Introduction to Proof Theory

Lecture 4
Cut-elimination

Marianna Girlando, Sonia Marin

Universtiy of Amsterdam, University of Birmingham

OPLSS 2023

Eugene, Oregon, June 26 - July 8, 2023

Outline

Today's goal

$$\text{cut}\ \dfrac{\begin{array}{c}\mathcal{D}_1\\ \Gamma \Rightarrow A\end{array} \qquad \begin{array}{c}\mathcal{D}_2\\ A, \Gamma \Rightarrow C\end{array}}{\Gamma \Rightarrow C}$$

Theorem (*Hauptsatz*, Gentzen 1935)

*Every theorem of **LJ** has a proof that does not use the cut rule.*

Corollary (Analyticity)

*Every theorem of **LJ** has a proof that contains only subformulas of it.*

Informal example

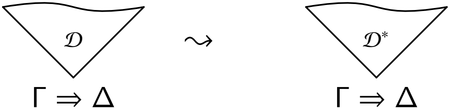
$$\text{cut}\;\dfrac{\wedge_R\,\dfrac{\overset{\mathcal{D}_1}{\Gamma \Rightarrow A} \qquad \overset{\mathcal{D}_2}{\Gamma \Rightarrow B}}{\Gamma \Rightarrow A \wedge B} \qquad \wedge_L\,\dfrac{\overset{\mathcal{D}_3}{A, B, \Gamma \Rightarrow C}}{A \wedge B, \Gamma \Rightarrow C}}{\Gamma \Rightarrow C}\; *$$

Let's eliminate the occurrence of cut marked by $*$

$$\text{cut}\;\dfrac{\overset{\mathcal{D}_1}{\Gamma \Rightarrow A} \qquad \text{cut}\;\dfrac{wk\;\dfrac{\overset{\mathcal{D}_2}{\Gamma \Rightarrow B}}{A, \Gamma \Rightarrow, B} \qquad \overset{\mathcal{D}_3}{A, B, \Gamma \Rightarrow C}}{A, \Gamma \Rightarrow C}}{\Gamma \Rightarrow C}$$

General strategy of the proof

LJ derivation $\rightsquigarrow$ cut-free **LJ** derivation



- Apply the cut on smaller formulas, until they disappear!
- Push the cuts upwards in the proof, until they disappear!
- We need a "measure" on formulas and on derivations, to ensure that the cut-elimination procedure terminates.

. . . The cut-elimination proof is quite complex.

We are going to sketch the proof for **LJ**.

References

Several proofs of cut-elimination exist in the literature, using slightly different procedures and for slightly different systems:

- [Buss, 1998]. *Handbook of Proof Theory*.
- [Troelstra and Schwichtenberg, 1996]. *Basic Proof Theory*.
- [Negri and von Plato, 2001]. *Structural Proof Theory*.
- ...

Introduction

Preliminary definitions and lemmas

The cut elimination theorem

Normalisation

From LJ to NJ and back

Measuring height

The height of $\mathcal{D}$, $\mathrm{ht}(\mathcal{D})$, is the length of its longest branch, minus one.

The level of a cut rule is the sum of heights of derivations of the two premisses of cut.

$$
m \left\{ \; \dfrac{\dfrac{\mathcal{D}_1}{\Gamma \Rightarrow A} \qquad \dfrac{\mathcal{D}_2}{A, \Gamma \Rightarrow C}}{\Gamma \Rightarrow C}\,\text{cut} \; \right\} n
$$

$m + n$

Measuring degree

The degree of a formula A, $\mathrm{deg}(A)$, is the number of logical connectives occurring in it.

$$\begin{aligned} \mathrm{deg}(p) &:= 0 \\ \mathrm{deg}(\bot) &:= 0 \\ \mathrm{deg}(A \star B) &:= \mathrm{deg}(A) + \mathrm{deg}(B) + 1 \\ & \text{for } \star \in \{\vee, \wedge, \to\} \end{aligned}$$

The rank of a cut rule is the degree of the cut formula A, plus 1.

$$\text{cut}\ \frac{\Gamma \Rightarrow A \quad A, \Gamma \Rightarrow C}{\Gamma \Rightarrow C}$$

rank

The rank of $\mathcal{D}$, $\mathrm{rk}(\mathcal{D})$, is the maximum of the cut formulas occurring in $\mathcal{D}$.

$\Gamma \Rightarrow^m_p C$ means *there is a derivation of $\Gamma \Rightarrow C$ of height **at most** m and rank **at most** p.*

Rank of derivations (more formally)

Height and rank can be inductively defined on the structure of $\mathcal{D}$:

$$\mathcal{D} = \text{init}\,\dfrac{}{\Gamma \Rightarrow C} \qquad \mathrm{rk}(\mathcal{D}) = 0$$

$$\mathcal{D} = \text{R}\,\dfrac{\begin{array}{c}\mathcal{D}_1\\ \Gamma_1 \Rightarrow C_1\end{array}}{\Gamma \Rightarrow C} \qquad \mathrm{rk}(\mathcal{D}) = \mathrm{rk}(\mathcal{D}_1)$$

$$\mathcal{D} = \text{R}\,\dfrac{\begin{array}{c}\mathcal{D}_1\\ \Gamma_1 \Rightarrow C_1\end{array} \qquad \begin{array}{c}\mathcal{D}_2\\ \Gamma_2 \Rightarrow C_2\end{array}}{\Gamma \Rightarrow C} \qquad \mathrm{rk}(\mathcal{D}) = \max(\mathrm{rk}(\mathcal{D}_1), \mathrm{rk}(\mathcal{D}_2))$$

$$\mathcal{D} = \text{cut}\,\dfrac{\begin{array}{c}\mathcal{D}_1\\ \Gamma \Rightarrow A\end{array} \qquad \begin{array}{c}\mathcal{D}_2\\ A, \Gamma \Rightarrow C\end{array}}{\Gamma \Rightarrow C} \qquad \mathrm{rk}(\mathcal{D}) = \max(\mathrm{rk}(\mathcal{D}_1), \mathrm{rk}(\mathcal{D}_2), \underline{\mathrm{deg}(A) + 1})$$

Some preliminary lemmas

$$\begin{array}{c} \Gamma \Rightarrow C \\ \vdots \\ \Gamma', \Gamma \Rightarrow C \end{array} \qquad \boxed{\dfrac{\Gamma \Rightarrow C}{A, \Gamma \Rightarrow C}\,wk}$$

$$\dfrac{A, A, \Gamma \Rightarrow C}{A, \Gamma \Rightarrow C}\,ctr$$

1. Lemma: Closure under weakening
   - If $\Gamma \Rightarrow^m_p C$, then $\Gamma', \Gamma \Rightarrow^m_p C$, for any Γ'.

   Proof. Easy induction on the height m, of the derivation.

2. Lemma: Closure under contraction
   - If $A, A, \Gamma \Rightarrow^m_p \overset{C}{\cancel{A}}$, then $A, \Gamma \Rightarrow^m_p C$.

   Proof. Induction on m, using weakening (and invertibility)

NB: all the above preserve height and rank of the derivation.

The plan

LJ derivation $\rightsquigarrow$ cut-free **LJ** derivation

$$\begin{array}{c} \mathcal{D} \\ \Gamma \Rightarrow_{p} C \end{array} \quad \rightsquigarrow \quad \begin{array}{c} \mathcal{D}^{*} \\ \Gamma \Rightarrow_{0} C \end{array}$$

- We show how to simulate instances of cut. Lemma
- We show how to eliminate all cuts occurring in a derivation, starting with topmost cuts having maximal rank. Theorem

Lemma (closure under cut)

If $\Gamma \Rightarrow^{m}_{0} A$ and $A, \Gamma \Rightarrow^{n}_{0} C$

$$\text{cut}\ \dfrac{\begin{array}{c}\mathcal{D}_1 \\ R_1 \\ \Gamma \Rightarrow^{m}_{0} A\end{array} \qquad \begin{array}{c}\mathcal{D}_2 \\ R_2 \\ A, \Gamma \Rightarrow^{n}_{0} C\end{array}}{\Gamma \Rightarrow_{1} C}$$

then we can construct the following derivation $\mathcal{D}^*$:

$$\begin{array}{c}\mathcal{D}^* \\ \Gamma \Rightarrow_{0} C\end{array}$$

Proof. Induction on $(\deg(A), m+n)$. We distinguish cases:

1. R_1 is init (R_2 is init)
2. A is principal in both R_1 and R_2
3. A is not principal in R_1 (A is not principal in R_2)

R_1 is init

$\boxed{h=0}$

$$\mathcal{D}_1 = \text{init}\,\frac{}{\underbrace{A,\Gamma'}_{\Gamma} \Rightarrow^{m}_{p} A} \qquad \mathrm{R}_2\,\frac{\begin{array}{c}\mathcal{D}_2\\ \Gamma'' \Rightarrow^{n-1}_{p} C''\end{array}}{A,\Gamma \Rightarrow^{n}_{p} C}$$

$A, \Gamma' \Rightarrow C$

with $\Gamma = A, \Gamma'$. We construct the following derivation $\mathcal{D}$ of $\Gamma \Rightarrow_p \Delta$:

$$\mathit{ctr}\,\frac{\mathrm{R}_2\,\dfrac{\begin{array}{c}\mathcal{D}_2\\ \Gamma'' \Rightarrow C''\end{array}}{A, A, \Gamma' \Rightarrow C}}{A, \Gamma' \Rightarrow C}$$

A is principal in both R_1 and R_2

$\boxed{p=0}$

R_1 is $\rightarrow_R$ and R_2 is $\rightarrow_L$

$$
\dfrac{\dfrac{\mathcal{D}_1}{A, \Gamma \Rightarrow_p^{m-1} B}}{\Gamma \Rightarrow_p^{m} A \rightarrow B}\,{\scriptstyle \rightarrow_R}
\qquad
\dfrac{\dfrac{\mathcal{D}_2'}{A \rightarrow B, \Gamma \Rightarrow_p^{n1} A} \quad \dfrac{\mathcal{D}_2''}{B, \Gamma \Rightarrow_p^{n2} C}}{A \rightarrow B, \Gamma \Rightarrow_p^{n} C}\,{\scriptstyle \rightarrow_L}
$$

$\Gamma \Rightarrow C$

with $n1, n2 < n$. We construct the following derivation $\mathcal{D}$ of $\Gamma \Rightarrow_p \Delta$:

$(deg(A \rightarrow B), m+n)$

$$
\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\mathcal{D}_1}{A, \Gamma \Rightarrow B}}{\Gamma \overset{m}{\Rightarrow} A \rightarrow B}\,{\scriptstyle \rightarrow_R} \quad \dfrac{\mathcal{D}_2'}{A \rightarrow B, \Gamma \overset{m1}{\Rightarrow} A}}{\Gamma \Rightarrow A}\,{\scriptstyle \text{cut}} \quad \dfrac{\mathcal{D}_1}{A, \Gamma \Rightarrow B}}{\Gamma \Rightarrow B}\,{\scriptstyle \text{cut}} \quad \dfrac{\mathcal{D}_2''}{B, \Gamma \Rightarrow C}}{\Gamma \Rightarrow C}\,{\scriptstyle \text{cut}}
$$

$(deg(A \rightarrow B),$
$m+n1)$

$m+n1 < m+n$

Cases for the other rules . . .

A is not principal in R_1

$n = 0$

R_1 is a one-premiss rule

$$
\begin{array}{c}\mathcal{D}_1\\ \dfrac{\Gamma' \Rightarrow_p^{m-1} A}{\Gamma \Rightarrow_p^{m} A}\,R_1\end{array}
\qquad\qquad
\begin{array}{c}\mathcal{D}_2\\ \dfrac{\Gamma'' \Rightarrow_p^{n-1} C''}{A, \Gamma \Rightarrow_p^{n} C}\,R_2\end{array}
$$

We construct the following derivation $\mathcal{D}$ of $\Gamma \Rightarrow_p \Delta$:

$$
\dfrac{\dfrac{\dfrac{\dfrac{\begin{array}{c}\mathcal{D}_1\\ \Gamma' \Rightarrow, A\end{array}}{\Gamma', \Gamma \Rightarrow A}\,wk \qquad \dfrac{\dfrac{\begin{array}{c}\mathcal{D}_2\\ \Gamma'' \Rightarrow C''\end{array}}{A, \Gamma \Rightarrow C}\,R_2}{A, \Gamma', \Gamma \Rightarrow C}\,wk}{\Gamma', \Gamma \Rightarrow C}\,cut}{\Gamma, \Gamma \Rightarrow C}\,R_1}{\Gamma \Rightarrow C}\,ctr
$$

R_1 is a two-premisses rule ...

End of the proof

Eliminating cut

Cut-elimination Theorem If we have a derivation $\mathcal{D}$ of $\Gamma \Rightarrow_p C$, we can construct a derivation $\mathcal{D}^*$ of $\Gamma \Rightarrow_0 C$, that is, a derivation **where** cut **does not occur.**

Proof. We apply the proof transformation detailed in the Lemma to the cuts occurring in $\mathcal{D}$, starting with topmost cuts of maximal rank. The Lemma ensures us that after every proof transformation one instance of cut is eliminated.
Therefore, in finitely many steps, we obtain a derivation $\mathcal{D}^*$ of $\Gamma \Rightarrow_0 C$, where the cut rule does not occur.

☞ How many steps? $\Gamma \Rightarrow^m_p C \;\rightsquigarrow\; \Gamma \Rightarrow_0^{4_p(\cancel{n}\, m)} C$

$$4_p(\cancel{n}\, m) \quad = \quad \underbrace{4^{4^{\cdot^{\cdot^{4^{m}}}}}}_{p}$$

Why all this work?

- analiticity

- $\vdash_{Hcp} A \rightsquigarrow \vdash_{LJ} \Rightarrow A$ completeness

  - all the axioms of Hilbert system are derivable in LJ
  - MP can be simulated in LJ

$$\dfrac{A \qquad A \to B}{B} \qquad\qquad \dfrac{\dfrac{\vdots}{\Rightarrow A} \qquad \dfrac{\vdots}{A \Rightarrow B}}{\Rightarrow B}\,\text{cut}$$

Outline

Detours

In NJ a <u>detour</u> is a gadget formed by an intro. rule followed by an elim. rule

↳ its major premiss is the conclusion of the intro

major

$$\cfrac{\cfrac{M : A \qquad N : B}{\langle M,N\rangle : A\wedge B}\,\wedge_I}{\boxed{\pi_1(\langle M,N\rangle)} : A}\,\wedge_{E_1}$$

major

$$\cfrac{\cfrac{L : A}{in_1(L) : \boxed{A\vee B}}\,\vee_I \qquad \begin{array}{c}[x : A]\\ \vdots\\ M : C\end{array} \qquad \begin{array}{c}[y : B]\\ \vdots\\ N : C\end{array}}{\boxed{case\,(in_1(L),\, x \Rightarrow M,\, y \Rightarrow N)} : C}\,\vee_E$$

major

$$\cfrac{\cfrac{\begin{array}{c}x : [A]\\ \vdots\\ M : B\end{array}}{\lambda x.M : \boxed{A \to B}}\,\to_I \qquad N : A}{\boxed{(\lambda x.M)\,N} : B}\,\to_E$$

β-reduction

eliminating detours in derivations in NJ

$\longleftrightarrow$ normalizing λ-terms (weak)

$$\boxed{(\lambda x.M)N \leadsto M[x/N]}$$