Course on Proof Theory - Lecture 4

A proof of the cut-elimination theorem

Gianluca Curzi, Marianna Girlando

University of Birmingham

Midlands Graduate School
Nottingham, 10-14 April 2022

The rules of LK

$$\text{init}\ \frac{}{p, \Gamma \vdash \Delta, p} \qquad \text{cut}\ \frac{\Gamma \vdash \Delta, A \quad A, \Gamma \vdash \Delta}{\Gamma \vdash \Delta}$$

$$\neg_{\mathbf{L}}\ \frac{\Gamma \vdash \Delta, A}{\neg A, \Gamma \vdash \Delta} \qquad \neg_{\mathbf{R}}\ \frac{A, \Gamma \vdash \Delta}{\Gamma \vdash \Delta, \neg A} \qquad \wedge_{\mathbf{L}}\ \frac{A, B, \Gamma \vdash \Delta}{A \wedge B, \Gamma \vdash \Delta} \qquad \wedge_{\mathbf{R}}\ \frac{\Gamma \vdash \Delta, A \quad \Gamma \vdash \Delta, B}{\Gamma \vdash \Delta, A \wedge B}$$

$$\vee_{\mathbf{L}}\ \frac{A, \Gamma \vdash \Delta \quad B, \Gamma \vdash \Delta}{A \vee B, \Gamma \vdash \Delta} \qquad \vee_{\mathbf{R}}\ \frac{\Gamma \vdash \Delta, A, B}{\Gamma \vdash \Delta, A \vee B} \qquad \rightarrow_{\mathbf{L}}\ \frac{\Gamma \vdash \Delta, A \quad B, \Gamma \vdash \Delta}{A \rightarrow B, \Gamma \vdash \Delta} \qquad \rightarrow_{\mathbf{R}}\ \frac{A, \Gamma \vdash \Delta, B}{\Gamma \vdash \Delta, A \rightarrow B}$$

$$\forall_{\mathbf{L}}\ \frac{A[t/x], \forall x.A, \Gamma \vdash \Delta}{\forall x.A, \Gamma \vdash \Delta} \qquad \forall_{\mathbf{R}}\ \frac{\Gamma \vdash \Delta, A[y/x]}{\Gamma \vdash \Delta, \forall x.A} * \qquad \exists_{\mathbf{L}}\ \frac{A[y/x], \Gamma \vdash \Delta}{\exists x.A, \Gamma \vdash \Delta} * \qquad \exists_{\mathbf{R}}\ \frac{\Gamma \vdash \Delta, A[t/x], \exists x.A}{\Gamma \vdash \Delta, \exists x.A}$$

$*$ y does not occur free in Γ, Δ, A

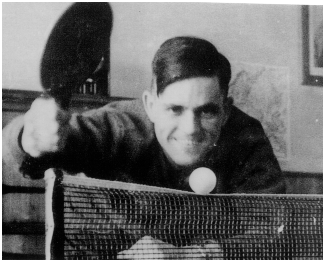
Derivation example

Is $(a \to b) \to ((a \to \bot) \vee b)$ a theorem of LK?

$$\text{init}\ \frac{}{p, \Gamma \vdash \Delta, p} \qquad \text{cut}\ \frac{\Gamma \vdash \Delta, {\color{red}A} \quad {\color{red}A}, \Gamma \vdash \Delta}{\Gamma \vdash \Delta}$$

$$\vee_{\mathbf{L}}\ \frac{A, \Gamma \vdash \Delta \quad B, \Gamma \vdash \Delta}{A \vee B, \Gamma \vdash \Delta} \qquad \vee_{\mathbf{R}}\ \frac{\Gamma \vdash \Delta, A, B}{\Gamma \vdash \Delta, A \vee B} \qquad \to_{\mathbf{L}}\ \frac{\Gamma \vdash \Delta, A \quad B, \Gamma \vdash \Delta}{A \to B, \Gamma \vdash \Delta} \qquad \to_{\mathbf{R}}\ \frac{A, \Gamma \vdash \Delta, B}{\Gamma \vdash \Delta, A \to B}$$

Today goal



Theorem (*Hauptsatz*, Gentzen 1934)

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

Corollary (Analyticity)

Every theorem of LK *has a proof that contains only subformulas of it (up to substitution of free variables).*

Informal example 1

$$\text{cut}\ \dfrac{\wedge_{\mathbf{R}}\ \dfrac{\overset{\mathcal{D}_1}{\vdots}\atop{\Gamma \vdash \Delta, A} \qquad \overset{\mathcal{D}_2}{\vdots}\atop{\Gamma \vdash \Delta, B}}{\Gamma \vdash \Delta, \color{red}{A \wedge B}} \qquad \wedge_{\mathbf{L}}\ \dfrac{\overset{\mathcal{D}_3}{\vdots}\atop{A, B, \Gamma \vdash \Delta}}{\color{red}{A \wedge B}, \Gamma \vdash \Delta}}{\Gamma \vdash \Delta}\ *$$

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

Informal example 2

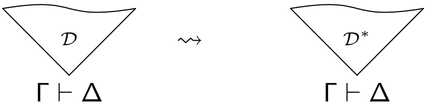
$$\text{cut}\;\dfrac{\forall_{\mathbf{R}}\,\dfrac{\begin{matrix}\mathcal{D}_1\\ \Gamma \vdash \Delta, B[x/y]\end{matrix}}{\Gamma \vdash \Delta, \color{red}{\forall x B}} \qquad \forall_{\mathbf{L}}\,\dfrac{\begin{matrix}\mathcal{D}_2\\ B[x/t], \forall x.B, \Gamma \vdash \Delta\end{matrix}}{\color{red}{\forall x.B}, \Gamma \vdash \Delta}}{\Gamma \vdash \Delta}\; *$$

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

$$\text{cut}\;\dfrac{\text{subst}\;\dfrac{\begin{matrix}\mathcal{D}_1\\ \Gamma \vdash \Delta, B[x/y]\end{matrix}}{\Gamma \vdash \Delta, \color{purple}{B[x/t]}} \qquad \text{cut}\;\dfrac{\text{wk}\;\dfrac{\forall_{\mathbf{R}}\,\dfrac{\begin{matrix}\mathcal{D}_1\\ \Gamma \vdash \Delta, B[x/y]\end{matrix}}{\Gamma \vdash \Delta, \forall x B}}{B[x/t], \Gamma \vdash \Delta, \color{red}{\forall x.B}} \qquad \begin{matrix}\mathcal{D}_2\\ B[x/t], \color{red}{\forall x.B}, \Gamma \vdash \Delta\end{matrix}}{\color{purple}{B[x/t]}, \Gamma \vdash \Delta}}{\Gamma \vdash \Delta}$$

General strategy of the proof

LK derivation cut-free LK derivation



- Apply the cut on smaller formulas, until they disappear
- Push the cuts upwards in the proof, and deal with them using IH
- 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 propositional LK (no quantifiers rules).

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*.
- ...

A measure of formulas

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

Inductive definition on the structure of the formula:

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

A measure of derivations

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

The rank of $\mathcal{D}$, $\mathrm{rk}(\mathcal{D})$, is the maximal degree of the cut formulas occurring in $\mathcal{D}$, plus 1.

We write

$$\Gamma \vdash^m_p \Delta$$

meaning

> *There is a derivation of $\Gamma \vdash \Delta$ of height* ***at most*** *m and rank* ***at most*** *p.*

A measure of derivations (more formally)

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

$$\mathcal{D} = \text{init}\,\frac{}{\Gamma \vdash \Delta} \qquad \mathrm{ht}(\mathcal{D}) = \mathrm{rk}(\mathcal{D}) = 0$$

$$\mathcal{D} = \text{R}\,\frac{\begin{array}{c}\mathcal{D}_1\\ \Gamma_1 \vdash \Delta_1\end{array}}{\Gamma \vdash \Delta} \qquad \mathrm{ht}(\mathcal{D}) = \mathrm{ht}(\mathcal{D}_1) + 1 \quad \mathrm{rk}(\mathcal{D}) = \mathrm{rk}(\mathcal{D}_1)$$

$$\mathcal{D} = \text{R}\,\frac{\begin{array}{c}\mathcal{D}_1\\ \Gamma_1 \vdash \Delta_1\end{array} \qquad \begin{array}{c}\mathcal{D}_2\\ \Gamma_2 \vdash \Delta_2\end{array}}{\Gamma \vdash \Delta} \qquad \begin{array}{l}\mathrm{ht}(\mathcal{D}) = \max(\mathrm{ht}(\mathcal{D}_1) + 1, \mathrm{ht}(\mathcal{D}_2) + 1)\\ \mathrm{rk}(\mathcal{D}) = \max(\mathrm{rk}(\mathcal{D}_1), \mathrm{rk}(\mathcal{D}_2))\end{array}$$

$$\mathcal{D} = \text{cut}\,\frac{\begin{array}{c}\mathcal{D}_1\\ \Gamma \vdash \Delta, A\end{array} \qquad \begin{array}{c}\mathcal{D}_2\\ A, \Gamma \vdash \Delta\end{array}}{\Gamma \vdash \Delta} \qquad \begin{array}{l}\mathrm{ht}(\mathcal{D}) = \max(\mathrm{ht}(\mathcal{D}_1) + 1, \mathrm{ht}(\mathcal{D}_2) + 1)\\ \mathrm{rk}(\mathcal{D}) = \max(\mathrm{rk}(\mathcal{D}_1), \mathrm{rk}(\mathcal{D}_2), \mathrm{deg}(\mathrm{A}) + 1)\end{array}$$

Some preliminary lemmas

1. Lemma: Closure under weakening

   ▷ If $\Gamma \vdash_p^m \Delta$, then $\Gamma', \Gamma \vdash_p^m \Delta, \Delta'$, for any Γ', Δ'.

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

2. Lemma: Invertibility All the rules are invertible:

   $(\wedge_L)$ If $A \wedge B, \Gamma \vdash_p^m \Delta$, then $A, B, \Gamma \vdash_p^m \Delta$.

   $(\wedge_R)$ If $\Gamma \vdash_p^m \Delta, A \wedge B$, then $\Gamma \vdash_p^m \Delta, A$ and $\Gamma \vdash_p^m \Delta, A$.

   (. . . and so on for all the rules)

   Proof. Induction on m, using closure under weakening.

3. Lemma: Closure under contraction

   ▷ If $A, A, \Gamma \vdash_p^m \Delta$, then $A, \Gamma \vdash_p^m \Delta, A$.

   ▷ If $\Gamma \vdash_p^m \Delta, A, A$, then $\Gamma \vdash_p^m \Delta, A$.

   Proof. Induction on m, using invertibility.

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

The plan

- Principal Lemma (most of the work)

  If $\Gamma \vdash_p^m \Delta, A$ and $A, \Gamma \vdash_p^n \Delta$ for $p = \mathrm{deg(A)}$, then $\Gamma \vdash_p^{m+n} \Delta$.

- Reduction Lemma (uses PrL)

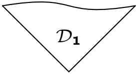
  If $\Gamma \vdash_{p+1}^m \Delta$, then $\Gamma \vdash_p^{2^m} \Delta$.

- Cut-elimination Theorem (uses RedL)

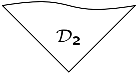
  If $\Gamma \vdash_p^m \Delta$ then $\Gamma \vdash_0^{2_p(m)} \Delta$.

Principal Lemma

Lemma Let $\Gamma \vdash_p^m \Delta, A$ and $A, \Gamma \vdash_p^n \Delta$ with $p = \mathrm{deg}(A)$:

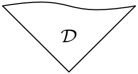


$\Gamma \vdash_p^m \Delta, A$



$A, \Gamma \vdash_p^n \Delta$

Then, we can construct a derivation $\Gamma \vdash_p^{m+n} \Delta$:



$\Gamma \vdash_p^{m+n} \Delta$

Proof of the Principal Lemma

$$\mathrm{R}_1\ \dfrac{\begin{array}{c}\mathcal{D}_1\\ \Gamma' \vdash_p^{m1} \Delta'\end{array}}{\Gamma \vdash_p^{m} \Delta, A} \qquad\qquad \mathrm{R}_2\ \dfrac{\begin{array}{c}\mathcal{D}_2\\ \Gamma'' \vdash_p^{n1} \Delta''\end{array}}{A, \Gamma \vdash_p^{n} \Delta}$$

Induction on $m + n$. We distinguish cases:

1. R_1 is init (R$_2$ is init, symmetric)
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, symmetric)

R_1 is init

$$\mathcal{D}_1 = \text{init}\ \frac{}{A, \Gamma' \vdash_p^m \Delta, A}$$

$$\begin{array}{c}\mathcal{D}_2 \\ \mathsf{R}_2\ \dfrac{\Gamma'' \vdash_p^{n-1} \Delta''}{A, \Gamma \vdash_p^n \Delta}\end{array}$$

with $\Gamma = A, \Gamma'$ We construct the following derivation $\mathcal{D}$ of $\Gamma \vdash_p^{m+n} \Delta$:

A is principal in both R_1 and R_2

R_1 is $\to_R$ and R_2 is $\to_L$

$$
\dfrac{\dfrac{\mathcal{D}_1}{B, \Gamma \vdash_p^{m-1} \Delta, C}}{\Gamma \vdash_p^{m} \Delta, B \to C}\ \to_R
\qquad\qquad
\dfrac{\dfrac{\mathcal{D}_2'}{\Gamma \vdash_p^{n1} \Delta, B} \qquad \dfrac{\mathcal{D}_2''}{C, \Gamma \vdash_p^{n2} \Delta}}{B \to C, \Gamma \vdash_p^{n} \Delta}\ \to_L
$$

with $n1, n2 < n$ and $n, m \geq 1$.

We construct the following derivation $\mathcal{D}$ of $\Gamma \vdash_p^{m+n} \Delta$:

A is not principal in R_1

R_1 is a one-premiss rule

$$\mathsf{R_1}\ \dfrac{\begin{array}{c}\mathcal{D}_\mathbf{1}\\ \Gamma' \vdash_p^{m-1} \Delta', A\end{array}}{\Gamma \vdash_p^{m} \Delta, A} \qquad\qquad \mathsf{R_2}\ \dfrac{\begin{array}{c}\mathcal{D}_\mathbf{2}\\ \Gamma'' \vdash_p^{n-1} \Delta''\end{array}}{A, \Gamma \vdash_p^{n} \Delta}$$

We construct the following derivation $\mathcal{D}$ of $\Gamma \vdash_p^{m+n} \Delta$:

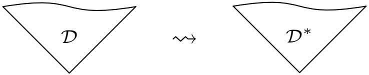
$$\mathsf{ctr}\ \dfrac{\mathsf{R_1}\ \dfrac{\mathsf{IH}\ \dfrac{\mathsf{wk}\ \dfrac{\begin{array}{c}\mathcal{D}_\mathbf{1}\\ \Gamma' \vdash_p^{m-1} \Delta', A\end{array}}{\Gamma', \Gamma \vdash_p^{m-1} \Delta, \Delta', A} \qquad \mathsf{wk}\ \dfrac{\mathsf{R_2}\ \dfrac{\begin{array}{c}\mathcal{D}_\mathbf{2}\\ \Gamma'' \vdash_p^{n-1} \Delta''\end{array}}{A, \Gamma \vdash_p^{n} \Delta}}{A, \Gamma', \Gamma \vdash_p^{n} \Delta, \Delta'}}{\Gamma', \Gamma \vdash_p^{(m-1)+n} \Delta, \Delta'}}{\Gamma, \Gamma \vdash_p^{m+n} \Delta, \Delta}}{\Gamma \vdash_p^{m+n} \Delta}$$

R_1 is a two-premisses rule ...

End of the proof of PrL $\square$

Reduction Lemma

Reduction Lemma If $\Gamma \vdash^{m}_{p+1} \Delta$, we can construct $\Gamma \vdash^{2^m}_{p} \Delta$.



$$\Gamma \vdash^{m}_{p+1} \Delta \qquad\qquad \Gamma \vdash^{2^m}_{p} \Delta$$

Proof. Induction on m

Base case Just set $\mathcal{D} = \mathcal{D}^*$

Induction step Case distinction according to the last rule R applied in $\mathcal{D}$.

We show just one case: R is cut.

Proof of the Reduction Lemma: key case

The last rule R applied in $\mathcal{D}$ is cut, with $\deg(A) = p$

$$\text{cut}\ \dfrac{\overset{\mathcal{D}_1}{\Gamma \vdash^{m1}_{p+1} \Delta, A} \qquad \overset{\mathcal{D}_2}{A, \Gamma \vdash^{m2}_{p+1} \Delta}}{\Gamma \vdash^{m}_{p+1} \Delta} \qquad \rightsquigarrow \qquad \overset{\mathcal{D}^*}{\Gamma \vdash^{2^m}_{p} \Delta}$$

We construct $\mathcal{D}^*$ as follows:

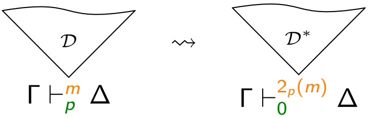
Since $m1, m2 < m$, by IH, we have:

$$\text{PrL}\ \dfrac{\overset{\mathcal{D}_1^*}{\Gamma \vdash^{2^{m1}}_{p} \Delta, A} \qquad \overset{\mathcal{D}_2^*}{A, \Gamma \vdash^{2^{m2}}_{p} \Delta}}{\Gamma \vdash^{2^{m1}+2^{m2}}_{p} \Delta}$$

We have constructed $\mathcal{D}^*$ of $\Gamma \vdash^{2^m}_{p} \Delta$. End of the proof of RedL $\square$

Cut-elimination Theorem

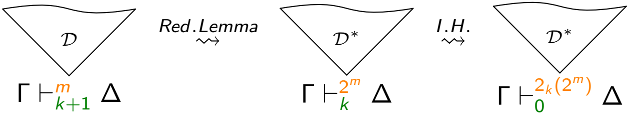
Cut-elimination Theorem If $\Gamma \vdash^{m}_{p} \Delta$, we can construct $\Gamma \vdash^{2_p(m)}_{0} \Delta$, that is, a derivation **where** cut **does not occur.**



Proof. Induction on p

Base case $\mathrm{p} = 0$ Just set $\mathcal{D} = \mathcal{D}^*$

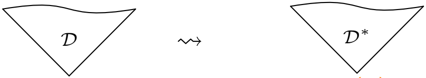
Induction step $\mathrm{p} = \mathrm{k}+1$



$\square$

Putting it all together

LK derivation $\qquad$ cut-free LK derivation



$$\Gamma \vdash_p^m \Delta \qquad\qquad \Gamma \vdash_0^{2_p(m)} \Delta$$

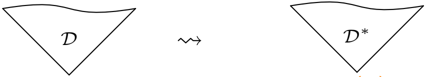
Proof sketch of the Cut-elimination Theorem

By induction on the rank od a proof:

- Identify the cuts of highest rank, say $p+1$, in the proof, and apply the Reduction Lemma to them.
- The Principal Lemma ensures that we might only introduce cuts of rank at most p in the process.
- Thus, the rank of the derivation decreases to p.
- We may conclude by the inductive hypothesis.

The cost of cut-elimination

LK derivation $\quad\quad$ cut-free LK derivation



$$\Gamma \vdash_{p}^{m} \Delta \quad \rightsquigarrow \quad \Gamma \vdash_{0}^{2_p(m)} \Delta$$

Eliminating cuts from a propositional LK derivation leads to a hyperexponential blow-up of the size of the proof.

Can we do better?

- For propositional LK: yes, we can get an exponential bound in proof size.
- For full LK (with quantifiers rules): no.

  Theorem (Statman '79, Orevkov '82). Cut-elimination for predicate logic necessarily has a non-elementary cost in proof size.

Cut-elimination for predicate logic

The method presented before can be extended to full LK, modulo assuming a renaming of variables and the following:

Substitution Lemma If $\Gamma \vdash^m_p \Delta$, then for each x variable and t term, $\Gamma[x/t] \vdash^m_p \Delta[x/t]$.

Proof. Easy induction on p.

The case of Principal Lemma in which the cut formula is $\forall x B$ and is principal in both subderivations is the following:

$$
\text{cut}\;\dfrac{
\text{subst}\;\dfrac{\dfrac{\mathcal{D}_1}{\Gamma \vdash \Delta, B[x/y]}}{\Gamma \vdash \Delta, B[x/t]}
\qquad
\text{IH}\;\dfrac{
\text{wk}\;\dfrac{\forall_R\;\dfrac{\dfrac{\mathcal{D}_1}{\Gamma \vdash \Delta, B[x/y]}}{\Gamma \vdash \Delta, \forall x B}}{B[x/t], \Gamma \vdash \Delta, \forall x.B}
\qquad
\dfrac{\mathcal{D}_2}{B[x/t], \forall x.B, \Gamma \vdash \Delta}
}{B[x/t], \Gamma \vdash \Delta}
}{\Gamma \vdash \Delta}
$$

Summing up: is it worth to eliminate cuts?

Drawbacks:

- Exponential or hyper-exponential blow-up of proof size w.r.t. input size
- Headache proof

Benefits:

- Analyticity: automated proof search
- Consistency
- Interpolation
- Herbrand's theorem
- ...

References

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

Exercises for Lecture 4

1. Prove that the following are derivable:
   - $(a \to b) \to ((a \wedge (b \to \bot)) \to \bot)$
   - $((a \wedge (b \to \bot)) \to \bot) \to ((a \to \bot) \vee b)$

2. Prove the following cases of the Principal Lemma:
   - The cut formula is $B \vee C$ and it is principal in both derivations:

$$\vee_{\mathbf{R}}\ \dfrac{\begin{array}{c}\mathcal{D}_1\\ \Gamma \vdash_p^{m-1} \Delta, B, C\end{array}}{\Gamma \vdash_p^{m} \Delta, B \vee C} \qquad \vee_{\mathbf{L}}\ \dfrac{\begin{array}{c}\mathcal{D}_2'\\ B, \Gamma \vdash_p^{n1} \Delta\end{array} \quad \begin{array}{c}\mathcal{D}_2''\\ C, \Gamma \vdash_p^{n2} \Delta\end{array}}{B \vee C, \Gamma \vdash_p^{n} \Delta}$$

- The cut formula is A and it is principal in the leftmost derivation, with $\deg(C) < p$:

$$\text{cut}\ \dfrac{\dfrac{\mathcal{D}'_1}{\Gamma \vdash_p^{m1} \Delta, A, C} \quad \dfrac{\mathcal{D}''_1}{C, \Gamma \vdash_p^{m2} \Delta, A}}{\Gamma \vdash_p^{m} \Delta, A} \qquad \mathrm{R}_1\ \dfrac{\dfrac{\mathcal{D}_2}{C, \Gamma' \vdash_p^{n-1} \Delta'}}{A, \Gamma \vdash_p^{n} \Delta}$$

- The cut formula is $\forall x B$ and it is principal in both R_1 and R_2

$$\forall_{\mathbf{R}}\ \dfrac{\dfrac{\mathcal{D}_1}{\Gamma \vdash \Delta, B[x/y]}}{\Gamma \vdash \Delta, \forall x B} \qquad \forall_{\mathbf{L}}\ \dfrac{\dfrac{\mathcal{D}_2}{B[x/t], \forall x.B, \Gamma \vdash \Delta}}{\forall x.B, \Gamma \vdash \Delta}$$