vec{\alpha}))$$

To witness $\exists X$, define

$$Z = F(b, \vec{a}, \vec{\alpha}) \leftrightarrow |Z| \leq b \wedge \forall i < b \phi(i, b, \vec{a}, \vec{\alpha})$$

Case II: string $\exists$-**right**

$$\frac{S_1}{S} = \quad \frac{\Gamma \rightarrow \Delta, \phi(\gamma)}{\Gamma \rightarrow \Delta, \exists X \phi(X)}$$

If γ occurs free in S, define

$$F(\vec{a}, \gamma, \vec{\alpha}, \vec{\beta}) = \gamma$$

Otherwise define $F(...) = \lambda$ (emptyset), and modify each witnessing function F_j for formulas in Δ by setting $\gamma = \lambda$ (in arguments for F_j).

Case III: cut

$$\frac{S_1 \quad S_2}{S} = \frac{\Gamma \to \Delta, \exists X \phi(X) \qquad \exists X \phi(X), \Gamma \to \Delta}{\Gamma \to \Delta}$$

No new witnessing functions required.

$$\frac{S_1' \quad S_2'}{S'} = \frac{\Gamma' \to \Delta', \phi(G) \qquad \phi(\beta), \Gamma' \to \Delta'}{\Gamma' \to \Delta'}$$

Let ..., F_j, ... be witness functions for Δ' in S_1'
Let ..., F_j', ... witness same formulas in S_2'.
Let G witness $\exists X$ in S_1'.
Let $\phi(\beta)$ correspond to $\exists X \phi(X)$ in S_2'.
Then $F_j'(\beta)$ has argument β. $F_j()$ does not.

Define corresponding witness function F_j'' in S' by

$$F_j''() = \begin{cases} F_j() & \text{if } \neg\phi(G) \\ F_j'(G) & \text{if } \phi(G) \end{cases}$$

Case IV: Induction (Ignore side formulas)

$$\frac{\exists X \leq r(b)\phi(b,X) \rightarrow \exists X \leq r(b+1)\phi(b+1,X)}{\exists X \leq r(0)\phi(0,X) \rightarrow \exists X \leq r(t)\phi(t,X)}$$

Let $F(b,\beta)$ witness top, so top becomes

$$|\beta| \leq r(b) \wedge \phi(b,\beta) \rightarrow \\ |F(b,\beta)| \leq r(b+1) \wedge \phi(b+1, F(b,\beta))$$

Define witness $\widehat{F}$ for bottom by recursion:

$$\begin{aligned} \widehat{F}(0,\beta) &= \beta \\ \widehat{F}(b+1,\beta) &= F(b,\widehat{F}(b,\beta)) \end{aligned}$$

(Note that by induction on b,
if $|\beta| \leq r(0)$ then

$$|\widehat{F}(b,\beta)| \leq r(b)$$

Then $\widehat{F}(t,\beta)$ witnesses bottom.

Other Cases...

Tomorrow:

• Quick introduction to propositional proof complexity

• Translating theorems of $\mathbf{V^0}, \mathbf{V^1}$ to families of propositional tautologies with short proofs.