Proof Theory

Gaisi Takeuti

June 2, 2021

Contents

1 First Order Predicate Calculus

In this chapter we shall present Gentzen's formulation of the first order predicate calculus **LK** (logistischer klassischer Kalkül). Intuitionisitic logic is known as **LJ** (logistischer intuitionistischer Kalkül)

1.1 Formalization of statements

Definition 1.1. **Terms** are defined inductively as follows:

1. Every individual constant is a term

2. Every free variable is a term

3. If f^i is a function constant with i argument-places and $t_1, \dots, t_i$ are terms, then $f^i(t_1, \dots, t_i)$ is a term

4. Terms are only those expressions obtained by 1-3.

Definition 1.2. **Formulas** are defined inductively as:

3. If A is a formula, a is a free variable and x is a bound variable not occurring in A, then $\forall x A'$ and $\exists x A'$ are formulas, where A' is the expression obtained from A by writing x in place of a at each occurrence of a in A

Definition 1.3. Let A be an expression, let $\tau_1, \dots, \tau_n$ be distinct primitive symbols, and let $\sigma_1, \dots, \sigma_n$ be any symbols. By

$$\left(A \frac{\tau_1, \dots, \tau_n}{\sigma_1, \dots, \sigma_n} \right)$$

we mean the expression obtained from A by writing $\sigma_1, \dots, \sigma_n$ in place of $\tau_1, \dots, \tau_n$ respectively at each occurrence of $\tau_1, \dots, \tau_n$. Such an operation is called the **(simultaneous) replacement of** $(\tau_1, \dots, \tau_n)$ **by** $(\sigma_1, \dots, \sigma_n)$ **in** A.

Proposition 1.4. 1. *If A contains none of $\tau_1, \dots, \tau_n$, then*

$$\left(A \frac{\tau_1, \dots, \tau_n}{\sigma_1, \dots, \sigma_n} \right)$$

is A itself

2. *If $\sigma_1, \dots, \sigma_n$ are distinct primitive symbols, then*

$$\left(\left(A \frac{\tau_1, \dots, \tau_n}{\sigma_1, \dots, \sigma_n} \right) \frac{\sigma_1, \dots, \sigma_n}{\theta_1, \dots, \theta_n} \right)$$

is identical with

$$\left(A \frac{\tau_1, \dots, \tau_n}{\theta_1, \dots, \theta_n} \right)$$

Definition 1.5. 1. Let A be a formula and $t_1, \dots, t_n$ be terms. If there is a formula B and n distinct free variables $b_1, \dots, b_n$ s.t. A is

$$\left(B \frac{b_1, \dots, b_n}{t_1, \dots, t_n} \right)$$

then for each $i (1 \leq i \leq n)$ the occurrences of t_1 resulting from the above replacement are said to be **indicated** in A, and this fact is also expressed by writing B as $B(b_1, \dots, b_n)$ and A as $B(t_1, \dots, t_n)$

2. A term t is **fully indicated** in A, or every occurrence of t in A is indicated, if every occurrence of t is obtained by such a replacement

Proposition 1.6. *If A is a formula (where a is not necessarily fully indicated) and x is a bound variable not occurring in $A(a)$, then $\forall x A(x)$ and $\exists x A(x)$ are formulas*

1.2 Formal proofs and related concepts

Definition 1.7. An **inference** is an expression of the form

$$\frac{S_1}{S} \text{ or } \frac{S_1 \qquad S_2}{S}$$

where S_1, S_2 and S are sequents. S_1 and S_2 are called the **upper sequents** and S is called the **lower sequent** of the inference

Definition 1.8. For arbitrary Γ and Δ in the above notation, $\Gamma \to \Delta$ is called a **sequent**. Γ and Δ are called the **antecedent** and **succedent**, respectively, of the sequent and each formula in Γ and Δ is called a **sequent-formula**

Structural rules

1. **Weakening**:

$$\text{left: } \frac{\Gamma \to \Delta}{D, \Gamma \to \Delta}; \quad \text{right: } \frac{\Gamma \to \Delta}{\Gamma \to \Delta, D}$$

   D is called the **weakening formula**

2. **Contraction**:

$$\text{left: } \frac{D, D, \Gamma \to \Delta}{D, \Gamma \to \Delta} \quad \text{right: } \frac{\Gamma \to \Delta, D, D}{\Gamma \to \Delta, D}$$

3. **Exchange**

$$\text{left: } \frac{\Gamma, C, D, \Pi \to \Delta}{\Gamma, D, C, \Pi \to \Delta} \quad \text{right: } \frac{\Gamma \to \Delta, C, D, \Lambda}{\Gamma \to \Delta, D, C, \Lambda}$$

We will refer to these three kinds of inferences as "weak inferences", while all others will be called "strong inferences"

4. **Cut**

$$\frac{\Gamma \to \Delta, D \qquad D, \Pi \to \Lambda}{\Gamma, \Pi \to \Delta, \Lambda}$$

   D is called the **cut formula** of this instance

Logical rules

1.

$$\neg : \text{left: } \frac{\Gamma \to \Delta, D}{\neg D, \Gamma \to \Delta}; \quad \neg : \text{right: } \frac{D, \Gamma \to \Delta}{\Gamma \to \Delta, \neg D}$$

   D and $\neg D$ are called the **auxiliary formula** and the **principal formula** respectively, of this inference

2.

$$\frac{C, \Gamma \to \Delta}{C \wedge D, \Gamma \to \Delta} \wedge\text{left} \quad \text{and} \quad \frac{D, \Gamma \to \Delta}{C \wedge D, \Gamma \to \Delta} \wedge\text{left}$$

$$\frac{\Gamma \to \Delta, C \qquad \Gamma \to \Delta, D}{\Gamma \to \Delta, C \wedge D} \wedge\text{right}$$

C and D are called the auxiliary formulas and $C \wedge D$ is called the principal formula of this inference

3.

$$\frac{C, \Gamma \to \Delta \qquad D, \Gamma \to \Delta}{C \vee D, \Gamma \to \Delta} \vee\text{left}$$

$$\frac{\Gamma \to \Delta, C}{\Gamma \to \Delta, C \vee D} \vee\text{right} \quad \text{and} \quad \frac{\Gamma \to \Delta, D}{\Gamma \to \Delta, C \vee D} \vee\text{right}$$

C and D are called the auxiliary formulas and $C \vee D$ the principal formula of this inference

4.

$$\frac{\Gamma \to \Delta, C \qquad D, \Pi \to \Lambda}{C \supset D, \Gamma, \Pi \to \Delta, \Lambda} \supset\text{left} \qquad \frac{C, \Gamma \to \Delta, D}{\Gamma \to \Delta, C \supset D} \supset\text{right}$$

C and D are called the auxiliary formulas and $C \supset D$ the principal formula

1-4 are called **propositional inferences**

5.

$$\frac{F(t), \Gamma \to \Delta}{\forall x F(x), \Gamma \to \Delta} \forall\text{left} \qquad \frac{\Gamma \to \Delta, F(a)}{\Gamma \to \Delta, \forall x F(x)} \forall\text{right}$$

where t is an arbitrary term, and a does not occur in the lower sequent. $F(t)$ and $F(a)$ are called the auxiliary formulas and $\forall x F(x)$ the principal formula. The a in $\forall$right is called the **eigenvariable** of this inference

In $\forall$right all occurrences of a in $F(a)$ are indicated. In $\forall$left, $F(t)$ and $F(x)$ are

$$\left(F(a)\frac{a}{t}\right) \quad \text{and} \quad \left(F(a)\frac{a}{t}\right)$$

respectively, so not every t in $F(t)$ is necessarily indicated

6. $$\frac{F(a), \Gamma \to \Delta}{\exists x F(x), \Gamma \to \Delta}\ \exists\text{left} \qquad \frac{\Gamma \to \Delta, F(t)}{\Gamma \to \Delta, \exists x F(x)}\ \exists\text{right}$$

   where a does not occur in the lower sequent, and t is an arbitrary term

   $F(a)$ and Ft are called the auxiliary formulas and $\exists x F(x)$ the principal formula. The a in $\exists$left is called the eigenvariable of this inference

In $\exists$left a is fully indicated

5 and 6 are called the **quantifier inferences**. The condition, that the eigenvariable must not occur in the lower sequent in $\forall$right and $\exists$left is called the **eigenvariable condition**

A sequent of the form $A \to A$ is called an **initial sequent** or axiom

Definition 1.9. A **proof** P (in **LK**), or **LK-proof**, is a tree of sequents satisfying the following conditions

1. The topmost sequents of P are initial sequents

2. Every sequent in P except the lowest one is an upper sequent of an inference whose lower sequent is also in P

Definition 1.10.
1. A sequence of sequents in a proof P is called a **thread** (of P) if the following conditions are satisfied

   (a) The sequence begins with an initial sequent and ends with the end-sequent

   (b) Every sequent in the sequence except the last is an upper sequent of an inference, and is immediately followed by the lower sequent of this inference

2. Let $S_1, S_2$and S_3 be sequents in a proof P. We say S_1 is **above** S_2or S_2 is **below** S_1 if there is a thread containing both $S_1$and S_2 where S_1 appears before S_2. If $S_1$is above $S_2$and S_2 is above S_3, we say S_2 is **between** S_1 and S_3

3. An inference in P is said to be **below a sequent** S if its lower sequent is below S

4. Let P be a proof. A part of P which itself is a proof is called a **subproof** of P. For any sequent S in P, that part of P which consists of all sequents which are either Sitself or which occur above $S$is called a subproof of P (with end-sequent S)

5. Let P_0 be a proof of the form

$$\begin{array}{c} \vdots \\ \Gamma \to \Theta \\ \vdots\,(^*) \end{array}$$

where (*) denotes the part of P_0 under $\Gamma \to \Theta$, and let Q be a proof ending with $\Gamma, D \to \Theta$. By a copy of P_0 from Q we mean a proof P of the form

$$\begin{array}{c} \vdots\,Q \\ \Gamma, D \to \Theta \\ \vdots\,(^{**}) \end{array}$$

where $(**)$ differs from $(*)$ only in that for each sequent in $(*)$, say $\Gamma \to \Lambda$, the corresponding sequent in $(**)$ has the form $\Pi, D \to \Lambda$.

6. Let $S(a)$ or $\Gamma(a) \to \Delta(a)$, denote a sequent of the form $A_1(a), \dots, A_m(a) \to B_1(a), \dots, B_n(a)$. Then $S(t)$, or $\Gamma(t) \to \Delta(t)$, denotes the sequent $A_1(t), \dots, A_m(t) \to B_1(t), \dots, B_n(t)$

Definition 1.11. A proof in **LK** is called **regular** if it satisfies the condition that all eigenvariables are distinct from one another and if a free variable a occurs as an eigenvariable in a sequent S of the proof, then a occurs only in sequents above S

Lemma 1.12. *1. Let $\Gamma(a) \to \Delta(a)$ be an (**LK**-)provable sequent in which a is fully indicated, and let $P(a)$ be a proof of $\Gamma(a) \to \Delta(a)$. Let b be a free variable not occurring in $P(a)$. Then the tree $P(b)$, obtained from $P(a)$ by replacing a by b at each occurrence of a in $P(a)$, is also a proof and its end-sequent is $\Gamma(b) \to \Delta(b)$*

*2. For an arbitrary **LK**-proof there exists a regular proof of the same end-sequent. Moreover, the required proof is obtained from the original proof simply by replacing free variables*

Proof. 1. By induction on the number of inference in $P(a)$. If $P(a)$ consists of simply an initial sequent $A(a) \to A(a)$, then $P(b)$ consists of the sequent $A(b) \to A(b)$.

Suppose that our proposition holds for proofs containing at most n inferences and suppose that $P(a)$ contains $n + 1$ inferences. We treat the possible cases according to the last inferences in $P(a)$. Since other cases can be treated similarly, we consider only the case where the last inference, say J, is a $\forall$right. Suppose the eigenvariable of J is a, and $P(a)$ is of the form

$$\dfrac{\begin{array}{c}\vdots\, Q(a)\\ \Gamma \to \Lambda, A(a)\end{array}}{\Gamma \to \Lambda, \forall x A(x)}\, J$$

where $Q(a)$ is the subproof of $P(a)$ ending with $\Gamma \to \Lambda, A(a)$. a doesnt occur in Γ, Λ or $A(x)$. By the induction hypotheses the result of replacing all a's in $Q(a)$ by b is a proof whose end-sequent is $\Gamma \to \Lambda, A(b)$. Γ and Λ contain no b's. Thus we can apply a $\forall$right to this sequent using b as its eigenvariable

$$\dfrac{\begin{array}{c}\vdots\, Q(b)\\ \Gamma \to \Lambda, A(b)\end{array}}{\Gamma \to \Lambda, \forall x A(x)}$$

and so $P(b)$ is a proof ending with $\Gamma \to \Lambda, \forall x A(x)$. If a is not the eigenvariable of J, $P(a)$ is of the form

$$\dfrac{\begin{array}{c}\vdots\, Q(a)\\ \Gamma(a) \to \Lambda(a), A(a,c)\end{array}}{\Gamma(a) \to \Lambda(a), \forall x A(a,x)}$$

By the induction hypothesis the result of replacing all a's in $Q(a)$ by $b$is a proof and its end-sequent is $\Gamma(b) \to \Lambda(b), A(b,c)$

Since by assumption b doesn't occur in $P(a)$, b is not c and so we can apply a $\forall$right to this sequent, with c as its eigenvariable

2. By mathematical induction on the number l of applications of $\forall$right and $\exists$left in a given proof P. If $l = 0$ then take P itself. Otherwise, P can be represented in the form

$$\begin{array}{c} P_1 \quad P_2 \ldots P_k \\ \vdots\,(*) \\ S \end{array}$$

where P_i is a subproof of P of the form

$$\dfrac{\begin{array}{c}\vdots\\ \Gamma_i \to \Delta_i, F_i(b_i)\end{array}}{\Gamma_i \to \Delta_i, \forall y_i F_i(y_i)}\ I_i \qquad \text{or} \qquad \dfrac{\begin{array}{c}\vdots\\ F_i(b_i), \Gamma_i \to \Delta_i\end{array}}{\exists y_i F_i(y_i), \Gamma_i \to \Delta_i}\ I_i$$

and I_i is a lowermost $\forall$right or $\exists$left in P

Let us deal with the case where I_i is $\forall$right. P_i has fewer applications of $\forall$right or $\exists$left than P, so by the induction hypothesis there is a regular proof P_i' of $\Gamma_i \to \Delta_i, F_i(b_i)$. Note that no free variable in $\Gamma_i \to \Delta_i, F_i(b_i)$ (including b_i) is used as an eigenvariable in P_i'. Suppose $c_1, \ldots, c_m$ are all the eigenvariables in all the P_i's which occur in P above $\Gamma_i \to \Delta_i, \forall y_i F_i(y_i), i = 1, \ldots, k$. Then change $c_1, \ldots, c_m$ to $d_1, \ldots, d_m$ respectively, where $d_1, \ldots, d_m$ are the first m variables which occur neither in P nor in P_i'. If b_i occurs in P below $\Gamma_i \to \Delta_i, \forall y_i F_i(y_i)$ then change it to d_{m+i}

Let P_i'' be the proof which is obtained from P_i' by the above replacement of varaibles. Then $P_1'', \ldots, P_k''$ are each regular

$$\begin{array}{c} P_1'' \ldots \dfrac{P_i''}{\Gamma_i \to \Delta_i, \forall y_i F_i(y_i)} \ldots P_n'' \\ \vdots\,(*) \\ S \end{array}$$

□

From now on we will assume that we are dealing with regular proofs whenever convenient

Lemma 1.13. *Let t be an arbitrary term. Let $\Gamma(a) \to \Delta(a)$ be a provable (in **LK**) sequent in which a is fully indicated, and let $P(a)$ be a proof ending with $\Gamma(a) \to \Delta(a)$ in which **every eigenvariable is different from** a **and not contained in** t. Then $P(t)$ is a proof whose end-sequent is $\Gamma(t) \to \Delta(t)$*

Lemma 1.14. *Let t be an arbitrary term. Let $\Gamma(a) \to \Delta(a)$ be a provable (in **LK**) sequent in which a is fully indicated, and let $P(a)$ be a proof of $\Gamma(a) \to \Delta(a)$. Let $P'(a)$ be a proof obtained from $P(a)$ by changing eigenvariables in such a way that in $P'(a)$ every eigenvariable is different from a and not contained in t. Then $P'(t)$ is a proof of $\Gamma(t) \to \Delta(t)$*

Proposition 1.15. *Let t be an arbitrary term and $S(a)$ a provable sequent in which a is fully indicated. Then $S(t)$ is also provable*

Proposition 1.16. *If a sequent is provable, then it is provable with a proof in which all the initial sequents consist of atmoic formulas. Furthermore, if a sequent is provable without cut, then it is provable without cut with a proof of the above sort*

Proof. It suffices to show that for an arbitrary formula A, $A \to A$ is provable without cut, starting with initial sequents consisting of atomic formulas. $\square$

Definition 1.17. Two formulas A and B are **alphabetical variants** if for some $x_1, \dots, x_n, y_1, \dots, y_n$

$$\left(A\frac{x_1, \dots, x_n}{z_1, \dots, z_n}\right)$$

is

$$\left(B\frac{y_1, \dots, y_n}{z_1, \dots, z_n}\right)$$

where $z_1, \dots, z_n$ are bound variables occurring neither in A nor in B. The fact that A and B are alphabetical variants will be expressed by $A \sim B$

1.3 A formulation of intuitionistic predicate calculus

Definition 1.18. We can formalize the intuitionistic predicate calculus as a subsystem of **LK** which we call **LJ** following Gentzen (**J** stands for "intuitionistic"). **LJ**is obtained from **LK** by modifying it as follows

1. A sequent in **LJ** is of the form $\Gamma \to \Delta$ where Δ consists of at most one formula

2. Inferences in **LJ** are those obtained from those in **LK** by imposing the restriction that the succedent of each upper and lower sequent consists of at most one formula; thus there are no inferences in **LJ**corresponding to contraction right or exchange right

Proposition 1.19. *If a sequent S of **LJ** is provable in **LJ**, then it is also provable in **LK***

1.4 Axiom systems

Definition 1.20. The basic system is **LK**

1. A finite or infinite set $\mathcal{A}$ of sentences is called an **axiom system**, and each of these sentences is called an **axiom** of $\mathcal{A}$. Sometimes an axiom system is called a **theory**

2. A finite (possibly empty) sequence of formulas consisting only of axioms of $\mathcal{A}$ is called an **axiom sequence** of $\mathcal{A}$

3. If there exists an axiom sequence Γ_0 of $\mathcal{A}$ s.t. $\Gamma_0, \Gamma \to \Delta$ is **LK**-provable, then $\Gamma \to \Delta$ is said to be **provable from** $\mathcal{A}$ (in **LK**). We express this by $\mathcal{A}, \Gamma \to \Delta$

4. $\mathcal{A}$ is **inconsistent** (with **LK**) if the empty sequent $\to$ is provable from $\mathcal{A}$ (in **LK**)

5. If all function constants and predicate constants in a formula A occur in $\mathcal{A}$, then A is said to be **dependent on** $\mathcal{A}$

6. A sentence A is **consistent** if the axiom system $\{A\}$ is consistent

7. $\mathbf{LK}_{\mathcal{A}}$ is the system obtained from **LK** by adding $\to A$ as initial sequents for all A in $\mathcal{A}$

Proposition 1.21. *Let $\mathcal{A}$ be an axiom system. Then the following are equivalent*

1. *$\mathcal{A}$ is inconsistent (with **LK**)*

2. *for every formula A, A is provable from $\mathcal{A}$*

3. *for some formula A, $A$and $\neg A$ are both provable from $\mathcal{A}$*

Proof. $1 \leftrightarrow 2$, $2 \leftrightarrow 3$. $\to A \vee B$ and $\to \neg A \vee B$ implies $\to B$ $\square$

Proposition 1.22. *Let $\mathcal{A}$ be an axiom system. Then a sequent $\Gamma \to \Delta$ is $\mathbf{LK}_{\mathcal{A}}$-provable iff $\Gamma \to \Delta$ is provable from $\mathcal{A}$ (in **LK**)*

Corollary 1.23. *An axiom system $\mathcal{A}$ is consistent (with **LK**) iff $\mathbf{LK}_{\mathcal{A}}$ is consistent*

These definitions and the propositions hold also for **LJ**

1.5 The cut-elimination theroem

Theorem 1.24 (the cut-elimination theroem: Gentzen)**.** *If a sequent is* (**LK**)*-provable, then it is* (**LK**)*-provable without a cut*

Let Abe a formula. An inference of the following form is called a **mix** (w.r.t. A):

$$\frac{\Gamma \to \Delta \qquad \Pi \to \Lambda}{\Gamma, \Pi^* \to \Delta^*, \Lambda}\ A$$

where both Δ and Π contain the formula A, and Δ^* and Π^* are obtained from Δ and Π respectively by deleting all the occurrences of A in them. We call A the mix formula of this inference.

Let's call the system which is obtained from **LK** by replacing the cut rule by the mix rule, **LK***.

Lemma 1.25. *__LK__ and __LK__* are equivalent, that is, a sequent S is __LK__-provable iff S is __LK__*-provable*

Theorem 1.26. *If a sequent is provable in __LK__*, then it's provable in __LK__* without a mix*

Lemma 1.27. *If P is a proof of S (in __LK__*) which contains (only) one mix, occurring as the last inference, then S is provable without a mix*

The **grade** of a formula A (denoted by $g(A)$) is the number of logical symbols contained in A. The grade of a mix is the grade of the mix formula. When a proof P has a mix as the last inference, we define the grade of P (denoted by $g(P)$) to be the grade of this mix.

Let P be a proof which contains a mix only as the last inference

$$J\ \frac{\Gamma \to \Delta \qquad \Pi \to \Lambda}{\Gamma, \Pi^* \to \Delta^*, \Lambda}\ (A)$$

We refer to the left and right upper sequents as S_1 and S_2 and to the lower sequent as S. We call a thread in P a **left (right) thread** if it contains the left (right) upper sequent of the mix J. The **rank** of a thread $\mathcal{F}$ in P is defined as follows: if $\mathcal{F}$ is a left (right) thread, then the rank of $\mathcal{F}$ is the number consecutive sequents, counting upward from the left (right) upper sequent of J, that contains the mix formula in its succedent (antecedent). The rank of a thread $\mathcal{F}$ in P is denoted by $\mathrm{rank}(\mathcal{F}; P)$. We define

$$\mathrm{rank}_l(P) = \max_{\mathcal{F}}(\mathrm{rank}(\mathcal{F}; P))$$

where $\mathcal{F}$ ranges over all the left threads in P, and

$$\mathrm{rank}_r(P) = \max_{\mathcal{F}}(\mathrm{rank}(\mathcal{F}; P))$$

where $\mathcal{F}$ ranges over all the right threads in P. The rank of P, rank(P), is defined as

$$\text{rank}(P) = \text{rank}_l(P) + \text{rank}_r(P)$$

Note that rank$(P) \geq 2$

Proof. We prove the Lemma by double induction on the grade g and rank r of the proof P (i.e. transfinite induction on $\omega \cdot g + r$). We divide the proof into two main cases, namely $r = 2$ and $r > 2$

1. $r = 2$, $\text{rank}_l(P) = \text{rank}_r(P) = 1$

   (a) The left upper sequent S_1 is an initial sequent. In this case we may assume P is of the form

   $$J\,\frac{A \to A \qquad \Pi \to \Lambda}{A, \Pi^* \to \Lambda}$$

   We can obtain the lower sequent without a mix

   $$\cfrac{\cfrac{\cfrac{\Pi \to \Lambda}{\text{some exchanges}}}{A, \dots, A, \Pi^* \to \Lambda}}{\cfrac{\text{some contractions}}{A, \Pi^* \to \Lambda}}$$

   (b) The right upper sequent S_2 is an initial sequent.

   (c) Neither S_1 nor S_2 is an initial sequent, and S_1 is the lower sequent of a structural inference J_1. Since $\text{rank}_l(P) = 1$, the formula A cannot appear in the succedent of the upper sequent of J_1. Hence

   $$\frac{\dfrac{\Gamma \to \Delta_1}{\Gamma \to \Delta_1, A}\,J_1 \qquad \Pi \to \Lambda}{\Gamma, \Pi^* \to \Delta_1, \Lambda}\,J$$

   where Δ_1 doesn't contain A. We can eliminate the mix as follows

   $$\cfrac{\cfrac{\cfrac{\Gamma \to \Delta_1}{\text{some weakenings}}}{\Pi^*, \Gamma \to \Delta_1, \Lambda}}{\cfrac{\text{some exchanges}}{\Gamma, \Pi^* \to \Delta_1, \Lambda}}$$

(d) None of 1.1-1.3 holds but S_2 is the lower sequent of a structural inference. Similarly

(e) Both $S_1$and $S_2$are the lower sequents of logical inferences. In this case, since $\text{rank}_l(P) = \text{rank}_r(P) = 1$, the mix formula on each side must be the principal formula of the logical inference. We use induction on the grade, distinguishing several cases according to the outermost logical symbol of A

i. The outermost logical symbol of A is $\wedge$

$$\dfrac{\dfrac{\Gamma \to \Delta_1, B \qquad \Gamma \to \Delta_1, C}{\Gamma \to \Delta_1, B \wedge C} \qquad \dfrac{B, \Pi_1 \to \Lambda}{B \wedge C, \Pi_1 \to \Lambda}}{\Gamma, \Pi_1 \to \Delta_1, \Lambda}\ (B \wedge C)$$

where by assumption none of the proofs ending with $\Gamma \to \Delta_1, B$;$\Gamma \to \Delta_1, C$ or $B, \Pi_1 \to \Lambda$ contain a mix. Consider the following

$$\dfrac{\Gamma \to \Delta_1, B \qquad B, \Pi_1 \to \Lambda}{\Gamma, \Pi_1'' \to \Delta_1'', \Lambda}\ (B)$$

This proof contains only one mix, a mix that occurs as its last inference. Furthermore the grade of the mix formula B is less than $g(A)$. So by induction hypothesis we can obtain a proof which contains no mixes and whose end-sequent is $\Gamma, \Pi_1'' \to \Delta_1'', \Lambda$. From this we can obtain a proof without a mix with end-sequent $\Gamma, \Pi_1 \to \Delta_1, \Lambda$

ii. The outermost logical symbol of A is $\forall$

$$\dfrac{\dfrac{\Gamma \to \Delta_1, F(a)}{\Gamma \to \Delta_1, \forall x F(x)} \qquad \dfrac{F(t), \Pi_1 \to \Lambda}{\forall x F(x), \Pi_1 \to \Lambda}}{\Gamma, \Pi_1 \to \Delta_1, \Lambda}$$

(a being fully indicated in $F(a)$). By the eigenvariable condition, a does not occur in Γ, Δ_1 or $F(x)$. Since by assumption the proof ending with $\Gamma \to \Delta_1, F(a)$ contains no mix, we can obtain a proof without a mix, ending with $\Gamma \to \Delta_1, F(t)$. Consider now

$$\dfrac{\Gamma \to \Delta_1, F(t) \qquad F(t), \Pi_1 \to \Lambda}{\Gamma, \Pi_1''' \to \Delta_1''', \Lambda}\ (F(t))$$

2. $r > 2$, i.e., $\text{rank}_l(P) > 1$ and/or $\text{rank}_r(P) > 1$

   The induction hypothesis is that every proof Q which contains a mix only as the last inference, and which satisfies either $g(Q) < g(P)$, or $g(Q) = g(P)$ and $\text{rank}(Q) < \text{rank}(P)$, we can eliminate the mix

   (a) $\text{rank}_r(P) > 1$

      i. Γ or Δ (in S_1) contains A. Construct a proof as follows

$$\dfrac{\dfrac{\dfrac{\dfrac{\vdots}{\Pi \to \Lambda}}{\text{exchanges/contractions}}}{A, \Pi^* \to \Lambda}}{\dfrac{\text{weakenings/exchanges}}{\Gamma, \Pi^* \to \Delta^*, \Lambda}} \qquad \dfrac{\dfrac{\dfrac{\dfrac{\vdots}{\Gamma \to \Delta}}{\text{exchanges/contractions}}}{\Gamma \to \Delta^*, A}}{\dfrac{\text{weakenings/exchanges}}{\Gamma, \Pi^* \to \Delta^*, \Lambda}}$$

      ii. S_2 is the lower sequent of an inference J_2, where J_2 is not a logical inference whose principal formula is A. The last part of P looks like this

$$\dfrac{\Gamma \to \Delta \qquad \dfrac{\Phi \to \Psi}{\Pi \to \Lambda}\, J_2}{\Gamma, \Pi^* \to \Delta^*, \Lambda}$$

      where the proofs $\Gamma \to \Delta$ and $\Phi \to \Psi$ contain no mixes and Φ contains at least one A. Consider the following proof P':

$$\dfrac{\Gamma \to \Delta \qquad \Phi \to \Psi}{\Gamma, \Phi^* \to \Delta^*, \Psi}\,(A)$$

      In P', the grade of the mix is equal to $g(P)$, $\text{rank}_l(P') = \text{rank}_l(P)$ and $\text{rank}_r(P') = \text{rank}_r(P) - 1$. Thus by induction hypothesis, $\Gamma, \Phi^* \to \Delta^*, \Psi$ is provable without a mix. Then we construct the proof

$$\dfrac{\dfrac{\dfrac{\Gamma, \Phi^* \to \Delta^*, \Psi}{\text{some exchanges}}}{\Phi^*, \Gamma \to \Delta^*, \Psi}}{\Pi^*, \Gamma \to \Delta^*, \Lambda}\, J_2$$

      iii. Γ contains no A's and S_2 is the lower sequent of a logical inference whose principal formula is A.

□

Theorem 1.28. *The cut-elimination theorem holds for **LJ***

1.6 Some consequences of the cut-elimination theorem

Definition 1.29. By a **subformula** of a formula A we mean a formula used in building up A.

Two formulas A and B are said to be **equivalent** in **LK**if $A \equiv B$ is provable in **LK**

In a formula A an occurrence of a logical symbol, say $\sharp$ is **in the scope** of an occurrences of a logical symbol, say $\natural$, if in the construction of A (from atomic formulas) the stage where $\sharp$ is the outermost logical symbol precedes the stage where $\natural$ is the outermost logical symbol. Further, a symbol $\sharp$ is said to be in the left scope of a $\supset$ if $\supset$ occurs in the form $B \supset C$ and $\sharp$ occurs in B

A formula is called **prenex** (in prenex form) if no quantifier in it is in the scope of a propositional connective.

A proof without a cut contains only subformulas of the formulas occurring in the end-sequent. A formula is provable iff it is provable by use of its subformulas only

Theorem 1.30 (consistency)**.** ***LK*** *and* ***LJ*** *are consistent*

Proof. Suppose $\rightarrow$ were provable in **LK**. Then by the cut-elimination theorem, it would be provable in **LK** without a cut. But this is impossible, by the subformula property of cut-free proofs □

Theorem 1.31. *In a cut-free proof in* ***LK*** *(or* ***LJ****) all the formulas which occur in it are subformulas of the formulas in the end-sequent*

Theorem 1.32 (Gentzen's midsequent theorem for **LK**)**.** *Let S be a sequent which consists of prenex formulas only and is provable in* ***LK****. Then there is a cut-free proof of S which contains a sequent (called a* ***midsequent****), say S', which satisfies the following*

1. *S' is quantifier-free*
2. *Every inference above S' is either structural or propositional*
3. *Every inference below S' is either structural or a quantifier inference*

Thus a midsequent splits the proof into an upper part, which contains the propositional inferences, and a lower part, which contains the quantifier inferences.

*The above holds reading "****LJ*** *without $\vee$left" in place of* ***LK*** *outline.* Combining Proposition 1.16 and the cut-elimination theorem we may assume that there is a cut-free proof of S, say P, in which all the initial sequents consist of atmoic formulas only ($_{\text{why}}$ do we need atomic formula_). Let I be a quantifier inference in P. The number of propositional inference under I is called the order of I. The sum of orders for all the quantifier inferences in $P$is called the order of P. The proof is carried out by induction on the order of P.

Case 1: The order of a proof P is 0. If there is a propositional inference, take the lowermost such, and call its lower sequent S_0. Above this sequent there is no quantifier inference. Therefore if there is a quantifier in or above S_0, then it is introduced by weakening. Since the proof is cut-free, the weakening formula is a subformula of one of the formulas in the end-sequent. Hence no propositional inferences apply to it. ($_{\text{don}}$'t understand_) We can thus eliminate these weakenings and obtain a sequent S'_0 corresponding to S_0. By adding some weakenings under S'_0 we derive S and S'_0 serves as the mid-sequent

If there is no propositional inference in P, then take the uppermost quantifier inferences. Its upper sequent serves as a midsequent

Case 2: The order of P is not 0. Then there is at least one propositional inference which is below a quantifier property. Moreover, there is a quantifier inference I with the following property: the uppermost logical inference under I is a propositional inference. Call it I'. We can lower the order by interchanging the positions of I and I'. Say I is $\forall$right, then proof P is

$$\dfrac{\dfrac{\vdots}{\Gamma \to \Theta, F(a)}}{\Gamma \to \Theta, \forall x F(x)}\, I$$

$$\vdots\,(*)$$

$$\dfrac{}{\Delta \to \Lambda}\, I'$$

where the (*)-part of P contains only structural inferences and Λ contains

$\forall x F(x)$ as a sequent-formula. Transform P into the following proof P':

$$\dfrac{\dfrac{\dfrac{\begin{array}{c}\Gamma \to \Theta, F(a) \\ \vdots \text{ structural inferences} \\ \Gamma \to F(a), \Theta, \forall x F(x) \\ \vdots \end{array}}{\Delta \to F(a), \Lambda}\, I'}{\Delta, \Lambda, \forall x F(x)}\, I}{\begin{array}{c}\Delta \to \Lambda \\ \vdots\end{array}}$$

It is obvious that the order of P' is less than that of P $\square$

For technical reasons we introduce the predicate symbol $\top$ with 0 argument places, and admit $\to \top$ as an additional initial sequent. The system which is obtained from **LK** thus extended is denoted by **LK#**

Lemma 1.33. *Let $\Gamma \to \Delta$ be **LK**-provable, and let (Γ_1, Γ_2) and Δ_1, Δ_2 be arbitrary partitions of Γ and Δ, respectively (including the cases that one or more of $\Gamma_1, \Gamma_2, \Delta_1, \Delta_2$ are empty). We denote such a partition by $[\{\Gamma_1; \Delta_1\}, \{\Gamma_2; \Delta_2\}]$ and call it a partition of the sequent $\Gamma \to \Delta$. Then there exists a formula C of **LK#** (called an **interpolant** of $[\{\Gamma_1; \Delta_1\}, \{\Gamma_2; \Delta_2\}]$) s.t.*

1. *$\Gamma_1 \to \Delta_1, C$ and $C, \Gamma_2 \to \Delta_2$ are both **LK#**-provable*
2. *All free variables and individual and predicate constants in C (apart from $\top$) occur both in $\Gamma_1 \cup \Delta_1$ and $\Gamma_2 \cup \Delta_2$*

Theorem 1.34 (Craig's interpolation theorem for **LK**). 1. *Let A and B be two formulas s.t. $A \supset B$ is **LK**-provable. If A and B have at least one predicate constant in common, then there exists a formula C, called an interpolant of $A \supset B$ s.t. C contains only those individual constants, predicate constants and free variables that occur in both A and B and s.t. $A \supset C$ and $C \supset B$ are **LK**-provable. If A and B contain no predicate constant in common, then either $A \to$ or $\to B$ is **LK**-provable*

2. *As above, with **LJ** inplace of **LK***

Proof. Assume that $A \supset B$, and hence $A \to B$ is provable, and A and B have at least one predicate constant in common. Then by Lemma 1.33, taking A

as Γ_1 and B as Δ_2 (with Γ_2 and Δ_1 empty), there exists a formula Csatisfying 1 and 2. So $A \to C$ and $C \to B$ are **LK#**-provable. Let R be predicate constant which is common to A and B and has k argument places. Let R' be $\forall y_1 \ldots \forall y_k R(y_1, \ldots, y_k)$, where $y_1, \ldots, y_k$ are new bound variables. By replacing $\top$ by $R' \supset R'$ we can transform C into a formula C' of the original language, s.t. $A \to C'$ and $C' \to B$ are **LK**-provable. C' is then the desired interpolant.

If there is no predicate common to $\Gamma_1 \cup \Delta_1$ and $\Gamma_2 \cup \Delta_2$ in the partition, then by Lemma 1.33 there is a C s.t. $\Gamma_1 \to \Delta_1, C$ and $C, \Gamma_2 \to \Delta_2$ are provable, and C consists of $\top$ and logical symbols only. Then it can easily be shown, by induction on the complexity of C, that either $\to C$ or $C \to$ is provable. Hence either $\Gamma_1 \to \Delta_1$ or $\Gamma_2 \to \Delta_2$ is provable. $\square$

*Lemma [**?**].* The lemma is proved by induction on the number of inferences k, in a cut-free proof of $\Gamma \to \Delta$. At each stage there are several cases to consider; we deal with some examples only.

1. $k = 0, \Gamma \to \Delta$ has the form $D \to D$. There are four cases: 1. $[\{D; D\}, \{;\}]$, 2. $[\{;\}, \{D; D\}]$, 3. $[\{D;\}, \{; D\}]$, 4. $\{; D\}, \{D;\}$. Take for $C : \neg\top$ in 1, $\top$ in 2, D in 3 and $\neg D$ in 4

2. $k > 0$ and the last inference is $\wedge$right:

$$\frac{\Gamma \to \Delta, A \qquad \Gamma \to \Delta, B}{\Gamma \to \Delta, A \wedge B}$$

   Suppose the partition is $[\{\Gamma_1; \Delta_1, A \wedge B\}, \{\Gamma_2; \Delta_2\}]$. Consider the induced partition of the upper sequents, viz $[\{\Gamma_1; \Delta_1, A\}, \{\Gamma_2; \Delta_2\}]$ and $[\{\Gamma_1; \Delta_1, B\}, \{\Gamma_2; \Delta_2\}]$ respectively. By the induction hypothesis applied to the subproofs of the upper sequents, there exists interpolants C_1 and C_2 so that $\Gamma_1 \to \Delta_1, A, C_1; C_1, \Gamma_2 \to \Delta_2; \Gamma_1 \to \Delta_1, B, C_2$ and $C_2, \Gamma_2 \to \Delta_2$ are all **LK#**-provable. From these sequents, $\Gamma_1 \to \Delta_1, A \wedge B, C_1 \vee C_2$ and $C_1 \vee C_2, \Gamma_2 \to \Delta_2$

3. $k > 0$ and the last inference is $\forall$left

$$\frac{F(s), \Gamma \to \Delta}{\forall x F(x), \Gamma \to \Delta}$$

   Suppose $b_1, \ldots, b_n$ are all the free variables and constants which occur in s. Suppose the partition is $[\{\forall x F(x), \Gamma_1; \Delta_1\}, \{\Gamma_2; \Delta_2\}]$. Consider the induced partition of the upper sequent and apply the induction hypothesis. So there exists and interpolant $C(b_1, \dots, b_n)$ so that

$$F(s), \Gamma_1 \to \Delta_1, C(b_1, \dots, b_n)$$
$$C(b_1, \dots, b_n), \Gamma_2 \to \Delta_2$$

are **LK#**-provable. Let $b_{i_1}, \dots, b_{i_m}$ be all the variables and constants among $b_1, \dots, b_n$ which do not occur in $\{F(x), \Gamma_1; \Delta_1\}$. Then

$$\forall y_1 \dots \forall y_m C(b_1, \dots, y_1, \dots, y_m, \dots, b_n)$$

where $b_{i_1}, \dots, b_{i_m}$ are replaced by the bound variables, serve as the required interpolant.

4. $k > 0$ and the last inference is $\forall$right

$$\frac{\Gamma \to \Delta, F(a)}{\Gamma \to \Delta, \forall x F(x)}$$

where a doesn't occur in the lower sequent.

Suppose the partition is $[\{\Gamma_1; \Delta_1, \forall x F(x)\}, \{\Gamma_2; \Delta_2\}]$. By the induction hypothesis there exists an interpolant C so that $\Gamma_1 \to \Delta_1, F(a), C$ and $C, \Gamma_2 \to \Delta_2$ are provable. Since C doesn't contain a, we can derive

$$\Gamma_1 \to \Delta_1, \forall x F(x), C$$

and hence C serves as the interpolant

$\square$

Exercise 1.6.1. Let A and B be prenex formulas which have only $\forall$ and $\wedge$ as logical symbols. Assume futhermore that there is at least one predicate constant common to A and B. Suppose $A \supset B$ is provable.

Show that there exists a formula C s.t.

1. $A \supset C$ and $C \supset B$ are provable
2. C is a prenex formula
3. the only logical symbols in C are $\forall$ and $\wedge$
4. the predicate constants in C are common to A and B

Definition 1.35. 1. A **semi-term** is an expression like a term, except that bound variables are allowed in its construction. Let t be a term and s a semi-term. We call s a **sub-semi-term** of t if

   (a) s contain a bound variable (s is not a term)

   (b) s is not a bound variable itself

   (c) some subterm of t is obtained from s by replacing all the bound variables in s by appropriate terms

2. A **semi-formula** is an expression like a formula, except that bound variables are (also) allowed to occur free in it

Theorem 1.36. *Let t be a term and S a provable sequent satisfying*

$$\textit{There is no sub-semi-term of } t \textit{ in } S \tag{1}$$

Then the sequent which is obtained from S by replacing all the occurrences of t in S by a free variable is also provable

Proof. Consider a cut-free regular proof of S, say P. If 1 holds for the lower sequent of an inference in P then it holds for the upper sequents. The theorem follows by mathematical induction on the number of inferences in P □

Definition 1.37. Let $R_1, \dots, R_m, R$ be predicate constants. Let $A(R, R_1, \dots, R_m)$ be a sentence in which all occurrences of $R, R_1, \dots, R_m$ are indicated. Let R' be a predicate constant with the same number of argument-places as R. Let B be $\forall x_1 \dots \forall x_k(R(x_1, \dots, x_k) \equiv R'(x_1, \dots, x_k))$, where the string of quantifiers is empty if $k = 0$. Let C be $A(R, R_1, \dots, R_m) \wedge A(R', R_1, \dots, R_m)$. We say that $A(R, R_1, \dots, R_m)$ **defines (in LK)** R **implicitly** in terms of $R_1, \dots, R_m$ if $C \supset B$ is (**LK**-)provable and we say that $A(R, R_1, \dots, R_m)$ **defines (in LK)** R **explicitly** in terms of $R_1, \dots, R_m$ and the individual constants in $A(R, R_1, \dots, R_m)$ if there exists a formula $F(a_1, \dots, a_k)$ containing only the predicate constants $R_1, \dots, R_m$ and the individual constants in $A(R, R_1, \dots, R_m)$ s.t.

$$A(R, R_1, \dots, R_m) \to \forall x_1 \dots \forall x_k(R(x_1, \dots, x_k) \equiv F(x_1, \dots, x_k))$$

is **LK**-provable

Proposition 1.38 (Beth's definability theorem for **LK**). *If a predicate constant R is defined implicitly in terms of $R_1, \dots, R_m$ by $A(R, R_1, \dots, R_m)$, then R can be defined explicitly in terms of $R_1, \dots, R_m$ and the individual constants in $A(R, R_1, \dots, R_m)$*

outline. Let $c_1, \dots, c_n$ be free variables not occurring in A. Then

$$A(R, R_1, \dots, R_m), A(R', R_1, \dots, R_m) \to R(c_1, \dots, c_n) \equiv R'(c_1, \dots, c_n)$$

and hence also

$$A(R, R_1, \dots, R_m) \wedge R(c_1, \dots, c_k) \to A(R', R_1, \dots, R_m) \supset R'(c_1, \dots, c_n)$$

are provable. Now apply Craig's theorem to the latter sequent. □

1.7 TODO ALL the problems

1.6