Five Lectures on Proof Analysis

Presented at the

Summer School on Proof Theory, Computation, and Complexity

Technische Universität Dresden, June 23 – July 4, 2003

Sara Negri

Department of Philosophy
P.O. Box 9
00014 University of Helsinki, Finland

Contents:

Preprint submitted to Elsevier Science *5 July 2004*

Summary

These five lectures on proof analysis cover the following topics: Lecture 1. Natural deduction is presented, with what are known as general elimination rules. These rules lead in a natural way to sequent calculus. Systems of classical and intuitionistic logic are given as contraction-free sequent calculi, what are called **G3** calculi. Lecture 2. Axioms of universal form are converted into rules by which **G3** calculi are extended. No structural rules (weakening, contraction, cut) are present in these extensions. As an example, predicate logic with equality is given. It is shown through proof analysis that the extension is conservative over sequents that do not contain equalities. A proof of Herbrand's theorem for universal theories as systems with rules is given. Lecture 3. Theories of equality, apartness, and order are presented as extensions of sequent calculus. It is shown that proof search within these theories can be limited to formulas with terms that appear in the endsequent to be derived (the subterm property). Contraction-freeness and the subterm property lead to the decidability of derivability with the rule of a theory. Lecture 4. The approach and results of Lectures 2 and 3 are generalized to what are called, in terms of category theory, geometric theories. As examples, Robinson arithmetic, ordered fields, and plane projective geometry are treated. Each contains existential axioms that can be converted into rules of contraction-free sequent calculus. It is noted that the well-known theorem of Barr follows immediately from the formulation as a rule system. Lecture 5. Variations of the method of converting axioms into rules of sequent calculus are studied. These include a rule scheme that acts on the succedent parts of sequents, instead of the antecedent parts, and a rule scheme for single succedent sequents. The latter is presented also in natural deduction style, and applied to lattice theory. The subterm property for lattice theory in proved through an argument using per-

mutability properties for the lattice rules. The method of permutation of rules is then extended to an independent-context system of sequent calculus with right rules for linear lattices.

Lecture 1: From natural deduction to sequent calculus

In this first lecture we shall outline a route from natural deduction to sequent calculus. In doing so, we shall depart from standard expositions of natural deduction.

Natural deduction and sequent calculus were both introduced by Gentzen. The original Gentzen systems are not isomorphic with respect to normalization and cut elimination (cf. Zucker 1974 and Pottinger 1977). By modifying natural deduction a good correspondence is found. Moreover, the rules of the modified system are supported by a neat meaning explanation of the logical constants. Also the systems of sequent calculus presented here depart from Gentzen's **LJ** and **LK**. Besides the regained correspondence, there are other reasons for such modifications that will emerge throughout the course.

1.1. Natural deduction with general elimination rules

A system of natural deduction is specified by giving, for each logical connective, introduction and elimination rules.

The **introduction rules** are given through the meaning explanation of the logical constants. The so-called "BHK-conditions" (for Brouwer-Heyting- Kolmogorov) give the explanations of logical operations of propositional logic in terms of **direct provability** of propositions

1. A direct proof of the proposition $A\&B$ consists of proofs of the propositions A and B.

2. A direct proof of the proposition $A \vee B$ consists of a proof of the proposition A or a proof of the proposition B.

3. A direct proof of the proposition $A \supset B$ consists of a proof of the proposition B from the assumption that there is a proof of the proposition A.

4. A direct proof of the proposition $\perp$ is impossible.

Here *proof* is an informal notion to be gradually replaced by the formal notion of derivability in a given system of rules. Rules of inference act by transforming a proof of the assumptions into a proof of the conclusion. Observe the hypothetical content of the last condition. On the formal level, a proof of the proposition A is no longer needed when drawing the inference to $A \supset B$, and the assumption A is *discharged* by the rule and put in square brackets.

The BHK-explanation justifies the introduction rules:

$$\frac{A \quad B}{A\&B}\,\&I \qquad \frac{A}{A \vee B}\,\vee I1 \qquad \frac{B}{A \vee B}\,\vee I2 \qquad \frac{\begin{matrix}[A]\\ \vdots \\ B\end{matrix}}{A \supset B}\,\supset I$$

There is no introduction rule for $\bot$.

Corresponding to introduction rules there are **elimination rules**. They have a proposition formed by the logical constants of conjunction, disjunction, implication, or falsity as a major premiss and derive their consequences. The elimination rules are found from the introduction rules through an *inversion principle.* In Prawitz (1965) the following principle is stated:

> **Prawitz' inversion principle:** *The conclusion of an elimination rule R with major premiss $A \star B$ is already contained in the assumptions used to derive $A{\star}B$ from the $\star$-introduction rules, together with the minor premisses of the rule.*

The principle asserts that elimination rules are inverse to the corresponding introduction rules insofar as nothing is gained if an introduction rule is followed by an elimination rule. Prawitz' inversion principle justifies but does not uniquely determine the elimination rules. The following inversion principle has been given in Negri and von Plato 2001 as a justification and unique determination of the elimination rules:

> **Generalized inversion principle:** *Whatever follows from the direct grounds for deriving a proposition, must follow from that proposition.*

For conjunction $A\&B$, the direct grounds are derivations of A and of B. Given that C follows when A and B are assumed, we thus find through the inversion principle the elimination rule

$$\frac{A\&B \qquad \begin{matrix}[A, B]\\ \vdots \\ C\end{matrix}}{C}\,\&E$$

If in a derivation the premisses A and B of the introduction rule have been derived and C has been derived from A and B, the derivation

$$\dfrac{\dfrac{\begin{matrix}\vdots & \vdots \\ A & B\end{matrix}}{A\&B}\,{\scriptstyle \&I} \qquad \begin{matrix}[A,B] \\ \vdots \\ C\end{matrix}}{C}\,{\scriptstyle \&E}$$

converts into a derivation of C without the introduction and elimination rules,

$$\begin{matrix}\vdots & & \vdots \\ A & & B \\ & \vdots & \\ & C & \end{matrix}$$

Therefore, if $\&I$ is followed by $\&E$, the derivation can be simplified.

For disjunction, we have two cases. Either $A \vee B$ has been derived from A, and C is derivable from assumption A, or it has been derived from B and C is derivable from assumption B. Taking into account that both cases are possible, we find the elimination rule

$$\dfrac{A \vee B \qquad \begin{matrix}[A] \\ \vdots \\ C\end{matrix} \qquad \begin{matrix}[B] \\ \vdots \\ C\end{matrix}}{C}\,{\scriptstyle \vee E}$$

Assume now that A or B has been derived. If it is the former and if C is derivable from A and C is derivable from B, the derivation

$$\dfrac{\dfrac{\begin{matrix}\vdots \\ A\end{matrix}}{A \vee B}\,{\scriptstyle \vee I1} \qquad \begin{matrix}[A] \\ \vdots \\ C\end{matrix} \qquad \begin{matrix}[B] \\ \vdots \\ C\end{matrix}}{C}\,{\scriptstyle \vee E}$$

converts to

$$\begin{matrix}\vdots \\ A \\ \vdots \\ C\end{matrix}$$

In the latter case, the conversion produces the derivation

$$\begin{matrix}\vdots \\ B \\ \vdots \\ C\end{matrix}$$

The elimination rule for implication is harder to find. The direct ground for deriving $A \supset B$ is the existence of a **hypothetical** derivation of B from the assumption A. A direct implementation of the principle would lead to a higher level rule (as in Schroeder-Heister 1984). Instead, we keep to a first-order level by observing that if we have $A \supset B$ then, for an arbitrary C,

If C follows from B, then it already follows from A.

We thus obtain the elimination rule

$$\frac{A \supset B \quad A \quad \begin{matrix}[B]\\ \vdots \\ C\end{matrix}}{C}\supset E$$

In addition to the major premiss $A \supset B$, there is the **minor premiss** A in rule $\supset E$. If B has been derived from A and C from B, the derivation

$$\frac{\dfrac{\begin{matrix}[A]\\ \vdots\\ B\end{matrix}}{A \supset B}\supset I \quad \begin{matrix}\vdots\\ A\end{matrix} \quad \begin{matrix}[B]\\ \vdots\\ C\end{matrix}}{C}\supset E$$

converts to

$$\begin{matrix}\vdots\\ A\\ \vdots\\ B\\ \vdots\\ C\end{matrix}$$

Finally, the grounds for deriving $\bot$ are empty, so as a limiting case of the inversion principle get the rule of **falsity elimination** ("ex falso quodlibet"):

$$\frac{\bot}{C}\bot E$$

So far, we have not said how to start derivations. The **rule of assumption** permits to begin a derivation with any formula.

In a given derivation tree the leaves are the assumptions. Among assumptions, *open assumptions* are those which are not discharged by any rule, and *closed assumptions* are discharged assumptions.

The **standard elimination rules** of natural deduction are obtained as special cases from the (general) elimination rules.

If $C = A$ and $C = B$ for $\&E$, and $C = B$ for $\supset E$, we obtain the rules:

$$\frac{A\&B}{A}\,\&E1 \qquad \frac{A\&B}{B}\,\&E2 \qquad \frac{A \supset B \quad A}{B}\,\supset E \text{ (modus ponens)}$$

Why general elimination rules? The reasons for employing general elimination rules rather than standard elimination rules in natural deduction are here of a specific didactic nature: They open up very easily the way to sequent calculus. More substantial reasons are presented in the course by Jan von Plato.

1.2. Natural deduction in sequent calculus style

By making explicit the derivability relation we can write a derivation of C from assumptions Γ as

$$\Gamma \vdash C$$

and obtain a system known as natural deduction in sequent calculus style. For instance, the introduction rule for & and the elimination for $\vee$ become, respectively

$$\frac{\Gamma \vdash A \quad \Delta \vdash B}{\Gamma \cup \Delta \vdash A\&B}\,\&I$$

and

$$\frac{\Gamma \vdash A \vee B \quad \Delta \cup \{A\} \vdash C \quad \Theta \cup \{B\} \vdash C}{\Gamma \cup \Delta \cup \Theta \vdash C}\,\vee E$$

1.3. Sequent calculus

Sequent calculus[1] has notation for keeping track of open assumptions; rules are local: Each formula C has the open assumptions Γ it depends on listed on the same line

$$\Gamma \Rightarrow C$$

Sequent calculus can be regarded as a formal theory of the **derivability relation**. In $\Gamma \Rightarrow C$, the left side Γ is called the **antecedent** and C the **succedent**.

[1] The use of the word "sequent" as a noun was begun by Kleene. His *Introduction to Metamathematics* of 1952 (p. 441) explains the origin of the term as follows: "Gentzen says 'Sequenz', which we translate as 'sequent', because we have already used 'sequence' for any succession of objects, where the German is 'Folge'." This is the standard terminology now; Kleene's usage has even been adopted to some other languages. But Mostowski (1965) for example uses the literal translation "sequence."

In Gentzen's original formulation of 1934–35, the assumptions Γ, Δ, Θ were finite sequences, or **lists** as we would now say. Thus Gentzen had rules permitting the exchange of order of formulas in a sequence. We instead consider assumptions **finite multisets**, that is, lists with multiplicity but no order.

The rules of natural deduction show only the active formulas, and the remaining open assumptions are left implicit. We can make the assumptions explicit, and, instead of

$$\frac{A \quad B}{A\&B}\,\&I$$

write

$$\frac{\begin{array}{cc}\Gamma & \Delta \\ \vdots & \vdots \\ A & B\end{array}}{A\&B}\,\&I$$

but the dots are informal.

When the derivability relation (the dots) are made formal, the introduction rules of natural deduction become **right rules** of sequent calculus, where a comma replaces multiset union:

$$\frac{\Gamma \Rightarrow A \quad \Delta \Rightarrow B}{\Gamma, \Delta \Rightarrow A\&B}\,R\&$$

$$\frac{A, \Gamma \Rightarrow B}{\Gamma \Rightarrow A \supset B}\,R\supset$$

$$\frac{\Gamma \Rightarrow A}{\Gamma \Rightarrow A \vee B}\,R\vee_1$$

$$\frac{\Gamma \Rightarrow B}{\Gamma \Rightarrow A \vee B}\,R\vee_2$$

The formula with the connective in a rule is the **principal** formula of that rule and its components in the premisses the **active** formulas. The Greek letters denote possible additional assumptions that are not active in a rule; They are called the **contexts** of the rules.

The elimination rules of natural deduction correspond to **left rules** of sequent calculus.

$$\frac{A, B, \Gamma \Rightarrow C}{A\&B, \Gamma \Rightarrow C}\,L\&$$

$$\frac{A, \Gamma \Rightarrow C \quad B, \Delta \Rightarrow C}{A \vee B, \Gamma, \Delta \Rightarrow C}\,L\vee$$

$$\frac{\Gamma \Rightarrow A \quad B, \Delta \Rightarrow C}{A \supset B, \Gamma, \Delta \Rightarrow C}\,L\supset$$

We can ask whether this is enough for giving a system of sequent calculus. The answer is negative for the following reasons:

The rules of natural deduction do not give complete instructions to draw derivations. In particular:

1. The same formula can act as assumption and conlcusion in a derivation

$$\dfrac{\overset{1.}{[A]}}{A \supset A}{\scriptstyle \supset I, 1.}$$

2. It is possible to discharge assumptions which have been not made (**vacuous discharge**) as in the following derivation

$$\dfrac{\dfrac{\overset{1.}{[A]}}{B \supset A}{\scriptstyle \supset I}}{A \supset (B \supset A)}{\scriptstyle \supset I, 1.}$$

3. It is possible to discharge more occurrences of the same formula at once (**multiple discharge**) as follows

$$\dfrac{\dfrac{\dfrac{\dfrac{\overset{2.}{[A \supset (A \supset B)]} \quad \overset{1.}{[A]}}{A \supset B}{\scriptstyle \supset E} \qquad \overset{1.}{[A]}}{B}{\scriptstyle \supset E}}{A \supset B}{\scriptstyle \supset I, 1.}}{(A \supset (A \supset B)) \supset (A \supset B)}{\scriptstyle \supset I, 2.}$$

4. It is possible to replace an assumption A in a derivation with a derivation of A and obtain a derivation (**substitution**) by "glueing" of two derivations as follows

$$\begin{array}{cc} & \Delta \\ & \vdots \\ \Gamma, & A \\ \vdots & \\ C & \end{array}$$

The **structural rules** of sequent calculus rules correspond to the natural deduction construction principles 2–4 (sometimes also 1 is included)

Weakening introduces an extra assumption in the antecedent:

$$\dfrac{\Gamma \Rightarrow C}{A, \Gamma \Rightarrow C}{\scriptstyle Wk}$$

In sequent calculus example 2 becomes

$$\dfrac{\dfrac{\dfrac{A \Rightarrow A}{A, B \Rightarrow A}\,{\scriptstyle Wk}}{A \Rightarrow B \supset A}\,{\scriptstyle R\supset}}{\Rightarrow A \supset (B \supset A)}\,{\scriptstyle R\supset}$$

showing that weakening corresponds to vacuous discharge in natural deduction.

Contraction is the rule:

$$\dfrac{A, A, \Gamma \Rightarrow C}{A, \Gamma \Rightarrow C}\,{\scriptstyle Ctr}$$

Example 3 becomes the sequent calculus derivation

$$\dfrac{\dfrac{\dfrac{\dfrac{A \Rightarrow A \quad \dfrac{A \Rightarrow A \quad B \Rightarrow B}{A \supset B, A \Rightarrow B}\,{\scriptstyle L\supset}}{A \supset (A \supset B), A, A \Rightarrow B}\,{\scriptstyle L\supset}}{A \supset (A \supset B), A \Rightarrow B}\,{\scriptstyle Ctr}}{A \supset (A \supset B) \Rightarrow A \supset B}\,{\scriptstyle R\supset}}{\Rightarrow (A \supset (A \supset B)) \supset (A \supset B)}\,{\scriptstyle R\supset}$$

where contraction replaces use of multiple discharge in natural deduction.

If assumptions are treated as sets instead of multisets, contraction is built into the system and is no longer expressed as a distinct rule. This innocent-looking change leads to the loss of syntactic control over derivations.

Cut is the rule:

$$\dfrac{\Gamma \Rightarrow A \quad A, \Delta \Rightarrow C}{\Gamma, \Delta \Rightarrow C}\,{\scriptstyle Cut}$$

The use of cut corresponds in natural deduction to substitution, but not only. Cut is needed in the translation from natural deduction to sequent calculus to express those instances of elimination rules in which the major premiss is derived, that is, not an assumption. These are called *non-normal* instances. Sometimes cut is explained through the familiar practice in mathematics of breaking proofs into lemmas.

The propositional part of the sequent calculus **G0i** is now completely determined.

Logical axiom:

$A \Rightarrow A$

Logical rules:

$$\frac{A, B, \Gamma \Rightarrow C}{A\&B, \Gamma \Rightarrow C}{\scriptstyle L\&} \qquad \frac{\Gamma \Rightarrow A \quad \Delta \Rightarrow B}{\Gamma, \Delta \Rightarrow A\&B}{\scriptstyle R\&}$$

$$\frac{A, \Gamma \Rightarrow C \quad B, \Delta \Rightarrow C}{A \vee B, \Gamma, \Delta \Rightarrow C}{\scriptstyle L\vee} \qquad \frac{\Gamma \Rightarrow A}{\Gamma \Rightarrow A \vee B}{\scriptstyle R\vee 1} \qquad \frac{\Gamma \Rightarrow B}{\Gamma \Rightarrow A \vee B}{\scriptstyle R\vee 2}$$

$$\frac{\Gamma \Rightarrow A \quad B, \Delta \Rightarrow C}{A \supset B, \Gamma, \Delta \Rightarrow C}{\scriptstyle L\supset} \qquad \frac{A, \Gamma \Rightarrow B}{\Gamma \Rightarrow A \supset B}{\scriptstyle R\supset}$$

$$\frac{}{\bot \Rightarrow C}{\scriptstyle L\bot}$$

Rules of weakening and contraction:

$$\frac{\Gamma \Rightarrow C}{A, \Gamma \Rightarrow C}{\scriptstyle Wk} \qquad \frac{A, A, \Gamma \Rightarrow C}{A, \Gamma \Rightarrow C}{\scriptstyle Ctr}$$

Table 1. The sequent calculus G0ip

Natural deduction for **classical** propositional logic is obtained by adding to the rules for intuitionistic logic a **rule of excluded middle**

$$\frac{\begin{matrix}[A] \\ \vdots \\ C\end{matrix} \qquad \begin{matrix}[\sim A] \\ \vdots \\ C\end{matrix}}{C}{\scriptstyle Em}$$

Both A and $\sim A$ are discharged at the inference. The law of excluded middle, $A \vee \sim A$, is derivable with the rule:

$$\frac{\dfrac{[A]}{A \vee \sim A}{\scriptstyle \vee I1} \qquad \dfrac{[\sim A]}{A \vee \sim A}{\scriptstyle \vee I2}}{A \vee \sim A}{\scriptstyle Em}$$

The rule of excluded middle is a generalization of the **rule of indirect proof** ("reductio ad absurdum"),

$$\frac{\begin{matrix}[\sim A] \\ \vdots \\ \bot\end{matrix}}{A}{\scriptstyle Raa}$$

Next we shall outline some desiderata on sequent calculus with a view to applications in proof search.

The rules of sequent calculus can have **independent** or **shared** contexts. The two styles in the right rule for conjunction give the rules

$$\frac{\Gamma \Rightarrow A \quad \Delta \Rightarrow B}{\Gamma, \Delta \Rightarrow A\&B} \qquad \frac{\Gamma \Rightarrow A \quad \Gamma \Rightarrow B}{\Gamma \Rightarrow A\&B}$$

Context-independent and context-sharing rules are easily seen to be equivalent in the presence of the structural rules, in the sense that each rule of one style is derivable from the corresponding rule of the other style. However, the two styles are not equivalent for proof search purposes. If the rules of sequent calculus are used to look for derivation root-first, from the sequent to be derived, application of context-independent rules leads to a combinatorial explosion due to the splitting of the context. With context-sharing rules, the premisses are uniquely determined once the principal formula is chosen.

Cut elimination is the best known desired property of sequent calculus: With the cut rule the **subformula property** is no longer guaranteed. Thus one of the main tasks of structural proof theory is the design of sequent calculi in which cut is an **eliminable** or **admissible** rule.

Contraction can be as "bad" as cut, as concerns a root-first search for a derivation of a given sequent: Formulas in antecedents can be multiplied with no end.

Weakening is easily avoided by modifying the axiom $A \Rightarrow A$ into the form $A, \Gamma \Rightarrow A$.

Classical logic: The rule obtained by a direct translation of the natural deduction rule of excluded middle,

$$\frac{A, \Gamma \Rightarrow C \quad \sim A, \Gamma \Rightarrow C}{\Gamma \Rightarrow C}$$

is not good for proof search purposes, since A is an arbitrary formula that *a priori* has no relation with the formulas in Γ, C. Nevertheless, it can be shown that A can be restricted to atomic formulas and indeed to atoms from C. This works, however, only for the propositional fragment of classical logic.

Alternatively we can extend the notion of a sequent into **multi-succedent** sequents (thus moving away from the intuition of natural deduction). Sequents are thus expressions of the form

$$\Gamma \Rightarrow \Delta$$

where Γ and Δ are both multisets of formulas.

By choosing a multisuccedent formulation, excluded middle is derivable as follows

$$\cfrac{\cfrac{A \Rightarrow A, \bot}{\Rightarrow A, \sim A}}{\Rightarrow A \vee \sim A}$$

The intuitionistic system of sequent calculus is then obtained as a special case of the classical system, by a restriction on the context in implication rules (and for predicate logic in the right rule for the universal quantifier as well). We have thus the advantage of a uniform formalism for intuitionistic and classical logic.

In Gentzen (1934–35), what is sometimes called the **denotational** interpretation of multisuccedent sequents was given: A sequent $\Gamma \Rightarrow \Delta$ expresses that the conjunction of the formulas in Γ implies the disjunction of the formulas in Δ.

The **operational** interpretation of single succedent sequents $\Gamma \Rightarrow C$ (from assumptions Γ, conclusion C can be derived) does not extend to multiple succedents.

Again in Gentzen (1938), the multisuccedent calculus is explained as the natural representation of the **division into cases** often found in mathematical proofs. Thus the antecedent Γ gives the **open assumptions** and the succedent Δ the **open cases**.

The logical rules change and combine open assumptions and cases: $L\&$ replaces the open assumptions A, B by the open assumption $A\&B$; the dual multisuccedent rule $R\vee$ changes the open cases A, B into the open case $A \vee B$, and so on. If there is just one case we have an ordinary conclusion from open assumptions. The other limiting case with no formula in the succedent, as in $\Gamma \Rightarrow$, correspond to the empty case, that is, impossibility.

Invertible rules are needed for decomposing root-first a sequent to be proved. We recall that a rule is invertible if from the derivability of its conclusion, the derivability of its premiss(es) follows.

For instance, the single invertible rule

$$\frac{A, B, \Gamma \Rightarrow \Delta}{A\&B, \Gamma \Rightarrow \Delta}$$

is better than the two equivalent noninvertible rules

$$\frac{A, \Gamma \Rightarrow \Delta}{A\&B, \Gamma \Rightarrow \Delta} \qquad \frac{B, \Gamma \Rightarrow \Delta}{A\&B, \Gamma \Rightarrow \Delta}$$

Summing up, the desiderata for our sequent calculus are: multi-succedent, with context-sharing rules, admissible structural rules, and invertible logical rules.

1.5. The **G3** *sequent calculi*

All the above mentioned desiderata are satisfied in the calculus **G3c**.

Logical axiom:

$P, \Gamma \Rightarrow \Delta, P$

Logical rules:

$$\frac{A, B, \Gamma \Rightarrow \Delta}{A\&B, \Gamma \Rightarrow \Delta}\,{\scriptstyle L\&} \qquad \frac{\Gamma \Rightarrow \Delta, A \quad \Gamma \Rightarrow \Delta, B}{\Gamma \Rightarrow \Delta, A\&B}\,{\scriptstyle R\&}$$

$$\frac{A, \Gamma \Rightarrow \Delta \quad B, \Gamma \Rightarrow \Delta}{A \vee B, \Gamma \Rightarrow \Delta}\,{\scriptstyle L\vee} \qquad \frac{\Gamma \Rightarrow \Delta, A, B}{\Gamma \Rightarrow \Delta, A \vee B}\,{\scriptstyle R\vee}$$

$$\frac{\Gamma \Rightarrow \Delta, A \quad B, \Gamma \Rightarrow \Delta}{A \supset B, \Gamma \Rightarrow \Delta}\,{\scriptstyle L\supset} \qquad \frac{A, \Gamma \Rightarrow \Delta, B}{\Gamma \Rightarrow \Delta, A \supset B}\,{\scriptstyle R\supset}$$

$$\frac{}{\bot, \Gamma \Rightarrow \Delta}\,{\scriptstyle L\bot}$$

Table 2. The sequent calculus G3cp

Observe that in the table of rules P is an atomic formula.

The corresponding intuitionistic system is obtained as a special case by a modification of the implication rules:

$$\frac{\Gamma \Rightarrow A \quad B, \Gamma \Rightarrow \Delta}{A \supset B, \Gamma \Rightarrow \Delta}\,{\scriptstyle L\supset} \qquad \frac{A, \Gamma \Rightarrow B}{\Gamma \Rightarrow \Delta, A \supset B}\,{\scriptstyle R\supset}$$

The resulting calculus will be denoted by **G3im** (where **m** stands for *multi-succedent.*

The **G3** calculi have a long history, dating back to the work of Ketonen (1944).

For a comprehensive historical account we refer to the notes for Chapers 2 and 3 in Negri and von Plato (2001).

In the results that follow we use the notation $\vdash_n \Gamma \Rightarrow \Delta$ to indicate that the sequent $\Gamma \Rightarrow \Delta$ has a derivation of height bounded by n, where the height of a derivation is its height as a tree.

Theorem 1.1: Height-preserving inversion. *All rules of* **G3cp** *are invertible, with height-preserving inversion.*

E.g.: *If* $\vdash_n \Gamma \Rightarrow \Delta, A\&B$, *then* $\vdash_n \Gamma \Rightarrow \Delta, A$ *and* $\vdash_n \Gamma \Rightarrow \Delta, B$.

Proof: By induction on n. If $\Gamma \Rightarrow \Delta, A\&B$ is an axiom or conclusion of $L\bot$, then, $A\&B$ not being atomic, also $\Gamma \Rightarrow \Delta, A$ and $\Gamma \Rightarrow \Delta, B$ are axioms or conclusions of $L\bot$. Assume height-preserving inversion up to height n, and let $\vdash_{n+1} \Gamma \Rightarrow \Delta, A\&B$. There are two cases:

If $A\&B$ is not principal in the last rule, it has one or two premisses $\Gamma' \Rightarrow \Delta', A\&B$ and $\Gamma'' \Rightarrow \Delta'', A\&B$, of derivation height $\leqslant n$, so by inductive hypothesis, $\vdash_n \Gamma' \Rightarrow \Delta', A$ and $\vdash_n \Gamma' \Rightarrow \Delta', B$ and $\vdash_n \Gamma'' \Rightarrow \Delta'', A$ and $\vdash_n \Gamma'' \Rightarrow \Delta'', B$. Now apply the last rule to these premisses to conclude $\Gamma \Rightarrow \Delta, A$ and $\Gamma \Rightarrow \Delta, B$ with a height of derivation $\leqslant n+1$.

If $A\&B$ is principal in the last rule, the premisses $\Gamma \Rightarrow \Delta, A$ and $\Gamma \Rightarrow \Delta, B$ have derivations of height $\leqslant n$. $\square$

Next we have admissibility of the structural rules:

$$\frac{\Gamma \Rightarrow \Delta}{A, \Gamma \Rightarrow \Delta}{\scriptstyle LW} \qquad \frac{\Gamma \Rightarrow \Delta}{\Gamma \Rightarrow \Delta, A}{\scriptstyle RW}$$

$$\frac{A, A, \Gamma \Rightarrow \Delta}{A, \Gamma \Rightarrow \Delta}{\scriptstyle LC} \qquad \frac{\Gamma \Rightarrow \Delta, A, A}{\Gamma \Rightarrow \Delta, A}{\scriptstyle RC}$$

Theorem 1.2: Height-preserving admissibility of weakening.

If $\vdash_n \Gamma \Rightarrow \Delta$, *then* $\vdash_n A, \Gamma \Rightarrow \Delta$.

If $\vdash_n \Gamma \Rightarrow \Delta$, *then* $\vdash_n \Gamma \Rightarrow \Delta, A$.

Theorem 1.3: Height-preserving admissibility of contraction.

If $\vdash_n A, A, \Gamma \Rightarrow \Delta$, *then* $\vdash_n A, \Gamma \Rightarrow \Delta$.

If $\vdash_n \Gamma \Rightarrow \Delta, A, A$, *then* $\vdash_n \Gamma \Rightarrow \Delta, A$.

The proof is by simultaneous induction on the height of the derivation for left and right contraction, using height-preserving invertibility of the rules.

Theorem 1.4 *The rule of cut,*

$$\frac{\Gamma \Rightarrow \Delta, D \quad D, \Gamma' \Rightarrow \Delta'}{\Gamma, \Gamma' \Rightarrow \Delta, \Delta'}\,_{Cut}$$

is admissible in **G3cp**.

For the proof, see Chapter 3 of Negri and von Plato (2001).

The proof of cut elimination for **G3c** shows a first advantage in the use of a contraction-free calculus: there is no need of *multicut*, a device that Gentzen had to use in order to cope with a case of nonreducibility in the proof of cut elimination, the one in which one of the premisses of cut is derived by contraction. Cut is a special case of multicut, so the elimination of multicut gives cut elimination as a corollary. However, by a deeper analysis, not just of the last rule applied in the premisses of cut, it is possible to prove cut elimination without using multicut even in the presence of the other structural rules. This was shown recently in Jan von Plato (2001).

Corollary 1.5. *Each formula in the derivation of* $\Gamma \Rightarrow \Delta$ *in* **G3cp** *is a subformula of* Γ, Δ.

Corollary 1.6. Consistency. *The sequent* $\Rightarrow$ *is not derivable.*

By admissibility of weakening, if $\Gamma \Rightarrow$ is derivable, then also $\Gamma \Rightarrow \bot$ is derivable. The converse is obtained by applying cut to $\Gamma \Rightarrow \bot$ and $\bot \Rightarrow$, thus, an empty succedent behaves like $\bot$.

Lecture 2: Extension of sequent calculi

2.1. Cut elimination in the presence of axioms

It is well known that cut elimination fails in the presence of proper axioms. A simple counterexample is given in Girard 1984: Let the axioms have the forms $\Rightarrow A \supset B$ and $\Rightarrow A$. The sequent $\Rightarrow B$ is derived from these axioms by

$$\frac{\Rightarrow A \qquad \dfrac{\Rightarrow A \supset B \qquad \dfrac{A \Rightarrow A \quad B \Rightarrow B}{A, A \supset B \Rightarrow B}\,_{L\supset}}{A \Rightarrow B}\,_{Cut}}{\Rightarrow B}\,_{Cut}$$

but there is no cut-free derivation of $\Rightarrow B$.

Observe however that if the axioms are converted into the *nonlogical rules*

$$\frac{B \Rightarrow C}{A \Rightarrow C} \qquad \frac{A \Rightarrow C}{\Rightarrow C}$$

then the sequent $\Rightarrow B$ has the cut-free derivation

$$\dfrac{\dfrac{B \Rightarrow B}{A \Rightarrow B}}{\Rightarrow B}$$

The above example shows only the idea of the conversion of axioms into rules. In order to make the idea precise we have to look carefully at the proof of admissibility of the structural rules for the **G3** sequent calculi. This inspection tells how to convert axioms into rules while maintaining admissibility of structural rules in the extended systems.

First, the rules that correspond to the Hilbert-style axioms have to be *logic-free*. The logical content of the nonlogical axioms is absorbed into the geometry of the sequent calculus nonlogical rules. As a consequence, only atomic formulas can appear as active and principal in nonlogical rules.

The rules must have an arbitrary multiset in the succedent.[2]

The most general scheme corresponding to this principle is

$$\frac{Q_1, \Gamma \Rightarrow \Delta \quad \dots \quad Q_n, \Gamma \Rightarrow \Delta}{P_1, \dots, P_m, \Gamma, \Rightarrow \Delta}\,\mathit{Reg}$$

where Γ, Δ are arbitrary multisets and $P_1, \dots, P_m, Q_1, \dots, Q_n$ are fixed atoms with $m, n \geq 0$. In particular, the rule can have zero premisses.

The full rule scheme corresponds to the formula $P_1 \& \dots \& P_m \supset Q_1 \vee \dots \vee Q_n$. In order to see better what forms of axioms the rule scheme covers, we write out a few cases, together with their corresponding axiomatic statements in Hilbert-style calculus. Omitting the contexts, the rules for axioms of the forms $Q \& R$, $Q \vee R$ and $P \supset Q$ are

$$\frac{Q \Rightarrow \Delta}{\Rightarrow \Delta}\ ,\ \frac{R \Rightarrow \Delta}{\Rightarrow \Delta} \qquad \frac{Q \Rightarrow \Delta \quad R \Rightarrow \Delta}{\Rightarrow \Delta} \qquad \frac{Q \Rightarrow \Delta}{P \Rightarrow \Delta}$$

[2] This is true for left rules. For right rules we have instead an arbitrary antecedent. Left rules were found first since they allow the treatment of non-Harrop intuitionistic theories, whereas left and right rules are equivalent as far as classical systems are concerned.

The rules for axioms of the forms Q, $\sim P$ and $\sim(P_1\&P_2)$ are:

$$\frac{Q \Rightarrow \Delta}{\Rightarrow \Delta} \qquad \frac{}{P \Rightarrow \Delta} \qquad \frac{}{P_1, P_2 \Rightarrow \Delta}$$

In order to deal with admissibility of contraction, we have to augment the rule scheme. Right contraction is unproblematic due to the arbitrary context Δ in the succedents of the rule scheme. In order to handle left contraction, we need a closer analysis. So assume we have a derivation of $A, A, \Gamma \Rightarrow \Delta$, and assume the last rule is nonlogical. Then the derivation of $A, A, \Gamma \Rightarrow \Delta$ can be of three different forms. First, neither occurrence of A is principal in the rule; second, one is principal; third, both are principal. The first case is handled by a straightforward induction, and the second case by the method, familiar from the work of Kleene and exemplified by the $L\supset$ rule of **G3ip**, of repeating the principal formulas of the conclusion in the premisses. Thus, the general rule scheme becomes

$$\frac{P_1, \ldots, P_m, Q_1, \Gamma \Rightarrow \Delta \quad \ldots \quad P_1, \ldots, P_m, Q_n, \Gamma \Rightarrow \Delta}{P_1, \ldots, P_m, \Gamma \Rightarrow \Delta}\,\text{\scriptsize\textit{Reg}}$$

Here $P_1, \ldots, P_m$ in the conclusion are principal in the rule, and $P_1, \ldots, P_m$ and $Q_1, \ldots, Q_n$ in the premisses are active in the rule. Repetitions in the premisses will make left contractions commute with rules following the scheme. For the remaining case, with both occurrences of formula A principal in the last rule, consider the situation with a Hilbert-style axiomatization. We have some axiom, say $\sim(a < b \;\&\; b < a)$ in the theory of strict linear order, and substitution of b with a produces $\sim(a < a \;\&\; a < a)$ that we routinely abbreviate to $\sim a < a$, irreflexivity of strict linear order. This is in fact a contraction. For systems with rules, the case where a substitution produces two identical formulas that are both principal in a nonlogical rule, is taken care of by the

Closure condition. *Given a system with nonlogical rules, if it has a rule where a substitution instance in the atoms produces a rule of the form*

$$\frac{P_1, \ldots, P_{m-2}, P, P, Q_1, \Gamma \Rightarrow \Delta \quad \ldots \quad P_1, \ldots, P_{m-2}, P, P, Q_n, \Gamma \Rightarrow \Delta}{P_1, \ldots, P_{m-2}, P, P, \Gamma \Rightarrow \Delta}\,\text{\scriptsize\textit{Reg}}$$

then it also has to contain the rule

$$\frac{P_1, \ldots, P_{m-2}, P, Q_1, \Gamma \Rightarrow \Delta \quad \ldots \quad P_1, \ldots, P_{m-2}, P, Q_n, \Gamma \Rightarrow \Delta}{P_1, \ldots, P_{m-2}, P, \Gamma \Rightarrow \Delta}\,\text{\scriptsize\textit{Reg}}$$

The condition is unproblematic, because the number of rules to be added to a given system of nonlogical rules is bounded. Often the closure condition is superfluous; For example, the rule expressing irreflexivity in the constructive theory of strict linear order is derivable from the other rules.

What axioms are representable as rules following the rule scheme?

For classical systems, the answer is unproblematic: All univeral axioms can have their propositional matrix converted to a conjunction of disjunctions of atoms and negations of atoms. Each conjunct can be converted into the classically equivalent form $P_1 \& \ldots \& P_m \supset Q_1 \vee \ldots \vee Q_n$ which is representable as a rule of inference. We therefore have

Proposition 2.1. *All classical quantifier-free axioms can be represented by rules following the rule scheme.*

The conversion to conjunctive normal form does not hold in intuitionistic logic, so for intuitionistic system we have a smaller class of axioms convertible into rules following the rule scheme. See 6.1(a) in Negri and von Plato (2001) for details.

The following result (proved in Section 6.2 of Negri and von Plato 2001) holds for all the extensions of **G3** sequent systems. We shall denote with **G3c*** (**G3im***) any extension of **G3** (**G3im**) with rules following the rule scheme:

Theorem 2.2. *All the structural rules (weakening, contraction, and cut) are admissible in* **G3c*** *and in* **G3im***. *Weakening and contraction are height-preserving admissible.*

Thus, to every classical quantifier-free theory, there is a corresponding sequent calculus with structural rules admissible.

In systems with nonlogical rules we have a *weak subformula property* :

Theorem 2.3. *If* $\Gamma \Rightarrow \Delta$ *is derivable in* **G3im*** *or* **G3c***, *then all formulas in the derivation are either subformulas of the endsequent or atomic formulas.*

The subformula property is weaker than that for purely logical systems, but sufficient for structural proof-analysis. In the applications to the theories of order and lattice theory we shall improve the property by establishing a *subterm property*, which will have the same consequences for the purpose of proof-search as a proper subformula property.

A simple test for consistency for theories convertible to rules can be performed by analyzing the only possible form of a derivation of $\Rightarrow\perp$ in **G3c*** or **G3im*** (detailed proof in 6.4.2 of Negri and von Plato 2001):

Theorem 2.4. *Assume a theory convertible to rules to be inconsistent. Then*

(i) *All rules in the derivation of* $\Rightarrow \perp$ *are nonlogical,*

(ii) *All sequents in the derivation have* $\bot$ *as succedent*,

(iii) *Each branch in the derivation begins with a nonlogical rule of the form*

$$\frac{}{P_1, \ldots, P_m \Rightarrow \bot}$$

(iv) *The last step in the derivation is a rule of the form*

$$\frac{Q_1 \Rightarrow \bot \quad \ldots \quad Q_n \Rightarrow \bot}{\Rightarrow \bot}$$

As a consequence we have that if an axiom system is inconsistent, then the constituents of their conjunctive normal forms contain negations, and atoms or disjunctions. Therefore, if there are neither atoms nor disjunctions, the axioms are consistent, and similarly if there are no negations.

2.2. Four approaches to extension by axioms

The method presented in the previous section is not the only way of extending sequent calculi for treating axiomatic theories. We can summarize the following four methods and their behaviour with respect to cut elimination and proof analysis:

1. Addition of axioms A into sequent calculus in the form of sequents $\Rightarrow A$ by which derivations can start. As shown in the example at the beginning, the method leads to failure of cut elimination.

2. Gentzen (1938, sec. 1.4): add "mathematical basic sequents" of the form

$$P_1, \ldots, P_m \Rightarrow Q_1, \ldots, Q_n.$$

By Gentzen's "Hauptsatz," the cut rule can be pushed into such basic sequents, and arbitrary cuts reduced to cuts on atoms. Weakening and contraction have to be added as explicit rules.

3. Gentzen's consistency proof of elementary arithmetic (1934, sec. IV.3): treat axioms as a context Γ, and prove results of the form $\Gamma \Rightarrow C$. Cut elimination applies but the resulting system is not contraction-free. Arbitrary instances of the axiom may appear in the antecedent.

4. Axioms as rules, our method.

All these approaches are equivalent, but the fourth is the one best suited for proof analysis.

A formal definition of each of these methods and a proof of their equivalence can be found in 6.3.1 and 6.3.2 in Negri and von Plato (2001).

2.3. Predicate logic with equality

Axiomatic presentations of predicate logic with equality assume a primitive relation $a = b$ with the axiom of **reflexivity**, $a = a$, and the **replacement scheme**, $a = b \,\&A(a/x) \supset A(b/x)$.

In the usual treatment in sequent calculus (as in Troelstra and Schwichtenberg 1996, p. 98), one permits derivations to start with sequents of the form

$$\Rightarrow a = a$$

$$a = b, P(a/x) \Rightarrow P(b/x)$$

where P is atomic. By Gentzen's "extended Hauptsatz" cuts can be reduced to cuts on axioms, but cut elimination fails. For instance, there is no cut-free derivation of symmetry. (Weakening and) contraction must be assumed.

By our method, cuts on equality axioms are avoided. We first restrict the replacement scheme to atomic predicates $P, Q, R, \ldots$ and then convert the axioms into rules,

$$\frac{a = a, \Gamma \Rightarrow \Delta}{\Gamma \Rightarrow \Delta}\,{\scriptstyle Ref} \qquad \frac{a = b, P(a/x), P(b/x), \Gamma \Rightarrow \Delta}{a = b, P(a/x), \Gamma \Rightarrow \Delta}\,{\scriptstyle Repl}$$

G3im and **G3c** plus *Ref* and *Repl* give **intuitionistic and classical predicate logic with equality**.

By the restriction to atomic predicates, both forms of rules follow the rule scheme. A case of duplication is produced in the conclusion of the replacement rule in case P is $x = b$. The rule where both duplications are contracted is an instance of the reflexivity rule so that the closure condition is satisfied. We therefore have, both for **G3im** and **G3c**, the

Theorem 2.5. *The rules of weakening, contraction, and cut are admissible in predicate logic with equality.*

Lemma 2.6. *The replacement axiom $a = b, A(a/x) \Rightarrow A(b/x)$ is derivable for arbitrary A.*

Theorem 2.7. *The replacement rule*

$$\frac{a = b, A(a/x), A(b/x), \Gamma \Rightarrow \Delta}{a = b, A(a/x), \Gamma \Rightarrow \Delta}{\scriptstyle Repl}$$

is admissible for arbitrary predicates A.

Our cut- and contraction-free calculus is equivalent to the usual calculi. But the formulation of equality axioms as rules permits proofs by induction on height of derivation. The conservativity of predicate logic with equality over predicate logic illustrates such proofs.

To prove the conservativity, we show that *Ref* can be eliminated from derivations of equality-free sequents.

As observed above, the rule of replacement has an instance with a duplication, and the closure condition is satisfied because the instance in which both duplications are contracted is an instance of reflexivity. For the proof of conservativity, in the absence of *Ref*, the closure condition is satisfied by the addition of the the contracted instance of *Repl*:

$$\frac{a = b, b = b, \Gamma \Rightarrow \Delta}{a = b, \Gamma \Rightarrow \Delta}{\scriptstyle Repl^*}$$

We have the immediate result:

Lemma 2.8. *If* $\Gamma \Rightarrow \Delta$ *has no equalities and is derivable in* **G3c**+Ref+Repl+Repl*, *no sequents in its derivation have equalities in the succedent.*

The following lemma contains the essential analysis in the proof of conservativity:

Lemma 2.9. *If* $\Gamma \Rightarrow \Delta$ *has no equalities and is derivable in* **G3c**+Ref+Repl+Repl* *it is derivable in* **G3c**+Repl+Repl*.

Proof: We show that all instances of *Ref* can be eliminated from a given derivation, by induction on the height of derivation of a topmost instance

$$\frac{a = a, \Gamma' \Rightarrow \Delta'}{\Gamma' \Rightarrow \Delta'}{\scriptstyle Ref}$$

If the premiss is an axiom also the conclusion is, since by the above lemma Δ' contains no equality, and the same if it is a conclusion of $L\bot$. If the premiss has been concluded by a logical rule, apply the inductive hypothesis to the premisses and then the rule.

If the premiss has been concluded by *Repl* there are two cases, according to whether $a = a$ is or is not principal. In the latter case the derivation is, with

$\Gamma' = P(b/x), \Gamma''$,

$$\cfrac{\cfrac{a = a, b = c, P(b/x), P(c/x), \Gamma'' \Rightarrow \Delta'}{a = a, b = c, P(b/x), \Gamma'' \Rightarrow \Delta'}\,{\scriptstyle Repl}}{b = c, P(b/x), \Gamma'' \Rightarrow \Delta'}\,{\scriptstyle Ref}$$

By permuting the two rules, the inductive hypothesis can be applied.

If $a = a$ is principal, the derivation is, with $\Gamma' = P(a/x), \Gamma''$,

$$\cfrac{\cfrac{a = a, P(a/x), P(a/x), \Gamma'' \Rightarrow \Delta'}{a = a, P(a/x), \Gamma'' \Rightarrow \Delta'}\,{\scriptstyle Repl}}{P(a/x), \Gamma'' \Rightarrow \Delta'}\,{\scriptstyle Ref}$$

By height-preserving contraction, there is a derivation of $a = a, P(a/x), \Gamma'' \Rightarrow \Delta'$ to which the inductive hypothesis applies, giving a derivation of $\Gamma' \Rightarrow \Delta'$ without rule *Ref*.

If the premiss of *Ref* has been concluded by *Repl** with $a = a$ not principal the derivation is

$$\cfrac{\cfrac{a = a, b = c, c = c, \Gamma' \Rightarrow \Delta'}{a = a, b = c, \Gamma' \Rightarrow \Delta'}\,{\scriptstyle Repl^*}}{b = c, \Gamma'' \Rightarrow \Delta'}\,{\scriptstyle Ref}$$

The rules are permuted and the inductive hypothesis applied.

If $a = a$ is principal the derivation is

$$\cfrac{\cfrac{a = a, a = a, \Gamma' \Rightarrow \Delta'}{a = a, \Gamma' \Rightarrow \Delta'}\,{\scriptstyle Repl^*}}{\Gamma' \Rightarrow \Delta'}\,{\scriptstyle Ref}$$

and we apply height-preserving contraction and the inductive hypothesis. □

Next, since the rules *Repl* and *Repl** have equalities in their conclusions, we obtain:

Theorem 2.10 *If* $\Gamma \Rightarrow \Delta$ *is derivable in* **G3c**+Ref+ Repl+Repl* *and if* Γ, Δ *contain no equality, then* $\Gamma \Rightarrow \Delta$ *is derivable in* **G3c**.

2.4. Herbrand's theorem for universal theories

Let **T** be a theory with a finite number of purely universal axioms and classical logic. We turn the theory **T** into a system of nonlogical rules by first removing the quantifiers from each axiom, then converting the remaining part into nonlogical rules. The resulting system will be denoted **G3cT**.

Theorem 2.11. Herbrand's theorem: *If the sequent* $\Rightarrow \forall x \exists y_1 \dots \exists y_k A$, *with* A *quantifier-free, is derivable in* **G3cT**, *then there are terms* t_{i_j} *with* $i \leqslant n, j \leqslant k$ *such that*

$$\bigvee_{i=1}^{n} A(t_{i_1}/y_1, \dots, t_{i_k}/y_k)$$

is derivable in **G3cT**.

Proof: Suppose that $k = 1$. The derivation of $\Rightarrow \forall x \exists y A$ ends with

$$\cfrac{\cfrac{\Rightarrow A(z/x, t_1/y), \exists y A(z/x)}{\Rightarrow \exists y A(z/x)}\, R\exists}{\Rightarrow \forall x \exists y A}\, R\forall$$

Every sequent in the derivation is of the form

$$\Gamma \Rightarrow \Delta, A(z/x, t_m/y), \dots, A(z/x, t_{m+l}/y), \exists y A(z/x)$$

where Γ, Δ consist of subformulas of $A(z/x, t_i/y)$, with $i < m$, and atomic formulas.

Consider the topsequents of the derivation. If they are axioms or conclusions of $L\bot$ they remain so after deletion of the formula $\exists y A(z/x)$. If they are conclusions of zero-premiss nonlogical rules, they remain so after the deletion because the right context in these rules is arbitrary. After deletion, every topsequent in the derivation is of the form

$$\Gamma \Rightarrow \Delta, A(z/x, t_m/y), \dots, A(z/x, t_{m+l}/y)$$

Making the propositional and nonlogical inferences as before, but without the formula $\exists y A(z/x)$ in the succedent, produces a derivation of

$$\Rightarrow A(z/x, t_1/y), \dots, A(z/x, t_{m-1}/y), A(z/x, t_m/y), \dots, A(z/x, t_n/y)$$

and the conclusion follows by applications of rule $R\vee$. $\square$

If the theory **T** is empty we have

Corollary 2.12. *If* $\Rightarrow \exists x A$ *is derivable in* **G3c**, *there are terms* $t_1, \dots, t_n$ *such that* $\Rightarrow A(t_1/x) \vee \dots \vee A(t_n/x)$ *is derivable.*

Lecture 3: Proof analysis in the theories of equality, apartness, and order

The extension of sequent calculi with rules presented in the previous lecture enjoys all the structural properties of the ground, purely logical, sequent calculus, i.e., the rules of weakening, contraction, and cut are admissible. In addition to being admissible, weakening and contraction are height-preserving admissible. The usual consequence of cut elimination, the subformula property, holds in a weaker form, because all the formulas in the derivations in such extensions are subformulas of the endsequent or atomic formulas. However, by analyzing *minimal* derivations (defined below) in specific theories, we can establish a *subterm property*, by which all terms in a derivation are terms in the conclusion.

Before continuing, we observe that a derivation in which a rule, read root-first, produces a duplication of an atom, can be shortened by applying height-preserving admissibility of contraction in place of the rule that introduces that atom. This justifies the following definition:

Definition 3.1. Minimal derivations. *A* minimal *derivation is a derivation in which shortenings through height-preserving admissibility of contraction are not possible, and sequents that can be concluded by zero-premiss rules appear only as topsequent.*

The subterm property, together with height-preserving admissibility of contraction, will give a bound on proof search for the theories under exam: In a minimal derivation, no new term can appear, nor any instantiations of rules that produce a duplication of formulas.

3.1. Theory of equality

The theory of equality has one basic relation $a = b$ with the axioms

EQ1. $a = a$,

EQ2. $a = b \;\&\; a = c \supset b = c$.

Symmetry of equality follows by substituting a for c in EQ2. Transitivity is directly an instance of the replacement axiom, with A equal to $x = c$.

Addition of the rules

$$\frac{a = a, \Gamma \Rightarrow \Delta}{\Gamma \Rightarrow \Delta}\,{\scriptstyle Ref} \qquad \frac{b = c, a = b, a = c, \Gamma \Rightarrow \Delta}{a = b, a = c, \Gamma \Rightarrow \Delta}\,{\scriptstyle Trans}$$

gives a calculus **G3im**+*Ref*+*Trans* the rules of which follow the rule scheme. A duplication in *Trans* is produced in case b is identical to c, but the corresponding contracted rule is an instance of rule *Ref*. The closure condition is satisfied and the structural rules admissible.

Occurrences of atoms of the form $a = a$ disappear in *Ref*, however...

Theorem 3.2. Subterm property. *All terms in a minimal derivation of* $\Gamma \Rightarrow \Delta$ *in* **G3EQ** *are terms in the conclusion.*

Proof: Consider an atom $a = a$ active in *Ref* and trace it up along the derivation. If it is never principal, it can be removed altogether, and *Ref* as well, with a subsequent shortening of the derivation. If it is principal in an axiom, a is found in Δ. If if is principal in *Trans*, it has the form

$$\frac{a = c, a = b, a = c, \Gamma \Rightarrow \Delta}{a = a, a = c, \Gamma \Rightarrow \Delta}\,{\scriptstyle Trans}$$

Apply height-preserving contraction to the premiss and get a shorter derivation. QED

3.2. Theory of apartness

The theory of apartness has one basic relation $a \neq b$ with the axioms

AP1. $\sim a \neq a$,

AP2. $a \neq b \supset a \neq c \;\vee\; b \neq c$.

The rules are

$$\frac{}{a \neq a, \Gamma \Rightarrow \Delta}\,{\scriptstyle Irref} \qquad \frac{a \neq c, a \neq b, \Gamma \Rightarrow \Delta \quad b \neq c, a \neq b, \Gamma \Rightarrow \Delta}{a \neq b, \Gamma \Rightarrow \Delta}\,{\scriptstyle Split}$$

Both rules follow the rule scheme, the closure condition does not arise because there is only one principal formula, and therefore structural rules are admissible in **G3im**+*Irref*+*Split*.

Exercise: Prove the subterm property for the theory of apartness. Hint: Replace AP2 by

$a \neq b \supset b \neq a$,

$a \neq b \supset a \neq c \;\vee\; c \neq b$, and trace up in the derivation the term removed by *Split.*

Other exercises on extensions and other topics treated in this series of lectures are available from the web page `prooftheory.helsinki.fi`

The elementary theories of equality and apartness can also be given in a single-succedent formulation based on extension of the calculus **G3i**, as in Negri (1999).

Corollary 3.3. Disjunction property for the theory of apartness. *If* $\Rightarrow A \vee B$ *is derivable in the single-succedent calculus for the theory of apartness, either* $\Rightarrow A$ *or* $\Rightarrow B$ *is derivable.*

Proof: Consider the last rule in a derivation. The rules for apartness cannot conclude a sequent with an empty antecedent and therefore the last rule must be rule $R\vee$ of **G3i**. $\square$

Compare to the treatment of axiom systems as a context: $\Gamma \Rightarrow A \vee B$ is derivable in **G3i**. Whenever Γ contains an instance of the "split" axiom it has a formula with a disjunction in the consequent of an implication. Therefore, Γ does not consist only of Harrop formulas, so the above corollary gives a proper extension of the disjunction property under hypotheses that are Harrop formulas. The definition of a Harrop system is recalled below.

3.3. Harrop systems

Definition 3.4 The class of *Harrop formulas* is defined inductively by the following clauses:

1. *Atoms* $P, Q, R, \ldots$, *and* $\perp$ *are Harrop formulas,*

2. *If* A *and* B *are Harrop formulas, then* $A \& B$ *is a Harrop formula,*

3. *If* B *is a Harrop formula, then* $A \supset B$ *is a Harrop formula.*

A Harrop theory *is a theory the axioms of which consist of Harrop formulas.*

A left Harrop system *is a system of mathematical rules obtained from the axioms of a Harrop theory by using the left rule scheme.*

The rules of a left Harrop system have *at most one* premiss, thus the derivations are linear, and therefore

Theorem 3.5. *If a sequent $\Gamma \Rightarrow \Delta$ is derivable in a left Harrop system, then $\Gamma \Rightarrow P$ is derivable for some atom P. If Δ contains atoms, the atom P can be chosen from Δ.*

Proof: Consider a derivation $\mathcal{D}$ of $\Gamma \Rightarrow \Delta$. If the topsequent is a logical axiom $P, \Gamma' \Rightarrow \Delta', P$, with $\Delta = \Delta', P$, the succedent can be changed into P. If the topsequent is a zero-premiss mathematical rule, any atom P can be put as the succedent and the derivation with the new succedent continued as with Δ. $\square$

3.4. Theory of partial order

The axioms of partial order are

PO1. $a \leqslant a$,

PO2. $a \leqslant b \,\&\, b \leqslant c \supset a \leqslant c$.

Equality is defined by $a = b \equiv a \leqslant b \,\&\, b \leqslant a$. (Thus, we are working with what are sometimes called quasiorderings.) Clearly, the equality defined is an equivalence relation and satisfies the principle of substitution of equals.

We defined **GPO** as the system with the rules

$$\frac{a \leqslant a, \Gamma \Rightarrow \Delta}{\Gamma \Rightarrow \Delta}\,Ref \qquad \frac{a \leqslant c, a \leqslant b, b \leqslant c, \Gamma \Rightarrow \Delta}{a \leqslant b, b \leqslant c, \Gamma \Rightarrow \Delta}\,Trans$$

The closure condition arises when $a \equiv b$ and $b \equiv c$ so the premiss of rule *Trans* to consider is

$$a \leqslant a, a \leqslant a, a \leqslant a, \Gamma \Rightarrow \Delta$$

The conclusion follows by rule *Ref* so that the closure condition is satisfied. Thus, the rules of weakening, contraction, and cut are admissible in **GPO**. Weakening and contraction are admissible and height preserving.

Proof analysis in GPO: There are exactly 2 kinds of derivation. To see what they are, assume that derivations are minimal (def. 4.1).

If $\Gamma \Rightarrow \Delta$ is derivable, the topsequent has the form $P, \Gamma' \Rightarrow \Delta', P$ with $\Delta', P = \Delta$, and we can delete Δ' from the topsequent.

The two kinds of derivations are:

1. Reflexivity derivations: $P \equiv a \leqslant a$.

The conclusion $\Gamma \Rightarrow a \leqslant a$ follows in one step from the logical axiom $a \leqslant a, \Gamma \Rightarrow a \leqslant a$ with one application of rule *Ref*:

$$\frac{a \leqslant a, \Gamma \Rightarrow a \leqslant a}{\Gamma \Rightarrow a \leqslant a}\,\text{Ref}$$

The context Γ is superfluous and can be deleted, thus, the conclusion becomes $\Rightarrow a \leqslant a$.

2. Transitivity derivations: The topsequent is $a_1 \leqslant a_n, \Gamma' \Rightarrow a_1 \leqslant a_n$.

The atom $a_1 \leqslant a_n$ must be the *removed atom* in a first step of transitivity or else the derivation can be shortened: If some other atom P were removed, with $\Gamma' \equiv P, \Gamma''$, the derivation could be shortened by starting with $a_1 \leqslant a_n, \Gamma'' \Rightarrow a_1 \leqslant a_n$ as topsequent.

There cannot be steps of reflexivity in this derivation: The reflexivity atom would be principal in a step of transitivity, else it could be removed tout court from the derivation with a subsequent shortening, thus there would be a step of the form

$$\frac{a \leqslant b, a \leqslant a, a \leqslant b, \Gamma \Rightarrow a_1 \leqslant a_n}{a \leqslant a, a \leqslant b, \Gamma \Rightarrow a_1 \leqslant a_n}\,\text{Trans}$$

By height-preserving admissibility of contraction the conclusion of this step could be obtained already from the premiss, without using transitivity, in one step less.

Two atoms $a_1 \leqslant a_2, a_2 \leqslant a_n$ are *activated* by the step of *Trans* removing $a_1 \leqslant a_n$ so that the topsequent is of the form

$$a_1 \leqslant a_n, a_1 \leqslant a_2, a_2 \leqslant a_n, \Gamma'' \Rightarrow a_1 \leqslant a_n$$

In the second step, one of the activated atoms must become the removed atom, with two new activated atoms, say $a_2 \leqslant a_3, a_3 \leqslant a_n$, or else the derivation can be shortened. The closure of the principal atom $a_1 \leqslant a_n$ with respect to the activation relation gives us a *chain* $a_1 \leqslant a_2, a_2 \leqslant a_3, \dots a_{n-1} \leqslant a_n$ in the topsequent. Deleting the atoms that have not been active in the derivation, we have a derivation of the form

$$\begin{array}{c}\Gamma''', a_1 \leqslant a_2, a_2 \leqslant a_3, \dots a_{n-1} \leqslant a_n \Rightarrow a_1 \leqslant a_n \\ \hline \vdots \\ \hline a_1 \leqslant a_2, a_2 \leqslant a_3, \dots a_{n-1} \leqslant a_n \Rightarrow a_1 \leqslant a_n\end{array}\,\text{Trans}$$

in which Γ''' consists of the removed atoms $a_1 \leqslant a_n, \dots$.

Sequents $\Gamma \Rightarrow \Delta$ *derivable in* **GPO** *are derivable as left and right weakenings of reflexivity and transitivity derivations.*

Proof search for a sequent $\Gamma \Rightarrow \Delta$ is effected by two controls:

Does Δ contain a reflexivity atom?

Does Γ contain a chain from a_1 to a_n with the atom $a_1 \leqslant a_n$ in Δ?

If so, the sequent $\Gamma \Rightarrow \Delta$ is derivable, otherwise it is underivable.

Nondegenerate partial order is obtained by adding the axiom

PO3. $\sim 1 \leqslant 0$

to PO1 and PO2.

The corresponding rule has zero premisses:

$$\frac{}{1 \leqslant 0, \Gamma \Rightarrow \Delta}{\scriptstyle Ndeg}$$

Derivations remain linear and the theorem on Harrop systems applies.

If the topsequent is an instance of *Ndeg*, the atom $1 \leqslant 0$ is removed by *Trans* (it cannot be removed by *Ref*). Steps of *Trans* hide the inconsistent assumption $1 \leqslant 0$, with the general form of conclusion

$$1 \leqslant a_1, a_1 \leqslant a_2, \ldots, a_{n-1} \leqslant 0 \Rightarrow a \leqslant b$$

with the chain in the antecedent being the closure of formulas activated by $1 \leqslant 0$ and $a \leqslant b$ in the succedent an arbitrary atom.

Nontrivial partial order has in addition

PO4. $0 \leqslant 1$.

The corresponding rule is

$$\frac{0 \leqslant 1, \Gamma \Rightarrow \Delta}{\Gamma \Rightarrow \Delta}{\scriptstyle Ntriv}$$

This rule commutes down with instances of *Ref* and *Trans*. The only interesting case is a transitivity derivation with a chain from which atoms $0 \leqslant 1$ have been removed by *Ntriv*.

3.5. Linear order and Szpilrajn's theorem

The theory of linear order is obtained by adding to partial order the *linearity axiom*

LO. $a \leqslant b \vee b \leqslant a$

The corresponding rule is

$$\frac{a \leqslant b, \Gamma \Rightarrow \Delta \quad b \leqslant a, \Gamma \Rightarrow \Delta}{\Gamma \Rightarrow \Delta}_{Lin}$$

The system of rules for linear order is designated **GLO**.

Theorem 3.6. Conservativity. *If* $\Gamma \Rightarrow P$ *is derivable in* **GLO***, it is derivable already in* **GPO**.

Proof: Consider a derivation with just one instance of *Lin*, as the last rule, and assume the derivation to be minimal. Thus, the premisses of *Lin* $c \leqslant d, \Gamma \Rightarrow P$ and $d \leqslant c, \Gamma \Rightarrow P$ are derivable in partial order. If P is a reflexivity atom in either topsequent, $\Gamma \Rightarrow P$ is derivable in one step of *Ref.* Otherwise, with $P \equiv a \leqslant b$, there will be two transitive closures of the removed atom $a \leqslant b$ in both derivations of the two premisses of *Lin*, and let them be $a \leqslant a_1, \ldots a_{m-1} \leqslant b$ and $a \leqslant b_1, \ldots b_{n-1} \leqslant b$. If $c \leqslant d$ is not an atom in the first chain, it can be deleted and a derivation of $\Gamma \Rightarrow P$ in partial order obtained, and similarly for $d \leqslant c$ in the second chain. Thus, we have the two chains

$a \leqslant a_1, \ldots, a_i \leqslant c, c \leqslant d, d \leqslant a_{i+1}, \ldots a_{m-1} \leqslant b$

$a \leqslant b_1, \ldots, b_j \leqslant d, d \leqslant c, c \leqslant b_{j+1}, \ldots b_{n-1} \leqslant b$

The chain

$a \leqslant a_1, \ldots, a_i \leqslant c, c \leqslant b_{j+1}, \ldots b_{n-1} \leqslant b$

appears in the antecedent of the topsequent, thus the sequent

$a \leqslant a_1, \ldots, a_i \leqslant c, c \leqslant b_{j+1}, \ldots b_{n-1} \leqslant b \Rightarrow a \leqslant b$

is derivable in partial order. $\square$

The conservativity theorem extends to nondegenerate nontrivial partial order:

Theorem 3.7. *If* $\Gamma \Rightarrow P$ *is derivable in nondegenerate* **GLO**, *it is derivable in nondegenerate nontrivial* **GPO** *already.*

(Note that a nondegenerate linear order is always nontrivial.)

Extension algorithm:

Definition 3.8. *An ordering* Σ *is* inconsistent *if* $\Gamma \Rightarrow 1 \leqslant 0$ *is derivable for some finite subset* Γ *of* Σ*, otherwise it is* consistent.

Corollary 3.9. Szpilrajn's theorem. *Given a set Σ of atoms in a consistent nondegenerate partial ordering, it can be extended to a consistent nondegenerate linear ordering.*

Proof: Let a, b be any two elements in Σ not ordered in Σ. We claim that either $\Sigma, a \leqslant b$ or $\Sigma, b \leqslant a$ is consistent in **GPO**. Let us assume the contrary, i.e., that there exists a finite subset Γ of Σ such that both $\Gamma, a \leqslant b \Rightarrow 1 \leqslant 0$ and $\Gamma, b \leqslant a \Rightarrow 1 \leqslant 0$ are derivable in **GPO**. Application of rule *Lin* gives the conclusion $\Gamma \Rightarrow 1 \leqslant 0$ in **GLO**. By the conservativity theorem, $\Gamma \Rightarrow 1 \leqslant 0$ is already derivable in **GPO**, contrary to the consistency assumption. Iteration of the procedure gives the desired extension. □

Remark 3.10. Constructive conservativity vs. nonconstructive extension.

The proof of the conservativity theorem is constructive, and effectivity of the extension depends on how the set Σ is given.

We observe a general phenomenon: Classical, nonconstructive set-theoretic extension results using non-constructive principles like Zorn's lemma are reformulated as constructive proof-theoretical conservativity results. An example is the constructive conservativity of linear order over partial order vs. the classical Szpilrajn theorem. Another example is the pointfree constructive Hahn-Banach theorem (Cederquist, Coquand, and Negri 1998) vs. the classical nonconstructive Hahn-Banach theorem.

Decidability of the order relation is often assumed, either explicitly or through the application of the law of excluded middle. Our approach does not impose any such requirement and therefore does not rule out a computational approach to order relations in continuous sets.

The law of excluded middle is avoided by considering extensions of the intuitionistic calculus **G3im** instead of the classical one.

Observe that the intuitionistic rules of implication do not permute down with mathematical rules if these latter have at least two premisses. In the case of Harrop theories, such as partial order or lattice theory, logical rules do permute down and derivations with mathematical rules can be considered in isolation. The separations of the logical and mathematical parts of derivations holds with no restrictions if classical propositional logic is used.

Lecture 4: Geometric theories

So far, we have dealt with proof analysis in universal theories, that is, theories expressed by purely universal axioms. For some classes of first-order axioms, the way quantifiers work can be expressed in a *logic-free* way, suitable for a treatment as a system with rules. The case of geometric axioms will be considered here. All the results in this lecture are presented with detailed proofs in Negri (2003).

Definition 4.1. *A formula in the language of first-order logic is called* geometric *if it does not contain* $\supset$ *or* $\forall$.

A geometric implication *is a sentence of the form*

$$\forall \overline{x}(A \supset B)$$

in which A and B are geometric formulas.

A geometric theory *is a theory axiomatized by geometric implications.*

Proposition 4.2. Canonical form for geometric implications: *Geometric implications can be reduced to conjunctions of formulas of the form*

$$\forall \overline{x}(P_1 \& \dots \& P_m \supset \exists y_1 M_1 \vee \dots \vee \exists y_n M_n) \quad GA$$

in which the P_i are atomic formulas, the M_j conjunctions of atomic formulas, and the variables y_j are not free in the P_i.

The **geometric rule scheme** that corresponds to geometric axioms has the form

$$\frac{\overline{Q}_1(y_1/x_1), \overline{P}, \Gamma \Rightarrow \Delta \quad \dots \quad \overline{Q}_n(y_n/x_n), \overline{P}, \Gamma \Rightarrow \Delta}{\overline{P}, \Gamma \Rightarrow \Delta}\,{\scriptstyle GRS}$$

The variables x_i are called the *replaced variables* of the scheme, and the variables y_i the *proper variables.*

The geometric rule scheme is subject to the following condition that expresses in a logic-free way the role of the existential quantifier in a geometric axiom:

Condition. *The proper variables must not be free in* $\overline{P}, \Gamma, \Delta$.

4.1. Elimination of structural rules

Definition 4.3 *Let T be a geometric theory. Then* **G3cT** (**G3imT**) *is the Gentzen system obtained by adding to* **G3c** (**G3im**) *the geometric rules cor-*

responding to the geometric axioms of T, together with the rules arising from the closure condition.

Proposition 4.4. Equivalence of axiomatic systems and rule systems. *A geometric axiom is derivable from the corresponding geometric rules. Conversely, a geometric rule is derivable from the corresponding geometric axiom in* **G3imT**+ *Contr*+ *Cut.*

Lemma 4.5. Inversion. *All the inversions of the propositional rules that hold for* **G3c** *and* **G3im** *hold for also their geometric extension.*

Lemma 4.6. Substitution. *Given a derivation of* $\Gamma \Rightarrow \Delta$ *in* **G3cT** (**G3imT**), *with x a free variable in Γ, Δ, t a term free for x in Γ, Δ and not containing any of the variables of the geometric rules in the derivation, we can find a derivation of* $\Gamma(t/x) \Rightarrow \Delta(t/x)$ *in* **G3cT** (**G3im**) *with the same height.*

Lemma 4.7. Inversion for quantifier rules.

(i) *If* $\vdash_n \exists x A, \Gamma \Rightarrow \Delta$ *and y is not among the variables of the geometric rules in the derivation, then* $\vdash_n A(y/x), \Gamma \Rightarrow \Delta$.

(ii) *If* $\vdash_n \Gamma \Rightarrow \Delta, \forall x A$ *and y is not among the variables of the geometric rules in the derivation, then* $\vdash_n \Gamma \Rightarrow \Delta, A(y/x)$.

Without loss of generality, we can assume that the following condition on variables is satisfied:

Disjointness condition. *In a derivation in* **G3cT** (**G3imT**), *the sets of proper variables of the geometric rules are pairwise disjoint.*

Theorem 4.8. *The rules of weakening*

$$\frac{\Gamma \Rightarrow \Delta}{A, \Gamma \Rightarrow \Delta}\,{\scriptstyle LW} \qquad \frac{\Gamma \Rightarrow \Delta}{\Gamma \Rightarrow \Delta, A}\,{\scriptstyle RW}$$

are admissible and height preserving in **G3cT** *and in* **G3imT**.

Theorem 4.9. *The rules of contraction*

$$\frac{A, A, \Gamma \Rightarrow \Delta}{A, \Gamma \Rightarrow \Delta}\,{\scriptstyle LC} \qquad \frac{\Gamma \Rightarrow \Delta, A, A}{\Gamma \Rightarrow \Delta, A}\,{\scriptstyle RC}$$

are admissible and height preserving in **G3cT** *and in* **G3imT**.

Theorem 4.10. *The rule of cut*

$$\frac{\Gamma \Rightarrow \Delta, A \qquad A, \Gamma' \Rightarrow \Delta'}{\Gamma, \Gamma' \Rightarrow \Delta, \Delta'}\,{\scriptstyle Cut}$$

is admissible in **G3cT** *and in* **G3imT**.

4.2. Examples of geometric theories

a. Robinson arithmetic

Language: constant 0, unary successor function s, binary functions $+$ and $\cdot$, relation $=$. Atomic formulas of the form $a = b$, for arbitrary terms a and b.

1. $\neg\, s(x) = 0$
2. $s(x) = s(y) \supset x = y$
3. $x = 0 \vee \exists y\ x = s(y)$
4. $x + 0 = x$
5. $x + s(y) = s(x + y)$
6. $x \cdot 0 = 0$
7. $x \cdot s(y) = x \cdot y + x$

The classically equivalent axiomatization with *3* replaced by

$3'$. $\neg\, x = 0 \supset \exists y\ x = s(y)$

is not geometric because it has an implication $x = 0 \supset \bot$ in the antecedent of an implication.

b. Ordered fields

I. Axioms for nondegenerate linear order
  1. $x \leqslant x$
  2. $x \leqslant y \vee y \leqslant x$
  3. $x \leqslant y \ \&\ y \leqslant z \supset x \leqslant z$
  4. $\neg\, 1 \leqslant 0$

II. Axioms for ordered additive group
  5. $(x + y) + z = x + (y + z)$
  6. $x + y = y + x$
  7. $x + 0 = x$
  8. $\exists y\ x + y = 0$
  9. $x \leqslant y \supset x + z \leqslant y + z$

III. Axioms for multiplication
  10. $(x \cdot y) \cdot z = x \cdot (y \cdot z)$
  11. $x \cdot y = y \cdot y$
  12. $x \cdot 1 = x$
  13. $x = 0 \vee \exists y\ x \cdot y = 1$
  14. $x \cdot (y + z) = x \cdot y + x \cdot z$
  15. $x \leqslant y \ \&\ 0 \leqslant z \supset x \cdot z \leqslant y \cdot z$

The classically equivalent axiomatization with

13′. $\neg\ x = 0 \supset \exists y\ x \cdot y = 1$

in place of *13* is not geometric.

Real closed fields: Add the axioms that state the existence of square roots and zeroes of polynomials of odd degree

16. $0 \leqslant x \supset \exists y\ x = y \cdot y$
17. $a_{2n+1} = 0 \vee \exists x\ a_{2n+1} \cdot x^{2n+1} + a_{2n} \cdot x^{2n} + \ldots a_1 \cdot x + a_0 = 0$

The classically equivalent axiomatization with *17* replaced by

17′. $\neg\ a_{2n+1} = 0 \supset \exists x\ a_{2n+1} \cdot x^{2n+1} + a_{2n} \cdot x^{2n} + \ldots a_1 \cdot x + a_0 = 0$

is not geometric.

c. Classical projective geometry with constructions:

Basic concepts: equality of points, equality of lines, and incidence between points and lines: $a = b, \quad l = m, \quad a \in l$. Constructions: connecting line $ln(a, b)$, intersection point $pt(l, m)$.

I. Axioms for equivalence relations
$a = a, \quad a = c \ \&\ b = c \supset a = b$
$l = l, \quad l = n \ \&\ m = n \supset l = m$
II. Axioms of incidence
$a = b \vee a \in ln(a, b), \qquad a = b \vee b \in ln(a, b)$
$l = m \vee pt(l, m) \in l, \qquad l = m \vee pt(l, m) \in m$
III. Uniqueness axiom
$a \in l \ \&\ b \in l \ \&\ a \in m \ \&\ b \in m \supset a = b \vee l = m$
IV. Substitution axioms
$a \in l \ \&\ a = b \supset b \in l$
$a \in l \ \&\ l = m \supset a \in m$
V. Existence of three noncollinear points
$\exists x \exists y \exists z (\sim x = y \ \&\ \sim z \in ln(x, y))$

The above is not a geometric theory. We obtain a geometric axiomatization by using apartness between points and lines instead of equality as basic notion, as follows:

d. Constructive projective geometry with constructions:

Basic concepts: $a \neq b, \quad l \neq m, \quad a \notin l$. Constructions: $ln(a, b), \quad pt(l, m)$

I. Axioms for apartness relations

$a = a \supset \bot, \quad a \neq b \supset a \neq c \vee b = c$
$l = l \supset \bot, \quad l \neq m \supset l \neq n \vee m = n$

II. Axioms of incidence

$a \neq b \;\&\; a \notin ln(a, b) \supset \bot$
$a \neq b \;\&\; b \notin ln(a, b) \supset \bot$
$l \neq m \;\&\; pt(l, m) \notin l \supset \bot$
$l \neq m \;\&\; pt(l, m) \notin m \supset \bot$

III. Uniqueness axiom

$a \neq b \;\&\; l \neq m \supset a \notin l \vee b \notin l \vee a \notin m \vee b \notin m$

IV. Substitution axioms

$a \notin l \supset a \neq b \vee b \notin l$
$a \notin l \supset l \neq m \vee a \notin m$

V. Existence of three noncollinear points

$\exists x \exists y \exists z (x \neq y \;\&\; z \notin ln(x, y))$

4.3. Barr's theorem

We apply here the method of extension with rules to a general result on geometric theories. The result states that if a geometric implication is provable classically in a geometric theory, then it is provable intuitionistically. This result is proved in topos theory by using a completeness theorem for geometric theories in Grothendieck topoi and the construction of a suitable Boolean topos out of a Grothendieck topos (cf. Johnstone 1977, Mac Lane and Moerdijk 1992).

Topos Theory: *For any Grothendieck topos* $\mathbf{E}$*, there is a Boolean topos* $\mathbf{B}$ *and a geometric morphism* $\gamma : \mathbf{B} \to \mathbf{E}$ *such that* γ^* *is faithful.*

In logical terms, the result states the following:

Logic: *If a geometric implication is classically derivable in a geometric theory, then it is intuitionistically derivable.*

In Palmgren (2002) the result is proved using the Dragalin-Friedman translation.

Our statement: *Let* T *be a geometric theory, and let* A *be a geometric implication. If* $\mathbf{G3cT} \vdash\; \Rightarrow A$*,* then $\mathbf{G3imT} \vdash\; \Rightarrow A$.

Our proof: There is nothing to prove, because a derivation of a geometric implication in **G3cT** is indeed a derivation in **G3imT**. □

Therefore, with our method, the result reduces to a proof-theoretical triviality:

A classical proof of a geometric implication in a geometric theory formulated as a sequent system with rules is already an intuitionistic proof. If we add the requirement that the geometric implication must not contain $\perp$ in the antecedent, then the classical proof is even a proof in minimal logic.

Lecture 5: Variations of proof analysis

In the previous lectures, we have performed proof analyses by using the sequent calculus **G3** as a logical calculus and the mathematical rules formulated in the form of a left rule scheme. The question arises whether we can change the basic calculus, or the form of the rule scheme, or both. The answer is positive, but some care is needed to guarantee the admissibility of the structural rules in the extended calculi. In general, the form of the rule scheme will have to be in harmony with the basic calculus. If, for instance, we modify the basic calculus in favour of context-independent rules, the rule scheme will have to be context-independent as well.

In this lecture we shall illustrate variations of proof analysis with applications to specific problems.

5.1. The right rule scheme

The *left rule scheme* for an axiom of the form $P_1 \& \dots \& P_m \supset Q_1 \vee \dots \vee Q_n$

$$\frac{Q_1, P_1, \dots, P_m, \Gamma \Rightarrow \Delta \quad \dots \quad Q_n, P_1, \dots, P_m, \Gamma \Rightarrow \Delta}{P_1, \dots, P_m, \Gamma \Rightarrow \Delta}$$

has a dual formulation as a *right rule scheme:*

$$\frac{\Gamma \Rightarrow \Delta, Q_1, \dots, Q_n, P_1 \quad \dots \quad \Gamma \Rightarrow \Delta, Q_1, \dots, Q_n, P_m}{\Gamma \Rightarrow \Delta, Q_1, \dots, Q_n}$$

We have repetition of the atoms Q_i in the premisses, to obtain admissibility of right contraction.

As for the left rule scheme, we have the following condition:

Closure condition. *If the atoms in a rule have an instance that makes two atoms in the conclusion identical, the contracted rule has to be added.*

Theorem 5.1 *The structural rules of left and right weakening and contraction and the rule of cut are admissible in extensions of* **G3c** *and* **G3im** *with rules* *following the right rule scheme and satisfying the closure condition.*

As an example, we consider the theory of linear order as a system with right rules. For the details see Negri, von Plato, and Coquand (2004):

Right rules for linear order

$$\frac{}{\Gamma \Rightarrow \Delta, a \leqslant b, b \leqslant a}\,{\scriptstyle Lin} \qquad \frac{}{\Gamma \Rightarrow \Delta, a \leqslant a}\,{\scriptstyle Ref}$$

$$\frac{\Gamma \Rightarrow \Delta, a \leqslant c, a \leqslant b, \quad \Gamma \Rightarrow \Delta, a \leqslant c, b \leqslant c}{\Gamma \Rightarrow \Delta, a \leqslant c}\,{\scriptstyle Trans}$$

Term b in rule *Trans* is called a *middle term.*

Theorem 5.2 *All terms in a minimal derivation of* $\Gamma \Rightarrow \Delta$ *in the right theory of linear order are terms in* Γ, Δ.

Corollary 5.3 *The quantifier-free theory of linear order is decidable.*

Proof: Application of rule *Trans* root-first with middle terms chosen from the conclusion can produce only a bounded number of distinct atoms in the premisses. Whenever a duplication is produced, proof search fails by the admissibility of contraction. □

5.2 Subterm property through permutation of rules

The extensions of the contraction-free sequent **G3** calculi analyzed in the previous lectures have shown some good properties: height-preserving invertibility of rules, height-preserving admissibility of weakening and contraction, subterm property for minimal derivations, feasibility of proof search. Proof analysis, on the other hand, may be difficult in such systems.

Extensions of natural deduction, instead, admit of a more flexible proof analysis but they are limited to Harrop theories.

We shall show this second approach at work in the example of lattice theory. For proofs and details we refer to Negri and von Plato (2002).

We consider a system of natural deduction rules for lattice theory with the meet and join operations:

$$\frac{}{a \leqslant a}\,Ref \qquad \frac{a \leqslant b \quad b \leqslant c}{a \leqslant c}\,Trans$$

$$\frac{}{a\wedge b \leqslant a}\,L\wedge_1 \qquad \frac{}{a\wedge b \leqslant b}\,L\wedge_2 \qquad \frac{c \leqslant a \quad c \leqslant b}{c \leqslant a\wedge b}\,R\wedge$$

$$\frac{}{a \leqslant a\vee b}\,R\vee_1 \qquad \frac{}{b \leqslant a\vee b}\,R\vee_2 \qquad \frac{a \leqslant c \quad b \leqslant c}{a\vee b \leqslant c}\,L\vee$$

Table 3. The system NDLT

Transitivity cannot be eliminated, but it can be reduced to instances in which the middle term is a subterm of the conclusion. Decidability of the derivability of an atom from given atomic assumptions then follows.

Definition 5.4. *A* **new term** *in a derivation tree is a term that is not a term or a subterm in an assumption or in the conclusion.*

Theorem 5.5. Subterm property for NDLT. *If an atom is derivable from atomic assumptions in* **NDLT**, *it has a derivation with no new terms.*

Proof: Consider a topmost instance of *Trans* removing a new term b:

$$\frac{\begin{matrix}\vdots \\ a \leqslant b\end{matrix} \quad \begin{matrix}\vdots \\ b \leqslant c\end{matrix}}{a \leqslant c}\,Trans \qquad (1)$$

1. First consider the derivation of the left premiss. If $a \leqslant b$ concluded by *Trans*, permute up the *Trans* removing b:

$$\frac{\dfrac{a \leqslant d \quad d \leqslant b}{a \leqslant b}\,Trans \qquad b \leqslant c}{a \leqslant c}\,Trans \quad \rightsquigarrow$$

$$\frac{a \leqslant d \qquad \dfrac{d \leqslant b \quad b \leqslant c}{d \leqslant c}\,Trans}{a \leqslant c}\,Trans \qquad (2)$$

Note that, by assumption, d is not a new term.

If $a \leqslant b$ concluded by $L\vee$, the term a has a form $a \equiv d\vee e$ and *Trans* is permuted up as follows:

$$\frac{\dfrac{d \leqslant b \quad e \leqslant b}{d\vee e \leqslant b}\,L\vee \qquad b \leqslant c}{d\vee e \leqslant c}\,Trans \quad \rightsquigarrow$$

$$\frac{\dfrac{d \leqslant b \quad b \leqslant c}{d \leqslant c}\,Trans \qquad \dfrac{e \leqslant b \quad b \leqslant c}{e \leqslant c}\,Trans}{d\vee e \leqslant c}\,L\vee \qquad (3)$$

The permutation of *Trans* removing b is repeated until a left premiss $d' \leqslant b$ not derived by *Trans* or $L\vee$. It can be derived by one of the following rules:

1.1. *Ref*: Then $d' \equiv b$, right premiss of *Trans* identical to the conclusion so b not a new term.

$L\wedge_1$: Then $d' \equiv b\wedge e$ so b not a new term.

$L\wedge_2$: Then $d' \equiv e\wedge b$ so b not a new term.

1.2. $R\vee_1$, we have $b \equiv d'\vee b'$ and the step

$$\begin{array}{c} \dfrac{\dfrac{}{d' \leqslant d'\vee b'}{\scriptstyle R\vee_1} \qquad d'\vee b' \leqslant c}{d' \leqslant c}{\scriptstyle Trans} \\ \vdots \\ a \leqslant c \end{array} \tag{4}$$

The case of $R\vee_2$ is similar.

1.3. $R\wedge$, we have some terms a' and d, e such that $b \equiv d\wedge e$ and

$$\begin{array}{c} \dfrac{\dfrac{a' \leqslant d \quad a' \leqslant e}{a' \leqslant d\wedge e}{\scriptstyle R\wedge} \qquad d\wedge e \leqslant c}{a' \leqslant c}{\scriptstyle Trans} \\ \vdots \\ a \leqslant c \end{array} \tag{5}$$

2. Consider the right premiss $b \leqslant c$ of (4) and (5). If concluded by rules *Trans* or $R\wedge$, permute as in (2) and (3).

Rules $R\vee_1, R\vee_2$ are excluded dually to the excluded rules $L\wedge_1, L\wedge_2$ in the left branch of (1).

This leaves two cases for (4) and also for (5):

2.1. In (4), the right premiss after permutation becomes $d'\vee b' \leqslant c'$ for some term c'

$$\begin{array}{c} \dfrac{\dfrac{}{d' \leqslant d'\vee b'}{\scriptstyle R\vee_1} \qquad \dfrac{\begin{array}{c}\vdots \\ d' \leqslant c' \quad b' \leqslant c'\end{array}}{d'\vee b' \leqslant c'}{\scriptstyle L\vee}}{d' \leqslant c'}{\scriptstyle Trans} \\ \vdots \\ a' \leqslant c \end{array}$$

is transformed into

$$\begin{array}{c} \vdots \\ d' \leqslant c' \\ \vdots \\ a \leqslant c \end{array}$$

with the transitivity step removed.

2.2

$$\cfrac{\cfrac{\begin{matrix}\vdots \\ a' \leqslant d\end{matrix} \quad a' \leqslant e}{a' \leqslant d \wedge e}\,R\wedge \qquad \cfrac{}{d \wedge e \leqslant c'}\,L\wedge_1}{\begin{matrix} a' \leqslant c' \\ \vdots \\ a \leqslant c \end{matrix}}\,\mathit{Trans} \tag{6}$$

Now $c' \equiv d$ so the derivation is transformed into

$$\begin{matrix} \vdots \\ a' \leqslant c \\ \vdots \\ a \leqslant c \end{matrix}$$

with the transitivity step removed. Rule $L\wedge_2$ is treated similarly. □

Corollary 5.6 *Lattice theory is conservative over partial order for universal formulas.*

Corollary 5.7. Decidability of universal formulas. *The derivability of universal formulas in lattice theory is decidable.*

Proof. Consider a universal formula in prenex form $\forall x \ldots \forall z A$ with A in conjunctive normal form. Each conjunct A_k is of the form $P_1 \& \ldots \& P_m \supset Q_1 \vee \ldots \vee Q_n$, with P_i, Q_j atoms. The lattice axioms have no disjunctions in positive parts and therefore (by Harrop's theorem, see, e.g., Negri and von Plato 2001) A_k is derivable if and only if $P_1 \& \ldots \& P_m \supset Q_j$ is derivable for some j. Apply theorem 5.5 to each of the Q_j. □

A proof-theoretic treatment of *relational lattice theory*, with existential axioms in place of the meet and join operation, is presented in Negri and von Plato (2004).

5.3. A multisuccedent generalization

The limitation of proof analysis in natural deduction to Harrop theories can be overcome by considering extensions of a sequent calculus that allows permutability of rules while relaxing the single-succedent limitation. Consider the following sequent calculus with independent contexts:

Logical axiom:

$A \Rightarrow A$

Logical rules:

$$\frac{A, B, \Gamma \Rightarrow \Delta}{A\&B, \Gamma \Rightarrow \Delta}{\scriptstyle L\&} \qquad \frac{\Gamma \Rightarrow \Delta, A \quad \Gamma' \Rightarrow \Delta', B}{\Gamma, \Gamma' \Rightarrow \Delta, \Delta', A\&B}{\scriptstyle R\&}$$

$$\frac{A, \Gamma \Rightarrow \Delta \quad B, \Gamma' \Rightarrow \Delta'}{A \vee B, \Gamma, \Gamma' \Rightarrow \Delta, \Delta'}{\scriptstyle L\vee} \qquad \frac{\Gamma \Rightarrow \Delta, A, B}{\Gamma \Rightarrow \Delta, A \vee B}{\scriptstyle R\vee}$$

$$\frac{\Gamma \Rightarrow \Delta, A \quad B, \Gamma' \Rightarrow \Delta'}{A \supset B, \Gamma, \Gamma' \Rightarrow \Delta, \Delta'}{\scriptstyle L\supset} \qquad \frac{A, \Gamma \Rightarrow \Delta, B}{\Gamma \Rightarrow \Delta, A \supset B}{\scriptstyle R\supset}$$

$$\frac{}{\bot \Rightarrow C}{\scriptstyle L\bot}$$

Rules of weakening:

$$\frac{\Gamma \Rightarrow \Delta}{A, \Gamma \Rightarrow \Delta}{\scriptstyle LW} \qquad \frac{\Gamma \Rightarrow \Delta}{\Gamma \Rightarrow \Delta, A}{\scriptstyle RW}$$

Rules of contraction:

$$\frac{A, A, \Gamma \Rightarrow \Delta}{A, \Gamma \Rightarrow \Delta}{\scriptstyle LC} \qquad \frac{\Gamma \Rightarrow \Delta, A, A}{\Gamma \Rightarrow \Delta, A}{\scriptstyle RC}$$

Table 4. The sequent calculus G0cp

The calculus **G0cp** has the following properties: All the logical rules are invertible, but invertibility is not height preserving. For precise statements and proofs see Lemma 5.1.3 in Negri and von Plato (2001). The rule of cut,

$$\frac{\Gamma \Rightarrow \Delta, D \quad D, \Gamma' \Rightarrow \Delta'}{\Gamma, \Gamma' \Rightarrow \Delta, \Delta'}{\scriptstyle Cut}$$

is admissible in **G0c** (Theorem 5.1.4 *ibidem*).

For a proof, see Negri and von Plato 2001, Chapter 5.

Extensions of **G0cp** with rules following the right rule scheme

$$\frac{\Gamma_1 \Rightarrow \Delta_1, P_1 \quad \dots \quad \Gamma_m \Rightarrow \Delta_m, P_m}{\Gamma_1, \dots, \Gamma_m \Rightarrow \Delta_1, \dots, \Delta_m, Q_1, \dots, Q_n}{\scriptstyle RRS}$$

that corresponds to axioms of the form

$$P_1 \& \dots P_m \supset Q_1 \vee \dots \vee Q_n$$

will be denoted by **G0c*** .

Observe that the atoms $Q_1, \ldots, Q_n$ need not be repeated in the premisses (as in section 5.1) because the calculus has explicit contraction.

Theorem 5.8.

Admissibility of cut holds in **G0c***.

In **G0c****, the logical and structural rules permute down with respect to the right rule scheme, and therefore it suffices in proof analysis to consider derivations of sequents with only atomic formulas.*

The weak subformula property holds.

Observe that, in some cases, permutation of logical rules down the mathematical rules produces multiplications of steps of inference. For instance, it the rule scheme is preceded by right contraction on its active formula P_1

$$\dfrac{\dfrac{\Gamma_1 \Rightarrow \Delta_1, P_1, P_1}{\Gamma_1 \Rightarrow \Delta_1, P_1}{\scriptstyle RContr} \qquad \Gamma_2 \Rightarrow \Delta_2, P_2 \quad \ldots \quad \Gamma_m \Rightarrow \Delta_m, P_m}{\Gamma_1, \ldots, \Gamma_m \Rightarrow \Delta_1, \ldots, \Delta_m, Q_1, \ldots, Q_n}{\scriptstyle RRS}$$

we permute as follows:

$$\dfrac{\dfrac{\dfrac{\Gamma_1 \Rightarrow \Delta_1, P_1, P_1 \quad \Gamma_2 \Rightarrow \Delta_2, P_2 \quad \ldots \quad \Gamma_m \Rightarrow \Delta_m, P_m}{\Gamma_1, \ldots, \Gamma_m \Rightarrow \Delta_1, \ldots, \Delta_m, P_1, Q_1, \ldots, Q_n}{\scriptstyle RRS} \quad \Gamma_2 \Rightarrow \Delta_2, P_2 \quad \ldots \quad \Gamma_m \Rightarrow \Delta_m, P_m}{\Gamma_1, \Gamma_2, \Gamma_2, \ldots, \Gamma_m, \Gamma_m \Rightarrow \Delta_1, \Delta_2, \Delta_2, \ldots, \Delta_m, \Delta_m, Q_1, Q_1, \ldots, Q_n, Q_n}{\scriptstyle RRS}}{\Gamma_1, \Gamma_2 \ldots, \Gamma_m \Rightarrow \Delta_1, \Delta_2 \ldots, \Delta_m, Q_1, \ldots, Q_n}{\scriptstyle Contr^*}$$

where *Contr** denotes repeated steps of left and right contractions.

Sequent calculi with independent contexts are closer to natural deduction, but allow extension to non-Harrop theories. An example is the theory of linear lattices (Negri 2003a).

5.4. Theory of linear lattices

The theory of linear lattices has a binary partial order relation $a \leqslant b$, and equality is defined by

$a = b \equiv a \leqslant b \;\&\; b \leqslant a.$

The axioms of linear lattices are

$a \leqslant a, \quad Ref, \qquad a \leqslant b \;\&\; b \leqslant c \supset a \leqslant c, \quad Trans, \quad a \leqslant b \vee b \leqslant a, \quad Lin,$

$a \wedge b \leqslant a \quad (L\wedge_1), \qquad a \wedge b \leqslant b \quad (L\wedge_2), \qquad c \leqslant a \;\&\; c \leqslant b \supset c \leqslant a \wedge b \quad (R\wedge),$

$$a \leqslant a \vee b \quad (R\vee_1), \qquad b \leqslant a \vee b \quad (R\vee_2), \qquad a \leqslant c \;\&\; b \leqslant c \supset a \vee b \leqslant c \quad (L\vee).$$

The principle of substitution of equals in the lattice operations can be proved, because equality is defined through the partial order relation.

As observed already, the theory of linear lattices is not a Harrop theory because of the linearity axiom *Lin*. Therefore it cannot be treated as a system with rules in the same fashion as the theory of lattices. In order to cover non-Harrop theories one would need a multi-conclusion system of natural deduction, however, natural deduction is inherently a single-conclusion system. Multi-conclusion rules and derivations cannot be written as two-dimensional trees, but the difficulty can be circumvented by using sequent systems.

For linear lattices, we distinguish between *ground terms* $p, q, r, \ldots$ that contain no meet or join operations, and arbitrary terms $a, b, c, \ldots$.

The (right) rules for our calculus for linear lattices are the following:

$$\frac{}{\Rightarrow p \leqslant p}\,{\scriptstyle Ref} \qquad \frac{}{\Rightarrow p \leqslant q, q \leqslant p}\,{\scriptstyle Lin} \qquad \frac{\Gamma_1 \Rightarrow \Delta_1, a \leqslant b \quad \Gamma_2 \Rightarrow \Delta_2, b \leqslant c}{\Gamma_1, \Gamma_2 \Rightarrow \Delta_1, \Delta_2, a \leqslant c}\,{\scriptstyle Trans}$$

$$\frac{}{\Rightarrow a \wedge b \leqslant a}\,{\scriptstyle L\wedge_1} \qquad \frac{}{\Rightarrow a \wedge b \leqslant b}\,{\scriptstyle L\wedge_2} \qquad \frac{\Gamma_1 \Rightarrow \Delta_1, c \leqslant a \quad \Gamma_2 \Rightarrow \Delta_2, c \leqslant b}{\Gamma_1, \Gamma_2 \Rightarrow \Delta_1, \Delta_2, c \leqslant a \wedge b}\,{\scriptstyle R\wedge}$$

$$\frac{}{\Rightarrow a \leqslant a \vee b}\,{\scriptstyle R\vee_1} \qquad \frac{}{\Rightarrow b \leqslant a \vee b}\,{\scriptstyle R\vee_2} \qquad \frac{\Gamma_1 \Rightarrow \Delta_1, a \leqslant c \quad \Gamma_2 \Rightarrow \Delta_2, b \leqslant c}{\Gamma_1, \Gamma_2 \Rightarrow \Delta_1, \Delta_2, a \vee b \leqslant c}\,{\scriptstyle L\vee}$$

Table 5. Rule system for linear lattices

In the rules, the formulas in Γ, Δ are part of the *context*. The atoms in the premisses which are not in the context are called *active*, those in the conclusion are called *principal*. Derivations start with *initial sequents* of the form $a \leqslant b \Rightarrow a \leqslant b$ and with instances of the *zero-premiss* rules. Of these rules, *Ref* and *Lin* are restricted to ground terms. It is seen that all multisuccedent sequents in derivations stem from instances of rule *Lin*.

Term b in rule *Trans* is a *middle term*. An inspection of the rules shows that middle terms in *Trans* are the only terms in premisses that need not be also terms in a conclusion. Because of permutability of logical rules past the mathematical rules observed above, we can consider derivations of sequents with only atomic formulas in antecedents and succedents.

The rules above give a complete system for the theory of linear lattices because we have:

Lemma 5.9. *For arbitrary terms a and b the sequents $\Rightarrow a \leqslant a$ and $\Rightarrow$*

$a \leqslant b, b \leqslant a$ are derivable in the rule system for linear lattices, that is reflexivity and linearity are derivable for arbitrary terms.

Proof: By induction on the length of the terms a, b. For ground terms the sequents are zero-premiss rules of the system, thus derivable. For a compound term a, for instance $a \equiv a_1{\wedge}a_2$, reflexivity follows from the meet rules: $L{\wedge}_1$ and $L{\wedge}_2$ give $a_1{\wedge}a_2 \leqslant a_1$ and $a_1{\wedge}a_2 \leqslant a_2$, and from these, by $R{\wedge}$, we obtain $a_1{\wedge}a_2 \leqslant a_1{\wedge}a_2$. If a is a join, the proof uses instead the rules for join.

For linearity, we have to analyze the form of a and b. If a and b are not both ground terms, there are 8 cases, reduced to 5 by symmetry. In all such cases, linearity is reduced to linearity on the components that is derivable by the inductive hypothesis. For instance, in the case $a \equiv a_1{\wedge}a_2$, $b \equiv b_1{\vee}b_2$, linearity on a, b is derived by applying $R{\wedge}$ to the sequents $\Rightarrow a_1{\wedge}a_2 \leqslant b_1{\vee}b_2, b_1{\vee}b_2 \leqslant a_1$ and $\Rightarrow a_1{\wedge}a_2 \leqslant b_1{\vee}b_2, b_1{\vee}b_2 \leqslant a_2$. The former is derived by $L{\vee}$ from

$$\dfrac{\dfrac{\dfrac{}{a_1{\wedge}a_2 \leqslant a_1}\,{\scriptstyle L\wedge_1} \quad \dfrac{}{a_1 \leqslant b_1, b_1 \leqslant a_1}\,{\scriptstyle Lin}}{a_1{\wedge}a_2 \leqslant b_1, b_1 \leqslant a_1}\,{\scriptstyle Trans} \qquad \dfrac{}{b_1 \leqslant b_1{\vee}b_2}\,{\scriptstyle R\vee_1}}{a_1{\wedge}a_2 \leqslant b_1{\vee}b_2, b_1 \leqslant a_1}\,{\scriptstyle Trans}$$

and

$$\dfrac{\dfrac{\dfrac{}{a_1{\wedge}a_2 \leqslant a_1}\,{\scriptstyle L\wedge_1} \quad \dfrac{}{a_1 \leqslant b_2, b_2 \leqslant a_1}\,{\scriptstyle Lin}}{a_1{\wedge}a_2 \leqslant b_2, b_2 \leqslant a_1}\,{\scriptstyle Trans} \qquad \dfrac{}{b_2 \leqslant b_1{\vee}b_2}\,{\scriptstyle R\vee_2}}{a_1{\wedge}a_2 \leqslant b_1{\vee}b_2, b_2 \leqslant a_1}\,{\scriptstyle Trans}$$

The latter is derived in a similar way. □

The definition of new terms (def. 5.4) is here extended to the more general setting of derivations with cases:

Definition 5.10. *A* new term *in a derivation of a sequent $\Gamma \Rightarrow \Delta$ is a term that is not a term or a subterm in Γ, Δ.*

Theorem 5.11. Subterm property. *If a sequent is derivable in the theory of linear lattices, it has a derivation with no new terms.*

Before proving the theorem, we need preliminary notions for defining a suitable weight that indicates the presence of new terms and their depth in the derivation. The theorem will be proved by giving transformations that reduce such weight, until it becomes zero, with the removal of all new terms from the derivation.

Terms are ordered lexicographically. Given any two terms a and b, either a precedes b in the ordering, or b precedes a or a and b are syntactically identical.

Given a derivation $\mathcal{D}$, consider all the occurrences of a new term which is maximal in the lexicographic ordering among the new terms of the derivation, and among such occurrences, consider those which are *downmost* in the derivation, that is, not followed along a branch by other occurrences of the same term. Downmost occurrences of maximal new terms appear in steps of transitivity removing them from the derivation. Each branch of the derivation contains at most one such downmost maximal new term occurrence. Branches $\mathcal{B}_i$ in a derivation are assigned weight zero if they do not contain such a term, else they have as weight $w(\mathcal{B}_i)$ the length of their segment measured from the leaf down to the last occurrence of the term. The weight of the derivation is given by the multiset of the weights of its branches $\mathcal{B}_1, \ldots, \mathcal{B}_n$

$$w(\mathcal{D}) \equiv \langle w(\mathcal{B}_1), \ldots, w(\mathcal{B}_n) \rangle$$

Weights of derivations are multisets on natural numbers ordered as follows: We put

$$\langle n_1, \ldots, n_k \rangle \leq \langle m_1, \ldots, m_h \rangle$$

if either $max_{i=1}^k n_i \leq max_{i=1}^h m_i$ or $max_{i=1}^k n_i = max_{i=1}^h m_i$ and the number of n_i equal to the maximum is less than the number of m_i equal to the maximum.

This ordering is well founded and it is equivalent to the widely used Dershowitz-Manna multiset ordering.

Proof of Theorem 5.10: We show how to transform derivations so that the weight of the derivation in the multiset ordering gets reduced.

Consider a step of transitivity removing a downmost maximal new term b:

$$\frac{\begin{array}{c}\vdots \\ \Gamma_1 \Rightarrow \Delta_1, a \leqslant b\end{array} \quad \begin{array}{c}\vdots \\ \Gamma_2 \Rightarrow \Delta_2, b \leqslant c\end{array}}{\Gamma_1, \Gamma_2 \Rightarrow \Delta_1, \Delta_2, a \leqslant c} \, \textit{Trans} \qquad (7)$$

Consider the derivations $\mathcal{D}_1$, $\mathcal{D}_2$ of the premisses of *Trans*.

If the atoms $a \leqslant b$ and $b \leqslant c$ are not themselves principal in the last rules of $\mathcal{D}_1$ or $\mathcal{D}_2$, they are found in the premisses of that rule, and transitivity can be permuted above the rule. For example, if the last rule of $\mathcal{D}_1$ is *Trans* with $a \leqslant b$ not principal in it and middle term e different from b, we have, with $\Gamma_1 \equiv \Gamma_{1_1}, \Gamma_{1_2}$, $\Delta_1 \equiv \Delta_{1_1}, \Delta_{1_2}$

$$\frac{\dfrac{\Gamma_{1_1} \Rightarrow \Delta_{1_1}, a \leqslant b, d \leqslant e \quad \Gamma_{1_2} \Rightarrow \Delta_{1_2}, e \leqslant f}{\Gamma_1 \Rightarrow \Delta_1, a \leqslant b, d \leqslant f} \, \textit{Trans} \qquad \Gamma_2 \Rightarrow \Delta_2, b \leqslant c}{\Gamma_1, \Gamma_2 \Rightarrow \Delta_1, \Delta_2, a \leqslant c, d \leqslant f} \, \textit{Trans}$$

We permute as follows

$$\frac{\dfrac{\Gamma_{1_1} \Rightarrow \Delta_{1_1}, a \leqslant b, d \leqslant e \quad \Gamma_2 \Rightarrow \Delta_2, b \leqslant c}{\Gamma_{1_1}, \Gamma_2 \Rightarrow \Delta_{1_1}, \Delta_2, a \leqslant c, d \leqslant e} \, \textit{Trans} \qquad \Gamma_{1_2} \Rightarrow \Delta_{1_2}, e \leqslant f}{\Gamma_1, \Gamma_2 \Rightarrow \Delta_1, \Delta_2, a \leqslant c, d \leqslant f} \, \textit{Trans}$$

Observe that the height of the left premiss of *Trans* removing b is shortened, so that one of the non-zero branches of the derivation has weight reduced by the transformation.

If the middle term e is identical to b we have a block of two consecutive transitivities with middle term b, and we do nothing.

If the last rule of $\mathcal{D}_1$ is *Trans* with $a \leqslant b$ principal, we have

$$\dfrac{\dfrac{\Gamma_{1_1} \Rightarrow \Delta_{1_1}, a \leqslant d \quad \Gamma_{1_2} \Rightarrow \Delta_{1_2}, d \leqslant b}{\Gamma_1 \Rightarrow \Delta_1, a \leqslant b}\,\scriptstyle{Trans} \qquad \Gamma_2 \Rightarrow \Delta_2, b \leqslant c}{\Gamma_1, \Gamma_2 \Rightarrow \Delta_1, \Delta_2, a \leqslant c}\,\scriptstyle{Trans}$$

Observe a subtlety here (which explains why we have chosen the lexicographic ordering on terms): since b is the downmost maximal new term occurrence, d is smaller or equal than b in the lexicographic ordering. In case d is strictly smaller, the transformed derivation is

$$\dfrac{\Gamma_{1_1} \Rightarrow \Delta_{1_1}, a \leqslant d \qquad \dfrac{\Gamma_{1_2} \Rightarrow \Delta_{1_2}, d \leqslant b \quad \Gamma_2 \Rightarrow \Delta_2, b \leqslant c}{\Gamma_{1_2}, \Gamma_2 \Rightarrow \Delta_{1_2}, \Delta_2, d \leqslant c}\,\scriptstyle{Trans}}{\Gamma_1, \Gamma_2 \Rightarrow \Delta_1, \Delta_2, a \leqslant c}\,\scriptstyle{Trans} \tag{8}$$

and the weight is reduced. Else d is identical to b, thus the original derivation has the form

$$\dfrac{\dfrac{\Gamma_{1_1} \Rightarrow \Delta_{1_1}, a \leqslant b \quad \Gamma_{1_2} \Rightarrow \Delta_{1_2}, b \leqslant b}{\Gamma_1 \Rightarrow \Delta_1, a \leqslant b}\,\scriptstyle{Trans} \qquad \Gamma_2 \Rightarrow \Delta_2, b \leqslant c}{\Gamma_1, \Gamma_2 \Rightarrow \Delta_1, \Delta_2, a \leqslant c}\,\scriptstyle{Trans}$$

and the transformed derivation with reduced weight is

$$\dfrac{\dfrac{\Gamma_{1_1} \Rightarrow \Delta_{1_1}, a \leqslant b \quad \Gamma_2 \Rightarrow \Delta_2, b \leqslant c}{\Gamma_{1_1}, \Gamma_2 \Rightarrow \Delta_{1_1}, \Delta_2, a \leqslant c}\,\scriptstyle{Trans}}{\Gamma_1, \Gamma_2 \Rightarrow \Delta_1, \Delta_2, a \leqslant c}\,\scriptstyle{Weak^*}$$

Observe that the original endsequent is restored by steps of weakening.

A similar permutation is performed in case the right premiss of transitivity is derived by another transitivity.

If $a \leqslant b$ has been concluded by $L\vee$, the term a has a form $a \equiv d \vee e$ and the derivation

$$\dfrac{\dfrac{\Gamma_{1_1} \Rightarrow \Delta_{1_1}, d \leqslant b \quad \Gamma_{1_2} \Rightarrow \Delta_{1_2}, e \leqslant b}{\Gamma_1 \Rightarrow \Delta_1, d \vee e \leqslant b}\,\scriptstyle{L\vee} \qquad \Gamma_2 \Rightarrow \Delta_2, b \leqslant c}{\Gamma_1, \Gamma_2 \Rightarrow \Delta_1, \Delta_2, d \vee e \leqslant c}\,\scriptstyle{Trans}$$

is transformed as follows, with *Trans* permuted up to the two premisses of $L\vee$

$$\dfrac{\dfrac{\Gamma_{1_1} \Rightarrow \Delta_{1_1}, d \leqslant b \quad \Gamma_2 \Rightarrow \Delta_2, b \leqslant c}{\Gamma_{1_1}, \Gamma_2 \Rightarrow \Delta_{1_1}, \Delta_2, d \leqslant c}\,\textit{Trans} \qquad \dfrac{\Gamma_{1_2} \Rightarrow \Delta_{1_2}, e \leqslant b \quad \Gamma_2 \Rightarrow \Delta_2, b \leqslant c}{\Gamma_{1_2}, \Gamma_2 \Rightarrow \Delta_{1_2}, \Delta_2, e \leqslant c}\,\textit{Trans}}{\Gamma_1, \Gamma_2, \Gamma_2 \Rightarrow \Delta_1, \Delta_2, \Delta_2, d \vee e \leqslant c}\,L\vee \tag{9}$$

A similar permutation is performed if the right premiss of *Trans* is derived by $R\wedge$. Observe that the permutation of *Trans* over $L\vee$ and $R\wedge$ produces duplications in the contexts.

If $b \equiv b_1 \vee b_2$ and the premisses of *Trans* are derived by $R\vee_1$ and $L\vee$, the derivation

$$\dfrac{\dfrac{}{\Rightarrow b_1 \leqslant b_1 \vee b_2}\,R\vee_1 \qquad \dfrac{\begin{matrix}\vdots \\ \Gamma_1 \Rightarrow \Delta_1, b_1 \leqslant c\end{matrix} \quad \begin{matrix}\vdots \\ \Gamma_2 \Rightarrow \Delta_2, b_2 \leqslant c\end{matrix}}{\Gamma_1, \Gamma_2 \Rightarrow \Delta_1, \Delta_2, b_1 \vee b_2 \leqslant c}\,L\vee}{\Gamma_1, \Gamma_2 \Rightarrow \Delta_1, \Delta_2, b_1 \leqslant c}\,\textit{Trans}$$

is transformed into

$$\begin{matrix}\vdots \\ \Gamma_1 \Rightarrow \Delta_1, b_1 \leqslant c\end{matrix}$$

with the maximal new term occurrence $b_1 \vee b_2$ removed and thus the weight of the derivation reduced.

Next we permute blocks of transitivities with middle term b with iterated conversions of the kind exemplified so far by analyzing the premisses of the rules applied above the topmost transitivities of each block.

Eventually we reach a point in which we have a derivation starting with initial sequents or zero-premiss rules immediately followed by a block of transitivities with middle term b.

The term containing b must be principal in the initial sequents or zero-premiss rules, else we can simplify the derivation because the conclusion of the first rule below them would be again an initial sequents or a zero-premiss rule. We can similarly rule out the possibility of one initial sequent above *Trans* being derived by *Ref*: the conclusion if *Trans* would be identical to the other premiss.

If one premiss of the block is an initial sequent, then b would be a term in the antecedent of the conclusion, contrary to the assumption of b being a new term.

If it is a zero-premiss lattice rule, then we would have the atoms $b \wedge c \leqslant b$ or $b \leqslant b \vee f$. Since b is a maximal new term, then $b \wedge c$, $b \vee f$ are not new term, but then their subterm b is not a new term either.

We are thus left with the possibility that b is a ground term q in a linearity axiom $\Rightarrow p \leqslant q, q \leqslant p$. Then there is a second occurrence of term q in $q \leqslant p$ that has to disappear from the derivation. Since q is a maximal new term, $q \leqslant p$ is active in an instance of transitivity, not in a lattice rule or else q would be subterm of another new term. But this would lead to an infinite derivation. □

References

Dragalin, A. (1988) *Mathematical Intuitionism: Introduction to Proof Theory*, American Mathematical Society, Providence, Rhode Island.

Gentzen, G. (1934-35) Untersuchungen über das logische Schliessen, *Mathematische Zeitschrift*, vol. 39, pp. 176-210 and 405-431.

Gentzen, G. (1969) *The Collected Papers of Gerhard Gentzen*, ed. M. Szabo, North-Holland, Amsterdam.

Johnstone (1977) *Topos Theory*, Academic Press, London.

Ketonen, O. (1944) *Untersuchungen zum Prädikatenkalkül*, (Annales Acad. Sci. Fenn, Ser. A.I. 23), Helsinki.

Kleene, S. (1952) *Introduction to Metamathematics*, North-Holland, Amsterdam.

Mac Lane, S. and I. Moerdijk (1992) *Sheaves in Geometry and Logic*, Springer-Verlag.

Negri, S. (2001) A sequent calculus for constructive ordered fields. In P. Schuster et al. (eds) *Reuniting the Antipodes. Constructive and Nonstandard Views of the Continuum*, Kluwer, Dordrecht.

Negri, S. (2003) Contraction-free sequent calculi for geometric theories, with an application to Barr's theorem, Archive for Mathematical Logic, vol. 42, pp. 389–401.

Negri, S. (2003a) Proof analysis in linear lattices, to appear.

Negri, S. and J. von Plato (1998) Cut elimination in the presence of axioms, *The Bulletin of Symbolic Logic*, vol. 4, pp. 418–435.

Negri, S. and J. von Plato (2001) *Structural Proof Theory*, Cambridge University Press.

Negri, S. and J. von Plato (2002) Permutability of rules in lattice theory, *Algebra Universalis*, vol. 48, pp. 473–477.

Negri, S. and J. von Plato (2004) Proof systems for lattice theory, *Mathematical Structures in Computer Science*, vol. 14, pp. –.

Negri, S., J. von Plato, and Th. Coquand (2004) Proof-theoretic analysis of order relations, Archive for Mathematical Logic, vol. 43, pp. 297–309.

Palmgren, E. (2002) An intuitionistic axiomatization of real closed fields, *Mathematical Logic Quarterly*, vol. 48, pp. 297–299.

von Plato, J. (2001) A proof of Gentzen's *Hauptsatz* without multicut, *Archive for Mathematical Logic*, vol. 40, pp. 9-18.

Pottinger, G. (1977) Normalization as a homomorphic image of cut-elimination, *Annals of Mathematical Logic*, vol. 12, pp. 323–357.

Prawitz, D. (1965) *Natural Deduction: A Proof-Theoretical Study.* Almqvist & Wicksell, Stockholm.

Schroeder-Heister, P. (1984) A natural extension of natural deduction, *The Journal of Symbolic Logic*, vol. 49, pp. 1284–1300.

Troelstra, A. and H. Schwichtenberg (1996) *Basic Proof Theory*, Cambridge University Press, second edition 2000.

Zucker, J. (1974) Cut-elimination and normalization, *Annals of Mathematical Logic*, vol. 7, pp. 1–112.