Lecture Notes on Cut Elimination

15-317: Constructive Logic
Frank Pfenning*

Lecture 10
October 8, 2015

1 Introduction

The identity theorem of the sequent calculus exhibits one connection between the judgments $A\ \mathit{left}$ and $A\ \mathit{right}$: If we assume $A\ \mathit{left}$ we can prove $A\ \mathit{right}$. In other words, the left rules of the sequent calculus are strong enough so that we can reconstitute a proof of A from the assumption A. So the identity theorem is a global version of the local completeness property for the elimination rules.

The cut theorem of the sequent calculus expresses the opposite: if we have a proof of $A\ \mathit{right}$ we are licensed to assume $A\ \mathit{left}$. This can be interpreted as saying the left rules are not too strong: whatever we can do with the antecedent $A\ \mathit{left}$ can also be deduced without that, if we know $A\ \mathit{right}$. Because $A\ \mathit{right}$ occurs only as a succedent, and $A\ \mathit{left}$ only as an antecedent, we must formulate this in a somewhat roundabout manner: If $\Gamma \Longrightarrow A\ \mathit{right}$ and $\Gamma, A\ \mathit{left} \Longrightarrow J$ then $\Gamma \Longrightarrow J$. In the sequent calculus for pure intuitionistic logic, the only conclusion judgment we are considering is $C\ \mathit{right}$, so we specialize the above property.

Because it is very easy to go back and forth between sequent calculus deductions of $A\ \mathit{right}$ and verifications of $A\!\uparrow$, we can use the cut theorem to show that every true proposition has a verification, which establishes a fundamental, global connection between truth and verifications. While the sequent calculus is a convenient intermediary (and was conceived as such

*With edits by André Platzer

by Gentzen [Gen35]), this theorem can also be established directly using verifications.

2 Admissibility of Cut

The cut theorem is one of the most fundamental properties of logic. Because of its central role, we will spend some time on its proof. In lecture we developed the proof and the required induction principle incrementally; here we present the final result as is customary in mathematics. The proof is amenable to formalization in a logical framework; details can be found in a paper by the instructor [Pfe00].

Theorem 1 (Cut) *If* $\Gamma \Longrightarrow A$ *and* $\Gamma, A \Longrightarrow C$ *then* $\Gamma \Longrightarrow C$*.*

Proof: By nested inductions on the structure of A, the derivation $\mathcal{D}$ of $\Gamma \Longrightarrow A$ and $\mathcal{E}$ of $\Gamma, A \Longrightarrow C$. More precisely, we appeal to the induction hypothesis either with a strictly smaller cut formula, or with an identical cut formula and two derivations, one of which is strictly smaller while the other stays the same. The proof is constructive, which means we show how to transform

$$\overset{\mathcal{D}}{\Gamma \Longrightarrow A} \quad \text{and} \quad \overset{\mathcal{E}}{\Gamma, A \Longrightarrow C} \quad \text{to} \quad \overset{\mathcal{F}}{\Gamma \Longrightarrow C}$$

The proof is divided into several classes of cases. More than one case may be applicable, which means that the algorithm for constructing the derivation of $\Gamma \Longrightarrow C$ from the two given derivations is naturally non-deterministic.

Case: $\mathcal{D}$ is an initial sequent.

$$\mathcal{D} = \frac{}{\Gamma', P \Longrightarrow P}\ \mathsf{init}$$

$\Gamma = (\Gamma', P)$	This case
$\Gamma', P, P \Longrightarrow C$	Deduction $\mathcal{E}$
$\Gamma', P \Longrightarrow C$	By Contraction (see Lecture 9)
$\Gamma \Longrightarrow C$	By equality

Case: $\mathcal{E}$ is an initial sequent using the cut formula.

$$\mathcal{E} = \frac{}{\Gamma, P \Longrightarrow P}\ \mathsf{init}$$

$A = P = C$	This case
$\Gamma \Longrightarrow A$	Deduction $\mathcal{D}$

Case: $\mathcal{E}$ is an initial sequent not using the cut formula.

$$\mathcal{E} = \frac{}{\Gamma', P, A \Longrightarrow P}\ \mathsf{init}$$

$\Gamma = (\Gamma', P)$	This case
$\Gamma', P \Longrightarrow P$	By rule init
$\Gamma \Longrightarrow P$	By equality

Case: A is the principal formula of the final inference in both $\mathcal{D}$ and $\mathcal{E}$. There are a number of subcases to consider, based on the last inference in $\mathcal{D}$ and $\mathcal{E}$. We show some of them.

Subcase: If A is of the form $A_1 \wedge A_2$, then

$$\mathcal{D} = \frac{\begin{array}{cc} \mathcal{D}_1 & \mathcal{D}_2 \\ \Gamma \Longrightarrow A_1 & \Gamma \Longrightarrow A_2 \end{array}}{\Gamma \Longrightarrow A_1 \wedge A_2}\ \wedge R$$

$$\text{and} \qquad \mathcal{E} = \frac{\begin{array}{c} \mathcal{E}_1 \\ \Gamma, A_1 \wedge A_2, A_1 \Longrightarrow C \end{array}}{\Gamma, A_1 \wedge A_2 \Longrightarrow C}\ \wedge L_1$$

$\Gamma, A_1 \Longrightarrow C$	By i.h. on $A_1 \wedge A_2$, $\mathcal{D}$ and $\mathcal{E}_1$
$\Gamma \Longrightarrow C$	By i.h. on A_1, $\mathcal{D}_1$, and previous line

Actually we have ignored a detail: in the first appeal to the induction hypothesis, $\mathcal{E}_1$ has an additionaly hypothesis, A_1, and therefore does not match the statement of the theorem precisely. However, we can always weaken $\mathcal{D}$ to include this additional hypothesis without changing the structure of $\mathcal{D}$ (see the Weakening Theorem in Lecture 9) and then appeal to the induction hypothesis. We will not be explicit about these trivial weakening steps in the remaining cases.

It is crucial for a well-founded induction that $\mathcal{E}_1$ is smaller than $\mathcal{E}$, so even if the same cut formula and same $\mathcal{D}$ is used, $\mathcal{E}_1$ got smaller. Note that we cannot directly appeal to induction hypothesis on A_1,$\mathcal{D}_1$ and $\mathcal{E}_1$ because the additional formula $A_1 \wedge A_2$ might still be used in $\mathcal{E}_1$, e.g., by a subsequent use of $\wedge L_2$.

Subcase:

$$\mathcal{D} = \dfrac{\overset{\mathcal{D}_2}{\Gamma, A_1 \Longrightarrow A_2}}{\Gamma \Longrightarrow A_1 \supset A_2} \supset\! R$$

$$\text{and} \quad \mathcal{E} = \dfrac{\overset{\mathcal{E}_1}{\Gamma, A_1 \supset A_2 \Longrightarrow A_1} \quad \overset{\mathcal{E}_2}{\Gamma, A_2 \supset A_2, A_2 \Longrightarrow C}}{\Gamma, A_1 \supset A_2 \Longrightarrow C} \supset\! L$$

$\Gamma \Longrightarrow A_1$	By i.h. on $A_1 \supset A_2$, $\mathcal{D}$ and $\mathcal{E}_1$
$\Gamma \Longrightarrow A_2$	By i.h. on A_1 from above and $\mathcal{D}_2$
$\Gamma, A_2 \Longrightarrow C$	By i.h. on $A_1 \supset A_2$, $\mathcal{D}$ and $\mathcal{E}_2$
$\Gamma \Longrightarrow C$	By i.h. on A_2 from above

Note that the proof constituents of the last step $\Gamma \Longrightarrow C$ are longer than the original deductions $\mathcal{D}, \mathcal{E}$. Hence, it is crucial for a well-founded induction that the cut formula A_2 is smaller than $A_1 \supset A_2$.

Case: A is not the principal formula of the last inference in $\mathcal{D}$. In that case $\mathcal{D}$ must end in a left rule and we can appeal to the induction hypothesis on one of its premises. We show some of the subcases.

Subcase: If $\mathcal{D}$ ended with an $\wedge L_1$:

$$\mathcal{D} = \dfrac{\overset{\mathcal{D}_1}{\Gamma', B_1 \wedge B_2, B_1 \Longrightarrow A}}{\Gamma', B_1 \wedge B_2, \Longrightarrow A} \wedge L_1$$

$\Gamma = (\Gamma', B_1 \wedge B_2)$	This case
$\Gamma', B_1 \wedge B_2, B_1 \Longrightarrow C$	By i.h. on A, $\mathcal{D}_1$ and $\mathcal{E}$
$\Gamma', B_1 \wedge B_2 \Longrightarrow C$	By rule $\wedge L_1$
$\Gamma \Longrightarrow C$	By equality

Subcase:

$$\mathcal{D} = \dfrac{\overset{\mathcal{D}_1}{\Gamma', B_1 \supset B_2 \Longrightarrow B_1} \quad \overset{\mathcal{D}_2}{\Gamma', B_1 \supset B_2, B_2 \Longrightarrow A}}{\Gamma', B_1 \supset B_2 \Longrightarrow A} \supset\! L$$

$\Gamma = (\Gamma', B_1 \supset B_2)$ This case
$\Gamma', B_1 \supset B_2, B_2 \Longrightarrow C$ By i.h. on A, $\mathcal{D}_2$ and $\mathcal{E}$
$\Gamma', B_1 \supset B_2 \Longrightarrow C$ By rule $\supset L$ on $\mathcal{D}_1$ and above
$\Gamma \Longrightarrow C$ By equality

Case: A is not the principal formula of the last inference in $\mathcal{E}$. This overlaps with the previous case, since A may not be principal on either side. In this case, we appeal to the induction hypothesis on the subderivations of $\mathcal{E}$ and directly infer the conclusion from the results. We show some of the subcases.

Subcase:

$$\mathcal{E} = \dfrac{\begin{array}{cc}\mathcal{E}_1 & \mathcal{E}_2 \\ \Gamma, A \Longrightarrow C_1 & \Gamma, A \Longrightarrow C_2\end{array}}{\Gamma, A \Longrightarrow C_1 \wedge C_2} \wedge R$$

$C = C_1 \wedge C_2$ This case
$\Gamma \Longrightarrow C_1$ By i.h. on A, $\mathcal{D}$ and $\mathcal{E}_1$
$\Gamma \Longrightarrow C_2$ By i.h. on A, $\mathcal{D}$ and $\mathcal{E}_2$
$\Gamma \Longrightarrow C_1 \wedge C_2$ By rule $\wedge R$ on above

Subcase:

$$\mathcal{E} = \dfrac{\begin{array}{c}\mathcal{E}_1 \\ \Gamma', B_1 \wedge B_2, B_1, A \Longrightarrow C\end{array}}{\Gamma', B_1 \wedge B_2, A \Longrightarrow C} \wedge L_1$$

$\Gamma = (\Gamma', B_1 \wedge B_2)$ This case
$\Gamma', B_1 \wedge B_2, B_1 \Longrightarrow C$ By i.h. on A, $\mathcal{D}$ and $\mathcal{E}_1$
$\Gamma', B_1 \wedge B_2 \Longrightarrow C$ By rule $\wedge L_1$ from above

□

3 Cut Elimination

Gentzen's original presentation of the sequent calculus included an inference rule for cut. The analogue in our system would be

$$\dfrac{\Gamma \Longrightarrow A \quad \Gamma, A \Longrightarrow C}{\Gamma \Longrightarrow C} \,\mathsf{cut}$$

The advantage of this calculus is that it more directly corresponds to natural deduction in its full generality, rather than verifications. The disadvantage is that it cannot easily be seen as capturing the meaning of the connectives by inference rules, because with the rule of cut the meaning of C might depend on the meaning of any other proposition A (possibly even including C as a subformula).

In order to clearly distinguish between the two kinds of calculi, the one we presented is sometimes called the *cut-free sequent calculus*, while Gentzen's calculus would be a *sequent calculus with cut*. The theorem connecting the two is called *cut elimination*: for any deduction in the sequent calculus with cut, there exists a cut-free deduction of the same sequent. The proof is a straightforward induction on the structure of the deduction, appealing to the cut theorem in one crucial place.

Theorem 2 (Cut Elimination) *If $\mathcal{D}$ is a deduction of $\Gamma \Longrightarrow C$ possibly using the* cut *rule, then there exists a cut-free deduction $\mathcal{D}'$ of $\Gamma \Longrightarrow C$.*

Proof: By induction on the structure of $\mathcal{D}$. In each case, we appeal to the induction hypothesis on all premises and then apply the same rule to the result. The only interesting case is when a cut rule is encountered.

Case:

$$\mathcal{D} = \dfrac{\begin{array}{c}\mathcal{D}_1\\ \Gamma \Longrightarrow A\end{array} \quad \begin{array}{c}\mathcal{D}_2\\ \Gamma, A \Longrightarrow C\end{array}}{\Gamma \Longrightarrow C}\ \mathsf{cut}$$

$\Gamma \Longrightarrow A$ without cut	By i.h. on $\mathcal{D}_1$
$\Gamma, A \Longrightarrow C$ without cut	By i.h. on $\mathcal{D}_2$
$\Gamma \Longrightarrow C$	By the Cut Theorem

□

Similarly, Gentzen also allowed initial sequents with a non-atomic principal formula. It is a straightforward exercise to show that any deduction that uses non-atomic initial sequents can be expanded into one that uses only atomic ones.

4 Quantification in Sequent Calculus

In natural deduction, we had two forms of hypotheses: A *true* and $c : \tau$ for parameters c. The latter form was introduced into deductions by the $\forall I$

and $\exists E$ rules. In the sequent calculus we make all assumptions explicit on the left-hand side of sequents. In order to model parameters we therefore need a second kind of judgment on the left that reads $c : \tau$. It is customary to collect all such hypotheses in a different context, denoted Σ for *signature*. A sequent then has the form

$$\underbrace{c_1{:}\tau_1, \ldots, c_m{:}\tau_m}_{\Sigma}\ ;\ \underbrace{A_1\ \mathsf{left}, \ldots, A_n\ \mathsf{left}}_{\Gamma} \Longrightarrow C\ \mathsf{right}$$

We assume that all parameters declared in a signature Σ are distinct. Sometimes this requires us to choose a parameter with a name that has not yet been used. When writing down a sequent $\Sigma; \Gamma \Longrightarrow C$ we presuppose that all parameters in Γ and C are declared in Σ. In the bottom-up construction of a deduction we make sure to maintain this.

The typing judgment for terms, $t : \tau$, can depend on the signature Σ but not on logical assumptions A *left*. We therefore write $\Sigma \vdash t : \tau$ to express that term t has type τ in signature Σ.

In all the propositional rules we have so far, the signature Σ is propagated unchanged from the conclusion of the rule to all premises. In order to derive the rules for the quantifiers, we reexamine verifications for guidance, as we did for the propositional rules in Lecture 9.

Universal quantification. We show the verification on the left and with the corresponding right rule.

$$\dfrac{\begin{array}{c}\overline{c : \tau}\\ \vdots \\ A(c)\uparrow\end{array}}{\forall x{:}\tau.\ A(x)\uparrow}\ \forall I^c \qquad\qquad \dfrac{\Sigma, c{:}\tau; \Gamma \Longrightarrow A(c)}{\Sigma; \Gamma \Longrightarrow \forall x.\ A(x)}\ \forall R$$

Our general assumption that the signature declares every parameter at most once means that c cannot occur in Σ already or the rule would not apply. Also note that Σ declares all parameters occurring in Γ, so c cannot occur there, either. Hence, proving from assumption Γ that $A(x)$ holds for all x of type τ amounts to proving $A(c)$ for a new generic c of that type.

The elimination rule that uses a universally quantified assumption corresponds to a left rule.

$$\dfrac{\forall x{:}\tau.\ A(x)\downarrow \quad t : \tau}{A(t)\downarrow}\ \forall E \qquad\qquad \dfrac{\Sigma \vdash t : \tau \quad \Sigma; \Gamma, \forall x{:}\tau.\ A(x), A(t) \Longrightarrow C}{\Sigma; \Gamma, \forall x{:}\tau.\ A(x) \Longrightarrow C}\ \forall L$$

If we assume that $A(x)$ holds for all x of type τ, we might as well assume that $A(t)$ also holds for term t of said type.

Existential quantification. Again, we derive the sequent calculus rules from the introduction and elimination rules.

$$\frac{t : \tau \quad A(t)\uparrow}{\exists x{:}\tau.\ A(x)\uparrow}\ \exists I \qquad \frac{\Sigma \vdash t : \tau \quad \Sigma; \Gamma \Longrightarrow A(t)}{\Sigma; \Gamma \Longrightarrow \exists x{:}\tau.\ A(x)}\ \exists R$$

Existence of an x of type τ for which $A(x)$ has been proved from assumptions Γ after $A(t)$ has been proved from assumptions Γ. The arbitrary term t is called witness.

As for disjunction elimination, the natural deduction rule already has somewhat of the flavor of the sequent calculs.

$$\frac{\exists x{:}\tau.\ A(x)\downarrow \qquad \begin{matrix} \overline{c : \tau} \quad \overline{A(c)\downarrow}\ u \\ \vdots \\ C\uparrow \end{matrix}}{C\uparrow}\ \exists E^{c,u} \qquad \frac{\Sigma, c{:}\tau; \Gamma, \exists x{:}\tau.\ A(x), A(c) \Longrightarrow C}{\Sigma; \Gamma, \exists x{:}\tau.\ A(x) \Longrightarrow C}\ \exists L$$

From the assumption that $A(x)$ holds for some x of type τ, C follows, if C follows from the additional assumption that $A(c)$ holds for a new generic c of that type. Of course, by the well-formedness assumptions on sequents, c will be new (so not in Σ, Γ or C), which is important, because we could hardly assume $A(c)$ to hold for our specific favorite c if all we assume is that $A(x)$ holds for some x.

5 Cut Elimination with Quantification

The proof of the cut theorem extends to the case where we add quantifiers. A crucial property we need is substitution for parameters, which corresponds to a similar substitution principle on natural deductions:

Lemma 3 (Parameter substitution) *If* $\Sigma \vdash t : \tau$ *and* $\Sigma, c : \tau; \Gamma \vdash A$ *then* $\Sigma; [t/c]\Gamma \vdash [t/c]A$.

This is proved by a straightforward induction over the structure of the second deduction, appealing to some elementary properties such as weakening where necessary.

We show only two cases of the extended proof of cut, where an existential (or universal) formula is cut and was just introduced on the right and left, respectively.

Subcase:

$$\mathcal{D} = \cfrac{\begin{array}{cc} \mathcal{T} & \mathcal{D}_1 \\ \Sigma \vdash t : \tau & \Sigma; \Gamma \Longrightarrow A_1(t) \end{array}}{\Sigma; \Gamma \Longrightarrow \exists x{:}\tau.\ A_1(x)}\ \exists R$$

$$\text{and} \quad \mathcal{E} = \cfrac{\begin{array}{c} \mathcal{E}_1 \\ \Sigma, c{:}\tau; \Gamma, \exists x{:}\tau.\ A_1(x), A_1(c) \Longrightarrow C \end{array}}{\Sigma; \Gamma, \exists x{:}\tau.\ A_1(x) \Longrightarrow C}\ \exists L$$

$\Sigma; \Gamma, \exists x{:}\tau.\ A_1(x), A_1(t) \Longrightarrow C$	By substitution $[t/c]\mathcal{E}_1$ using $\mathcal{T}$
$\Sigma; \Gamma, A_1(t) \Longrightarrow C$	By i.h. on $\exists x.\ A_1(x)$, $\mathcal{D}$, and $[t/c]\mathcal{E}_1$
$\Sigma; \Gamma \Longrightarrow C$	By i.h. on $A_1(t)$, $\mathcal{D}_1$, and above

The induction requires that $A_1(t)$ is considered smaller than $\exists x.\ A_1(x)$. Formally, this can be justified by counting the number of quantifiers and logical connectives in a proposition and noting that the term t does not contain any. A similar remark applies to check that the proof $[t/c]\mathcal{E}_1$ is smaller than $\mathcal{E}$. Also note how the side condition that c must be a new parameter in the $\exists L$ rule is required in the substitution step to conclude that $[t/c]\Gamma = \Gamma$, $[t/c]A_1(c) = A(t) = [t/x]A_1(x)$, and $[t/c]C = C$.

Subcase:

$$\mathcal{D} = \cfrac{\begin{array}{c} \mathcal{D}_1 \\ \Sigma, c{:}\tau; \Gamma \Longrightarrow A_1(c) \end{array}}{\Sigma; \Gamma \Longrightarrow \forall x{:}\tau.\ A_1(x)}\ \forall R$$

$$\text{and} \quad \mathcal{E} = \cfrac{\begin{array}{cc} \mathcal{T} & \mathcal{E}_1 \\ \Sigma \vdash t : \tau & \Sigma; \Gamma, \forall x{:}\tau.\ A_1(x), A_1(t) \Longrightarrow C \end{array}}{\Sigma; \Gamma, \forall x{:}\tau.\ A_1(x) \Longrightarrow C}\ \forall L$$

$\Sigma; \Gamma, A_1(t) \Longrightarrow C$	By i.h. on $\forall x.\ A_1(x)$, $\mathcal{D}$, $\mathcal{E}_1$
$\Sigma; \Gamma \Longrightarrow A_1(t)$	By substitution $[t/c]\mathcal{D}_1$ using $\mathcal{T}$
$\Sigma; \Gamma \Longrightarrow C$	By i.h. on $A_1(t)$ and the above two

It is again important that $A_1(t)$ is considered smaller than $\forall x.\ A_1(x)$. Again note that side condition that c must be a new parameter in $\forall R$ is needed to ensure that the substitution leaves Γ unchanged and that $[t/c]A_1(c) = A_1(t)$.

References

[Gen35] Gerhard Gentzen. Untersuchungen über das logische Schließen. *Mathematische Zeitschrift*, 39:176–210, 405–431, 1935. English translation in M. E. Szabo, editor, *The Collected Papers of Gerhard Gentzen*, pages 68–131, North-Holland, 1969.

[Pfe00] Frank Pfenning. Structural cut elimination I. Intuitionistic and classical logic. *Information and Computation*, 157(1/2):84–141, March 2000.