Constructive Logics. Part I: A Tutorial on Proof Systems and Typed λ-Calculi

Jean Gallier*
Department of Computer and Information Science
University of Pennsylvania
200 South 33rd St.
Philadelphia, PA 19104, USA
e-mail: jean@saul.cis.upenn.edu

October 22, 1992

Abstract. The purpose of this paper is to give an exposition of material dealing with constructive logic, typed λ-calculi, and linear logic. The emergence in the past ten years of a coherent field of research often named "logic and computation" has had two major (and related) effects: firstly, it has rocked vigorously the world of mathematical logic; secondly, it has created a new computer science discipline, which spans from what is traditionally called theory of computation, to programming language design. Remarkably, this new body of work relies heavily on some "old" concepts found in mathematical logic, like natural deduction, sequent calculus, and λ-calculus (but often viewed in a different light), and also on some newer concepts. Thus, it may be quite a challenge to become initiated to this new body of work (but the situation is improving, there are now some excellent texts on this subject matter). This paper attempts to provide a coherent and hopefully "gentle" initiation to this new body of work. We have attempted to cover the basic material on natural deduction, sequent calculus, and typed λ-calculus, but also to provide an introduction to Girard's linear logic, one of the most exciting developments in logic these past six years. The first part of these notes gives an exposition of background material (with some exceptions, such as "contraction-free" systems for intuitionistic propositional logic and the Girard-translation of classical logic into intuitionistic logic, which is new). The second part is devoted to more current topics such as linear logic, proof nets, the geometry of interaction, and unified systems of logic (LU).

*This research was partially supported by ONR Grant NOOO14-88-K-0593.

Contents

1 Introduction **3**

2 Natural Deduction, Simply-Typed λ-Calculus **4**

3 Adding Conjunction, Negation, and Disjunction **10**

4 Gentzen's Sequent Calculi **14**

5 Definition of the Transformation $\mathcal{N}$ from $\mathcal{G}_i$ to $\mathcal{N}_i$ **18**

6 Definition of the Transformation $\mathcal{G}$ from $\mathcal{N}_i$ to $\mathcal{G}_i$ **25**

7 First-Order Quantifiers **28**

8 Gentzen's Cut Elimination Theorem **36**

9 Invertible Rules **43**

10 A Proof-Term Calculus for $\mathcal{G}_i^{\supset,\wedge,\vee,\forall,\exists,\perp,cut}$ **48**

11 The Gentzen Systems $\mathcal{LJ}$ and $\mathcal{LK}$ **52**

12 Cut Elimination in $\mathcal{LK}$ (and $\mathcal{LJ}$) **55**

13 Reductions of Classical Logic to Intuitionistic Logic **70**

1 Introduction

The purpose of this paper is to give an exposition of material dealing with constructive logics, typed λ-calculi, and linear logic. During the last fifteen years, a significant amount of research in the areas of programming language theory, automated deduction, and more generally logic and computation, has relied heavily on concepts and results found in the fields of constructive logics and typed λ-calculi. However, there are very few comprehensive and introductory presentations of constructive logics and typed λ-calculi for noninitiated researchers, and many people find it quite frustrating to become acquainted to this type of research. Our motivation in writing this paper is to help fill this gap. We have attempted to cover the basic material on natural deduction, sequent calculus, and typed λ-calculus, but also to provide an introduction to Girard's linear logic [12], one of the most exciting developments in logic these past six years. As a consequence, we discovered that the amount of background material necessary for a good understanding of linear logic was quite extensive, and we found it convenient to break this paper into two parts. The first part gives an exposition of background material (with some exceptions, such as "contraction-free" systems for intuitionistic propositional logic and the Girard-translation of classical logic into intuitionistic logic, which is new [14]). The second part is devoted to more current topics such as linear logic, proof nets, the geometry of interaction, and unified systems of logic (LU).

In our presentation of background material, we have tried to motivate the introduction of various concepts by showing that they are indispensable to achieve certain natural goals. For pedagogical reasons, it seems that it is best to begin with proof systems in natural deduction style (originally due to Gentzen [8] and thoroughly investigated by Prawitz [23] in the sixties). This way, it is fairly natural to introduce the distinction between intuitionistic and classical logic. By adopting a description of natural deduction in terms of judgements, as opposed to the tagged trees used by Gentzen and Prawitz, we are also led quite naturally to the encoding of proofs as certain typed λ-terms, and to the correspondence between proof normalization and β-conversion (the *Curry/Howard isomorphism* [16]). Sequent calculi can be motivated by the desire to obtain more "symmetric" systems, but also systems in which proof search is easier to perform (due to the subformula property). At first, the cut rule is totally unnecessary and even undesirable, since we are trying to design systems as deterministic as possible. We then show how every proof in the sequent calculus ($\mathcal{G}_i$) can be converted into a natural deduction proof (in $\mathcal{N}_i$). In order to provide a transformation in the other direction, we introduce the cut rule. But then, we observe that there is a mismatch, since we have a transformation $\mathcal{N}: \mathcal{G}_i \rightarrow \mathcal{N}_i$ on cut-free proofs, whereas $\mathcal{G}: \mathcal{N}_i \rightarrow \mathcal{G}_i^{cut}$ maps to proofs possibly with cuts. The mismatch is resolved by Gentzen's fundamental *cut elimination theorem*, which in turn singles out the crucial role played by the *contraction rule*. Indeed, the contraction rule plays a crucial role in the proof of the cut elimination theorem, and furthermore it cannot be dispensed with in traditional systems for intuitionistic logic (however, in the case of intuitionistic propositional logic, it is possible to design contraction-free systems, see section 9 for details). We are thus setting the stage for linear logic, in which contraction (and weakening) are dealt with in a very subtle way. We then investigate a number of sequent calculi that allow us to prove the decidability of provability in propositional classical logic and in propositional intuitionistic logic. In particular, we discuss some "contraction-free" systems for intuitionistic propositional logic for which proof search always terminates. Such systems were discovered in the early fifties by Vorob'ev [35, 36]. Interest in such systems has been revived recently due to some work in automated theorem proving by Dyckhoff [5], on the embedding of intuitionistic logic into linear logic by Lincoln, Scedrov and Shankar [20], and on the complexity of cut-elimination by Hudelmaier [17]. The cut elimination theorem is proved in full for the Gentzen system $\mathcal{LK}$ using Tait's induction measure [29] and some twists due to Girard [13]. We conclude with a fairly extensive discussion of the reduction of classical logic to intuitionistic logic. Besides the standard translations due to Gödel, Gentzen, and Kolmogorov, we present an improved translation due to Girard [14] (based on the notion of polarity of a formula).

In writing this paper, we tried to uncover some of the intuitions that may either have been lost or obscured in advanced papers on the subject, but we have also tried to present relatively sophisticated material, because this is more exciting for the reader. Thus, we have assumed that the reader has a certain familiarity with logic and the lambda calculus. If the reader does not feel sufficiently comfortable with these topics, we suggest consulting Girard, Lafont, Taylor [9] or Gallier [6] for background on logic, and Barendregt [2], Hindley and Seldin [15], or Krivine [19] for background on the lambda calculus. For an in-depth study of constructivism in mathematics, we highly recommend Troelstra and van Dalen [32].

2 Natural Deduction, Simply-Typed λ-Calculus

We first consider a syntactic variant of the natural deduction system for implicational propositions due to Gentzen [8] and Prawitz [23].

In the natural deduction system of Gentzen and Prawitz, a deduction consists in deriving a proposition from a finite number of packets of assumptions, using some predefined inference rules. Technically, packets are multisets of propositions. During the course of a deduction, certain packets of assumptions can be "closed", or "discharged". A proof is a deduction such that all the assumptions have been discharged. In order to formalize the concept of a deduction, one faces the problem of describing rigorously the process of discharging packets of assumptions. The difficulty is that one is allowed to discharge any number of occurrences of the same proposition in a single step, and this requires some form of tagging mechanism. At least two forms of tagging techniques have been used.

- The first one, used by Gentzen and Prawitz, consists in viewing a deduction as a tree whose nodes are labeled with propositions (for a lucid presentation, see van Dalen [34]). One is allowed to tag any set of occurrences of some proposition with a natural number, which also tags the inference that triggers the simultaneous discharge of all the occurrences tagged by that number.

- The second solution consists in keeping a record of all undischarged assumptions at every stage of the deduction. Thus, a deduction is a tree whose nodes are labeled with expressions of the form $\Gamma \vdash A$, called *sequents*, where A is a proposition, and Γ is a record of all undischarged assumptions at the stage of the deduction associated with this node.

Although the first solution is perhaps more natural from a human's point of view and more economical, the second one is mathematically easier to handle. In the sequel, we adopt the second solution. It is convenient to tag packets of assumptions with labels, in order to discharge the propositions in these packets in a single step. We use variables for the labels, and a packet labeled with x consisting of occurrences of the proposition A is written as $x: A$. Thus, in a sequent $\Gamma \vdash A$, the expression Γ is any finite set of the form $x_1: A_1, \ldots, x_m: A_m$, where the x_i are pairwise distinct (but the A_i need not be distinct). Given $\Gamma = x_1: A_1, \ldots, x_m: A_m$, the notation $\Gamma, x: A$ is only well defined when $x \neq x_i$ for all i, $1 \leq i \leq m$, in which case it denotes the set $x_1{:}\, A_1, \ldots, x_m{:}\, A_m, x{:}\, A$. We have the following axioms and inference rules.

Definition 2.1 The axioms and inference rules of the system $\mathcal{N}_m^{\supset}$ (implicational logic) are listed below:

$$\Gamma, x{:}\, A \vdash A$$

$$\frac{\Gamma, x{:}\, A \vdash B}{\Gamma \vdash A \supset B} \quad (\supset\textit{-intro})$$

$$\frac{\Gamma \vdash A \supset B \quad \Gamma \vdash A}{\Gamma \vdash B} \quad (\supset\textit{-elim})$$

In an application of the rule ($\supset$-*intro*), we say that the proposition A which appears as a hypothesis of the deduction is *discharged* (or *closed*).[1] It is important to note that the ability to label packets consisting of occurrences of the same proposition with different labels is essential, in order to be able to have control over which groups of packets of assumptions are discharged simultaneously. Equivalently, we could avoid tagging packets of assumptions with variables if we assumed that in a sequent $\Gamma \vdash C$, the expression Γ, also called a *context*, is a *multiset* of propositions. The following two examples illustrate this point.

Example 2.2 Let

$$\Gamma = x{:}\, A \supset (B \supset C), y{:}\, A \supset B, z{:}\, A.$$

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\Gamma \vdash A \supset (B \supset C) \qquad \Gamma \vdash A}{\Gamma \vdash B \supset C} \qquad \dfrac{\Gamma \vdash A \supset B \qquad \Gamma \vdash A}{\Gamma \vdash B}}{x{:}\, A \supset (B \supset C), y{:}\, A \supset B, z{:}\, A \vdash C}}{x{:}\, A \supset (B \supset C), y{:}\, A \supset B \vdash A \supset C}}{x{:}\, A \supset (B \supset C) \vdash (A \supset B) \supset (A \supset C)}}{\vdash (A \supset (B \supset C)) \supset ((A \supset B) \supset (A \supset C))}$$

In the above example, two occurrences of A are discharged simultaneously. Compare with the example below where these occurrences are discharged in two separate steps.

Example 2.3 Let

$$\Gamma = x{:}\, A \supset (B \supset C), y{:}\, A \supset B, z_1{:}\, A, z_2{:}\, A.$$

[1] In this system, the packet of assumptions A is *always* discharged. This is not so in Prawitz's system (as presented for example in van Dalen [34]), but we also feel that this is a slightly confusing aspect of Prawitz's system.

$$
\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\Gamma \vdash A \supset (B \supset C) \qquad \Gamma \vdash A}{\Gamma \vdash B \supset C} \qquad \dfrac{\Gamma \vdash A \supset B \qquad \Gamma \vdash A}{\Gamma \vdash B}}{x\colon A \supset (B \supset C), y\colon A \supset B, z_1\colon A, z_2\colon A \vdash C}}{x\colon A \supset (B \supset C), y\colon A \supset B, z_1\colon A \vdash A \supset C}}{x\colon A \supset (B \supset C), z_1\colon A \vdash (A \supset B) \supset (A \supset C)}}{z_1\colon A \vdash (A \supset (B \supset C)) \supset ((A \supset B) \supset (A \supset C))}}{\vdash A \supset \big((A \supset (B \supset C)) \supset ((A \supset B) \supset (A \supset C))\big)}
$$

For the sake of comparison, we show what these two natural deductions look like in the system of Gentzen and Prawitz, where packets of assumptions discharged in the same inference are tagged with a natural number. Example 2.2 corresponds to the following tree:

Example 2.4

$$
\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{(A \supset (B \supset C))^3 \qquad A^1}{B \supset C} \qquad \dfrac{(A \supset B)^2 \qquad A^1}{B}}{C}}{A \supset C}\,1}{(A \supset B) \supset (A \supset C)}\,2}{(A \supset (B \supset C)) \supset ((A \supset B) \supset (A \supset C))}\,3
$$

and Example 2.3 to the following tree:

Example 2.5

$$
\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{(A \supset (B \supset C))^3 \qquad A^1}{B \supset C} \qquad \dfrac{(A \supset B)^2 \qquad A^4}{B}}{C}}{A \supset C}\,1}{(A \supset B) \supset (A \supset C)}\,2}{(A \supset (B \supset C)) \supset ((A \supset B) \supset (A \supset C))}\,3}{A \supset \big((A \supset (B \supset C)) \supset ((A \supset B) \supset (A \supset C))\big)}\,4
$$

It is clear that a context (the Γ in a sequent $\Gamma \vdash A$) is used to tag packets of assumptions and to record the time at which they are discharged. From now on, we stick to the presentation of natural deduction using sequents.

Proofs may contain redundancies, for example when an elimination immediately follows an introduction, as in the following example in which $\mathcal{D}_1$ denotes a deduction with conclusion $\Gamma, x{:}\,A \vdash B$ and $\mathcal{D}_2$ denotes a deduction with conclusion $\Gamma \vdash A$.

$$\dfrac{\dfrac{\begin{array}{c}\mathcal{D}_1\\ \Gamma, x{:}\,A \vdash B\end{array}}{\Gamma \vdash A \supset B} \qquad \begin{array}{c}\mathcal{D}_2\\ \Gamma \vdash A\end{array}}{\Gamma \vdash B}$$

Intuitively, it should be possible to construct a deduction for $\Gamma \vdash B$ from the two deductions $\mathcal{D}_1$ and $\mathcal{D}_2$ without using at all the hypothesis $x{:}\,A$. This is indeed the case. If we look closely at the deduction $\mathcal{D}_1$, from the shape of the inference rules, assumptions are never created, and the leaves must be labeled with expressions of the form $\Gamma', \Delta, x{:}\,A, y{:}\,C \vdash C$ or $\Gamma, \Delta, x{:}\,A \vdash A$, where $y \neq x$ and either $\Gamma = \Gamma'$ or $\Gamma = \Gamma', y{:}\,C$. We can form a new deduction for $\Gamma \vdash B$ as follows: in $\mathcal{D}_1$, wherever a leaf of the form $\Gamma, \Delta, x{:}\,A \vdash A$ occurs, replace it by the deduction obtained from $\mathcal{D}_2$ by adding Δ to the premise of each sequent in $\mathcal{D}_2$. Actually, one should be careful to first make a fresh copy of $\mathcal{D}_2$ by renaming all the variables so that clashes with variables in $\mathcal{D}_1$ are avoided. Finally, delete the assumption $x{:}\,A$ from the premise of every sequent in the resulting proof. The resulting deduction is obtained by a kind of substitution and may be denoted as $\mathcal{D}_1[\mathcal{D}_2/x]$, with some minor abuse of notation. Note that the assumptions $x{:}\,A$ occurring in the leaves of the form $\Gamma', \Delta, x{:}\,A, y{:}\,C \vdash C$ were never used anyway. This illustrates the fact that not all assumptions are necessarily used. This will not be the case in linear logic [12]. Also, the same assumption may be used more than once, as we can see in the ($\supset$-*elim*) rule. Again, this will not be the case in linear logic, where every assumption is used exactly once, unless specified otherwise by an explicit mechanism. The step which consists in transforming the above redundant proof figure into the deduction $\mathcal{D}_1[\mathcal{D}_2/x]$ is called a *reduction step* or *normalization step*.

We now show that the simply-typed λ-calculus provides a natural notation for proofs in natural deduction, and that β-conversion corresponds naturally to proof normalization. The trick is to annotate inference rules with terms corresponding to the deductions being built, by placing these terms on the righthand side of the sequent, so that the conclusion of a sequent appears to be the "type of its proof". This way, inference rules have a reading as "type-checking rules". This discovery due to Curry and Howard is known as the *Curry/Howard isomorphism*, or *formulae-as-types principle* [16]. An early occurrence of this correspondence can be found in Curry and Feys [3] (1958), Chapter 9E, pages 312-315. Furthermore, and this is the deepest aspect of the Curry/Howard isomorphism, proof normalization corresponds to term reduction in the λ-calculus associated with the proof system.

Definition 2.6 The type-checking rules of the λ-calculus $\lambda^{\supset}$ (simply-typed λ-calculus) are listed below:

$$\Gamma, x{:}\,A \vdash x{:}\,A$$

$$\dfrac{\Gamma, x{:}\,A \vdash M{:}\,B}{\Gamma \vdash (\lambda x{:}\,A.\, M){:}\,A \supset B} \quad (abstraction)$$

$$\frac{\Gamma \vdash M\colon A \supset B \quad \Gamma \vdash N\colon A}{\Gamma \vdash (MN)\colon B} \quad (\mathit{application})$$

Now, sequents are of the form $\Gamma \vdash M\colon A$, where M is a simply-typed λ-term representing a deduction of A from the assumptions in Γ. Such sequents are also called *judgements*, and Γ is called a *type assignment* or *context*.

The example of redundancy is now written as follows:

$$\frac{\dfrac{\Gamma, x\colon A \vdash M\colon B}{\Gamma \vdash (\lambda x\colon A.\, M)\colon A \supset B} \qquad \Gamma \vdash N\colon A}{\Gamma \vdash (\lambda x\colon A.\, M)N\colon B}$$

Now, $\mathcal{D}_1$ is incorporated in the deduction as the term M, and $\mathcal{D}_2$ is incorporated in the deduction as the term N. The great bonus of this representation is that $\mathcal{D}_1[\mathcal{D}_2/x]$ corresponds to $M[N/x]$, the result of performing a β-reduction step on $(\lambda x\colon A.\, M)N$.

Example 2.7

$$\frac{\dfrac{\dfrac{x\colon P \supset (Q \supset P), u\colon P \vdash u\colon P}{x\colon P \supset (Q \supset P) \vdash \lambda u\colon P.\, u\colon (P \supset P)}}{\vdash \lambda x\colon (P \supset (Q \supset P)).\, \lambda u\colon P.\, u\colon (P \supset (Q \supset P)) \supset (P \supset P)} \qquad \dfrac{\dfrac{y\colon P, z\colon Q \vdash y\colon P}{y\colon P \vdash \lambda z\colon Q.\, y\colon (Q \supset P)}}{\vdash \lambda y\colon P.\, \lambda z\colon Q.\, y\colon P \supset (Q \supset P)}}{\vdash (\lambda x\colon (P \supset (Q \supset P)).\, \lambda u\colon P.\, u)\lambda y\colon P.\, \lambda z\colon Q.\, y\colon (P \supset P)}$$

The term $(\lambda x\colon (P \supset (Q \supset P)).\, \lambda u\colon P.\, u)\lambda y\colon P.\, \lambda z\colon Q.\, y$ reduces to $\lambda u\colon P.\, u$, which is indeed the term representation of the natural deduction proof

$$\frac{u\colon P \vdash P}{\vdash (P \supset P)}$$

Thus, the simply-typed λ-calculus arises as a natural way to encode natural deduction proofs, and β-reduction corresponds to proof normalization. The correspondence between proof normalization and term reduction is the deepest and most fruitful aspect of the Curry/Howard isomorphism. Indeed, using this correspondence, results about the simply-typed λ-calculus can be translated into the framework of natural deduction proofs, a very nice property. On the other hand, one should not be too dogmatic (or naive) about the Curry/Howard isomorphism and make it into some kind of supreme commandment (as we say in French, "prendre ses désirs pour des réalités"). In the functional style of programming, λ-reduction corresponds to parameter-passing, but more is going on, in particular recursion. Thus, although it is fruitful to view a program as a proof, the specification of a program as the proposition proved by that proof, and the execution of a program as proof normalization (or cut elimination, but it is confusing to say that, since in most cases we are dealing with a natural deduction system), it is abusive to claim that this is what programming is all about. In fact, I believe that statements to that effect are detrimental to our field. There are plenty of smart people who are doing research in the theory of programming and programming language design, and such statements will only make them skeptical (at best). Programming cannot be reduced to the Curry/Howard isomorphism.

When we deal with the calculus $\lambda^{\supset}$, rather than using $\supset$, we usually use $\rightarrow$, and thus, the calculus is denoted as $\lambda^{\rightarrow}$. In order to avoid ambiguities, the delimiter used to separate the lefthand side from the righthand side of a judgement $\Gamma \vdash M : A$ will be $\triangleright$, so that judgements are written as $\Gamma \triangleright M : A$.

Before moving on to more fascinating topics, we cannot resist a brief digression on notation (at least, we will spare the reader the moralistic lecture that we have inflicted upon students over more than fourteen years!). Notation is supposed to help us, but the trouble is that it can also be a handicap. This is because there is a very delicate balance between the explicit and the implicit. Our philosophy is that the number of symbols used should be minimized, and that notation should *help* remembering what things are, rather than *force* remembering what things are. The most important thing is that notation should be as unambiguous as possible. Furthermore, we should allow ourselves dropping certain symbols as long as no serious ambiguities arise, and we should avoid using symbols that already have a standard meaning, although this is nearly impossible.

Lambda-abstraction and substitution are particularly spicy illustrations. For example, the notation $\lambda x\colon\sigma M$ together with (MN) for application is unambiguous. However, when we see the term $(\lambda x\colon\sigma MN)$, we have to think a little (in fact, too much) to realize that this is indeed the application of $\lambda x\colon\sigma M$ to N, and not the abstraction $\lambda x\colon\sigma(MN)$. This is even worse if we look at the term $\lambda x\colon\sigma MN$ where the parentheses have been dropped. So, we may consider introducing extra markers, just to help readability, although they are not strictly necessary. For example, we can add a dot between σ and M: abstraction is then written as $\lambda x\colon\sigma.\, M$. Similarly, universally quantified formulae are written as $\forall x\colon\sigma.\, A$. Now, $\lambda x\colon\sigma.\, MN$ is a little better, but still requires an effort. Thus, we will add parentheses around the lambda abstraction and write $(\lambda x\colon\sigma.\, M)N$. Yes, we are using more symbols than we really need, but we feel that we have removed the potential confusion with $\lambda x\colon\sigma.MN$ (which should really be written as $\lambda x\colon\sigma.(MN)$). Since we prefer avoiding subscripts or superscripts unless they are really necessary, we favor the notation $\lambda x\colon\sigma.\, M$ over the (slightly old-fashion) $\lambda x^{\sigma}.\, M$ (we do not find the economy of one symbol worth the superscript).[2] Now, let us present another choice of notation, a choice that we consider poor since it forces us to remember something rather than help us. In this choice, abstraction is written as $[x\colon\sigma]M$, and universal quantification as $(x\colon\sigma)A$. The problem is that the reader needs to remember which kind of bracket corresponds to abstraction or to (universal) quantification. Since additional parentheses are usually added when applications arise, we find this choice quite confusing. The argument that this notation corresponds to some form of machine language is the worst that can be given. Humans are not machines, and thus should not be forced to read machine code! An interesting variation on the notations $\lambda x\colon\sigma.\, M$ and $\forall x\colon\sigma.\, A$ is $\lambda(x\colon\sigma)M$ and $\forall(x\colon\sigma)A$, which is quite defendable. Substitution is an even more controversial subject! Our view is the following. After all, a substitution is a function whose domain is a set of variables and which is the identity except on a finite set. Furthermore, substitutions can be composed. But **beware**: composition of substitutions is **not** function composition (indeed, a substitution φ induces a homomorphism $\widehat{\varphi}$, and the composition of two substitutions φ and ψ is the function composition of $\widehat{\varphi}$ and ψ, and **not** of φ and ψ). Thus, the

[2]The notation $\lambda x^{\sigma}.\, M$ seems to appear mostly in systems where contexts are not used, but instead where it is assumed that each variable has been preassigned a type.

choice of notation for composition of substitutions has an influence on the notation for substitution. If we choose to denote composition of substitution in the order $\varphi\,;\psi$, then it is more convenient to denote the result of applying a substitution φ to a term M as $M\varphi$, or $(M)\varphi$, or as we prefer as $M[\varphi]$. Indeed, this way, $M[\varphi][\psi]$ is equal to $M[\varphi\,;\psi]$. Now, since a substitution is a function with domain a finite set of variables, it can be denoted as $[x_1 \mapsto M_1, \ldots, x_n \mapsto M_n]$. In retrospect, we regret not having adopted this notation. If this was the case, applying a substitution to M would be denoted as $M[x_1 \mapsto M_1, \ldots, x_n \mapsto M_n]$. Instead, we use the notation $[t_1/x_1, \ldots, t_n/x_n]$ which has been used for some time in automated theorem proving. Then, applying a substitution to M is denoted as $M[t_1/x_1, \ldots, t_n/x_n]$ (think for just a second of the horrible clash if this notation was used with $[x{:}\,\sigma]M$ for abstraction!). Other authors denote substitutions as $[x_1{:=}\, M_1, \ldots, x_n{:=}\, M_n]$. Personally, we would prefer switching to $[x_1 \mapsto M_1, \ldots, x_n \mapsto M_n]$, because $:=$ is also used for denoting a function f whose value at some argument x is redefined to be a, as in $f[x{:=}\,a]$. Finally, a word about sequents and judgements. To us, the turnstile symbol $\vdash$ means provability. A sequent consists of two parts Γ and Δ, and some separator is needed between them. In principle, anything can do, and if the arrow $\rightarrow$ was not already used as a type-constructor, we would adopt the notation $\Gamma \rightarrow \Delta$. Some authors denote sequents as $\Gamma \vdash \Delta$. A problem then arises when we want to say that a sequent is provable, since this is written as $\vdash \Gamma \vdash \Delta$. The ideal is to use symbols of different size for the two uses of $\vdash$. In fact, we noticed that Girard himself has designed his own $\vdash$ which has a thicker but smaller (in height) foot: $\vdash$. Thus, we will use the "Girardian turnstile" $\vdash$ in writing sequents as $\Gamma \vdash \Delta$. Judgements have three parts, Γ, M, and σ. Our view is that Γ and M actually come together to form what we have called elsewhere a "declared term" (thinking of the context Γ as a declaration of the variables). Again we need a way to put together Γ and M, and we use the symbol $\triangleright$, thus forming $\Gamma \triangleright M$. Then, a declared term may have a type σ, and such a judgement is written as $\Gamma \triangleright M{:}\,\sigma$. To say that a judgement is provable, we write $\vdash \Gamma \triangleright M{:}\,\sigma$. We find this less confusing than the notation $\vdash \Gamma \vdash M{:}\,\sigma$, and this is why we favor $\Gamma \triangleright M{:}\,\sigma$ over $\Gamma \vdash M{:}\,\sigma$ (but some authors use $\triangleright$ for the reduction relation! We use $\longrightarrow$). And please, avoid the notation $\vdash \Gamma \vdash M \in \sigma$, which we find terribly confusing and cruel to $\in$. But we have indulged too long into this digression, and now back to more serious business.

3 Adding Conjunction, Negation, and Disjunction

First, we present the natural deduction systems, and then the corresponding extensions of the simply-typed λ-calculus. As far as proof normalization is concerned, conjunction does not cause any problem, but as we will see, negation and disjunction are more problematic. In order to add negation, we add the new constant $\perp$ (**false**) to the language, and define negation $\neg A$ as an abbreviation for $A \supset \perp$.

Definition 3.1 The axioms and inference rules of the system $\mathcal{N}_i^{\supset,\wedge,\vee,\perp}$ (intuitionistic propositional logic) are listed below:

$$\Gamma, x{:}\,A \vdash A$$

$$\frac{\Gamma \vdash \perp}{\Gamma \vdash A} \quad (\perp\text{-}elim)$$

$$\frac{\Gamma, x{:}\,A \vdash B}{\Gamma \vdash A \supset B} \quad (\supset\text{-}intro)$$

$$\frac{\Gamma \vdash A \supset B \quad \Gamma \vdash A}{\Gamma \vdash B} \quad (\supset\text{-}elim)$$

$$\frac{\Gamma \vdash A \quad \Gamma \vdash B}{\Gamma \vdash A \wedge B} \quad (\wedge\text{-}intro)$$

$$\frac{\Gamma \vdash A \wedge B}{\Gamma \vdash A} \quad (\wedge\text{-}elim) \qquad \frac{\Gamma \vdash A \wedge B}{\Gamma \vdash B} \quad (\wedge\text{-}elim)$$

$$\frac{\Gamma \vdash A}{\Gamma \vdash A \vee B} \quad (\vee\text{-}intro) \qquad \frac{\Gamma \vdash B}{\Gamma \vdash A \vee B} \quad (\vee\text{-}intro)$$

$$\frac{\Gamma \vdash A \vee B \quad \Gamma, x\colon A \vdash C \quad \Gamma, y\colon B \vdash C}{\Gamma \vdash C} \quad (\vee\text{-}elim)$$

Since the rule ($\perp$-*elim*) is trivial (does nothing) when $A = \perp$, from now on, we will assume that $A \neq \perp$. *Minimal propositional logic* $\mathcal{N}_m^{\supset,\wedge,\vee,\perp}$ is obtained by dropping the ($\perp$-*elim*) rule. In order to obtain the system of *classical propositional logic*, denoted $\mathcal{N}_c^{\supset,\wedge,\vee,\perp}$, we add to $\mathcal{N}_m^{\supset,\wedge,\vee,\perp}$ the following inference rule corresponding to the principle of proof by contradiction (*by-contra*) (also called *reductio ad absurdum*).

$$\frac{\Gamma, x\colon \neg A \vdash \perp}{\Gamma \vdash A} \quad (by\text{-}contra)$$

Several useful remarks should be made.

(1) In classical propositional logic ($\mathcal{N}_c^{\supset,\wedge,\vee,\perp}$), the rule

$$\frac{\Gamma \vdash \perp}{\Gamma \vdash A} \quad (\perp\text{-}elim)$$

can be derived, since if we have a deduction of $\Gamma \vdash \perp$, then for any arbitrary A we have a deduction $\Gamma, x\colon \neg A \vdash \perp$, and thus a deduction of $\Gamma \vdash A$ by applying the (*by-contra*) rule.

(2) The proposition $A \supset \neg\neg A$ is derivable in $\mathcal{N}_m^{\supset,\wedge,\vee,\perp}$, but the reverse implication $\neg\neg A \supset A$ is not derivable, even in $\mathcal{N}_i^{\supset,\wedge,\vee,\perp}$. On the other hand, $\neg\neg A \supset A$ is derivable in $\mathcal{N}_c^{\supset,\wedge,\vee,\perp}$:

$$\cfrac{\cfrac{\cfrac{x\colon \neg\neg A, y\colon \neg A \vdash \neg\neg A \qquad x\colon \neg\neg A, y\colon \neg A \vdash \neg A}{x\colon \neg\neg A, y\colon \neg A \vdash \perp}}{x\colon \neg\neg A \vdash A}\ (by\text{-}contra)}{\vdash \neg\neg A \supset A}$$

(3) Using the (*by-contra*) inference rule together with ($\supset$-*elim*) and ($\vee$-*intro*), we can prove $\neg A \vee A$ (that is, $(A \supset \perp) \vee A$). Let

$$\Gamma = x\colon ((A \supset \perp) \vee A) \supset \perp .$$

We have the following proof for $(A \supset \perp) \vee A$ in $\mathcal{N}_c^{\supset,\wedge,\vee,\perp}$:

$$
\dfrac{\dfrac{\Gamma \vdash ((A \supset \perp) \vee A) \supset \perp \qquad \dfrac{\dfrac{\dfrac{\Gamma, y\colon A \vdash ((A \supset \perp) \vee A) \supset \perp \qquad \dfrac{\Gamma, y\colon A \vdash A}{\Gamma, y\colon A \vdash (A \supset \perp) \vee A}}{\Gamma, y\colon A \vdash \perp}}{\Gamma \vdash A \supset \perp}}{\Gamma \vdash (A \supset \perp) \vee A}}{\Gamma \vdash \perp}}{\vdash (A \supset \perp) \vee A}\ (by\text{-}contra)
$$

As in (2), $\neg A \vee A$ is not derivable in $\mathcal{N}_i^{\supset,\wedge,\vee,\perp}$. The reader might wonder how one shows that $\neg\neg A \supset A$ and $\neg A \vee A$ are not provable in $\mathcal{N}_i^{\supset,\wedge,\vee,\perp}$. In fact, this is not easy to prove directly. One method is to use the fact (given by theorem 3.4 and theorem 3.5) that every proof-term reduces to a unique normal form. Then, argue that if the above propositions have a proof in normal form, this leads to a contradiction. Another even simpler method is to use cut-free Gentzen systems, to be discussed in sections 4, 8, and 9.

The typed λ-calculus $\lambda^{\rightarrow,\times,+,\perp}$ corresponding to $\mathcal{N}_i^{\supset,\wedge,\vee,\perp}$ is given in the following definition.

Definition 3.2 The typed λ-calculus $\lambda^{\rightarrow,\times,+,\perp}$ is defined by the following rules.

$$\Gamma, x\colon A \triangleright x\colon A$$

$$\dfrac{\Gamma \triangleright M\colon \perp}{\Gamma \triangleright \nabla_A(M)\colon A}\quad (\perp\text{-}elim)$$

with $A \neq \perp$,

$$\dfrac{\Gamma, x\colon A \triangleright M\colon B}{\Gamma \triangleright (\lambda x\colon A.\, M)\colon A \rightarrow B}\quad (abstraction)$$

$$\dfrac{\Gamma \triangleright M\colon A \rightarrow B \qquad \Gamma \triangleright N\colon A}{\Gamma \triangleright (MN)\colon B}\quad (application)$$

$$\dfrac{\Gamma \triangleright M\colon A \qquad \Gamma \triangleright N\colon B}{\Gamma \triangleright \langle M, N\rangle\colon A \times B}\quad (pairing)$$

$$\dfrac{\Gamma \triangleright M\colon A \times B}{\Gamma \triangleright \pi_1(M)\colon A}\quad (projection) \qquad \dfrac{\Gamma \triangleright M\colon A \times B}{\Gamma \triangleright \pi_2(M)\colon B}\quad (projection)$$

$$\dfrac{\Gamma \triangleright M\colon A}{\Gamma \triangleright \mathtt{inl}(M)\colon A + B}\quad (injection) \qquad \dfrac{\Gamma \triangleright M\colon B}{\Gamma \triangleright \mathtt{inr}(M)\colon A + B}\quad (injection)$$

$$\dfrac{\Gamma \triangleright P\colon A + B \qquad \Gamma, x\colon A \triangleright M\colon C \qquad \Gamma, y\colon B \triangleright N\colon C}{\Gamma \triangleright \mathtt{case}(P, \lambda x\colon A.\, M, \lambda y\colon B.\, N)\colon C}\quad (by\text{-}cases)$$

A syntactic variant of $\texttt{case}(P, \lambda x{:}\,A.\,M, \lambda y{:}\,B.\,N)$ often found in the literature is

$$\texttt{case}\ P\ \texttt{of}\ \texttt{inl}(x{:}\,A) \Rightarrow M \mid \texttt{inr}(y{:}\,B) \Rightarrow N,$$

or even

$$\texttt{case}\ P\ \texttt{of}\ \texttt{inl}(x) \Rightarrow M \mid \texttt{inr}(y) \Rightarrow N,$$

and the (*by-cases*) rule can be written as

$$\frac{\Gamma \triangleright P{:}\,A+B \quad \Gamma, x{:}\,A \triangleright M{:}\,C \quad \Gamma, y{:}\,B \triangleright N{:}\,C}{\Gamma \triangleright (\texttt{case}\ P\ \texttt{of}\ \texttt{inl}(x{:}\,A) \Rightarrow M \mid \texttt{inr}(y{:}\,B) \Rightarrow N){:}\,C} \quad (\textit{by-cases})$$

We also have the following reduction rules.

Definition 3.3 The reduction rules of the system $\lambda^{\to,\times,+,\perp}$ are listed below:

$$\begin{aligned}
(\lambda x{:}\,A.\,M)N &\longrightarrow M[N/x],\\
\pi_1(\langle M, N\rangle) &\longrightarrow M,\\
\pi_2(\langle M, N\rangle) &\longrightarrow N,\\
\texttt{case}(\texttt{inl}(P), \lambda x{:}\,A.\,M, \lambda y{:}\,B.\,N) &\longrightarrow M[P/x], \quad \text{or}\\
\texttt{case}\ \texttt{inl}(P)\ \texttt{of}\ \texttt{inl}(x{:}\,A) \Rightarrow M \mid \texttt{inr}(y{:}\,B) \Rightarrow N &\longrightarrow M[P/x],\\
\texttt{case}(\texttt{inr}(P), \lambda x{:}\,A.\,M, \lambda y{:}\,B.\,N) &\longrightarrow N[P/y], \quad \text{or}\\
\texttt{case}\ \texttt{inr}(P)\ \texttt{of}\ \texttt{inl}(x{:}\,A) \Rightarrow M \mid \texttt{inr}(y{:}\,B) \Rightarrow N &\longrightarrow N[P/y],\\
\nabla_{A\to B}(M)N &\longrightarrow \nabla_B(M),\\
\pi_1(\nabla_{A\times B}(M)) &\longrightarrow \nabla_A(M),\\
\pi_2(\nabla_{A\times B}(M)) &\longrightarrow \nabla_B(M),\\
\texttt{case}(\nabla_{A+B}(P), \lambda x{:}\,A.\,M, \lambda y{:}\,B.\,N) &\longrightarrow \nabla_C(P).
\end{aligned}$$

Alternatively, as suggested by Ascánder Suárez, we could replace the rules for $\texttt{case}$ by the rules

$$\begin{aligned}
\texttt{case}(\texttt{inl}(P), M, N) &\longrightarrow MP,\\
\texttt{case}(\texttt{inr}(P), M, N) &\longrightarrow NP,\\
\texttt{case}(\nabla_{A+B}(P), M, N) &\longrightarrow \nabla_C(P).
\end{aligned}$$

A fundamental result about natural deduction is the fact that every proof (term) reduces to a normal form, which is unique up to α-renaming. This result was first proved by Prawitz [24] for the system $\mathcal{N}_i^{\supset,\wedge,\vee,\perp}$.

Theorem 3.4 [Church-Rosser property, Prawitz (1971)] *Reduction in $\lambda^{\to,\times,+,\perp}$ (specified in Definition 3.3) is confluent. Equivalently, conversion in $\lambda^{\to,\times,+,\perp}$ is Church-Rosser.*

A proof can be given by adapting the method of Tait and Martin-Löf [21] using a form of parallel reduction (see also Barendregt [2], Hindley and Seldin [15], or Stenlund [27]).

Theorem 3.5 [Strong normalization property, Prawitz (1971)] *Reduction in $\lambda^{\rightarrow,\times,+,\perp}$ (as in Definition 3.3) is strongly normalizing.*

A proof can be given by adapting Tait's reducibility method [28], [30], as done in Girard [10] (1971), [11] (1972) (see also Gallier [7]).

If one looks at the rules of the systems $\mathcal{N}^{\supset,\wedge,\vee,\perp}$ (or $\lambda^{\rightarrow,\times,+,\perp}$), one notices a number of unpleasant features:

(1) There is an *asymmetry* between the lefthand side and the righthand side of a sequent (or judgement): the righthand side must consist of a single formula, but the lefthand side may have any finite number of assumptions. This is typical of intuitionistic logic (and it is one of the major characteristics of its sequent-calculus formulations, see section 4) but it is also a defect.

(2) Negation is very badly handled, only in an indirect fashion.

(3) The ($\supset$-*intro*) rule and the ($\vee$-*elim*) rule are global rules requiring the discharge of assumptions.

(4) Worse of all, the ($\vee$-*elim*) rule contains the parasitic formula C which has nothing to do with the disjunction being eliminated.

Finally, note that it is quite difficult to search for proofs in such a system. Gentzen's sequent systems remedy some of these problems.

4 Gentzen's Sequent Calculi

The main idea is that now, a sequent $\Gamma \vdash \Delta$ consists of two finite multisets Γ and Δ of formulae, and that rather than having introduction and elimination rules, we have rules introducing a connective on the left or on the right of a sequent. A first version of such a system for classical propositional logic is given next. In these rules Γ and Δ stand for possibly empty finite multisets of propositions.

Definition 4.1 The axioms and inference rules of the system $\mathcal{G}_c^{\supset,\wedge,\vee,\neg}$ for classical propositional logic are given below.

$$A, \Gamma \vdash \Delta, A$$

$$\frac{A, A, \Gamma \vdash \Delta}{A, \Gamma \vdash \Delta}\ (\textit{contrac: left}) \qquad \frac{\Gamma \vdash \Delta, A, A}{\Gamma \vdash \Delta, A}\ (\textit{contrac: right})$$

$$\frac{A, B, \Gamma \vdash \Delta}{A \wedge B, \Gamma \vdash \Delta}\ (\wedge\text{: } \textit{left}) \qquad \frac{\Gamma \vdash \Delta, A \quad \Gamma \vdash \Delta, B}{\Gamma \vdash \Delta, A \wedge B}\ (\wedge\text{: } \textit{right})$$

$$\frac{A, \Gamma \vdash \Delta \quad B, \Gamma \vdash \Delta}{A \vee B, \Gamma \vdash \Delta}\ (\vee\text{: } \textit{left}) \qquad \frac{\Gamma \vdash \Delta, A, B}{\Gamma \vdash \Delta, A \vee B}\ (\vee\text{: } \textit{right})$$

$$\frac{\Gamma \vdash \Delta, A \quad B, \Gamma \vdash \Delta}{A \supset B, \Gamma \vdash \Delta}\ (\supset\text{: } \textit{left}) \qquad \frac{A, \Gamma \vdash \Delta, B}{\Gamma \vdash \Delta, A \supset B}\ (\supset\text{: } \textit{right})$$

$$\frac{\Gamma \vdash \Delta, A}{\neg A, \Gamma \vdash \Delta}\ (\neg\text{: } \textit{left}) \qquad \frac{A, \Gamma \vdash \Delta}{\Gamma \vdash \Delta, \neg A}\ (\neg\text{: } \textit{right})$$

Note the perfect symmetry of the left and right rules. If one wants to deal with the extended language containing also $\perp$, one needs to add the axiom

$$\perp, \Gamma \vdash \Delta.$$

One might be puzzled and even concerned about the presence of the contraction rule. Indeed, one might wonder whether the presence of this rule will not cause provability to be undecidable. This would certainly be quite bad, since we are only dealing with propositions! Fortunately, it can be shown that the contraction rule is redundant for classical propositional logic (see section 8). But then, why include it in the first place? The main reason is that it cannot be dispensed with in traditional systems for intuitionistic logic, or in the case of quantified formulae (however, in the case of propositional intuitionistic logic, it is possible to formulate contraction-free systems which easily yield the decidability of provability, see section 9). Since we would like to view intuitionistic logic as a subsystem of classical logic, we cannot eliminate the contraction rule from the presentation of classical systems. Another important reason is that the contraction rule plays an important role in cut elimination. Although it is possible to hide it by dealing with sequents viewed as pairs of sets rather than multisets, we prefer to deal with it explicitly. Finally, the contraction rule plays a crucial role in linear logic, and in the understanding of the correspondence between proofs and computations, in particular strict versus lazy evaluation (see Abramsky [1]).

In order to obtain a system for intuitionistic logic, we restrict the righthand side of a sequent to consist of *at most one formula*. We also modify the ($\supset$: *left*) rule and the ($\vee$: *right*) rule which splits into two rules. The (*contrac*: *right*) rule disappears, and it is also necessary to add a rule of weakening on the right, to mimic the ($\perp$-*elim*) rule.

Definition 4.2 The axioms and inference rules of the system $\mathcal{G}_i^{\supset,\wedge,\vee,\neg}$ for intuitionistic propositional logic are given below.

$$A, \Gamma \vdash A$$

$$\frac{\Gamma \vdash}{\Gamma \vdash A} \quad (\textit{weakening: right})$$

$$\frac{A, A, \Gamma \vdash \Delta}{A, \Gamma \vdash \Delta} \quad (\textit{contrac: left})$$

$$\frac{A, B, \Gamma \vdash \Delta}{A \wedge B, \Gamma \vdash \Delta} \quad (\wedge\text{: } \textit{left}) \qquad \frac{\Gamma \vdash A \quad \Gamma \vdash B}{\Gamma \vdash A \wedge B} \quad (\wedge\text{: } \textit{right})$$

$$\frac{A, \Gamma \vdash \Delta \quad B, \Gamma \vdash \Delta}{A \vee B, \Gamma \vdash \Delta} \quad (\vee\text{: } \textit{left})$$

$$\frac{\Gamma \vdash A}{\Gamma \vdash A \vee B} \quad (\vee\text{: } \textit{right}) \qquad \frac{\Gamma \vdash B}{\Gamma \vdash A \vee B} \quad (\vee\text{: } \textit{right})$$

$$\frac{\Gamma \vdash A \quad B, \Gamma \vdash \Delta}{A \supset B, \Gamma \vdash \Delta} \quad (\supset\text{: } \textit{left}) \qquad \frac{A, \Gamma \vdash B}{\Gamma \vdash A \supset B} \quad (\supset\text{: } \textit{right})$$

$$\frac{\Gamma \vdash A}{\neg A, \Gamma \vdash} \quad (\neg\text{: } \textit{left}) \qquad \frac{A, \Gamma \vdash}{\Gamma \vdash \neg A} \quad (\neg\text{: } \textit{right})$$

In the above rules, Δ contains at most one formula. If one wants to deal with the extended language containing also $\perp$, one simply needs to add the axiom

$$\perp, \Gamma \vdash \Delta,$$

where again, Δ contains at most one formula. If we choose the language restricted to formulae over $\wedge, \supset, \vee$, and $\perp$, then negation $\neg A$ is viewed as an abbreviation for $A \supset \perp$. Such a system can be simplified a little bit if we observe that the axiom $\perp, \Gamma \vdash \Delta$ implies that the rule

$$\frac{\Gamma \vdash \perp}{\Gamma \vdash A}$$

with $A \neq \perp$ is derivable. Indeed, assume that we have the axiom $\perp, \Gamma \vdash \Delta$. If $\Gamma \vdash \perp$ is provable, inspection of the inference rules shows that the proof must contain some leaf nodes of the form $\Gamma' \vdash \perp$. Since these leaves are axioms, we must have $\perp \in \Gamma'$, in which case $\Gamma' \vdash A$ is also an axiom. A simple induction shows that we obtain a proof of $\Gamma \vdash A$ by replacing all occurrences of $\perp$ on the righthand side of $\vdash$ by A. We can also prove that the converse almost holds. Since $\perp, \Gamma \vdash \perp$ is an axiom, using the rule

$$\frac{\perp, \Gamma \vdash \perp}{\perp, \Gamma \vdash A}$$

we see that $\perp, \Gamma \vdash A$ is provable. The reason why this is not exactly the converse is that $\perp, \Gamma \vdash$ is not provable in this system. This suggests to consider sequents of the form $\Gamma \vdash A$ where A consists *exactly of a single formula*. In this case, the axiom $\perp, \Gamma \vdash A$ is equivalent to the rule

$$\frac{\Gamma \vdash \perp}{\Gamma \vdash A} \quad (\perp\colon \mathit{right})$$

(with $A \neq \perp$). We have the following system.

Definition 4.3 The axioms and inference rules of the system $\mathcal{G}_i^{\supset,\wedge,\vee,\perp}$ for intuitionistic propositional logic are given below.

$$A, \Gamma \vdash A$$

$$\frac{\Gamma \vdash \perp}{\Gamma \vdash A} \quad (\perp\colon \mathit{right})$$

with $A \neq \perp$,

$$\frac{A, A, \Gamma \vdash C}{A, \Gamma \vdash C} \quad (\mathit{contrac}\colon \mathit{left})$$

$$\frac{A, B, \Gamma \vdash C}{A \wedge B, \Gamma \vdash C} \quad (\wedge\colon \mathit{left}) \qquad \frac{\Gamma \vdash A \quad \Gamma \vdash B}{\Gamma \vdash A \wedge B} \quad (\wedge\colon \mathit{right})$$

$$\frac{A, \Gamma \vdash C \quad B, \Gamma \vdash C}{A \vee B, \Gamma \vdash C} \quad (\vee\colon \mathit{left})$$

$$\frac{\Gamma \vdash A}{\Gamma \vdash A \vee B} \quad (\vee\colon \mathit{right}) \qquad \frac{\Gamma \vdash B}{\Gamma \vdash A \vee B} \quad (\vee\colon \mathit{right})$$

$$\frac{\Gamma \vdash A \quad B, \Gamma \vdash C}{A \supset B, \Gamma \vdash C} \quad (\supset\colon \mathit{left}) \qquad \frac{A, \Gamma \vdash B}{\Gamma \vdash A \supset B} \quad (\supset\colon \mathit{right})$$

There is a close relationship between the natural deduction system $\mathcal{N}_i^{\supset,\wedge,\vee,\perp}$ and the Gentzen system $\mathcal{G}_i^{\supset,\wedge,\vee,\perp}$. In fact, there is a procedure $\mathcal{N}$ for translating every proof in $\mathcal{G}_i^{\supset,\wedge,\vee,\perp}$ into a deduction in $\mathcal{N}_i^{\supset,\wedge,\vee,\perp}$. The procedure $\mathcal{N}$ has the remarkable property that $\mathcal{N}(\Pi)$ is a deduction in normal form for every proof Π. Since there are deductions in $\mathcal{N}_i^{\supset,\wedge,\vee,\perp}$ that are not in normal form, the function $\mathcal{N}$ is *not* surjective. The situation can be repaired by adding a new rule to $\mathcal{G}_i^{\supset,\wedge,\vee,\perp}$, the *cut rule*. Then, there is a procedure $\mathcal{N}$ mapping every proof in $\mathcal{G}_i^{\supset,\wedge,\vee,\perp}$ to a deduction in $\mathcal{N}_i^{\supset,\wedge,\vee,\perp}$, and a procedure $\mathcal{G}$ mapping every deduction in $\mathcal{N}_i^{\supset,\wedge,\vee,\perp}$ to a proof in $\mathcal{G}_i^{\supset,\wedge,\vee,\perp,cut}$.

In order to close the loop, we would need to show that every proof in $\mathcal{G}_i^{\supset,\wedge,\vee,\perp,cut}$ can be transformed into a proof in $\mathcal{G}_i^{\supset,\wedge,\vee,\perp}$, that is, a cut-free proof. It is an extremely interesting and deep fact that the system $\mathcal{G}_i^{\supset,\wedge,\vee,\perp,cut}$ and the system $\mathcal{G}_i^{\supset,\wedge,\vee,\perp}$ are indeed equivalent. This fundamental result known as the *cut elimination theorem* was first proved by Gentzen in 1935 [8]. The proof actually gives an algorithm for converting a proof with cuts into a cut-free proof. The main difficulty is to prove that this algorithm terminates. Gentzen used a fairly complex induction measure which was later simplified by Tait [29].

The contraction rule plays a crucial role in the proof of this theorem, and it is therefore natural to believe that this rule cannot be dispensed with. This is indeed true for the intuitionistic system $\mathcal{G}_i^{\supset,\wedge,\vee,\perp}$ (but it can be dispensed with in the classical system $\mathcal{G}_c^{\supset,\wedge,\vee,\neg}$). If we delete the contraction rule from the system $\mathcal{G}_i^{\supset,\wedge,\vee,\perp}$ (or $\mathcal{G}_i^{\supset,\wedge,\vee,\neg}$), certain formulae are no longer provable. For example, $\vdash \neg\neg(P \vee \neg P)$ is provable in $\mathcal{G}_i^{\supset,\wedge,\vee,\neg}$, but it is impossible to build a cut-free proof for it without using (*contrac*: *left*). Indeed, the only way to build a cut-free proof for $\vdash \neg\neg(P \vee \neg P)$ without using (*contrac*: *left*) is to proceed as follows:

$$\dfrac{\dfrac{\vdash P \vee \neg P}{\neg(P \vee \neg P) \vdash}}{\vdash \neg\neg(P \vee \neg P)}$$

Since the only rules that could yield a cut-free proof of $\vdash P \vee \neg P$ are the ($\vee$: *right*) rules and neither $\vdash P$ nor $\vdash \neg P$ is provable, it is clear that there is no cut-free proof of $\vdash P \vee \neg P$.

However, $\vdash \neg\neg(P \vee \neg P)$ is provable in $\mathcal{G}_i^{\supset,\wedge,\vee,\neg}$, as shown by the following proof (the same example can be worked out in $\mathcal{G}_i^{\supset,\wedge,\vee,\perp}$):

Example 4.4

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{P \vdash P}{P \vdash P \vee \neg P}}{P, \neg(P \vee \neg P) \vdash}}{\neg(P \vee \neg P) \vdash \neg P}}{\neg(P \vee \neg P) \vdash P \vee \neg P}}{\neg(P \vee \neg P), \neg(P \vee \neg P) \vdash}\ (\textit{contrac}\text{: } \textit{left})}{\neg(P \vee \neg P) \vdash}}{\vdash \neg\neg(P \vee \neg P)}$$

Nevertheless, it is possible to formulate a cut-free system $\mathcal{GK}_i^{\supset,\wedge,\vee,\perp}$ which is equivalent to $\mathcal{G}_i^{\supset,\wedge,\vee,\perp}$ (see section 8). Such a system due to Kleene [18] has no contraction rule, and the premise of every sequent can be interpreted as a set as opposed to a multiset (furthermore, in the case of intuitionistic propositional logic, it is possible to design contraction-free systems which yield easily the decidability of provability, see section 9 for details).

5 Definition of the Transformation $\mathcal{N}$ from $\mathcal{G}_i$ to $\mathcal{N}_i$

The purpose of this section is to give a procedure $\mathcal{N}$ mapping every proof in $\mathcal{G}_i^{\supset,\wedge,\vee,\perp}$ to a deduction in $\mathcal{N}_i^{\supset,\wedge,\vee,\perp}$. The procedure $\mathcal{N}$ is defined by induction on the structure of proof trees and requires some preliminary definitions.

Definition 5.1 A proof tree Π in $\mathcal{G}_i^{\supset,\wedge,\vee,\perp}$ with root node $\Gamma \vdash C$ is denoted as

$$\begin{array}{c}\Pi\\ \Gamma \vdash C\end{array}$$

and similarly a deduction $\mathcal{D}$ in $\mathcal{N}_i^{\supset,\wedge,\vee,\perp}$ with root node $\Gamma \vdash C$ is denoted as

$$\begin{array}{c}\mathcal{D}\\ \Gamma \vdash C\end{array}$$

A proof tree Π whose last inference is

$$\frac{\Gamma \vdash B}{\Delta \vdash D}$$

is denoted as

$$\begin{array}{c}\Pi_1\\ \dfrac{\Gamma \vdash B}{\Delta \vdash D}\end{array}$$

where Π_1 is the immediate subproof of Π whose root is $\Gamma \vdash B$, and a proof tree Π whose last inference is

$$\frac{\Gamma \vdash B \quad \Gamma \vdash C}{\Delta \vdash D}$$

is denoted as

$$\frac{\begin{array}{c}\Pi_1\\ \Gamma \vdash B\end{array} \qquad\qquad \begin{array}{c}\Pi_2\\ \Gamma \vdash C\end{array}}{\Delta \vdash D}$$

where Π_1 and Π_2 are the immediate subproofs of Π whose roots are $\Gamma \vdash B$ and $\Gamma \vdash C$, respectively. A similar notation applies to deductions.

Given a proof tree Π with root node $\Gamma \vdash C$,

$$\begin{array}{c}\Pi\\ \Gamma \vdash C\end{array}$$

$\mathcal{N}$ yields a deduction $\mathcal{N}(\Pi)$ of C from the set of assumptions Γ^+,

$$\begin{array}{c}\mathcal{N}(\Pi)\\ \Gamma^+ \vdash C\end{array}$$

where Γ^+ is obtained from the multiset Γ. However, one has to exercise some care in defining Γ^+ so that $\mathcal{N}$ is indeed a function. This can be achieved as follows. We can assume that we have a fixed total order $\leq_p$ on the set of all propositions so that they can be enumerated as $P_1, P_2, \ldots$, and a fixed total order $\leq_v$ on the set of all variables so that they can be enumerated as $x_1, x_2, \ldots$.

Definition 5.2 Given a multiset $\Gamma = A_1, \ldots, A_n$, since $\{A_1, \ldots, A_n\} = \{P_{i_1}, \ldots, P_{i_n}\}$ where $P_{i_1} \leq_p P_{i_2} \leq_p \ldots \leq_p P_{i_n}$ (where $P_1, P_2, \ldots$, is the enumeration of all propositions and where $i_j = i_{j+1}$ is possible since Γ is a multiset), we define Γ^+ as the set $\Gamma^+ = x_1 : P_{i_1}, \ldots, x_n : P_{i_n}$.

We will also need the following concepts and notation.

Definition 5.3 Given a deduction

$$\begin{array}{c}\mathcal{D}\\ \Gamma \vdash C\end{array}$$

the deduction obtained by adding the additional assumptions Δ to the lefthand side of every sequent of $\mathcal{D}$ is denoted as $\Delta + \mathcal{D}$, and it is only well defined provided that $\mathrm{dom}(\Gamma') \cap \mathrm{dom}(\Delta) = \emptyset$ for every sequent $\Gamma' \vdash A$ occurring in $\mathcal{D}$.[3] Similarly, given a sequential proof

$$\begin{array}{c}\Pi\\ \Gamma \vdash \Delta\end{array}$$

we define the proof $\Lambda + \Pi$ by adding Λ to the lefthand side of every sequent of Π, and we define the proof $\Pi + \Theta$ by adding Θ to the righthand side of every sequent of Π.

We also need a systematic way of renaming the variables in a deduction.

Definition 5.4 Given a deduction $\mathcal{D}$ with root node $\Delta \vdash C$ the deduction $\mathcal{D}'$ obtained from $\mathcal{D}$ by *rectification* is defined inductively as follows:

Given a context $\Delta = y_1 : A_1, \ldots, y_m : A_m$, define the total order $<$ on Δ as follows:

$$y_i : A_i < y_j : A_j \quad \text{iff} \quad \begin{cases} A_i <_p A_j, & \text{or} \\ A_i = A_j \quad \text{and} \quad y_i <_v y_j. \end{cases}$$

[3]Given a context $\Gamma = x_1 : A_1, \ldots, x_n : A_n$, we let $\mathrm{dom}(\Gamma) = \{x_1, \ldots, x_n\}$.

The order $<$ on $y_1\colon A_1, \ldots, y_m\colon A_m$ defines the permutation σ such that

$$y_{\sigma(1)}\colon A_{\sigma(1)} < y_{\sigma(2)}\colon A_{\sigma(2)} < \ldots < y_{\sigma(m-1)}\colon A_{\sigma(m-1)} < y_{\sigma(m)}\colon A_{\sigma(m)}.$$

If $\mathcal{D}$ consists of the single node $y_1\colon A_1, \ldots, y_m\colon A_m \vdash C$, let $\Delta' = x_1\colon A_{\sigma(1)}, \ldots, x_m\colon A_{\sigma(m)}$, and define $\mathcal{D}'$ as $\Delta' \vdash C$. The permutation σ induces a bijection between $\{x_1, \ldots, x_m\}$ and $\{y_1, \ldots, y_m\}$, namely $x_i \mapsto y_{\sigma(i)}$.

If $\mathcal{D}$ is of the form

$$\frac{\begin{array}{c}\mathcal{D}_1 \\ y_1\colon A_1, y_2\colon A_2, \ldots, y_m\colon A_m \vdash B\end{array}}{y_2\colon A_2, \ldots, y_m\colon A_m \vdash A_1 \supset B}$$

by induction, we have the rectified deduction

$$\begin{array}{c}\mathcal{D}_1' \\ x_1\colon A_{\sigma(1)}, \ldots, x_{j-1}\colon A_{\sigma(j-1)}, x_j\colon A_1, x_{j+1}\colon A_{\sigma(j+1)}, \ldots, x_m\colon A_{\sigma(m)} \vdash B\end{array}$$

where x_j corresponds to y_1 in the bijection between $\{x_1, \ldots, x_m\}$ and $\{y_1, \ldots, y_m\}$ (in fact, $j = \sigma^{-1}(1)$ since $A_1 = A_{\sigma(j)}$). Then, apply the substitution $[x_m/x_j, x_j/x_{j+1}, \ldots, x_{m-1}/x_m]$ to the deduction $\mathcal{D}_1'$, and form the deduction

$$\frac{\begin{array}{c}\mathcal{D}_1'[x_m/x_j, x_j/x_{j+1}, \ldots, x_{m-1}/x_m] \\ x_1\colon A_{\sigma(1)}, \ldots, x_{j-1}\colon A_{\sigma(j-1)}, x_m\colon A_1, x_j\colon A_{\sigma(j+1)}, \ldots, x_{m-1}\colon A_{\sigma(m)} \vdash B\end{array}}{x_1\colon A_{\sigma(1)}, \ldots, x_{j-1}\colon A_{\sigma(j-1)}, x_j\colon A_{\sigma(j+1)}, \ldots, x_{m-1}\colon A_{\sigma(m)} \vdash A_1 \supset B}$$

A similar construction applies to the rule ($\vee$-*elim*) and is left as an exercise to the reader. The other inference rules do not modify the lefthand side of sequents, and $\mathcal{D}'$ is obtained by rectifying the immediate subtree(s) of $\mathcal{D}$.

Note that for any deduction $\mathcal{D}$ with root node $y_1\colon A_1, \ldots, y_m\colon A_m \vdash C$, the rectified deduction $\mathcal{D}'$ has for its root node the sequent $\Gamma^+ \vdash C$, where Γ^+ is obtained from the multiset $\Gamma = A_1, \ldots, A_m$ as in Definition 5.2.

The procedure $\mathcal{N}$ is defined by induction on the structure of the proof tree Π.

- An axiom $\Gamma, A \vdash A$ is mapped to the deduction $(\Gamma, A)^+ \vdash A$.
- A proof Π of the form

$$\frac{\begin{array}{c}\Pi_1 \\ \Gamma \vdash \perp\end{array}}{\Gamma \vdash A}$$

is mapped to the deduction

$$\dfrac{\begin{array}{c}\mathcal{N}(\Pi_1)\\ \Gamma^+ \vdash \perp\end{array}}{\Gamma^+ \vdash A}$$

- A proof Π of the form

$$\dfrac{\begin{array}{c}\Pi_1\\ A, A, \Gamma \vdash B\end{array}}{A, \Gamma \vdash B}$$

is mapped to a deduction as follows. First map Π_1 to the deduction $\mathcal{N}(\Pi_1)$

$$\begin{array}{c}\mathcal{N}(\Pi_1)\\ x\colon A, y\colon A, \Gamma^* \vdash B\end{array}$$

Next, replace every occurrence of "$x\colon A, y\colon A$" in $\mathcal{N}(\Pi_1)$ by "$z\colon A$" where z is a new variable not occurring in $\mathcal{N}(\Pi_1)$, and finally rectify the resulting tree.

Before we proceed any further, a sticky point needs to be clarified regarding the context $x\colon A, y\colon A, \Gamma^*$. In the above transformation, the multiset A, A, Γ is mapped to $(A, A, \Gamma)^+$ under the operation $^+$. Unfortunately, we cannot assume in general that $(A, A, \Gamma)^+ = x\colon A, y\colon A, \Gamma^+$, because in the enumeration of the propositions forming the multiset A, A, Γ, the two occurrences of A may not appear as the last two elements. Thus, $(A, A, \Gamma)^+$ is of the form $x\colon A, y\colon A, \Gamma^*$ for some Γ^* not necessarily equal to Γ^+. This point being cleared up, we will use the notation Γ^* in the rest of the construction without any further comments.

- A proof Π of the form

$$\dfrac{\begin{array}{c}\Pi_1\\ \Gamma \vdash A\end{array} \qquad \begin{array}{c}\Pi_2\\ \Gamma \vdash B\end{array}}{\Gamma \vdash A \wedge B}$$

is mapped to the deduction

$$\dfrac{\begin{array}{c}\mathcal{N}(\Pi_1)\\ \Gamma^+ \vdash A\end{array} \qquad \begin{array}{c}\mathcal{N}(\Pi_2)\\ \Gamma^+ \vdash B\end{array}}{\Gamma^+ \vdash A \wedge B}$$

- A proof Π of the form

$$\dfrac{\begin{array}{c}\Pi_1\\ A, B, \Gamma \vdash C\end{array}}{A \wedge B, \Gamma \vdash C}$$

is mapped to a deduction obtained as follows. First, map Π_1 to $\mathcal{N}(\Pi_1)$

$$\begin{array}{c} \mathcal{N}(\Pi_1) \\ x: A, y: B, \Gamma^* \vdash C \end{array}$$

Next, replace every leaf of the form $x: A, y: B, \Delta, \Gamma^* \vdash A$ in $\mathcal{N}(\Pi_1)$ by the subtree

$$\frac{z: A \wedge B, \Delta, \Gamma^* \vdash A \wedge B}{z: A \wedge B, \Delta, \Gamma^* \vdash A}$$

and every leaf of the form $x: A, y: B, \Delta, \Gamma^* \vdash B$ in $\mathcal{N}(\Pi_1)$ by the subtree

$$\frac{z: A \wedge B, \Delta, \Gamma^* \vdash A \wedge B}{z: A \wedge B, \Delta, \Gamma^* \vdash B}$$

where z is new, replace "$x: A, y: B$" by "$z: A \wedge B$" in every antecedent of the resulting deduction, and rectify this last tree.

- A proof Π of the form

$$\frac{\begin{array}{c} \Pi_1 \\ A, \Gamma \vdash B \end{array}}{\Gamma \vdash A \supset B}$$

is mapped to the deduction

$$\frac{\begin{array}{c} \mathcal{N}(\Pi_1) \\ x: A, \Gamma^* \vdash B \end{array}}{\Gamma^* \vdash A \supset B}$$

which is then rectified.

- A proof Π of the form

$$\frac{\begin{array}{c} \Pi_1 \\ \Gamma \vdash A \end{array} \qquad \begin{array}{c} \Pi_2 \\ B, \Gamma \vdash C \end{array}}{A \supset B, \Gamma \vdash C}$$

is mapped to a deduction as follows. First map Π_1 and Π_2 to deductions $\mathcal{N}(\Pi_1)$

$$\begin{array}{c} \mathcal{N}(\Pi_1) \\ \Gamma^+ \vdash A \end{array}$$

and $\mathcal{N}(\Pi_2)$

$$\begin{array}{c} \mathcal{N}(\Pi_2) \\ x: B, \Gamma^* \vdash C \end{array}$$

Modify $\mathcal{N}(\Pi_1)$ so that it becomes a deduction $\mathcal{N}(\Pi_1)'$ with conclusion $\Gamma^* \vdash A$.

Next, form the deduction $\mathcal{D}$

$$\dfrac{z\colon A \supset B, \Gamma^* \vdash A \supset B \qquad \begin{array}{c} z\colon A \supset B + \mathcal{N}(\Pi_1)' \\ z\colon A \supset B, \Gamma^* \vdash A \end{array}}{z\colon A \supset B, \Gamma^* \vdash B}$$

and modify $\mathcal{N}(\Pi_2)$ as follows: replace every leaf of the form $x\colon B, \Delta, \Gamma^* \vdash B$ by the deduction which itself is obtained from $\Delta + \mathcal{D}$ by replacing "$x\colon B$" by "$z\colon A \supset B$" in the lefthand side of every sequent. Finally, rectify this last deduction.

- A proof Π of the form

$$\dfrac{\begin{array}{c} \Pi_1 \\ \Gamma \vdash A \end{array}}{\Gamma \vdash A \vee B}$$

is mapped to the deduction

$$\dfrac{\begin{array}{c} \mathcal{N}(\Pi_1) \\ \Gamma^+ \vdash A \end{array}}{\Gamma^+ \vdash A \vee B}$$

and similarly for the other case of the ($\vee$: *right*) rule.

- A proof Π of the form

$$\dfrac{\begin{array}{c} \Pi_1 \\ A, \Gamma \vdash C \end{array} \qquad \begin{array}{c} \Pi_2 \\ B, \Gamma \vdash C \end{array}}{A \vee B, \Gamma \vdash C}$$

is mapped to a deduction as follows. First map Π_1 and Π_2 to deductions $\mathcal{N}(\Pi_1)$

$$\begin{array}{c} \mathcal{N}(\Pi_1) \\ x\colon A, \Gamma_1^* \vdash C \end{array}$$

and $\mathcal{N}(\Pi_2)$

$$\begin{array}{c} \mathcal{N}(\Pi_2) \\ y\colon B, \Gamma_2^* \vdash C \end{array}$$

Since Γ_1^* and Γ_2^* may differ, construct deductions $\mathcal{N}(\Pi_1)'$ and $\mathcal{N}(\Pi_2)'$ with conclusions $x\colon A, \Gamma^* \vdash C$ and $y\colon B, \Gamma^* \vdash C$ for the same Γ^*. Next, form the deduction

$$\dfrac{z\colon A \vee B, \Gamma^* \vdash A \vee B \qquad \begin{array}{c} z\colon A \vee B + \mathcal{N}(\Pi_1)' \\ z\colon A \vee B, x\colon A, \Gamma^* \vdash C \end{array} \qquad \begin{array}{c} z\colon A \vee B + \mathcal{N}(\Pi_2)' \\ z\colon A \vee B, y\colon B, \Gamma^* \vdash C \end{array}}{z\colon A \vee B, \Gamma^* \vdash C}$$

and rectify this last tree.

This concludes the definition of the procedure $\mathcal{N}$. Note that the contraction rule can be stated in the system of natural deduction as follows:

$$\frac{x\colon A, y\colon A, \Gamma \vdash B}{z\colon A, \Gamma \vdash B}$$

where z is a new variable. The following remarkable property of $\mathcal{N}$ is easily shown.

Lemma 5.5 [Gentzen (1935), Prawitz (1965)] *For every proof* Π *in* $\mathcal{G}_i^{\supset,\wedge,\vee,\perp}$, $\mathcal{N}(\Pi)$ *is a deduction in normal form (in* $\mathcal{N}_i^{\supset,\wedge,\vee,\perp}$*).*

Since there are deductions in $\mathcal{N}_i^{\supset,\wedge,\vee,\perp}$ that are not in normal form, the function $\mathcal{N}$ is not surjective. It is interesting to observe that the function $\mathcal{N}$ is not injective either. What happens is that $\mathcal{G}_i^{\supset,\wedge,\vee,\perp}$ is more sequential than $\mathcal{N}_i^{\supset,\wedge,\vee,\perp}$, in the sense that the order of application of inferences is strictly recorded. Hence, two proofs in $\mathcal{G}_i^{\supset,\wedge,\vee,\perp}$ of the same sequent may differ for bureaucratic reasons: independent inferences are applied in different orders. In $\mathcal{N}_i^{\supset,\wedge,\vee,\perp}$, these differences disappear. The following example illustrates this point. The sequent $\vdash (P \wedge P') \supset ((Q \wedge Q') \supset (P \wedge Q))$ has the following two sequential proofs

$$\cfrac{\cfrac{\cfrac{\cfrac{\cfrac{P, P', Q, Q' \vdash P \qquad\qquad P, P', Q, Q' \vdash Q}{P, P', Q, Q' \vdash P \wedge Q}}{P \wedge P', Q, Q' \vdash P \wedge Q}}{P \wedge P', Q \wedge Q' \vdash P \wedge Q}}{P \wedge P' \vdash (Q \wedge Q') \supset (P \wedge Q)}}{\vdash (P \wedge P') \supset ((Q \wedge Q') \supset (P \wedge Q))}$$

and

$$\cfrac{\cfrac{\cfrac{\cfrac{\cfrac{P, P', Q, Q' \vdash P \qquad\qquad P, P', Q, Q' \vdash Q}{P, P', Q, Q' \vdash P \wedge Q}}{P, P', Q \wedge Q' \vdash P \wedge Q}}{P \wedge P', Q \wedge Q' \vdash P \wedge Q}}{P \wedge P' \vdash (Q \wedge Q') \supset (P \wedge Q)}}{\vdash (P \wedge P') \supset ((Q \wedge Q') \supset (P \wedge Q))}$$

Both proofs are mapped to the deduction

$$\dfrac{\dfrac{x\colon P\wedge P', y\colon Q\wedge Q' \vdash P\wedge P'}{x\colon P\wedge P', y\colon Q\wedge Q' \vdash P} \qquad \dfrac{x\colon P\wedge P', y\colon Q\wedge Q' \vdash Q\wedge Q'}{x\colon P\wedge P', y\colon Q\wedge Q' \vdash Q}}{\dfrac{x\colon P\wedge P', y\colon Q\wedge Q' \vdash P\wedge Q}{\dfrac{x\colon P\wedge P' \vdash (Q\wedge Q')\supset (P\wedge Q)}{\vdash (P\wedge P')\supset ((Q\wedge Q')\supset (P\wedge Q))}}}$$

6 Definition of the Transformation $\mathcal{G}$ from $\mathcal{N}_i$ to $\mathcal{G}_i$

We now show that if we add a new rule, the cut rule, to the system $\mathcal{G}_i^{\supset,\wedge,\vee,\perp}$, then we can define a procedure $\mathcal{G}$ mapping every deduction in $\mathcal{N}_i^{\supset,\wedge,\vee,\perp}$ to a proof in $\mathcal{G}_i^{\supset,\wedge,\vee,\perp,cut}$.

Definition 6.1 The system $\mathcal{G}_i^{\supset,\wedge,\vee,\perp,cut}$ is obtained from the system $\mathcal{G}_i^{\supset,\wedge,\vee,\perp}$ by adding the following rule, known as the *cut rule*:

$$\frac{\Gamma \vdash A \quad A, \Gamma \vdash C}{\Gamma \vdash C} \quad (cut)$$

The system $\mathcal{G}_c^{\supset,\wedge,\vee,\neg,cut}$ is obtained from $\mathcal{G}_c^{\supset,\wedge,\vee,\neg}$ by adding the following rule, also known as the *cut rule*:

$$\frac{\Gamma \vdash \Delta, A \quad A, \Gamma \vdash \Delta}{\Gamma \vdash \Delta} \quad (cut)$$

Next, we define the procedure $\mathcal{G}$ mapping every deduction in $\mathcal{N}_i^{\supset,\wedge,\vee,\perp}$ to a proof in $\mathcal{G}_i^{\supset,\wedge,\vee,\perp,cut}$. The procedure $\mathcal{G}$ is defined by induction on the structure of deduction trees. Given a deduction tree $\mathcal{D}$ of C from the assumptions Γ,

$$\begin{array}{c}\mathcal{D}\\ \Gamma \vdash C\end{array}$$

$\mathcal{G}$ yields a proof $\mathcal{G}(\mathcal{D})$ of the sequent $\Gamma^- \vdash C$

$$\begin{array}{c}\mathcal{G}(\mathcal{D})\\ \Gamma^- \vdash C\end{array}$$

where Γ^- is the multiset $A_1, \ldots, A_n$ obtained from the context $\Gamma = x_1\colon A_1, \ldots, x_n\colon A_n$ by erasing $x_1, \ldots, x_n$.

- The deduction $\Gamma, x\colon A \vdash A$ is mapped to the axiom $\Gamma^-, A \vdash A$.
- A deduction $\mathcal{D}$ of the form

$$\begin{array}{c}\mathcal{D}_1\\ \dfrac{\Gamma \vdash \perp}{\Gamma \vdash A}\end{array}$$

is mapped to the proof

$$\dfrac{\begin{array}{c}\mathcal{G}(\mathcal{D}_1)\\ \Gamma^- \vdash \perp\end{array}}{\Gamma^- \vdash A}$$

- A deduction $\mathcal{D}$ of the form

$$\dfrac{\begin{array}{c}\mathcal{D}_1\\ \Gamma \vdash A\end{array} \qquad \begin{array}{c}\mathcal{D}_2\\ \Gamma \vdash B\end{array}}{\Gamma \vdash A \wedge B}$$

is mapped to the proof

$$\dfrac{\begin{array}{c}\mathcal{G}(\mathcal{D}_1)\\ \Gamma^- \vdash A\end{array} \qquad \begin{array}{c}\mathcal{G}(\mathcal{D}_2)\\ \Gamma^- \vdash B\end{array}}{\Gamma^- \vdash A \wedge B}$$

- A deduction $\mathcal{D}$ of the form

$$\dfrac{\begin{array}{c}\mathcal{D}_1\\ \Gamma \vdash A \wedge B\end{array}}{\Gamma \vdash A}$$

is mapped to the proof

$$\dfrac{\begin{array}{c}\mathcal{G}(\mathcal{D}_1)\\ \Gamma^- \vdash A \wedge B\end{array} \qquad \dfrac{A, B, \Gamma^- \vdash A}{A \wedge B, \Gamma^- \vdash A}}{\Gamma^- \vdash A}\ (cut)$$

and similarly for the symmetric rule.

- A deduction $\mathcal{D}$ of the form

$$\dfrac{\begin{array}{c}\mathcal{D}_1\\ x\colon A, \Gamma \vdash B\end{array}}{\Gamma \vdash A \supset B}$$

is mapped to the proof

$$\dfrac{\begin{array}{c}\mathcal{G}(\mathcal{D}_1)\\ A, \Gamma^- \vdash B\end{array}}{\Gamma^- \vdash A \supset B}$$

- A deduction $\mathcal{D}$ of the form

$$\dfrac{\begin{array}{c}\mathcal{D}_1\\ \Gamma \vdash A \supset B\end{array} \qquad \begin{array}{c}\mathcal{D}_2\\ \Gamma \vdash A\end{array}}{\Gamma \vdash B}$$

is mapped to the proof

$$\dfrac{\begin{array}{c}\mathcal{G}(\mathcal{D}_1)\\ \Gamma^- \vdash A \supset B\end{array} \qquad \dfrac{\begin{array}{c}\mathcal{G}(\mathcal{D}_2)\\ \Gamma^- \vdash A\end{array} \qquad B, \Gamma^- \vdash B}{A \supset B, \Gamma^- \vdash B}}{\Gamma^- \vdash B}\ (cut)$$

• A deduction $\mathcal{D}$ of the form

$$\dfrac{\begin{array}{c}\mathcal{D}_1\\ \Gamma \vdash A\end{array}}{\Gamma \vdash A \vee B}$$

is mapped to the proof

$$\dfrac{\begin{array}{c}\mathcal{G}(\mathcal{D}_1)\\ \Gamma^- \vdash A\end{array}}{\Gamma^- \vdash A \vee B}$$

and similarly for the symmetric rule.

• A deduction $\mathcal{D}$ of the form

$$\dfrac{\begin{array}{c}\mathcal{D}_1\\ \Gamma \vdash A \vee B\end{array} \qquad \begin{array}{c}\mathcal{D}_2\\ x\colon A, \Gamma \vdash C\end{array} \qquad \begin{array}{c}\mathcal{D}_3\\ y\colon B, \Gamma \vdash C\end{array}}{\Gamma \vdash C}$$

is mapped to the proof

$$\dfrac{\begin{array}{c}\mathcal{G}(\mathcal{D}_1)\\ \Gamma^- \vdash A \vee B\end{array} \qquad \dfrac{\begin{array}{c}\mathcal{G}(\mathcal{D}_2)\\ A, \Gamma^- \vdash C\end{array} \qquad \begin{array}{c}\mathcal{G}(\mathcal{D}_3)\\ B, \Gamma^- \vdash C\end{array}}{A \vee B, \Gamma^- \vdash C}}{\Gamma^- \vdash C}\ (cut)$$

This concludes the definition of the procedure $\mathcal{G}$.

For the sake of completeness, we also extend the definition of the function $\mathcal{N}$ which is presently defined on the set of sequential proofs of the system $\mathcal{G}_i^{\supset,\wedge,\vee,\perp}$ to proofs with cuts, that is, to proofs in the system $\mathcal{G}_i^{\supset,\wedge,\vee,\perp,cut}$. A proof Π of the form

$$\dfrac{\begin{array}{c}\Pi_1\\ \Gamma \vdash A\end{array} \qquad\qquad \begin{array}{c}\Pi_2\\ A, \Gamma \vdash C\end{array}}{\Gamma \vdash C}$$

is mapped to the deduction obtained as follows: First, construct

$$\begin{array}{c}\mathcal{N}(\Pi_1)\\ \Gamma^+ \vdash A\end{array}$$

and

$$\begin{array}{c}\mathcal{N}(\Pi_2)\\ x\colon A, \Gamma^* \vdash C\end{array}$$

Modify $\mathcal{N}(\Pi_1)$ to a deduction $\mathcal{N}(\Pi_1)'$ with conclusion $\Gamma^* \vdash A$. Then, replace every leaf $x\colon A, \Delta, \Gamma^* \vdash A$ in $\mathcal{N}(\Pi_2)$ by $\Delta + \mathcal{N}(\Pi_1)'$, delete "$x\colon A$" from the antecedent in every sequent, and rectify this last tree.

7 First-Order Quantifiers

We extend the systems $\mathcal{N}_i^{\supset,\wedge,\vee,\perp}$ and $\mathcal{G}_i^{\supset,\wedge,\vee,\perp,cut}$ to deal with the quantifiers.

Definition 7.1 The axioms and inference rules of the system $\mathcal{N}_i^{\supset,\wedge,\vee,\forall,\exists,\perp}$ for intuitionistic first-order logic are listed below:

$$\Gamma, x\colon A \vdash A$$

$$\frac{\Gamma \vdash \perp}{\Gamma \vdash A} \quad (\perp\text{-}elim)$$

with $A \neq \perp$,

$$\frac{\Gamma, x\colon A \vdash B}{\Gamma \vdash A \supset B} \quad (\supset\text{-}intro)$$

$$\frac{\Gamma \vdash A \supset B \quad \Gamma \vdash A}{\Gamma \vdash B} \quad (\supset\text{-}elim)$$

$$\frac{\Gamma \vdash A \quad \Gamma \vdash B}{\Gamma \vdash A \wedge B} \quad (\wedge\text{-}intro)$$

$$\frac{\Gamma \vdash A \wedge B}{\Gamma \vdash A} \quad (\wedge\text{-}elim) \qquad \frac{\Gamma \vdash A \wedge B}{\Gamma \vdash B} \quad (\wedge\text{-}elim)$$

$$\frac{\Gamma \vdash A}{\Gamma \vdash A \vee B} \quad (\vee\text{-}intro) \qquad \frac{\Gamma \vdash B}{\Gamma \vdash A \vee B} \quad (\vee\text{-}intro)$$

$$\frac{\Gamma \vdash A \vee B \quad \Gamma, x\colon A \vdash C \quad \Gamma, y\colon B \vdash C}{\Gamma \vdash C} \quad (\vee\text{-}elim)$$

$$\frac{\Gamma \vdash A[u/t]}{\Gamma \vdash \forall t A} \quad (\forall\text{-}intro) \qquad \frac{\Gamma \vdash \forall t A}{\Gamma \vdash A[\tau/t]} \quad (\forall\text{-}elim)$$

where in ($\forall$-*intro*), u does not occur free in Γ or $\forall t A$;

$$\frac{\Gamma \vdash A[\tau/t]}{\Gamma \vdash \exists t A} \quad (\exists\text{-}intro) \qquad \frac{\Gamma \vdash \exists t A \quad z\colon A[u/t], \Gamma \vdash C}{\Gamma \vdash C} \quad (\exists\text{-}elim)$$

where in ($\exists$-*elim*), u does not occur free in Γ, $\exists t A$, or C.

The variable u is called the *eigenvariable* of the inference.

One should observe that we are now using two kinds of variables: term (or package) variables $(x, y, z, \ldots)$, and individual (or type) variables $(t, u, \ldots)$.

The typed λ-calculus $\lambda^{\to,\times,+,\forall,\exists,\perp}$ corresponding to $\mathcal{N}_i^{\supset,\wedge,\vee,\forall,\exists,\perp}$ is given in the following definition.

Definition 7.2 The typed λ-calculus $\lambda^{\to,\times,+,\forall,\exists,\perp}$ is defined by the following rules.

$$\Gamma, x\colon A \rhd x\colon A$$

$$\frac{\Gamma \rhd M\colon \perp}{\Gamma \rhd \bigtriangledown_A(M)\colon A} \quad (\perp\text{-}elim)$$

with $A \neq \perp$,

$$\frac{\Gamma, x\colon A \rhd M\colon B}{\Gamma \rhd (\lambda x\colon A.\, M)\colon A \to B} \quad (abstraction)$$

$$\frac{\Gamma \rhd M\colon A \to B \quad \Gamma \rhd N\colon A}{\Gamma \rhd (MN)\colon B} \quad (application)$$

$$\frac{\Gamma \rhd M\colon A \quad \Gamma \rhd N\colon B}{\Gamma \rhd \langle M, N\rangle\colon A \times B} \quad (pairing)$$

$$\frac{\Gamma \rhd M\colon A \times B}{\Gamma \rhd \pi_1(M)\colon A} \quad (projection) \qquad \frac{\Gamma \rhd M\colon A \times B}{\Gamma \rhd \pi_2(M)\colon B} \quad (projection)$$

$$\frac{\Gamma \rhd M\colon A}{\Gamma \rhd \mathtt{inl}(M)\colon A + B} \quad (injection) \qquad \frac{\Gamma \rhd M\colon B}{\Gamma \rhd \mathtt{inr}(M)\colon A + B} \quad (injection)$$

$$\frac{\Gamma \rhd P\colon A + B \quad \Gamma, x\colon A \rhd M\colon C \quad \Gamma, y\colon B \rhd N\colon C}{\Gamma \rhd \mathtt{case}(P, \lambda x\colon A.\, M, \lambda y\colon B.\, N)\colon C} \quad (by\text{-}cases)$$

or

$$\frac{\Gamma \rhd P\colon A + B \quad \Gamma, x\colon A \rhd M\colon C \quad \Gamma, y\colon B \rhd N\colon C}{\Gamma \rhd (\mathtt{case}\ P\ \mathtt{of}\ \mathtt{inl}(x\colon A) \Rightarrow M \mid \mathtt{inr}(y\colon B) \Rightarrow N)\colon C} \quad (by\text{-}cases)$$

$$\frac{\Gamma \rhd M\colon A[u/t]}{\Gamma \rhd (\lambda u\colon \iota.\, M)\colon \forall t A} \quad (\forall\text{-}intro)$$

where u does not occur free in Γ or $\forall t A$;

$$\frac{\Gamma \triangleright M\colon \forall t A}{\Gamma \triangleright M\tau\colon A[\tau/t]} \quad (\forall\text{-}elim)$$

$$\frac{\Gamma \triangleright M\colon A[\tau/t]}{\Gamma \triangleright \mathtt{inx}(\tau, M)\colon \exists t A} \quad (\exists\text{-}intro)$$

$$\frac{\Gamma \triangleright M\colon \exists t A \quad \Gamma, x\colon A[u/t] \triangleright N\colon C}{\Gamma \triangleright \mathtt{casex}(M, \lambda u\colon \iota.\ \lambda x\colon A[u/t].\ N)\colon C} \quad (\exists\text{-}elim)$$

where u does not occur free in Γ, $\exists t A$, or C.

In the term $(\lambda u\colon \iota.\ M)$, the type ι stands for the type of individuals. Note that

$$\Gamma \triangleright \lambda u\colon \iota.\ \lambda x\colon A[u/t].\ N\colon \forall u(A[u/t] \rightarrow C).$$

The term $\lambda u\colon \iota.\ \lambda x\colon A[u/t].\ N$ contains the type $A[u/t]$ which is a *dependent* type, since it usually contains occurrences of u. Observe that $(\lambda u\colon \iota.\ \lambda x\colon A[u/t].\ N)\tau$ reduces to $\lambda x\colon A[\tau/t].\ N[\tau/u]$, in which the type of x is now $A[\tau/t]$. The term $\mathtt{casex}(M, \lambda u\colon \iota.\ \lambda x\colon A[u/t].\ N)$ is also denoted as $\mathtt{casex}\ M\ \mathtt{of}\ \mathtt{inx}(u\colon \iota, x\colon A[u/t]) \Rightarrow N$, or even $\mathtt{casex}\ M\ \mathtt{of}\ \mathtt{inx}(u, x) \Rightarrow N$, and the $(\exists\text{-}elim)$ rule as

$$\frac{\Gamma \triangleright M\colon \exists t A \quad \Gamma, x\colon A[u/t] \triangleright N\colon C}{\Gamma \triangleright (\mathtt{casex}\ M\ \mathtt{of}\ \mathtt{inx}(u\colon \iota, x\colon A[u/t]) \Rightarrow N)\colon C} \quad (\exists\text{-}elim)$$

where u does not occur free in Γ, $\exists t A$, or C.

Such a formalism can be easily generalized to many sorts (base types), if quantified formulae are written as $\forall t\colon \sigma.\ A$ and $\exists t\colon \sigma.\ A$, where σ is a sort (base type). A further generalization would be to allow higher-order quantification as in Girard's system F_ω (see Girard [11] or Gallier [7]). We also have the following reduction rules.

Definition 7.3 The reduction rules of the system $\lambda^{\rightarrow,\times,+,\forall,\exists,\perp}$ are listed below:

$$\begin{aligned}
(\lambda x\colon A.\ M)N &\longrightarrow M[N/x], \\
\pi_1(\langle M, N\rangle) &\longrightarrow M, \\
\pi_2(\langle M, N\rangle) &\longrightarrow N, \\
\mathtt{case}(\mathtt{inl}(P), M, N) &\longrightarrow MP, \quad \text{or} \\
\mathtt{case}\ \mathtt{inl}(P)\ \mathtt{of}\ \mathtt{inl}(x\colon A) \Rightarrow M \mid \mathtt{inr}(y\colon B) \Rightarrow N &\longrightarrow M[P/x], \\
\mathtt{case}(\mathtt{inr}(P), M, N) &\longrightarrow NP, \quad \text{or} \\
\mathtt{case}\ \mathtt{inr}(P)\ \mathtt{of}\ \mathtt{inl}(x\colon A) \Rightarrow M \mid \mathtt{inr}(y\colon B) \Rightarrow N &\longrightarrow N[P/y], \\
\nabla_{A\rightarrow B}(M)N &\longrightarrow \nabla_B(M), \\
\pi_1(\nabla_{A\times B}(M)) &\longrightarrow \nabla_A(M), \\
\pi_2(\nabla_{A\times B}(M)) &\longrightarrow \nabla_B(M), \\
(\lambda t\colon \iota.\ M)\tau &\longrightarrow M[\tau/t], \\
\nabla_{\forall t A}(M)\tau &\longrightarrow \nabla_{A[\tau/t]}(M),
\end{aligned}$$

$$\begin{aligned}
\texttt{case}(\bigtriangledown_{A+B}(P), M, N) &\longrightarrow \bigtriangledown_C(P), \\
\texttt{casex}(\texttt{inx}(\tau, P), M) &\longrightarrow (M\tau)P, \quad \text{or} \\
\texttt{casex}\ \texttt{inx}(\tau, P)\ \texttt{of}\ \texttt{inx}(t{:}\,\iota, x{:}\,A) \Rightarrow N &\longrightarrow N[\tau/t, P/x], \\
\texttt{casex}(\bigtriangledown_{\exists t A}(P), M) &\longrightarrow \bigtriangledown_C(P).
\end{aligned}$$

A fundamental result about natural deduction is the fact that every proof (term) reduces to a normal form, which is unique up to α-renaming. This result was first proved by Prawitz [24] for the system $\mathcal{N}_i^{\supset,\wedge,\vee,\forall,\exists,\perp}$.

Theorem 7.4 [Church-Rosser property, Prawitz (1971)] *Reduction in $\lambda^{\to,\times,+,\forall,\exists,\perp}$ (specified in Definition 7.3) is confluent. Equivalently, conversion in $\lambda^{\to,\times,+,\forall,\exists,\perp}$ is Church-Rosser.*

A proof can be given by adapting the method of Tait and Martin-Löf [21] using a form of parallel reduction (see also Barendregt [2], Hindley and Seldin [15], or Stenlund [27]).

Theorem 7.5 [Strong normalization property, Prawitz (1971)] *Reduction in $\lambda^{\to,\times,+,\forall,\exists,\perp}$ is strongly normalizing.*

A proof can be given by adapting Tait's reducibility method [28], [30], as done in Girard [10] (1971), [11] (1972) (see also Gallier [7]).

If one looks carefully at the structure of proofs, one realizes that it is not unreasonable to declare other proofs as being redundant, and thus to add some additional reduction rules. For example, the proof term $\langle \pi_1(M), \pi_2(M) \rangle$ can be identified with M itself. Similarly, if x is not free in M, the term $\lambda x{:}\, A.\,(Mx)$ can be identified with M. Thus, we have the following additional set of reduction rules:

$$\begin{aligned}
\lambda x{:}\, A.\,(Mx) &\longrightarrow M, && \text{if } x \notin FV(M), \\
\langle \pi_1(M), \pi_2(M) \rangle &\longrightarrow M, && \\
\texttt{case}\ M\ \texttt{of}\ \texttt{inl}(x{:}\,A) \Rightarrow \texttt{inl}(x) \mid \texttt{inr}(y{:}\,B) \Rightarrow \texttt{inr}(y) &\longrightarrow M, && \\
\lambda t{:}\, \iota.\,(Mt) &\longrightarrow M, && \text{if } t \notin FV(M), \\
\texttt{casex}\ M\ \texttt{of}\ \texttt{inx}(u{:}\,\iota, x{:}\,A[u/t]) \Rightarrow \texttt{inx}(u, x) &\longrightarrow M, && \text{if } u \notin FV(M).
\end{aligned}$$

These rules are important in setting up categorical semantics for intuitionistic logic. However, a discussion of this topic would take us far beyond the scope of this paper. Actually, in order to salvage some form of subformula property ruined by the introduction of the connectives $\vee$, $\exists$, and $\perp$, one can add further conversions known as "commuting conversions" (or "permutative conversions"). A lucid discussion of the necessity for such rules can be found in Girard [9]. Theorem 7.4 and theorem 7.5 can be extended to cover the reduction rules of definition 7.3 together with the new reductions rules, but at the cost of rather tedious and rather noninstructive technical complications. Due to the lack of space, we will not elaborate any further on this subject and simply refer the interested reader to Prawitz [23], Girard [11], or Girard [9] for details.

A sequent-calculus formulation for intuitionistic first-order logic is given in the next definition.

Definition 7.6 The axioms and inference rules of the system $\mathcal{G}_i^{\supset,\wedge,\vee,\forall,\exists,\perp,cut}$ for intuitionistic first-order logic are given below.

$$A, \Gamma \vdash A$$

$$\frac{\Gamma \vdash \perp}{\Gamma \vdash A} \quad (\perp\colon \mathit{right})$$

with $A \neq \perp$,

$$\frac{A, A, \Gamma \vdash C}{A, \Gamma \vdash C} \quad (\mathit{contrac}\colon \mathit{left})$$

$$\frac{\Gamma \vdash A \quad A, \Gamma \vdash C}{\Gamma \vdash C} \quad (\mathit{cut})$$

$$\frac{A, B, \Gamma \vdash C}{A \wedge B, \Gamma \vdash C} \quad (\wedge\colon \mathit{left}) \qquad \frac{\Gamma \vdash A \quad \Gamma \vdash B}{\Gamma \vdash A \wedge B} \quad (\wedge\colon \mathit{right})$$

$$\frac{A, \Gamma \vdash C \quad B, \Gamma \vdash C}{A \vee B, \Gamma \vdash C} \quad (\vee\colon \mathit{left})$$

$$\frac{\Gamma \vdash A}{\Gamma \vdash A \vee B} \quad (\vee\colon \mathit{right}) \qquad \frac{\Gamma \vdash B}{\Gamma \vdash A \vee B} \quad (\vee\colon \mathit{right})$$

$$\frac{\Gamma \vdash A \quad B, \Gamma \vdash C}{A \supset B, \Gamma \vdash C} \quad (\supset\colon \mathit{left}) \qquad \frac{A, \Gamma \vdash B}{\Gamma \vdash A \supset B} \quad (\supset\colon \mathit{right})$$

$$\frac{A[\tau/x], \Gamma \vdash C}{\forall x A, \Gamma \vdash C} \quad (\forall\colon \mathit{left}) \qquad \frac{\Gamma \vdash A[y/x]}{\Gamma \vdash \forall x A} \quad (\forall\colon \mathit{right})$$

where in ($\forall$: *right*), y does not occur free in the conclusion;

$$\frac{A[y/x], \Gamma \vdash C}{\exists x A, \Gamma \vdash C} \quad (\exists\colon \mathit{left}) \qquad \frac{\Gamma \vdash A[\tau/x]}{\Gamma \vdash \exists x A} \quad (\exists\colon \mathit{right})$$

where in ($\exists$: *left*), y does not occur free in the conclusion.

The variable y is called the *eigenvariable* of the inference.

A variation of the system $\mathcal{G}_i^{\supset,\wedge,\vee,\forall,\exists,\perp,cut}$ in which negation appears explicitly is obtained by replacing the rule

$$\frac{\Gamma \vdash \perp}{\Gamma \vdash A} \quad (\perp\colon \mathit{right})$$

by the rule

$$\frac{\Gamma \vdash}{\Gamma \vdash A} \quad (\mathit{weakening}\colon \mathit{right})$$

and adding the following negation rules:

$$\frac{\Gamma \vdash A}{\neg A, \Gamma \vdash} \quad (\neg\colon \mathit{left}) \qquad \frac{A, \Gamma \vdash}{\Gamma \vdash \neg A} \quad (\neg\colon \mathit{right})$$

The resulting system is denoted as $\mathcal{G}_i^{\supset,\wedge,\vee,\forall,\exists,\neg,cut}$.

The system $\mathcal{G}_c^{\supset,\wedge,\vee,\forall,\exists,\perp,cut}$ of classical logic is shown in the next definition.

Definition 7.7 The axioms and inference rules of the system $\mathcal{G}_c^{\supset,\wedge,\vee,\forall,\exists,\perp,cut}$ for classical first-order logic are given below.

$$A, \Gamma \vdash \Delta, A$$

$$\frac{A, A, \Gamma \vdash \Delta}{A, \Gamma \vdash \Delta} \quad (\textit{contrac: left}) \qquad \frac{\Gamma \vdash \Delta, A, A}{\Gamma \vdash \Delta, A} \quad (\textit{contrac: right})$$

$$\frac{\Gamma \vdash \Delta, \perp}{\Gamma \vdash \Delta, A} \quad (\perp\textit{: right})$$

with $A \neq \perp$,

$$\frac{\Gamma \vdash \Delta, A \quad A, \Gamma \vdash \Delta}{\Gamma \vdash \Delta} \quad (\textit{cut})$$

$$\frac{A, B, \Gamma \vdash \Delta}{A \wedge B, \Gamma \vdash \Delta} \quad (\wedge\textit{: left}) \qquad \frac{\Gamma \vdash \Delta, A \quad \Gamma \vdash \Delta, B}{\Gamma \vdash \Delta, A \wedge B} \quad (\wedge\textit{: right})$$

$$\frac{A, \Gamma \vdash \Delta \quad B, \Gamma \vdash \Delta}{A \vee B, \Gamma \vdash \Delta} \quad (\vee\textit{: left}) \qquad \frac{\Gamma \vdash \Delta, A, B}{\Gamma \vdash \Delta, A \vee B} \quad (\vee\textit{: right})$$

$$\frac{\Gamma \vdash \Delta, A \quad B, \Gamma \vdash \Delta}{A \supset B, \Gamma \vdash \Delta} \quad (\supset\textit{: left}) \qquad \frac{A, \Gamma \vdash \Delta, B}{\Gamma \vdash \Delta, A \supset B} \quad (\supset\textit{: right})$$

$$\frac{A[\tau/x], \Gamma \vdash \Delta}{\forall x A, \Gamma \vdash \Delta} \quad (\forall\textit{: left}) \qquad \frac{\Gamma \vdash \Delta, A[y/x]}{\Gamma \vdash \Delta, \forall x A} \quad (\forall\textit{: right})$$

where in ($\forall$: *right*), y does not occur free in the conclusion;

$$\frac{A[y/x], \Gamma \vdash \Delta}{\exists x A, \Gamma \vdash \Delta} \quad (\exists\textit{: left}) \qquad \frac{\Gamma \vdash \Delta, A[\tau/x]}{\Gamma \vdash \Delta, \exists x A} \quad (\exists\textit{: right})$$

where in ($\exists$: *left*), y does not occur free in the conclusion.

A variation of the system $\mathcal{G}_c^{\supset,\wedge,\vee,\forall,\exists,\perp,cut}$ in which negation appears explicitly is obtained by deleting the rule

$$\frac{\Gamma \vdash \Delta, \perp}{\Gamma \vdash \Delta, A} \quad (\perp\textit{: right})$$

and adding the following negation rules:

$$\frac{\Gamma \vdash \Delta, A}{\neg A, \Gamma \vdash \Delta} \quad (\neg\textit{: left}) \qquad \frac{A, \Gamma \vdash \Delta}{\Gamma \vdash \Delta, \neg A} \quad (\neg\textit{: right})$$

Indeed, it is easy to see that the rule

$$\frac{\Gamma \vdash \Delta}{\Gamma \vdash \Delta, A} \quad (\textit{weakening: right})$$

is derivable (using the axioms $A, \Gamma \vdash \Delta, A$). The resulting system is denoted as $\mathcal{G}_c^{\supset,\wedge,\vee,\forall,\exists,\neg,cut}$.

We now extend the functions $\mathcal{N}$ and $\mathcal{G}$ to deal with the quantifier rules. The procedure $\mathcal{N}$ is extended to $\mathcal{G}_i^{\supset,\wedge,\vee,\forall,\exists,\perp,cut}$ as follows.

- A proof Π of the form

$$\dfrac{\begin{array}{c}\Pi_1\\ A[\tau/x], \Gamma \vdash C\end{array}}{\forall x A, \Gamma \vdash C}$$

is mapped to a deduction obtained as follows. First, map Π_1 to

$$\begin{array}{c}\mathcal{N}(\Pi_1)\\ y\colon A[\tau/x], \Gamma^* \vdash C\end{array}$$

Next, replace every leaf of the form $y\colon A[\tau/x], \Delta, \Gamma^* \vdash A[\tau/x]$ in $\mathcal{N}(\Pi_1)$ by the subtree

$$\dfrac{y\colon \forall x A, \Delta, \Gamma^* \vdash \forall x A}{y\colon \forall x A, \Delta, \Gamma^* \vdash A[\tau/x]}$$

replace every occurrence of "$y\colon A[\tau/x]$" in the resulting tree by "$y\colon \forall x A$", and rectify this last tree.

- A proof Π of the form

$$\dfrac{\begin{array}{c}\Pi_1\\ \Gamma \vdash A[y/x]\end{array}}{\Gamma \vdash \forall x A}$$

is mapped to the deduction

$$\dfrac{\begin{array}{c}\mathcal{N}(\Pi_1)\\ \Gamma^+ \vdash A[y/x]\end{array}}{\Gamma^+ \vdash \forall x A}$$

- A proof Π of the form

$$\dfrac{\begin{array}{c}\Pi_1\\ A[y/x], \Gamma \vdash C\end{array}}{\exists x A, \Gamma \vdash C}$$

is mapped to the deduction

$$\dfrac{u\colon \exists x A, \Gamma^* \vdash \exists x A \qquad \begin{array}{c}u\colon \exists x A + \mathcal{N}(\Pi_1)\\ u\colon \exists x A, v\colon A[y/x], \Gamma^* \vdash C\end{array}}{u\colon \exists x A, \Gamma^* \vdash C}$$

and rectify this last tree.

- A proof Π of the form

$$\frac{\begin{matrix}\Pi_1\\ \Gamma \vdash A[\tau/x]\end{matrix}}{\Gamma \vdash \exists x A}$$

is mapped to the deduction

$$\frac{\begin{matrix}\mathcal{N}(\Pi_1)\\ \Gamma^+ \vdash A[\tau/x]\end{matrix}}{\Gamma^+ \vdash \exists x A}$$

It is easily seen that Lemma 5.5 generalizes to quantifiers.

Lemma 7.8 [Gentzen (1935), Prawitz (1965)] *For every proof* Π *in* $\mathcal{G}_i^{\supset,\wedge,\vee,\forall,\exists,\perp}$, $\mathcal{N}(\Pi)$ *is a deduction in normal form (in* $\mathcal{N}_i^{\supset,\wedge,\vee,\forall,\exists,\perp}$*).*

Next, we extend the procedure $\mathcal{G}$ to $\mathcal{N}_i^{\supset,\wedge,\vee,\forall,\exists,\perp}$.

- A deduction $\mathcal{D}$ of the form

$$\frac{\begin{matrix}\mathcal{D}_1\\ \Gamma \vdash A[y/x]\end{matrix}}{\Gamma \vdash \forall x A}$$

is mapped to the proof

$$\frac{\begin{matrix}\mathcal{G}(\mathcal{D}_1)\\ \Gamma^- \vdash A[y/x]\end{matrix}}{\Gamma^- \vdash \forall x A}$$

- A deduction $\mathcal{D}$ of the form

$$\frac{\begin{matrix}\mathcal{D}_1\\ \Gamma \vdash \forall x A\end{matrix}}{\Gamma \vdash A[\tau/x]}$$

is mapped to the proof

$$\frac{\begin{matrix}\mathcal{G}(\mathcal{D}_1)\\ \Gamma^- \vdash \forall x A\end{matrix} \qquad \dfrac{A[\tau/x], \Gamma^- \vdash A[\tau/x]}{\forall x A, \Gamma^- \vdash A[\tau/x]}}{\Gamma^- \vdash A[\tau/x]} \quad (cut)$$

- A deduction $\mathcal{D}$ of the form

$$\frac{\begin{array}{c}\mathcal{D}_1\\ \Gamma \vdash A[\tau/x]\end{array}}{\Gamma \vdash \exists x A}$$

is mapped to the proof

$$\frac{\begin{array}{c}\mathcal{G}(\mathcal{D}_1)\\ \Gamma^- \vdash A[\tau/x]\end{array}}{\Gamma^- \vdash \exists x A}$$

- A deduction $\mathcal{D}$ of the form

$$\frac{\begin{array}{c}\mathcal{D}_1\\ \Gamma \vdash \exists x A\end{array} \qquad \begin{array}{c}\mathcal{D}_2\\ z\colon A[y/x], \Gamma \vdash C\end{array}}{\Gamma \vdash C}$$

is mapped to the proof

$$\frac{\begin{array}{c}\mathcal{G}(\mathcal{D}_1)\\ \Gamma^- \vdash \exists x A\end{array} \qquad \dfrac{\begin{array}{c}\mathcal{G}(\mathcal{D}_2)\\ A[y/x], \Gamma^- \vdash C\end{array}}{\exists x A, \Gamma^- \vdash C}}{\Gamma^- \vdash C} \quad (cut)$$

We now turn to cut elimination.

8 Gentzen's Cut Elimination Theorem

As we said earlier before presenting the function $\mathcal{G}$ from $\mathcal{N}_i^{\supset,\wedge,\vee,\forall,\exists,\perp}$ to $\mathcal{G}_i^{\supset,\wedge,\vee,\forall,\exists,\perp,cut}$, it is possible to show that the system $\mathcal{G}_i^{\supset,\wedge,\vee,\forall,\exists,\perp,cut}$ is equivalent to the seemingly weaker system $\mathcal{G}_i^{\supset,\wedge,\vee,\forall,\exists,\perp}$.

We have the following fundamental result.

Theorem 8.1 [Cut Elimination Theorem, Gentzen (1935)] *There is an algorithm which, given any proof* Π *in* $\mathcal{G}_i^{\supset,\wedge,\vee,\forall,\exists,\perp,cut}$ *produces a cut-free proof* Π' *in* $\mathcal{G}_i^{\supset,\wedge,\vee,\forall,\exists,\perp}$. *There is an algorithm which, given any proof* Π *in* $\mathcal{G}_c^{\supset,\wedge,\vee,\forall,\exists,\perp,cut}$ *produces a cut-free proof* Π' *in* $\mathcal{G}_c^{\supset,\wedge,\vee,\forall,\exists,\perp}$.

Proof. The proof is quite involved. It consists in pushing up cuts towards the leaves, and in breaking cuts involving compound formulae into cuts on smaller subformulae. Full details are given for the system $\mathcal{LK}$ in Section 12. Interestingly, the need for the contraction rule arises when a cut involves an axiom. The typical example is as follows. The proof

$$\cfrac{A, \Gamma \vdash A \qquad \begin{array}{c}\Pi_1 \\ A, A, \Gamma \vdash C\end{array}}{A, \Gamma \vdash C}$$

is equivalent to a (*contrac*: *left*), and it is eliminated by forming the proof

$$\cfrac{\begin{array}{c}\Pi_1 \\ A, A, \Gamma \vdash C\end{array}}{A, \Gamma \vdash C} \quad (contrac\text{: } left)$$

□

If we are interested in cut-free proofs, except in *classical propositional logic*, the contraction rules *cannot* be dispensed with. We already saw in Example 4.4 that $\neg\neg(P \vee \neg P)$ is a proposition which is not provable without contractions in $\mathcal{G}_i^{\supset,\wedge,\vee,\neg}$. Another example involving quantifiers is the sequent $\forall x \exists y(Py \wedge \neg Px) \vdash$ which is not provable without contractions in $\mathcal{G}_i^{\supset,\wedge,\vee,\forall,\exists,\neg}$ or even in $\mathcal{G}_c^{\supset,\wedge,\vee,\forall,\exists,\neg}$. This sequent has the following proof in $\mathcal{G}_i^{\supset,\wedge,\vee,\forall,\exists,\neg}$:

Example 8.2

$$\cfrac{\cfrac{\cfrac{\cfrac{\cfrac{\cfrac{\cfrac{\cfrac{Pu, \neg Px, Pv \vdash Pu}{Pu, \neg Px, Pv, \neg Pu \vdash}}{Pu, \neg Px, (Pv \wedge \neg Pu) \vdash}}{(Pu \wedge \neg Px), (Pv \wedge \neg Pu) \vdash}}{(Pu \wedge \neg Px), \exists y(Py \wedge \neg Pu) \vdash}}{(Pu \wedge \neg Px), \forall x \exists y(Py \wedge \neg Px) \vdash}}{\exists y(Py \wedge \neg Px), \forall x \exists y(Py \wedge \neg Px) \vdash}}{\forall x \exists y(Py \wedge \neg Px), \forall x \exists y(Py \wedge \neg Px) \vdash}}{\forall x \exists y(Py \wedge \neg Px) \vdash} \quad (contrac\text{: } left)$$

It is an interesting exercise to find a deduction of $\forall x \exists y(Py \wedge \neg Px) \supset \perp$ in $\mathcal{N}_i^{\supset,\wedge,\vee,\forall,\exists,\perp}$.

For classical logic, it is possible to show that the contraction rules are only needed to permit an unbounded number of applications of the ($\forall$: *left*)-rule and the ($\exists$: *right*)-rule (see lemma 8.7). For example, the formula $\exists x \forall y(Py \supset Px)$ is provable in $\mathcal{G}_c^{\supset,\wedge,\vee,\forall,\exists,\perp}$, but not without the rule (*contrac*: *right*). The cut-free system $\mathcal{G}_c^{\supset,\wedge,\vee,\forall,\exists,\perp}$ can be modified to obtain another system $\mathcal{GK}_c^{\supset,\wedge,\vee,\forall,\exists,\perp}$ in which the contraction rules are deleted and the quantifier rules are slightly changed to incorporate contraction.

Definition 8.3 The axioms and inference rules of the cut-free system $\mathcal{GK}_c^{\supset,\wedge,\vee,\forall,\exists,\perp}$ for classical first-order logic are given below.

$$A, \Gamma \vdash \Delta, A$$

$$\perp, \Gamma \vdash \Delta, A$$

$$\frac{A, B, \Gamma \vdash \Delta}{A \wedge B, \Gamma \vdash \Delta} \quad (\wedge\colon \textit{left}) \qquad \frac{\Gamma \vdash \Delta, A \quad \Gamma \vdash \Delta, B}{\Gamma \vdash \Delta, A \wedge B} \quad (\wedge\colon \textit{right})$$

$$\frac{A, \Gamma \vdash \Delta \quad B, \Gamma \vdash \Delta}{A \vee B, \Gamma \vdash \Delta} \quad (\vee\colon \textit{left}) \qquad \frac{\Gamma \vdash \Delta, A, B}{\Gamma \vdash \Delta, A \vee B} \quad (\vee\colon \textit{right})$$

$$\frac{\Gamma \vdash \Delta, A \quad B, \Gamma \vdash \Delta}{A \supset B, \Gamma \vdash \Delta} \quad (\supset\colon \textit{left}) \qquad \frac{A, \Gamma \vdash \Delta, B}{\Gamma \vdash \Delta, A \supset B} \quad (\supset\colon \textit{right})$$

$$\frac{\forall x A, A[\tau/x], \Gamma \vdash \Delta}{\forall x A, \Gamma \vdash \Delta} \quad (\forall\colon \textit{left}) \qquad \frac{\Gamma \vdash \Delta, A[y/x]}{\Gamma \vdash \Delta, \forall x A} \quad (\forall\colon \textit{right})$$

where in ($\forall$: *right*), y does not occur free in the conclusion;

$$\frac{A[y/x], \Gamma \vdash \Delta}{\exists x A, \Gamma \vdash \Delta} \quad (\exists\colon \textit{left}) \qquad \frac{\Gamma \vdash \Delta, \exists x A, A[\tau/x]}{\Gamma \vdash \Delta, \exists x A} \quad (\exists\colon \textit{right})$$

where in ($\exists$: *left*), y does not occur free in the conclusion.

The above system is inspired from Kleene [18] (see system $G3$, page 481). Note that contraction steps have been incorporated in the ($\forall$: *left*)-rule and the ($\exists$: *right*)-rule. As noted in the discussion before definition 4.3, if we consider sequents in which the righthand side is nonempty, using axioms of the form

$$\perp, \Gamma \vdash \Delta, A$$

is equivalent to using the rule

$$\frac{\Gamma \vdash \Delta, \perp}{\Gamma \vdash \Delta, A} \quad (\perp\colon \textit{right})$$

with $A \neq \perp$. However, the axioms $\perp, \Gamma \vdash \Delta, A$ are technically simpler to handle in proving the next three lemmas, and thus we prefer them to the rule ($\perp$: *right*). Accordingly, from now on, we will also assume that $\mathcal{G}_c^{\supset,\wedge,\vee,\forall,\exists,\perp}$ has been formulated using the axioms $\perp, \Gamma \vdash \Delta, A$ rather than the rule ($\perp$: *right*).

The equivalence of the systems $\mathcal{G}_c^{\supset,\wedge,\vee,\forall,\exists,\perp}$ and $\mathcal{GK}_c^{\supset,\wedge,\vee,\forall,\exists,\perp}$ is shown using two lemmas inspired from Kleene [18] (1952). First, it is quite easy to see that every proof in $\mathcal{GK}_c^{\supset,\wedge,\vee,\forall,\exists,\perp}$ can be converted to a proof in $\mathcal{G}_c^{\supset,\wedge,\vee,\forall,\exists,\perp}$.

For the converse, we warn the reader that some of the lemmas given in an earlier version were incorrect. The proof of the converse is quite tricky, and we are grateful to Peter Baumann for pointing out the earlier errors and helping me in working out the new proof. The plan of attack is to show that the weakening and contraction rules are derived rules of the system $\mathcal{GK}_c^{\supset,\wedge,\vee,\forall,\exists,\perp}$. For the weakening rules, this follows immediately by induction, the crucial fact being that the axioms are "fat", that is of the form $A, \Gamma \vdash \Delta, A$. For technical reasons (in the proof of lemma 8.5), we will need the fact that in $\mathcal{GK}_c^{\supset,\wedge,\vee,\forall,\exists,\perp}$, every provable sequent has a proof in which every axiom has a special form. Such axioms $A, \Gamma \vdash \Delta, A$ are such that the formula A itself is atomic, all formulae in Γ are atomic or universal, and all formulae in Δ are atomic or existential. Such axioms will be called *atomic axioms*.

Lemma 8.4 *Every proof* Π *in* $\mathcal{GK}_c^{\supset,\wedge,\vee,\forall,\exists,\perp}$ *of a sequent* $A_1, \ldots, A_m \vdash B_1, \ldots, B_n$ *can be transformed into a proof with atomic axioms.*

Proof. In constructing a proof, whenever the rule ($\forall$: *left*) of definition 8.3 is used, let us mark the occurrence of $\forall x A$ recopied in the premise, and similarly mark the occurrence of $\exists x A$ recopied in the premise when ($\exists$: *right*) is used. The lemma is then shown by induction on $|A_1| + \ldots + |A_m| + |B_1| + \ldots + |B_n|$, the sum of the sizes of the *unmarked* formulae $A_1, \ldots, A_m$, $B_1, \ldots, B_n$ in a sequent. □

We now prove a useful permutation lemma. Given an inference of $\mathcal{GK}_c^{\supset,\wedge,\vee,\forall,\exists,\perp}$, note that the inference creates a new occurrence of a formula called the *principal formula*.

Lemma 8.5 *Given a proof* Π *with atomic axioms in* $\mathcal{GK}_c^{\supset,\wedge,\vee,\forall,\exists,\perp}$ *of a sequent* $\Gamma \vdash \Delta$, *for every selected occurrence of a formula of the form* $A \wedge B$, $A \vee B$, *or* $A \supset B$ *in* Γ *or* Δ, *or* $\exists x A$ *in* Γ, *or* $\forall x A$ *in* Δ, *there is another proof* Π' *whose last inference has the specified occurrence of the formula as its principal formula. Furthermore,* $depth(\Pi') \leq depth(\Pi)$.

Proof. The proof is by induction on the structure of the proof tree. There are a number of cases depending on what the last inference is. □

Lemma 8.5 does not hold for an occurrence of a formula $\forall x A$ in Γ or for a formula $\exists x A$ in Δ, because the inference that creates it involves a term τ, and moving this inference down in the proof may cause a conflict with the side condition on the eigenvariable y involved in the rules ($\forall$: *right*) or ($\exists$: *left*). As shown by the following example, Lemma 8.5 also fails for intuitionistic logic. The sequent $P, (P \supset Q), (R \supset S) \vdash Q$ has the following proof:

$$\frac{P, (R \supset S) \vdash P \qquad\qquad P, (R \supset S), Q \vdash Q}{P, (P \supset Q), (R \supset S) \vdash Q}$$

On the other hand, the following tree is *not* a proof:

$$\frac{\dfrac{P \vdash P \qquad P, Q \vdash R}{P, (P \supset Q) \vdash R} \qquad\qquad \dfrac{P, S \vdash P \qquad P, S, Q \vdash Q}{P, (P \supset Q), S \vdash Q}}{P, (P \supset Q), (R \supset S) \vdash Q}$$

This shows that in searching for a proof, one has to be careful not to stop after the first failure. Since the contraction rule cannot be dispensed with, it is not obvious at all that provability of an intuitionistic propositional sequent is a decidable property. In fact, it is, but proving it requires a fairly subtle argument. We will present an argument due to Kleene. For the time being, we return to classical logic.

Lemma 8.6 *Given any formula* A, *any pairwise disjoint sets of variables* $\{x_1, \ldots, x_n\}$, $\{y_1, \ldots, y_n\}$ *and* $\{z_1, \ldots, z_n\}$, *and any proof* Π *with atomic axioms of a sequent* $A[y_1/x_1, \ldots, y_n/x_n], A[z_1/x_1, \ldots, z_n/x_n], \Gamma \vdash \Delta$ *in* $\mathcal{GK}_c^{\supset,\wedge,\vee,\forall,\exists,\perp}$ *(resp. of a sequent* $\Gamma \vdash \Delta, A[y_1/x_1, \ldots, y_n/x_n], A[z_1/x_1, \ldots, z_n/x_n]$*), if the variables* y_i *and* z_j *are not free in* Γ, Δ, *or* A, *then there is a proof of the sequent* $A[y_1/x_1, \ldots, y_n/x_n], \Gamma \vdash \Delta$ *and a proof of the sequent* $A[z_1/x_1, \ldots, z_n/x_n], \Gamma \vdash \Delta$ *(resp. there is a proof of the sequent* $\Gamma \vdash \Delta, A[y_1/x_1, \ldots, y_n/x_n]$ *and a proof of the sequent* $\Gamma \vdash \Delta, A[z_1/x_1, \ldots, z_n/x_n]$*).*

Proof. We prove the slightly more general claim:

Claim. Given any formula A, any pairwise disjoint sets of variables $\{x_1, \ldots, x_n\}$, $\{y_1, \ldots, y_n\}$ and $\{z_1, \ldots, z_n\}$, and any proof Π with atomic axioms of a sequent $A[y_1/x_1, \ldots, y_n/x_n], A[z_1/x_1, \ldots, z_n/x_n], \Gamma \vdash \Delta$ in $\mathcal{GK}_c^{\supset,\wedge,\vee,\forall,\exists,\perp}$ (resp. of a sequent $\Gamma \vdash \Delta, A[y_1/x_1, \ldots, y_n/x_n], A[z_1/x_1, \ldots, z_n/x_n]$), if the y_i and z_j are not free in A, then there is a proof of $A[y_1/x_1, \ldots, y_n/x_n], \Gamma[y_1/z_1, \ldots, y_n/z_n] \vdash \Delta[y_1/z_1, \ldots, y_n/z_n]$ and a proof of $A[z_1/x_1, \ldots, z_n/x_n], \Gamma[z_1/y_1, \ldots, z_n/y_n] \vdash \Delta[z_1/y_1, \ldots, z_n/y_n]$ (resp. there is a proof of $\Gamma[y_1/z_1, \ldots, y_n/z_n] \vdash \Delta[y_1/z_1, \ldots, y_n/z_n], A[y_1/x_1, \ldots, y_n/x_n]$ and a proof of $\Gamma[z_1/y_1, \ldots, z_n/y_n] \vdash \Delta[z_1/y_1, \ldots, z_n/y_n], A[z_1/x_1, \ldots, z_n/x_n]$).

The proof of the claim is by induction on the structure of the proof tree and uses lemma 8.5. □

We can now prove that in classical logic, the contraction rules are only needed for the quantifier-rules.

Lemma 8.7 [Contraction elimination] *The contraction rules are derivable in the system* $\mathcal{GK}_c^{\supset,\wedge,\vee,\forall,\exists,\perp}$.

Proof. First, by lemma 8.4, we know that it is sufficient to consider proofs with atomic axioms. We establish the following claim:

Claim: Given any m formulae $A_1, \ldots, A_m$ and any n formulae $B_1, \ldots, B_n$, every proof Π with atomic axioms in $\mathcal{GK}_c^{\supset,\wedge,\vee,\forall,\exists,\perp}$ of the sequent $A_1, A_1, \ldots, A_m, A_m, \Gamma \vdash B_1, B_1, \ldots, B_n, B_n, \Delta$ can be converted to a proof of the sequent $A_1, \ldots, A_m, \Gamma \vdash B_1, \ldots, B_n, \Delta$.

The claim is proved by induction on the pairs

$$\langle \{|A_1|, \ldots, |A_m|, |B_1|, \ldots, |B_n|\}, h \rangle,$$

where $\{|A_1|, \ldots, |A_m|, |B_1|, \ldots, |B_n|\}$ is the multiset of the sizes of the formulae $A_1, \ldots, A_m, B_1, \ldots, B_n$, and h is the depth of Π. When A_1 is of the form $A \wedge B$, $A \vee B$, $A \supset B$ or $\forall x A$, or B_1 is of the form $A \wedge B$, $A \vee B$, $A \supset B$ or $\exists x A$, we use the induction hypothesis and lemma 8.5. When A_1 is of the form $\exists x A$ or B_1 is of the form $\forall x A$, we use lemma 8.5 and lemma 8.6. In this case, the multiset component is always reduced. □

Since the weakening and the contraction rules are derivable in $\mathcal{GK}_c^{\supset,\wedge,\vee,\forall,\exists,\perp}$, it is possible to convert every proof in $\mathcal{G}_c^{\supset,\wedge,\vee,\forall,\exists,\perp}$ into a proof in $\mathcal{GK}_c^{\supset,\wedge,\vee,\forall,\exists,\perp}$. Thus, the systems $\mathcal{G}_c^{\supset,\wedge,\vee,\forall,\exists,\perp}$ and $\mathcal{GK}_c^{\supset,\wedge,\vee,\forall,\exists,\perp}$ are equivalent. Note in passing that a semantic proof can also be given. Indeed, it is possible to show that $\mathcal{GK}_c^{\supset,\wedge,\vee,\forall,\exists,\perp}$ is complete (see Gallier [6]). Also, as suggested by Peter Baumann, it is possible to prove directly that every proof in $\mathcal{G}_c^{\supset,\wedge,\vee,\forall,\exists,\perp}$ can be converted to a proof in $\mathcal{GK}_c^{\supset,\wedge,\vee,\forall,\exists,\perp}$. Given a proof Π in $\mathcal{G}_c^{\supset,\wedge,\vee,\forall,\exists,\perp}$, the idea is to eliminate one by one all top-level instances of contraction rules in Π. For every such subproof Π', every inference above the root is not a contraction. This allows us to work with contraction-free proofs in $\mathcal{G}_c^{\supset,\wedge,\vee,\forall,\exists,\perp}$, and to prove lemmas analogous to lemma 8.5 and lemma 8.6.

We now present a cut-free system for intuitionistic logic which does not include any explicit contraction rules and in which the premise of every sequent can be interpreted as a set. Using this system $\mathcal{GK}_i$ due to Kleene (see system $G3a$, page 481, in [18]), we can give a very nice proof of the decidability of provability in intuitionistic propositional logic. The idea behind this system is to systematically keep a copy of the principal formula in the premise(s) of every left-rule. Since Lemma 8.7 *fails* for intuitionistic logic, such a system is of interest.

Definition 8.8 The axioms and inference rules of the system $\mathcal{GK}_i^{\supset,\wedge,\vee,\forall,\exists,\perp}$ for intuitionistic first-order logic are given below.

$$A, \Gamma \vdash A$$

$$\perp, \Gamma \vdash A$$

$$\frac{A \wedge B, A, B, \Gamma \vdash C}{A \wedge B, \Gamma \vdash C} \quad (\wedge\colon \mathit{left}) \qquad \frac{\Gamma \vdash A \quad \Gamma \vdash B}{\Gamma \vdash A \wedge B} \quad (\wedge\colon \mathit{right})$$

$$\frac{A \vee B, A, \Gamma \vdash C \quad A \vee B, B, \Gamma \vdash C}{A \vee B, \Gamma \vdash C} \quad (\vee\colon \mathit{left})$$

$$\frac{\Gamma \vdash A}{\Gamma \vdash A \vee B} \quad (\vee\colon \mathit{right}) \qquad \frac{\Gamma \vdash B}{\Gamma \vdash A \vee B} \quad (\vee\colon \mathit{right})$$

$$\frac{A \supset B, \Gamma \vdash A \quad A \supset B, B, \Gamma \vdash C}{A \supset B, \Gamma \vdash C} \quad (\supset\colon \mathit{left}) \qquad \frac{A, \Gamma \vdash B}{\Gamma \vdash A \supset B} \quad (\supset\colon \mathit{right})$$

$$\frac{\forall x A, A[\tau/x], \Gamma \vdash C}{\forall x A, \Gamma \vdash C} \quad (\forall\colon \mathit{left}) \qquad \frac{\Gamma \vdash A[y/x]}{\Gamma \vdash \forall x A} \quad (\forall\colon \mathit{right})$$

where in ($\forall$: *right*), y does not occur free in the conclusion;

$$\frac{\exists x A, A[y/x], \Gamma \vdash C}{\exists x A, \Gamma \vdash C} \quad (\exists\colon \mathit{left}) \qquad \frac{\Gamma \vdash A[\tau/x]}{\Gamma \vdash \exists x A} \quad (\exists\colon \mathit{right})$$

where in ($\exists$: *left*), y does not occur free in the conclusion.

The variable y is called the *eigenvariable* of the inference.

As noted in the discussion before definition 4.3, if we consider sequents in which the righthand side is nonempty, using axioms of the form

$$\perp, \Gamma \vdash A$$

is equivalent to using the rule

$$\frac{\Gamma \vdash \perp}{\Gamma \vdash A} \quad (\perp\colon \mathit{right})$$

with $A \neq \perp$. However, the axioms $\perp, \Gamma \vdash A$ are technically simpler to handle, and thus we prefer them to the rule ($\perp$: *right*). Thus, from now on, we will assume that $\mathcal{G}_i^{\supset,\wedge,\vee,\forall,\exists,\perp}$ has been formulated using the axioms $\perp, \Gamma \vdash A$ rather than the rule ($\perp$: *right*).

The following lemma shows that $\mathcal{GK}_i^{\supset,\wedge,\vee,\forall,\exists,\perp}$ is equivalent to $\mathcal{G}_i^{\supset,\wedge,\vee,\forall,\exists,\perp}$, and also that the premise of every sequent of $\mathcal{GK}_i^{\supset,\wedge,\vee,\forall,\exists,\perp}$ can be viewed as a set.

Lemma 8.9 *For every sequent $\Gamma \vdash C$, every proof Π in $\mathcal{G}_i^{\supset,\wedge,\vee,\forall,\exists,\perp}$ can be transformed into a proof Π' of $\Gamma \vdash C$ in $\mathcal{GK}_i^{\supset,\wedge,\vee,\forall,\exists,\perp}$. Furthermore, a proof Π' can be found such that every formula occurring on the left of any sequent in Π' occurs exactly once. In other words, for every sequent $\Gamma \vdash C$ in Π', the premise Γ can be viewed as a set.*

Proof. The proof is by induction on the structure of Π. The case where the last inference (at the root of the tree) is a contraction follows by the induction hypothesis. Otherwise, the sequent to be proved is either of the form $\Gamma \vdash D$ where Γ is a set, or it is of the form $\Delta, A, A \vdash D$. The first case reduces to the second since Γ can be written as Δ, A, and from a proof of $\Delta, A \vdash D$, we easily obtain a proof of $\Delta, A, A \vdash D$. If the last inference applies to a formula in Δ or D, the induction hypothesis yields the desired result. If the last inference applies to one of the two A's, we apply the induction hypothesis and observe that the rules of $\mathcal{GK}_i$ have been designed to automatically contract the two occurrences of A that would normally be created. For example, if $A = B \wedge C$, the induction hypothesis would yield a proof of $\Delta, B \wedge C, B, C \vdash D$ considered as a set, and the ($\wedge$: *left*)-rule of $\mathcal{GK}_i$ yields $\Delta, B \wedge C \vdash D$ considered as a set. $\square$

As a corollary of Lemma 8.9 we obtain the fact that provability is decidable for intuitionistic propositional logic. Similarly, Lemma 8.7 implies that provability is decidable for classical propositional logic.

Theorem 8.10 *It is decidable whether a proposition is provable in* $\mathcal{N}_i^{\supset,\wedge,\vee,\perp}$. *It is decidable whether a proposition is provable in* $\mathcal{G}_c^{\supset,\wedge,\vee,\perp,cut}$.

Proof. By the existence of the functions $\mathcal{N}$ and $\mathcal{G}$, there is a proof of a proposition A in $\mathcal{N}_i^{\supset,\wedge,\vee,\perp}$ iff there is a proof of the sequent $\vdash A$ in $\mathcal{G}_i^{\supset,\wedge,\vee,\perp,cut}$. By the cut elimination theorem (Theorem 8.1), there is a proof in $\mathcal{G}_i^{\supset,\wedge,\vee,\perp,cut}$ iff there is a proof in $\mathcal{G}_i^{\supset,\wedge,\vee,\perp}$. By Lemma 8.9, there is a proof in $\mathcal{G}_i^{\supset,\wedge,\vee,\perp}$ iff there is a proof in $\mathcal{GK}_i^{\supset,\wedge,\vee,\perp}$. Call a proof *irredundant* if for every sequent $\Gamma \vdash C$ in this proof, Γ is a set, and no sequent occurs twice on any path. If a proof contains a redundant sequent $\Gamma \vdash C$ occurring at two locations on a path, it is clear that this proof can be shortened by replacing the subproof rooted at the lower (closest to the root) location of the repeating sequent $\Gamma \vdash C$ by the smaller subproof rooted at the higher location of the sequent $\Gamma \vdash C$. Thus, a redundant proof in $\mathcal{GK}_i^{\supset,\wedge,\vee,\perp}$ can always be converted to an irredundant proof of the same sequent. Now, by lemma 8.9, we can assume that for every proof in $\mathcal{GK}_i^{\supset,\wedge,\vee,\perp}$ of a given sequent, every node is labeled with a sequent whose lefthand side is a set. Furthermore, since we are considering cut-free proofs, only subformulae of the formulae occurring in the original sequent to be proved can occur in this proof. Since the original sequent if finite, the number of all subformulae of the formulae occurring in this sequent is also finite, and thus there is a uniform bound on the size of every irredundant proof for this sequent. Thus, one simply has to search for an irredundant proof of the given sequent.

By the cut elimination theorem (Theorem 8.1), there is a proof in $\mathcal{G}_c^{\supset,\wedge,\vee,\perp,cut}$ iff there is a proof in $\mathcal{G}_c^{\supset,\wedge,\vee,\perp}$. By Lemma 8.7, there is a proof in $\mathcal{G}_c^{\supset,\wedge,\vee,\perp}$ iff there is a proof in $\mathcal{GK}_c^{\supset,\wedge,\vee,\perp}$. To conclude, note that every inference of $\mathcal{GK}_c^{\supset,\wedge,\vee,\perp}$ decreases the total number of connectives in the sequent. Thus, given a sequent, there are only finitely many proofs for it. $\square$

As an exercise, the reader can show that the proposition

$$((P \supset Q) \supset P) \supset P,$$

known as *Pierce's law*, is not provable in $\mathcal{N}_i^{\supset,\wedge,\vee,\perp}$, but is provable classically in $\mathcal{N}_c^{\supset,\wedge,\vee,\perp}$.

The fact that in any cut-free proof (intuitionistic or classical) of a propositional sequent only subformulae of the formulae occurring in that sequent can occur is an important property called the *subformula property*. The subformula property is not preserved by the quantifier rules, and this suggests that provability in first-order intuitionistic logic or classical logic is undecidable. This can indeed be shown.

9 Invertible Rules

If one is interested in algorithmic proof search for a certain logic, then a cut-free sequent-calculus formulation of this logic is particularly well suited because of the subformula property. In this case, the property of *invertibility* of rules is crucial. Given that the inference rules for sequent calculi are of the form

$$\frac{\Gamma \vdash \Delta}{\Gamma' \vdash \Delta'} \quad \text{(a)}$$

or

$$\frac{\Gamma_1 \vdash \Delta_1 \quad \Gamma_2 \vdash \Delta_2}{\Gamma' \vdash \Delta'} \quad \text{(b)}$$

we say that a rule of type (a) is *invertible* when $\Gamma' \vdash \Delta'$ is provable iff $\Gamma \vdash \Delta$ is provable, and that a rule of type (b) is *invertible* when $\Gamma' \vdash \Delta'$ is provable iff both $\Gamma_1 \vdash \Delta_1$ and $\Gamma_2 \vdash \Delta_2$ are provable. For usual inference rules, we can only claim that the conclusion of a rule is provable if the premises are provable, but the converse does not necessarily hold. When a cut-free sequent calculus has invertible rules and there is some measure of complexity such that the complexity of each premise is strictly smaller than the complexity of the conclusion, then we have a decidable proof system. This is the case of the propositional system $\mathcal{GK}_c^{\supset,\wedge,\vee,\perp}$ or the system $\mathcal{GK}_c^{\supset,\wedge,\vee,\neg}$ obtained from the system of definition 8.3 by deleting the axiom $\perp, \Gamma \vdash \Delta, A$ and adding the negation rules

$$\frac{\Gamma \vdash \Delta, A}{\neg A, \Gamma \vdash \Delta} \quad (\neg\colon \textit{left}) \qquad \frac{A, \Gamma \vdash \Delta}{\Gamma \vdash \Delta, \neg A} \quad (\neg\colon \textit{right})$$

(the full systems also have invertible rules, but the complexity of the premises of quantifier rules is not smaller than the complexity of the conclusion). Systems of invertible rules for classical logic are used systematically in Gallier [6] (see section 3.4 and 5.4). In our opinion, such systems are best suited for presenting completeness proofs in the most transparent fashion.

One of the major differences between Gentzen systems for intuitionistic and classical logic presented so far, is that in intuitionistic systems, sequents are restricted to have at most one formula on the righthand side of $\vdash$. This asymmetry causes the ($\supset$: *left*) and ($\vee$: *right*) rules of intuitionistic logic to be different from their counterparts in classical logic, and in particular, the intuitionistic rules cause some *loss* of information. These rules are **no longer** invertible. For instance, the intuitionistic ($\supset$: *left*)-rule is

$$\frac{\Gamma \vdash A \quad B, \Gamma \vdash C}{A \supset B, \Gamma \vdash C} \quad (\supset\colon \textit{left})$$

whereas its classical version is

$$\frac{\Gamma \vdash A, C \quad B, \Gamma \vdash C}{A \supset B, \Gamma \vdash C} \quad (\supset\colon \textit{left})$$

Note that C is dropped in the left premise of the intuitionistic version of the rule. Similarly, the intuitionistic ($\vee$: *right*)-rules are

$$\frac{\Gamma \vdash A}{\Gamma \vdash A \vee B} \quad (\vee\colon \textit{right}) \qquad \frac{\Gamma \vdash B}{\Gamma \vdash A \vee B} \quad (\vee\colon \textit{right})$$

whereas the classical version is

$$\frac{\Gamma \vdash A, B}{\Gamma \vdash A \lor B} \quad (\lor\colon \mathit{right})$$

Again, either A or B is dropped in the premise of the intuitionistic version of the rule. This loss of information is responsible for the fact that in searching for a proof of a sequent in $\mathcal{G}_i^{\supset,\land,\lor,\forall,\exists,\bot}$, one cannot stop after having found a deduction tree which is not a proof (*i.e.* a deduction tree in which some leaf is not labeled with an axiom). The rules may have been tried in the wrong order, and it is necessary to make sure that *all* attempts have failed to be sure that a sequent is not provable. In fact, proof search should be conducted in the system $\mathcal{GK}_i^{\supset,\land,\lor,\forall,\exists,\bot}$, since we know that searching for an irredundant proof terminates in the propositional case (see theorem 8.10).

Takeuti [31] has made an interesting observation about invertibility of rules in intuitionistic cut-free sequent calculi, but before discussing this observation, we shall discuss other contraction-free systems for intuitionistic propositional logic for which decidability of provability is immediate. Such systems were discovered in the early fifties by Vorob'ev [35, 36]. Interest in such systems has been revived recently due to some work in automated theorem proving by Dyckhoff [5], on the embedding of intuitionistic logic into linear logic (without using the exponentials) by Lincoln, Scedrov and Shankar [20], and on the complexity of cut-elimination by Hudelmaier [17]. In order to simplify the discussion, we first consider propositional intuitionistic logic based on the connective $\supset$. The system $\mathcal{GK}_i^{\supset}$ of Definition 8.8 restricted to propositions built up only from $\supset$ is shown below:

$$A, \Gamma \vdash A$$

$$\frac{A \supset B, \Gamma \vdash A \quad A \supset B, B, \Gamma \vdash C}{A \supset B, \Gamma \vdash C} \quad (\supset\colon \mathit{left}) \qquad \frac{A, \Gamma \vdash B}{\Gamma \vdash A \supset B} \quad (\supset\colon \mathit{right})$$

This system is contraction-free, but it is not immediately obvious that provability is decidable, since $A \supset B$ is recopied in the premises of the ($\supset$: *left*)-rule. First, note that because the systems $\mathcal{GK}_i^{\supset,\land,\lor,\forall,\exists,\bot}$ and $\mathcal{G}_i^{\supset,\land,\lor,\forall,\exists,\bot}$ are equivalent and because cut-elimination holds for $\mathcal{G}_i^{\supset,\land,\lor,\forall,\exists,\bot}$, then cut-elimination also holds for $\mathcal{GK}_i^{\supset,\land,\lor,\forall,\exists,\bot}$. In fact, it is easy to verify that cut-elimination holds for $\mathcal{GK}_i^{\supset}$. Now, it is easy to see that we can require A to be atomic in an axiom, and to see that we can drop $A \supset B$ from the right premise and obtain an equivalent system. The new rule is

$$\frac{A \supset B, \Gamma \vdash A \quad B, \Gamma \vdash C}{A \supset B, \Gamma \vdash C} \quad (\supset\colon \mathit{left})$$

Indeed, if we have a proof of $A \supset B, B, \Gamma \vdash C$, since $B, \Gamma \vdash A \supset B$ is provable, by a cut we obtain that the sequent $B, \Gamma \vdash C$ is provable. Now, the difficulty is to weaken the hypothesis $A \supset B$ in the left premise. What is remarkable is that when A itself is an implication, that is when $A \supset B$ is of the form $(A' \supset B') \supset B$, then $((A' \supset B') \supset B) \vdash (B' \supset B)$ is provable. Furthermore, $A', (B' \supset B) \vdash ((A' \supset B') \supset B)$ is also provable. As a consequence, for any Γ and any propositions A', B', B, D, the sequent $\Gamma, ((A' \supset B') \supset B) \vdash (A' \supset D)$ is provable iff the sequent $\Gamma, (B' \supset B) \vdash (A' \supset D)$ is provable. Then, it can be shown that $B' \supset B$ does indeed work. Also, when A is atomic, the rule ($\supset$: *left*) is simplified to a one-premise rule. The above discussion is intended as a motivation for the system $\mathcal{LJT}^{\supset}$ presented next, but we warn the reader that the equivalence of the new system $\mathcal{LJT}^{\supset}$ with the previous system $\mathcal{GK}_i^{\supset}$ is not as simple as it seems. The new system $\mathcal{LJT}^{\supset}$ (which is a subsystem of Dyckhoff's system [5]), is the following:

$$P, \Gamma \vdash P$$

$$\frac{P, B, \Gamma \vdash C}{P, P \supset B, \Gamma \vdash C} \quad (\supset: \mathit{left}) \qquad \frac{B \supset C, \Gamma \vdash A \supset B \quad C, \Gamma \vdash D}{(A \supset B) \supset C, \Gamma \vdash D} \quad (\supset: \mathit{left})$$

$$\frac{A, \Gamma \vdash B}{\Gamma \vdash A \supset B} \quad (\supset: \mathit{right})$$

where P is atomic.

As we said earlier, the equivalence of the new system $\mathcal{LJT}^{\supset}$ with the previous system $\mathcal{GK}_i^{\supset}$ is a bit challenging. A nice proof is given in Dyckhoff [5].

As an interesting application of the system $\mathcal{LJT}^{\supset}$, we observe that the sequent

$$(((P \supset Q) \supset P) \supset P) \supset Q \vdash Q$$

is provable. However, this sequent is not provable in $\mathcal{G}_i^{\supset}$ without using a contraction. Indeed, the only cut-free and contraction-free proof of this sequent would require proving $(((P \supset Q) \supset P) \supset P)$ in $\mathcal{G}_i^{\supset}$. However, this proposition is nonother than Pierce's law, and we leave it to the reader to verify that it is not provable in the system $\mathcal{LJT}^{\supset}$.

In showing that intuitionistic propositional logic can be embedded in linear logic (without using the exponentials), Lincoln, Scedrov and Shankar [20] use a system $\mathcal{ILL}^*$ almost identical to $\mathcal{LJT}^{\supset}$, except that instead of the rule

$$\frac{P, B, \Gamma \vdash C}{P, P \supset B, \Gamma \vdash C} \quad (\supset: \mathit{left})$$

they use the rule

$$\frac{\Gamma \vdash P \quad B, \Gamma \vdash C}{P \supset B, \Gamma \vdash C} \quad (\supset: \mathit{left})$$

with P atomic. This second rule is obviously sound, and it is trivial that the rule

$$\frac{P, B, \Gamma \vdash C}{P, P \supset B, \Gamma \vdash C} \quad (\supset: \mathit{left})$$

is derivable from the rule

$$\frac{\Gamma \vdash P \quad B, \Gamma \vdash C}{P \supset B, \Gamma \vdash C} \quad (\supset: \mathit{left})$$

Thus, the proof of the equivalence of $\mathcal{LJT}^{\supset}$ and $\mathcal{GK}_i^{\supset}$ directly implies the equivalence of $\mathcal{ILL}^*$ and $\mathcal{GK}_i^{\supset}$ (but the converse is not obvious). However, as far as proof search goes, the system $\mathcal{LJT}^{\supset}$ is superior since it avoids work done on trying to derive P in favour of waiting until it is obvious. On the other hand, $\mathcal{ILL}^*$ is the right system for the translation into linear logic.

Actually, it is possible to formulate a contraction-free system $\mathcal{LJT}^{\supset,\wedge,\vee,\perp}$ for the whole of intuitionistic propositional logic. Such a system given in Dyckhoff [5] is shown below.

Definition 9.1 The axioms and inference rules of the system $\mathcal{LJT}^{\supset,\wedge,\vee,\perp}$ are given below.

$$P, \Gamma \vdash P$$

where P is atomic;

$$\perp, \Gamma \vdash A$$

$$\frac{A, B, \Gamma \vdash C}{A \wedge B, \Gamma \vdash C} \quad (\wedge\colon \mathit{left}) \qquad \frac{\Gamma \vdash A \quad \Gamma \vdash B}{\Gamma \vdash A \wedge B} \quad (\wedge\colon \mathit{right})$$

$$\frac{A, \Gamma \vdash C \quad B, \Gamma \vdash C}{A \vee B, \Gamma \vdash C} \quad (\vee\colon \mathit{left})$$

$$\frac{\Gamma \vdash A}{\Gamma \vdash A \vee B} \quad (\vee\colon \mathit{right}) \qquad \frac{\Gamma \vdash B}{\Gamma \vdash A \vee B} \quad (\vee\colon \mathit{right})$$

$$\frac{P, B, \Gamma \vdash C}{P, P \supset B, \Gamma \vdash C} \quad (\supset\colon \mathit{left}_1)$$

where P is atomic;

$$\frac{A \supset (B \supset C), \Gamma \vdash D}{(A \wedge B) \supset C, \Gamma \vdash D} \quad (\supset\colon \mathit{left}_2)$$

$$\frac{A \supset C, B \supset C, \Gamma \vdash D}{(A \vee B) \supset C, \Gamma \vdash D} \quad (\supset\colon \mathit{left}_3)$$

$$\frac{B \supset C, \Gamma \vdash A \supset B \quad C, \Gamma \vdash D}{(A \supset B) \supset C, \Gamma \vdash D} \quad (\supset\colon \mathit{left}_4)$$

$$\frac{A, \Gamma \vdash B}{\Gamma \vdash A \supset B} \quad (\supset\colon \mathit{right})$$

Among the ($\supset$: *left*) rules, notice that only ($\supset$: left_4) is not invertible. The equivalence of $\mathcal{LJT}^{\supset,\wedge,\vee,\perp}$ and $\mathcal{GK}_i^{\supset,\wedge,\vee,\perp}$ is shown in Dyckhoff [5].

A nice feature of the system $\mathcal{LJT}^{\supset,\wedge,\vee,\perp}$ is that it yields easily the decidability of provability. Note that under the multiset ordering, the complexity of the premises of each rule decreases strictly (we consider the multiset of the weights $w(A)$ of the formulae A occurring in each sequent. We have $w(\perp) = 1$, $w(P) = 1$ for an atom, $w(A \vee B) = w(A \supset B) = w(A) + w(B) + 1$, and $w(A \wedge B) = w(A) + w(B) + 2$). For example, in rule ($\supset$: left_4), $(A \supset B) \supset C$ is replaced by $B \supset C$ and $A \supset B$, both of (strictly) smaller complexity. Thus, this system requires no test of circularity (test for the repetition of sequents to find irredundant proofs), unlike in the system $\mathcal{GK}_i^{\supset,\wedge,\vee,\perp}$.

Pitts [22] reports on applications of the system $\mathcal{LJT}$ to intuitionistic logic with quantification over propositional letters, with interesting applications to the theory of Heyting algebras. An in-depth study of invertibility and admissibility of rules in intuitionistic logic can be found in Paul Rozière's elegant thesis [25].

We now come back to Takeuti's observation [31] (see Chapter 1, paragraph 8). The crucial fact about intuitionistic systems is not so much the fact that sequents are restricted so that righthand sides have at most one formula, but that the application of the rules ($\supset$: *right*) and ($\forall$: *right*) should be *restricted* so that the righthand side of the conclusion of such a rule consists of a *single* formula (and similarly for ($\neg$: *right*) if $\neg$ is not treated as an abbreviation). The intuitive reason is that the rule ($\supset$: *right*) moves some formula from the lefthand side to the righthand side of a sequent (and similarly for ($\neg$: *right*)), and ($\forall$: *right*) involves a side condition. Now, we can view a classical sequent $\Gamma \vdash B_1, \ldots, B_n$ as the corresponding intuitionistic sequent $\Gamma \vdash B_1 \vee \ldots \vee B_n$. With this in mind, we can show the following result.

Lemma 9.2 *Let $\mathcal{GT}_i^{\supset,\wedge,\vee,\forall,\exists,\perp}$ be the system $\mathcal{G}_c^{\supset,\wedge,\vee,\forall,\exists,\perp}$ where the application of the rules ($\supset$: right) and ($\forall$: right) is restricted to situations in which the conclusion of the inference is a sequent whose righthand side has a single formula. Then, $\Gamma \vdash B_1, \ldots, B_n$ is provable in $\mathcal{GT}_i^{\supset,\wedge,\vee,\forall,\exists,\perp}$ iff $\Gamma \vdash B_1 \vee \ldots \vee B_n$ is provable in $\mathcal{G}_i^{\supset,\wedge,\vee,\forall,\exists,\perp}$.*

Proof. The proof is by induction on the structure of proofs. In the case of an axiom $A, \Gamma \vdash \Delta, A$, letting D be the disjunction of the formulae in Δ, we easily obtain a proof of $A, \Gamma \vdash D \vee A$ in $\mathcal{G}_i^{\supset,\wedge,\vee,\forall,\exists,\perp}$ by applications of ($\vee$: *right*) to the axiom $A, \Gamma \vdash A$. Similarly, $\perp, \Gamma \vdash D \vee A$ is provable since it is an axiom. It is also necessary to show that a number of intuitionistic sequents are provable. For example, we need to show that the following sequents are intuitionistically provable:

$$\begin{aligned}
A \supset B, (A \vee D) \wedge (B \supset D) &\vdash D, \\
(A \vee D) \wedge (B \vee D) &\vdash (A \wedge B) \vee D, \\
A[\tau/x] \vee D &\vdash \exists x A \vee D, \\
D \vee \perp &\vdash D \vee A, \\
C \vee C \vee D &\vdash C \vee D.
\end{aligned}$$

Going from $\mathcal{GT}_i^{\supset,\wedge,\vee,\forall,\exists,\perp}$ to $\mathcal{G}_i^{\supset,\wedge,\vee,\forall,\exists,\perp}$, it is much easier to assume that the cut rule can be used in $\mathcal{G}_i$, and then use cut elimination. For example, if the last inference is

$$\frac{\Gamma \vdash A, \Delta \quad B, \Gamma \vdash \Delta}{A \supset B, \Gamma \vdash \Delta} \quad (\supset: \textit{left})$$

letting D be the disjunction of the formulae in Δ, by the induction hypothesis, we have proofs in $\mathcal{G}_i^{\supset,\wedge,\vee,\forall,\exists,\perp}$ of $\Gamma \vdash A \vee D$ and $B, \Gamma \vdash D$. It is obvious that we also have proofs of $A \supset B, \Gamma \vdash A \vee D$ and $A \supset B, \Gamma \vdash B \supset D$, and thus a proof of $A \supset B, \Gamma \vdash (A \vee D) \wedge (B \supset D)$. Since the sequent $A \supset B, \Gamma, (A \vee D) \wedge (B \supset D) \vdash D$ is provable, using a cut, we obtain that $A \supset B, \Gamma \vdash D$ is provable, as desired. The reader should be aware that the special case where Δ is empty can arise and deserves special treatment. In this case, D corresponds to $\perp$. We only treated the case where Δ is nonempty. The case where Δ is empty is actually simpler and is left to the reader. The other cases are similar. $\square$

We can also adapt the system $\mathcal{GK}_c^{\supset,\wedge,\vee,\forall,\exists,\perp}$ to form a system $\mathcal{GKT}_i^{\supset,\wedge,\vee,\forall,\exists,\perp}$ having the same property as $\mathcal{GT}_i^{\supset,\wedge,\vee,\forall,\exists,\perp}$. In this system, it turns out that it is only necessary to recopy the principal formula in the rule ($\supset$: *left*), and of course in the rules ($\forall$: *left*), ($\exists$: *right*). Such a system can be shown to be complete *w.r.t.* Kripke semantics, and can be used to show the existence of a finite counter-model in the case of a refutable proposition. This system is given in the next definition.

Definition 9.3 The axioms and inference rules of the system $\mathcal{GKT}_i^{\supset,\wedge,\vee,\forall,\exists,\perp}$ are given below.

$$A, \Gamma \vdash \Delta, A$$

$$\perp, \Gamma \vdash \Delta, A$$

$$\frac{A, B, \Gamma \vdash \Delta}{A \wedge B, \Gamma \vdash \Delta} \quad (\wedge\colon \mathit{left}) \qquad \frac{\Gamma \vdash \Delta, A \quad \Gamma \vdash \Delta, B}{\Gamma \vdash \Delta, A \wedge B} \quad (\wedge\colon \mathit{right})$$

$$\frac{A, \Gamma \vdash \Delta \quad B, \Gamma \vdash \Delta}{A \vee B, \Gamma \vdash \Delta} \quad (\vee\colon \mathit{left}) \qquad \frac{\Gamma \vdash \Delta, A, B}{\Gamma \vdash \Delta, A \vee B} \quad (\vee\colon \mathit{right})$$

$$\frac{A \supset B, \Gamma \vdash A, \Delta \quad B, \Gamma \vdash \Delta}{A \supset B, \Gamma \vdash \Delta} \quad (\supset\colon \mathit{left}) \qquad \frac{A, \Gamma \vdash B}{\Gamma \vdash A \supset B, \Delta} \quad (\supset\colon \mathit{right})$$

$$\frac{\forall x A, A[\tau/x], \Gamma \vdash \Delta}{\forall x A, \Gamma \vdash \Delta} \quad (\forall\colon \mathit{left}) \qquad \frac{\Gamma \vdash A[y/x]}{\Gamma \vdash \forall x A, \Delta} \quad (\forall\colon \mathit{right})$$

where in ($\forall$: *right*), y does not occur free in the conclusion;

$$\frac{A[y/x], \Gamma \vdash \Delta}{\exists x A, \Gamma \vdash \Delta} \quad (\exists\colon \mathit{left}) \qquad \frac{\Gamma \vdash \Delta, A[\tau/x], \exists x A}{\Gamma \vdash \Delta, \exists x A} \quad (\exists\colon \mathit{right})$$

where in ($\exists$: *left*), y does not occur free in the conclusion.

In the system $\mathcal{GKT}_i^{\supset,\wedge,\vee,\forall,\exists,\perp}$, the application of the rules ($\supset$: *right*) and ($\forall$: *right*) is restricted to premises whose righthand side have a single formula. Using lemmas analogous to lemma 8.5, lemma 8.6, and lemma 8.7, it is possible to show that (*contraction*) and (*weakening*) are derived rules, and that this system is equivalent to $\mathcal{G}_i^{\supset,\wedge,\vee,\forall,\exists,\perp}$, in the sense that a sequent $\Gamma \vdash B_1, \ldots, B_n$ is provable in $\mathcal{GTK}_i^{\supset,\wedge,\vee,\forall,\exists,\perp}$ iff $\Gamma \vdash B_1 \vee \ldots \vee B_n$ is provable in $\mathcal{G}_i^{\supset,\wedge,\vee,\forall,\exists,\perp}$. However, lemma 8.5 now fails for propositions of the form $A \supset B$ or $\forall x A$. However, it can be shown that all the rules are invertible except ($\supset$: *right*) and ($\forall$: *right*). These rules are responsible for the crucial nondeterminism arising in proof search procedures. However, it can be shown that the strategy consisting in alternating phases in which invertible rules are fully applied, and then one of the rules ($\supset$: *right*) or ($\forall$: *right*) is applied once, is complete. As a matter of fact, this strategy amounts to building a Kripke model in the shape of a tree. Failure to complete this model corresponds to provability.[4] In the propositional case, it is also possible to formulate a contraction-free system similar to the system $\mathcal{LJT}^{\supset,\wedge,\vee,\perp}$. Such a system is given in Dyckhoff [5].

10 A Proof-Term Calculus for $\mathcal{G}_i^{\supset,\wedge,\vee,\forall,\exists,\perp,cut}$

Before we move on to the sequent calculi $\mathcal{LK}$ and $\mathcal{LJ}$ and a detailed proof of cut-elimination for these systems, it is worth describing a term calculus corresponding to the sequent calculus $\mathcal{G}_i^{\supset,\wedge,\vee,\forall,\exists,\perp,cut}$. The idea behind the design of the term calculus of this section arose from inspiring conversations with Val Breazu-Tannen, whoom I gladly thank. In this calculus, a sequent $\Gamma \vdash A$

[4] The management of quantified formulae is actually more complicated. For details, see Takeuti [31], Chapter 1, section 8.

becomes a judgement $\Gamma^* \triangleright M\colon A$, such that, if $\Gamma = A_1, \ldots, A_n$ then $\Gamma^* = x_1\colon A_1, \ldots, x_n\colon A_n$ is a context in which the x_i are distinct variables and M is a proof term. Since the sequent calculus has rules for introducing formulae on the left of a sequent as well as on the right of a sequent, we will have to create new variables to tag the newly created formulae, and some new term constructors. The reader should pay particular attention to the use of the `let` construct as a mechanism for suspending substitution.

Definition 10.1 The term calculus associated with $\mathcal{G}_i^{\supset,\wedge,\vee,\forall,\exists,\perp,cut}$ is defined as follows.

$$\Gamma, x\colon A \triangleright x\colon A$$

$$\frac{x\colon A, y\colon A, \Gamma \triangleright M\colon B}{z\colon A, \Gamma \triangleright \texttt{let}\ z\ \texttt{be}\ x\colon A@y\colon A\ \texttt{in}\ M\colon B} \quad (\textit{contrac: left})$$

$$\frac{\Gamma \triangleright M\colon \perp}{\Gamma \triangleright \nabla_A(M)\colon A} \quad (\perp\colon \textit{right})$$

with $A \neq \perp$,

$$\frac{\Gamma \triangleright N\colon A \quad x\colon A, \Gamma \triangleright M\colon C}{\Gamma \triangleright \texttt{let}\ N\ \texttt{be}\ x\colon A\ \texttt{in}\ M\colon C} \quad (\textit{cut})$$

$$\frac{x\colon A, y\colon B, \Gamma \triangleright M\colon C}{z\colon A \wedge B, \Gamma \triangleright \texttt{let}\ z\ \texttt{be}\ \langle x\colon A, y\colon B\rangle\ \texttt{in}\ M\colon C} \quad (\wedge\colon \textit{left})$$

$$\frac{\Gamma \triangleright M\colon A \quad \Gamma \triangleright N\colon B}{\Gamma \triangleright \langle M, N\rangle\colon A \wedge B} \quad (\wedge\colon \textit{right})$$

$$\frac{x\colon A, \Gamma \triangleright M\colon C \quad y\colon B, \Gamma \triangleright N\colon C}{z\colon A \vee B, \Gamma \triangleright \texttt{case}\ z\ \texttt{of}\ \texttt{inl}(x\colon A) \Rightarrow M \mid \texttt{inr}(y\colon B) \Rightarrow N\colon C} \quad (\vee\colon \textit{left})$$

$$\frac{\Gamma \triangleright M\colon A}{\Gamma \triangleright \texttt{inl}(M)\colon A \vee B} \quad (\vee\colon \textit{right}) \qquad \frac{\Gamma \triangleright M\colon B}{\Gamma \triangleright \texttt{inr}(M)\colon A \vee B} \quad (\vee\colon \textit{right})$$

$$\frac{\Gamma \triangleright M\colon A \quad x\colon B, \Gamma \triangleright N\colon C}{z\colon A \supset B, \Gamma \triangleright \texttt{let}\ zM\ \texttt{be}\ x\colon B\ \texttt{in}\ N\colon C} \quad (\supset\colon \textit{left})$$

$$\frac{x\colon A, \Gamma \triangleright M\colon B}{\Gamma \triangleright (\lambda x\colon A.\ M)\colon A \supset B} \quad (\supset\colon \textit{right})$$

$$\frac{x\colon A[\tau/t], \Gamma \triangleright M\colon C}{z\colon \forall t A, \Gamma \triangleright \texttt{let}\ z\tau\ \texttt{be}\ x\colon A[\tau/t]\ \texttt{in}\ M\colon C} \quad (\forall\colon \textit{left})$$

$$\frac{\Gamma \triangleright M\colon A[u/t]}{\Gamma \triangleright (\lambda u\colon \iota.\ M)\colon \forall t A} \quad (\forall\colon \textit{right})$$

where u does not occur free in Γ or $\forall t A$;

$$\frac{x\colon A[u/t], \Gamma \triangleright M\colon C}{z\colon \exists t A, \Gamma \triangleright \texttt{casex}\ z\ \texttt{of}\ \texttt{inx}(u\colon \iota, x\colon A[u/t]) \Rightarrow M\colon C} \quad (\exists\colon \textit{left})$$

where u does not occur free in Γ, $\exists t A$, or C;

$$\frac{\Gamma \triangleright M\colon A[\tau/t]}{\Gamma \triangleright \mathtt{inx}(\tau, M)\colon \exists t A} \quad (\exists\colon \textit{right})$$

The use of the `let` construct in the cut rule and the rules ($\supset$: *left*) and ($\forall$: *left*) should be noted. The effect of the `let` construct is to *suspend* substitution, and thus to allow more reductions to take place. Most presentations use the following alternate rules in which substitution takes place immediately:

$$\frac{\Gamma \triangleright N\colon A \quad x\colon A, \Gamma \triangleright M\colon C}{\Gamma \triangleright M[N/x]\colon C} \quad (\textit{cut})$$

$$\frac{\Gamma \triangleright M\colon A \quad x\colon B, \Gamma \triangleright N\colon C}{z\colon A \supset B, \Gamma \triangleright N[(zM)/x]\colon C} \quad (\supset\colon \textit{left})$$

$$\frac{x\colon A[\tau/t], \Gamma \triangleright M\colon C}{z\colon \forall t A, \Gamma \triangleright M[(z\tau)/x]\colon C} \quad (\forall\colon \textit{left})$$

Thus, in some sense, a reduction strategy has already been imposed. This (usual) presentation facilitates the comparison with natural deduction, but makes it impossible to describe general cut-elimination rules. For the sake of historical accuracy, note that essentially the same system appears in section 2 of Girard's classic paper [10]. With our system, it is possible to write reduction rules that correspond to cut-elimination steps (see section 12). For example,

$$\mathtt{let}\ \lambda x\colon A.\, M_1\ \mathtt{be}\ z\colon (A \supset B)\ \mathtt{in}\ (\mathtt{let}\ zM_2\ \mathtt{be}\ y\colon B\ \mathtt{in}\ N)$$
$$\longrightarrow \mathtt{let}\ (\mathtt{let}\ M_2\ \mathtt{be}\ x\colon A\ \mathtt{in}\ M_1)\ \mathtt{be}\ y\colon B\ \mathtt{in}\ N,$$

$$\mathtt{let}\ \langle M, N\rangle\ \mathtt{be}\ \langle x\colon A, y\colon B\rangle\ \mathtt{in}\ P \longrightarrow \mathtt{let}\ M\ \mathtt{be}\ x\colon A\ \mathtt{in}\ (\mathtt{let}\ N\ \mathtt{be}\ y\colon B\ \mathtt{in}\ P),$$

$$\mathtt{let}\ \langle M, N\rangle\ \mathtt{be}\ \langle x\colon A, y\colon B\rangle\ \mathtt{in}\ P \longrightarrow \mathtt{let}\ N\ \mathtt{be}\ y\colon B\ \mathtt{in}\ (\mathtt{let}\ M\ \mathtt{be}\ x\colon A\ \mathtt{in}\ P),$$

$$\mathtt{let}\ M\ \mathtt{be}\ u\colon A\ \mathtt{in}\ (\mathtt{let}\ z\ \mathtt{be}\ \langle x\colon C, y\colon D\rangle\ \mathtt{in}\ N) \longrightarrow \mathtt{let}\ z\ \mathtt{be}\ \langle x\colon C, y\colon D\rangle\ \mathtt{in}\ (\mathtt{let}\ M\ \mathtt{be}\ u\colon A\ \mathtt{in}\ N),$$

$$\mathtt{let}\ M\ \mathtt{be}\ x\colon A\ \mathtt{in}\ \langle N_1, N_2\rangle \longrightarrow \langle \mathtt{let}\ M\ \mathtt{be}\ x\colon A\ \mathtt{in}\ N_1, \mathtt{let}\ M\ \mathtt{be}\ x\colon A\ \mathtt{in}\ N_2\rangle,$$

$$\mathtt{let}\ M\ \mathtt{be}\ x\colon A\ \mathtt{in}\ x \longrightarrow M.$$

It can be shown that

$$\mathtt{let}\ N\ \mathtt{be}\ x\colon A\ \mathtt{in}\ M \longrightarrow^{+} M[N/x].$$

However, the reduction rules corresponding to cut elimination are finer than β-conversion. This is the reason why it is quite difficult to prove a strong version of cut-elimination where all reduction sequences terminate. Such a proof was given by Dragalin [4]. If the alternate rules

$$\frac{\Gamma \triangleright N\colon A \quad x\colon A, \Gamma \triangleright M\colon C}{\Gamma \triangleright M[N/x]\colon C} \quad (\textit{cut})$$

$$\frac{\Gamma \triangleright M\colon A \quad x\colon B, \Gamma \triangleright N\colon C}{z\colon A \supset B, \Gamma \triangleright N[(zM)/x]\colon C} \quad (\supset\colon \textit{left})$$

$$\frac{x\colon A[\tau/t], \Gamma \triangleright M\colon C}{z\colon \forall t A, \Gamma \triangleright M[(z\tau)/x]\colon C} \quad (\forall\colon \textit{left})$$

are used, then the reduction rules take a different form, For example,

$$\texttt{let}\ \langle M, N\rangle\ \texttt{be}\ \langle x\colon A, y\colon B\rangle\ \texttt{in}\ P \longrightarrow P[M/x, N/y].$$

This amounts to imposing certain strategies on the reductions in our calculus. In fact, it is possible to specify reduction rules imposing certain strategies, for example, *eager* or *lazy* evaluation. Such reduction strategies have been considered in a similar setting for linear logic by Abramsky [1].

The above proof-term assignment has the property that if $\Gamma \triangleright M\colon A$ is derivable and $\Gamma \subseteq \Delta$, then $\Delta \triangleright M\colon A$ is also derivable. This is because the axioms are of the form $\Gamma, x\colon A \triangleright x\colon A$. We can design a term assignment system for an $\mathcal{LJ}$-style system. In such a system, the axioms are of the form

$$x\colon A \triangleright x\colon A$$

and the proof-term assignment for weakening is as follows:

$$\frac{\Gamma \triangleright M\colon B}{z\colon A, \Gamma \triangleright \texttt{let}\ z\ \texttt{be}\ \_\ \texttt{in}\ M\colon B} \quad (\textit{weakening}\colon \textit{left})$$

Note that the above proof-term assignment has the property that if $\Gamma \triangleright M\colon A$ is provable and $\Gamma \subseteq \Delta$, then $\Delta \triangleright N\colon A$ is also derivable for some N easily obtainable from M.

If instead of the above $(\wedge\colon \textit{left})$ rule, we use the two $\mathcal{LJ}$-style rules

$$\frac{A, \Gamma \vdash C}{A \wedge B, \Gamma \vdash C} \quad (\wedge\colon \textit{left}) \qquad \frac{B, \Gamma \vdash C}{A \wedge B, \Gamma \vdash C} \quad (\wedge\colon \textit{left})$$

then we have the following proof-term assignment:

$$\frac{x\colon A, \Gamma \triangleright M\colon C}{z\colon A \wedge B, \Gamma \triangleright \texttt{let}\ z\ \texttt{be}\ \langle x\colon A, \_\rangle\ \texttt{in}\ M\colon C} \quad (\wedge\colon \textit{left})$$

$$\frac{y\colon B, \Gamma \triangleright M\colon C}{z\colon A \wedge B, \Gamma \triangleright \texttt{let}\ z\ \texttt{be}\ \langle \_, y\colon B\rangle\ \texttt{in}\ M\colon C} \quad (\wedge\colon \textit{left})$$

It is then natural to write the normalization rules as

$$\texttt{let}\ \langle M, N\rangle\ \texttt{be}\ \langle x\colon A, \_\rangle\ \texttt{in}\ P \longrightarrow \texttt{let}\ M\ \texttt{be}\ x\colon A\ \texttt{in}\ P,$$
$$\texttt{let}\ \langle M, N\rangle\ \texttt{be}\ \langle \_, y\colon B\rangle\ \texttt{in}\ P \longrightarrow \texttt{let}\ N\ \texttt{be}\ y\colon B\ \texttt{in}\ P.$$

We note that for these new rules, the reduction is *lazy*, in the sense that it is unnecessary to normalize N (or M) since it is discarded. With the old rules, the reduction is generally *eager* since both M and N will have to be normalized, unless x or y do not appear in P. Such aspects of lazy or eager evaluation become even more obvious in linear logic, as stressed by Abramsky [1].

We now consider some equivalent Gentzen systems.

11 The Gentzen Systems $\mathcal{LJ}$ and $\mathcal{LK}$

Axioms of the form $A, \Gamma \vdash \Delta, A$ are very convenient for searching for proofs backwards, but for logical purity, it may be desirable to consider axioms of the form $A \vdash A$. We can redefine axioms to be of this simpler form, but to preserve exactly the same notion of provability, we need to add the following rules of *weakening* (also called *thinning*).

Definition 11.1 The rules of *weakening* (or *thinning*) are

$$\frac{\Gamma \vdash \Delta}{A, \Gamma \vdash \Delta} \quad (\textit{weakening}\text{: } \textit{left}) \qquad \frac{\Gamma \vdash \Delta}{\Gamma \vdash \Delta, A} \quad (\textit{weakening}\text{: } \textit{right})$$

In the case of intuitionistic logic, we require that Δ be empty in (*weakening*: *right*).

One can also observe that in order to make the ($\wedge$: *left*) rule and the ($\vee$: *right*) rule analogous to the corresponding introduction rules in natural deduction, we can introduce the rules

$$\frac{A, \Gamma \vdash \Delta}{A \wedge B, \Gamma \vdash \Delta} \quad (\wedge\text{: } \textit{left}) \qquad \frac{B, \Gamma \vdash \Delta}{A \wedge B, \Gamma \vdash \Delta} \quad (\wedge\text{: } \textit{left})$$

and

$$\frac{\Gamma \vdash \Delta, A}{\Gamma \vdash \Delta, A \vee B} \quad (\vee\text{: } \textit{right}) \qquad \frac{\Gamma \vdash \Delta, B}{\Gamma \vdash \Delta, A \vee B} \quad (\vee\text{: } \textit{right})$$

They are equivalent to the old rules provided that we add (*contrac*: *left*), (*contrac*: *right*), (*weakening*: *left*) and (*weakening*: *right*). This leads us to the systems $\mathcal{LJ}$ and $\mathcal{LK}$ defined and studied by Gentzen [8] (except that Gentzen also had an explicit exchange rule, but we assume that we are dealing with multisets).

Definition 11.2 The axioms and inference rules of the system $\mathcal{LJ}$ for intuitionistic first-order logic are given below.

Axioms:

$$A \vdash A$$

Structural Rules:

$$\frac{\Gamma \vdash \Delta}{A, \Gamma \vdash \Delta} \quad (\textit{weakening}\text{: } \textit{left}) \qquad \frac{\Gamma \vdash}{\Gamma \vdash A} \quad (\textit{weakening}\text{: } \textit{right})$$

$$\frac{A, A, \Gamma \vdash \Delta}{A, \Gamma \vdash \Delta} \quad (\textit{contrac}\text{: } \textit{left})$$

$$\frac{\Gamma \vdash A \quad A, \Lambda \vdash \Theta}{\Gamma, \Lambda \vdash \Theta} \quad (\textit{cut})$$

Logical Rules:

$$\frac{A, \Gamma \vdash \Delta}{A \wedge B, \Gamma \vdash \Delta} \quad (\wedge\text{: } \textit{left}) \qquad \frac{B, \Gamma \vdash \Delta}{A \wedge B, \Gamma \vdash \Delta} \quad (\wedge\text{: } \textit{left})$$

$$\frac{\Gamma \vdash A \quad \Gamma \vdash B}{\Gamma \vdash A \wedge B} \quad (\wedge\colon \mathit{right})$$

$$\frac{A, \Gamma \vdash \Delta \quad B, \Gamma \vdash \Delta}{A \vee B, \Gamma \vdash \Delta} \quad (\vee\colon \mathit{left})$$

$$\frac{\Gamma \vdash A}{\Gamma \vdash A \vee B} \quad (\vee\colon \mathit{right}) \qquad \frac{\Gamma \vdash B}{\Gamma \vdash A \vee B} \quad (\vee\colon \mathit{right})$$

$$\frac{\Gamma \vdash A \quad B, \Gamma \vdash \Delta}{A \supset B, \Gamma \vdash \Delta} \quad (\supset\colon \mathit{left}) \qquad \frac{A, \Gamma \vdash B}{\Gamma \vdash A \supset B} \quad (\supset\colon \mathit{right})$$

$$\frac{\Gamma \vdash A}{\neg A, \Gamma \vdash} \quad (\neg\colon \mathit{left}) \qquad \frac{A, \Gamma \vdash}{\Gamma \vdash \neg A} \quad (\neg\colon \mathit{right})$$

In the logical rules above, $A \vee B$, $A \wedge B$, $A \supset B$, and $\neg A$ are called the *principal formulae* and A, B the *side formulae* of the inference.

$$\frac{A[\tau/x], \Gamma \vdash \Delta}{\forall x A, \Gamma \vdash \Delta} \quad (\forall\colon \mathit{left}) \qquad \frac{\Gamma \vdash A[y/x]}{\Gamma \vdash \forall x A} \quad (\forall\colon \mathit{right})$$

where in ($\forall$: *right*), y does not occur free in the conclusion;

$$\frac{A[y/x], \Gamma \vdash \Delta}{\exists x A, \Gamma \vdash \Delta} \quad (\exists\colon \mathit{left}) \qquad \frac{\Gamma \vdash A[\tau/x]}{\Gamma \vdash \exists x A} \quad (\exists\colon \mathit{right})$$

where in ($\exists$: *left*), y does not occur free in the conclusion.

In the above rules, Δ and Θ consist of at most one formula. The variable y is called the *eigenvariable* of the inference. The condition that the eigenvariable does not occur free in the conclusion of the rule is called the *eigenvariable condition*. The formula $\forall x A$ (or $\exists x A$) is called the *principal formula* of the inference, and the formula $A[\tau/x]$ (or $A[y/x]$) the *side formula* of the inference.

Definition 11.3 The axioms and inference rules of the system $\mathcal{LK}$ for classical first-order logic are given below.

Axioms:

$$A \vdash A$$

Structural Rules:

$$\frac{\Gamma \vdash \Delta}{A, \Gamma \vdash \Delta} \quad (\mathit{weakening}\colon \mathit{left}) \qquad \frac{\Gamma \vdash \Delta}{\Gamma \vdash \Delta, A} \quad (\mathit{weakening}\colon \mathit{right})$$

$$\frac{A, A, \Gamma \vdash \Delta}{A, \Gamma \vdash \Delta} \quad (\mathit{contrac}\colon \mathit{left}) \qquad \frac{\Gamma \vdash \Delta, A, A}{\Gamma \vdash \Delta, A} \quad (\mathit{contrac}\colon \mathit{right})$$

$$\frac{\Gamma \vdash \Delta, A \quad A, \Lambda \vdash \Theta}{\Gamma, \Lambda \vdash \Delta, \Theta} \quad (\mathit{cut})$$

Logical Rules:

$$\frac{A, \Gamma \vdash \Delta}{A \land B, \Gamma \vdash \Delta} \quad (\land\colon \textit{left}) \qquad \frac{B, \Gamma \vdash \Delta}{A \land B, \Gamma \vdash \Delta} \quad (\land\colon \textit{left})$$

$$\frac{\Gamma \vdash \Delta, A \quad \Gamma \vdash \Delta, B}{\Gamma \vdash \Delta, A \land B} \quad (\land\colon \textit{right})$$

$$\frac{A, \Gamma \vdash \Delta \quad B, \Gamma \vdash \Delta}{A \lor B, \Gamma \vdash \Delta} \quad (\lor\colon \textit{left})$$

$$\frac{\Gamma \vdash \Delta, A}{\Gamma \vdash \Delta, A \lor B} \quad (\lor\colon \textit{right}) \qquad \frac{\Gamma \vdash \Delta, B}{\Gamma \vdash \Delta, A \lor B} \quad (\lor\colon \textit{right})$$

$$\frac{\Gamma \vdash \Delta, A \quad B, \Gamma \vdash \Delta}{A \supset B, \Gamma \vdash \Delta} \quad (\supset\colon \textit{left}) \qquad \frac{A, \Gamma \vdash \Delta, B}{\Gamma \vdash \Delta, A \supset B} \quad (\supset\colon \textit{right})$$

$$\frac{\Gamma \vdash \Delta, A}{\neg A, \Gamma \vdash \Delta} \quad (\neg\colon \textit{left}) \qquad \frac{A, \Gamma \vdash \Delta}{\Gamma \vdash \Delta, \neg A} \quad (\neg\colon \textit{right})$$

In the logical rules above, $A \lor B$, $A \land B$, $A \supset B$, and $\neg A$ are called the *principal formulae* and A, B the *side formulae* of the inference.

$$\frac{A[\tau/x], \Gamma \vdash \Delta}{\forall x A, \Gamma \vdash \Delta} \quad (\forall\colon \textit{left}) \qquad \frac{\Gamma \vdash \Delta, A[y/x]}{\Gamma \vdash \Delta, \forall x A} \quad (\forall\colon \textit{right})$$

where in ($\forall$: *right*), y does not occur free in the conclusion;

$$\frac{A[y/x], \Gamma \vdash \Delta}{\exists x A, \Gamma \vdash \Delta} \quad (\exists\colon \textit{left}) \qquad \frac{\Gamma \vdash \Delta, A[\tau/x]}{\Gamma \vdash \Delta, \exists x A} \quad (\exists\colon \textit{right})$$

where in ($\exists$: *left*), y does not occur free in the conclusion.

The variable y is called the *eigenvariable* of the inference. The condition that the eigenvariable does not occur free in the conclusion of the rule is called the *eigenvariable condition*. The formula $\forall x A$ (or $\exists x A$) is called the *principal formula* of the inference, and the formula $A[\tau/x]$ (or $A[y/x]$) the *side formula* of the inference.

One will note that the cut rule (*multiplicative* version)

$$\frac{\Gamma \vdash \Delta, A \quad A, \Lambda \vdash \Theta}{\Gamma, \Lambda \vdash \Delta, \Theta} \quad (\textit{cut})$$

(with Δ empty in the intuitionistic case and Θ at most one formula) differs from the cut rule (*additive* version)

$$\frac{\Gamma \vdash \Delta, A \quad A, \Gamma \vdash \Delta}{\Gamma \vdash \Delta} \quad (\textit{cut})$$

used in $\mathcal{G}_c^{\supset,\land,\lor,\forall,\exists,\neg,cut}$ or in $\mathcal{G}_i^{\supset,\land,\lor,\forall,\exists,\bot,cut}$, in that the premises do not require the contexts Γ, Λ to coincide, and the contexts Δ, Θ to coincide. These rules are equivalent using contraction and weakening. Similarly, the other logical rules of $\mathcal{LK}$ (resp. $\mathcal{LJ}$) and $\mathcal{G}_c^{\supset,\land,\lor,\forall,\exists,\neg,cut}$ (resp. $\mathcal{G}_i^{\supset,\land,\lor,\forall,\exists,\neg,cut}$) are equivalent using contraction and weakening.

12 Cut Elimination in $\mathcal{LK}$ (and $\mathcal{LJ}$)

The cut elimination theorem also applies to $\mathcal{LK}$ and $\mathcal{LJ}$. Historically, this is the version of the cut elimination theorem proved by Gentzen [8] (1935). Gentzen's proof was later simplified by Tait [29] and Girard [13] (especially the induction measure). A simplified version of Tait's proof is nicely presented by Schwichtenberg [26]. The proof given here combines ideas from Tait and Girard. The induction measure used is due to Tait [29] (the cut-rank), but the explicit transformations are adapted from Girard [13], [9]. We need to define the cut-rank of a formula, the depth of a proof, and the logical depth of a proof.

Definition 12.1 The *degree* $|A|$ of a formula A is the number of logical connectives in A. Let Π be an $\mathcal{LK}$-proof. The *cut-rank* $c(\Pi)$ of Π is defined inductively as follows. If Π is an axiom, then $c(\Pi) = 0$. If Π is not an axiom, the last inference has either one or two premises. In the first case, the premise of that inference is the root of a subtree Π_1. In the second case, the left premise is the root of a subtree Π_1, and the right premise is the root of a subtree Π_2. If the last inference is not a cut, then if it has a single premise, $c(\Pi) = c(\Pi_1)$, else $c(\Pi) = max(c(\Pi_1), c(\Pi_2))$. If the last inference is a cut with cut formula A, then $c(\Pi) = max(\{|A| + 1, c(\Pi_1), c(\Pi_2)\})$. The *depth* of a proof tree Π, denoted as $d(\Pi)$, is defined inductively as follows: $d(\Pi) = 0$, when Π is an axiom. If the root of Π is a single-premise rule, then $d(\Pi) = d(\Pi_1) + 1$. If the root of Π is a two-premise rule, then $d(\Pi) = max(d(\Pi_1), d(\Pi_2)) + 1$. We also define the *logical depth* of a proof tree Π, denoted as $l(\Pi)$, inductively as follows: $l(\Pi) = 0$, when Π is an axiom. If the root of Π is a single-premise rule, then if the lowest rule is structural, $l(\Pi) = l(\Pi_1)$, else $l(\Pi) = l(\Pi_1) + 1$. If the root of Π is a two-premise rule, then $l(\Pi) = max(l(\Pi_1), l(\Pi_2)) + 1$.

Thus, for an atomic formula, $|A| = 0$. Note that $c(\Pi) = 0$ iff Π is cut free, and that if Π contains cuts, then $c(\Pi)$ is $1 +$ the maximum of the degrees of cut formulae in Π. The difference between the depth $d(\Pi)$ and the logical depth $l(\Pi)$ is that structural rules are not counted in $l(\Pi)$. Both are needed in lemma 12.3, $d(\Pi)$ as an induction measure, and $l(\Pi)$ to bound the increase on the size of a proof. We also need the definition of the function $exp(m, n, p)$.

$$exp(m, 0, p) = p;$$
$$exp(m, n + 1, p) = m^{exp(m,n,p)}.$$

This function grows extremely fast in the argument n. Indeed, $exp(m, 1, p) = m^p$, $exp(m, 2, p) = m^{m^p}$, and in general, $exp(m, n, p)$ is an iterated stack of exponentials of height n, topped with a p:

$$exp(m, n, p) = m^{m^{m^{\cdot^{\cdot^{\cdot^{m^p}}}}}} \Big\} n$$

The main idea is to move the cuts "upward", until one of the two premises involved is an axiom. In attempting to design transformations for converting an $\mathcal{LK}$-proof into a cut-free $\mathcal{LK}$-proof, we have to deal with the case in which the cut formula A is contracted in some premise. A transformation to handle this case is given below.

$$
\begin{prooftree}
\AxiomC{π_1}
\noLine
\UnaryInfC{$\Gamma \vdash \Delta, A, A$}
\UnaryInfC{$\Gamma \vdash \Delta, A$}
\AxiomC{π_2}
\noLine
\UnaryInfC{$A, \Lambda \vdash \Theta$}
\BinaryInfC{$\Gamma, \Lambda \vdash \Delta, \Theta$}
\end{prooftree}
$$

$$\Longrightarrow$$

$$
\begin{prooftree}
\AxiomC{π_1}
\noLine
\UnaryInfC{$\Gamma \vdash \Delta, A, A$}
\AxiomC{π_2}
\noLine
\UnaryInfC{$A, \Lambda \vdash \Theta$}
\BinaryInfC{$\Gamma, \Lambda \vdash \Delta, \Theta, A$}
\AxiomC{π_2}
\UnaryInfC{$A, \Lambda \vdash \Theta$}
\BinaryInfC{$\Gamma, \Lambda, \Lambda \vdash \Delta, \Theta, \Theta$}
\doubleLine
\UnaryInfC{$\Gamma, \Lambda \vdash \Delta, \Theta$}
\end{prooftree}
$$

The symmetric rule in which a contraction takes place in the right subtree is not shown. However, there is a problem with this transformation. The problem is that it yields infinite reduction sequences. Consider the following two transformation steps:

$$
\begin{prooftree}
\AxiomC{π_1}
\noLine
\UnaryInfC{$\Gamma \vdash \Delta, C, C$}
\UnaryInfC{$\Gamma \vdash \Delta, C$}
\AxiomC{π_2}
\noLine
\UnaryInfC{$C, C, \Lambda \vdash \Theta$}
\UnaryInfC{$C, \Lambda \vdash \Theta$}
\RightLabel{(cut)}
\BinaryInfC{$\Gamma, \Lambda \vdash \Delta, \Theta$}
\end{prooftree}
$$

$$\Longrightarrow$$

$$
\begin{prooftree}
\AxiomC{π_1}
\noLine
\UnaryInfC{$\Gamma \vdash \Delta, C, C$}
\AxiomC{π_2}
\noLine
\UnaryInfC{$C, C, \Lambda \vdash \Theta$}
\UnaryInfC{$C, \Lambda \vdash \Theta$}
\BinaryInfC{$\Gamma, \Lambda \vdash \Delta, \Theta, C$}
\AxiomC{π_2}
\noLine
\UnaryInfC{$C, C, \Lambda \vdash \Theta$}
\UnaryInfC{$C, \Lambda \vdash \Theta$}
\BinaryInfC{$\Gamma, \Lambda, \Lambda \vdash \Delta, \Theta, \Theta$}
\doubleLine
\UnaryInfC{$\Gamma, \Lambda \vdash \Delta, \Theta$}
\end{prooftree}
$$

$$\Longrightarrow$$

$$
\begin{prooftree}
\AxiomC{π_1}
\noLine
\UnaryInfC{$\Gamma \vdash \Delta, C, C$}
\AxiomC{π_1}
\noLine
\UnaryInfC{$\Gamma \vdash \Delta, C, C$}
\AxiomC{π_2}
\noLine
\UnaryInfC{$C, C, \Lambda \vdash \Theta$}
\BinaryInfC{$C, \Gamma, \Lambda \vdash \Delta, \Theta, C$}
\BinaryInfC{$\Gamma, \Gamma, \Lambda \vdash \Delta, \Delta, \Theta, C, C$}
\doubleLine
\UnaryInfC{$\Gamma, \Lambda \vdash \Delta, \Theta, C, C$}
\UnaryInfC{$\Gamma, \Lambda \vdash \Delta, \Theta, C$}
\AxiomC{π_2}
\noLine
\UnaryInfC{$C, C, \Lambda \vdash \Theta$}
\UnaryInfC{$C, \Lambda \vdash \Theta$}
\BinaryInfC{$\Gamma, \Lambda, \Lambda \vdash \Delta, \Theta, \Theta$}
\doubleLine
\UnaryInfC{$\Gamma, \Lambda \vdash \Delta, \Theta$}
\end{prooftree}
$$

The pattern with contractions on the left and on the right is repeated.

One solution is to consider a more powerful kind of cut rule. In the sequel, the multiset Γ, nA denotes the multiset consisting of all occurrences of $B \neq A$ in Γ and of $m+n$ occurrences of A where m is the number of occurrences of A in Γ.

Definition 12.2 (*Extended cut rule*)

$$\frac{\Gamma \vdash \Delta, mA \qquad nA, \Lambda \vdash \Theta}{\Gamma, \Lambda \vdash \Delta, \Theta}$$

where $m, n > 0$.

This rule coincides with the standard cut rule when $m = n = 1$, and it is immediately verified that it can be simulated by an application of the standard cut rule and some applications of the contraction rules. Thus, the system $\mathcal{LK}^+$ obtained from $\mathcal{LK}$ by replacing the cut rule by the extended cut rule is equivalent to $\mathcal{LK}$. From now on, we will be working with $\mathcal{LK}^+$. The problem with contraction is then resolved, since we have the following transformation:

$$\frac{\dfrac{\dfrac{\pi_1}{\Gamma \vdash \Delta, (m-1)A, A, A}}{\Gamma \vdash \Delta, mA} \qquad \dfrac{\pi_2}{nA, \Lambda \vdash \Theta}}{\Gamma, \Lambda \vdash \Delta, \Theta}$$

$$\Longrightarrow$$

$$\frac{\dfrac{\pi_1}{\Gamma \vdash \Delta, (m+1)A} \qquad \dfrac{\pi_2}{nA, \Lambda \vdash \Theta}}{\Gamma, \Lambda \vdash \Delta, \Theta}$$

We now prove the main lemma, for which a set of transformations will be needed.

Lemma 12.3 [Reduction Lemma, Tait, Girard] *Let Π_1 be an $\mathcal{LK}^+$-proof of $\Gamma \vdash \Delta, mA$, and Π_2 an $\mathcal{LK}^+$-proof of $nA, \Lambda \vdash \Theta$, where $m, n > 0$, and assume that $c(\Pi_1), c(\Pi_2) \leq |A|$. An $\mathcal{LK}^+$-proof Π of $\Gamma, \Lambda \vdash \Delta, \Theta$ can be constructed, such that $c(\Pi) \leq |A|$. We also have $l(\Pi) \leq 2(l(\Pi_1) + l(\Pi_2))$, and if the rules for $\supset$ are omitted, then $l(\Pi) \leq l(\Pi_1) + l(\Pi_2)$.*

Proof. It proceeds by induction on $d(\Pi_1) + d(\Pi_2)$,[5] where Π_1 and Π_2 are the immediate subtrees of the proof tree

$$\frac{\dfrac{\Pi_1}{\Gamma \vdash \Delta, mA} \qquad \dfrac{\Pi_2}{nA, \Lambda \vdash \Theta}}{\Gamma, \Lambda \vdash \Delta, \Theta}$$

There are several (non-mutually exclusive) cases depending on the structure of the immediate subtrees Π_1 and Π_2.

(1) The root of Π_1 and the root of Π_2 is the conclusion of some logical inference having some occurrence of the cut formula A as principal formula. We say that A is *active*.

[5] The reader is warned that earlier versions of this proof used the wrong measure $l(\Pi_1) + l(\Pi_2)$. Indeed, $l(\Pi_1) + l(\Pi_2)$ does **not** decrease in the case of the structural rules (in cases (2) and (4) of the forthcoming proof).

Every transformation comes in two versions. The first version corresponds to the case of an application of the standard cut rule. The other version, called the "cross-cuts" version, applies when the extended cut rule is involved.

(i) ($\wedge$: *right*) and ($\wedge$: *left*)

$$\dfrac{\dfrac{\overset{\pi_1}{\Gamma \vdash \Delta, B} \qquad \overset{\pi_2}{\Gamma \vdash \Delta, C}}{\Gamma \vdash \Delta, B \wedge C} \qquad \dfrac{\overset{\pi_3}{B, \Lambda \vdash \Theta}}{B \wedge C, \Lambda \vdash \Theta}}{\Gamma, \Lambda \vdash \Delta, \Theta}$$

$$\Longrightarrow$$

$$\dfrac{\overset{\pi_1}{\Gamma \vdash \Delta, B} \qquad \overset{\pi_3}{B, \Lambda \vdash \Theta}}{\Gamma, \Lambda \vdash \Delta, \Theta}$$

By the hypothesis $c(\Pi_1), c(\Pi_2) \leq |A|$, and it is clear that for the new proof Π we have $c(\Pi) \leq |A|$, since $c(\Pi) = max(\{|B| + 1, c(\pi_1), c(\pi_3)\})$, $|B| + 1 \leq |A|$ (since $A = B \wedge C$), $c(\pi_1) \leq c(\Pi_1)$, $c(\pi_2) \leq c(\Pi_1)$, and $c(\pi_3) \leq c(\Pi_2)$. It is also easy to establish the upper bound on $l(\Pi)$.[6]

Cross-cuts version. Some obvious simplifications apply when either $m = 0$ or $n = 0$, and we only show the main case where $m, n > 0$. Let $A = B \wedge C$.

$$\dfrac{\dfrac{\overset{\pi_1}{\Gamma \vdash \Delta, mA, B} \qquad \overset{\pi_2}{\Gamma \vdash \Delta, mA, C}}{\Gamma \vdash \Delta, (m+1)A} \qquad \dfrac{\overset{\pi_3}{B, nA, \Lambda \vdash \Theta}}{(n+1)A, \Lambda \vdash \Theta}}{\Gamma, \Lambda \vdash \Delta, \Theta}$$

Let Π_1' be the proof tree obtained by applying the induction hypothesis to

$$\dfrac{\overset{\pi_1}{\Gamma \vdash \Delta, mA, B} \qquad \dfrac{\overset{\pi_3}{B, nA, \Lambda \vdash \Theta}}{(n+1)A, \Lambda \vdash \Theta}}{\Gamma, \Lambda \vdash \Delta, \Theta, B}$$

and Π_2' the proof tree obtained by applying the induction hypothesis to

$$\dfrac{\dfrac{\overset{\pi_1}{\Gamma \vdash \Delta, mA, B} \qquad \overset{\pi_2}{\Gamma \vdash \Delta, mA, C}}{\Gamma \vdash \Delta, (m+1)A} \qquad \overset{\pi_3}{B, nA, \Lambda \vdash \Theta}}{B, \Gamma, \Lambda \vdash \Delta, \Theta}$$

and finally let Π be

[6] The simple fact that $max(a, b) \leq a + b$ for $a, b \geq 0$ is used here and in the next cases.

$$\dfrac{\dfrac{\begin{array}{c}\Pi_1' \\ \Gamma, \Lambda \vdash \Delta, \Theta, B\end{array} \qquad \begin{array}{c}\Pi_2' \\ B, \Gamma, \Lambda \vdash \Delta, \Theta\end{array}}{\Gamma, \Gamma, \Lambda, \Lambda \vdash \Delta, \Delta, \Theta, \Theta}}{\overline{\Gamma, \Lambda \vdash \Delta, \Theta}}$$

Since $c(\pi_1) \leq c(\Pi_1)$, $c(\pi_2) \leq c(\Pi_1)$, and $c(\pi_3) \leq c(\Pi_2)$, by the induction hypothesis, we have $c(\Pi_1'), c(\Pi_2') \leq |A|$, and it is clear that for the new proof Π we have $c(\Pi) \leq |A|$, since $c(\Pi) = max(\{|B| + 1, c(\Pi_1'), c(\Pi_2')\})$, and $|B| + 1 \leq |A|$ (since $A = B \wedge C$). It is also easy to establish the upper bound on $l(\Pi)$.

(ii) ($\vee$: *right*) and ($\vee$: *left*)

$$\dfrac{\dfrac{\begin{array}{c}\pi_1 \\ \Gamma \vdash \Delta, B\end{array}}{\Gamma \vdash \Delta, B \vee C} \qquad \dfrac{\begin{array}{c}\pi_2 \\ B, \Lambda \vdash \Theta\end{array} \qquad \begin{array}{c}\pi_3 \\ C, \Lambda \vdash \Theta\end{array}}{B \vee C, \Lambda \vdash \Theta}}{\Gamma, \Lambda \vdash \Delta, \Theta}$$

$$\Longrightarrow$$

$$\dfrac{\begin{array}{c}\pi_1 \\ \Gamma \vdash \Delta, B\end{array} \qquad \begin{array}{c}\pi_2 \\ B, \Lambda \vdash \Theta\end{array}}{\Gamma, \Lambda \vdash \Delta, \Theta}$$

By the hypothesis $c(\Pi_1), c(\Pi_2) \leq |A|$, it is clear that for the new proof Π we have $c(\Pi) \leq |A|$, since $c(\Pi) = max(\{|B| + 1, c(\pi_1), c(\pi_2)\})$, $|B| + 1 \leq |A|$ (since $A = B \vee C$), $c(\pi_1) \leq c(\Pi_1)$, $c(\pi_2) \leq c(\Pi_2)$, and $c(\pi_3) \leq c(\Pi_2)$. It is also easy to establish the upper bound on $l(\Pi)$.

Cross-cuts version: Similar to (i) (Some obvious simplifications apply when either $m = 0$ or $n = 0$).

(iii) ($\supset$: *right*) and ($\supset$: *left*)

Left as an exercise.

(iv) ($\neg$: *right*) and ($\neg$: *left*)

$$\dfrac{\dfrac{\begin{array}{c}\pi_1 \\ A, \Gamma \vdash \Delta\end{array}}{\Gamma \vdash \Delta, \neg A} \qquad \dfrac{\begin{array}{c}\pi_2 \\ \Lambda \vdash \Theta, A\end{array}}{\neg A, \Lambda \vdash \Theta}}{\Gamma, \Lambda \vdash \Delta, \Theta}$$

$$\Longrightarrow$$

$$\dfrac{\begin{array}{c}\pi_2 \\ \Lambda \vdash \Theta, A\end{array} \qquad \begin{array}{c}\pi_1 \\ A, \Gamma \vdash \Delta\end{array}}{\Gamma, \Lambda \vdash \Delta, \Theta}$$

By the hypothesis $c(\Pi_1), c(\Pi_2) \leq |\neg A|$, it is clear that for the new proof Π we have $c(\Pi) \leq |\neg A|$, since $c(\Pi) = max(\{|A| + 1, c(\pi_1), c(\pi_2)\})$, $c(\pi_1) \leq c(\Pi_1)$, $c(\pi_2) \leq c(\Pi_2)$. It is also easy to establish the upper bound on $l(\Pi)$.

Cross-cuts version (Some obvious simplifications apply when either $m = 0$ or $n = 0$).

$$\dfrac{\dfrac{\stackrel{\pi_1}{A, \Gamma \vdash \Delta, m(\neg A)}}{\Gamma \vdash \Delta, (m+1)(\neg A)} \qquad \dfrac{\stackrel{\pi_2}{n(\neg A), \Lambda \vdash \Theta, A}}{(n+1)(\neg A), \Lambda \vdash \Theta}}{\Gamma, \Lambda \vdash \Delta, \Theta}$$

Let Π'_1 be the proof tree obtained by applying the induction hypothesis to

$$\dfrac{\dfrac{\stackrel{\pi_1}{A, \Gamma \vdash \Delta, m(\neg A)}}{\Gamma \vdash \Delta, (m+1)(\neg A)} \qquad \stackrel{\pi_2}{n(\neg A), \Lambda \vdash \Theta, A}}{\Gamma, \Lambda \vdash \Delta, \Theta, A}$$

and Π'_2 the proof tree obtained by applying the induction hypothesis to

$$\dfrac{\stackrel{\pi_1}{A, \Gamma \vdash \Delta, m(\neg A)} \qquad \dfrac{\stackrel{\pi_2}{n(\neg A), \Lambda \vdash \Theta, A}}{(n+1)(\neg A), \Lambda \vdash \Theta}}{A, \Gamma, \Lambda \vdash \Delta, \Theta}$$

and finally let Π be

$$\dfrac{\dfrac{\stackrel{\Pi'_1}{\Gamma, \Lambda \vdash \Delta, \Theta, A} \qquad \stackrel{\Pi'_2}{A, \Gamma, \Lambda \vdash \Delta, \Theta}}{\Gamma, \Gamma, \Lambda, \Lambda \vdash \Delta, \Delta, \Theta, \Theta}}{\Gamma, \Lambda \vdash \Delta, \Theta}$$

Since $c(\pi_1) \leq c(\Pi_1)$, $c(\pi_2) \leq c(\Pi_2)$, by the induction hypothesis $c(\Pi'_1), c(\Pi'_2) \leq |\neg A|$, and it is clear that for the new proof Π we have $c(\Pi) \leq |\neg A|$, since $c(\Pi) = max(\{|A| + 1, c(\Pi'_1), c(\Pi'_2)\})$. It is also easy to establish the upper bound on $l(\Pi)$.

(v) ($\forall$: *right*) and ($\forall$: *left*)

$$\dfrac{\dfrac{\stackrel{\pi_1}{\Gamma \vdash \Delta, B[y/x]}}{\Gamma \vdash \Delta, \forall x B} \qquad \dfrac{\stackrel{\pi_2}{B[t/x], \Lambda \vdash \Theta}}{\forall x B, \Lambda \vdash \Theta}}{\Gamma, \Lambda \vdash \Delta, \Theta}$$

$$\Longrightarrow$$

$$\dfrac{\stackrel{\pi_1[t/y]}{\Gamma \vdash \Delta, B[t/x]} \qquad \stackrel{\pi_2}{B[t/x], \Lambda \vdash \Theta}}{\Gamma, \Lambda \vdash \Delta, \Theta}$$

In the above, it may be necessary to rename some eigenvariables in π_1 so that they are distinct from all the variables in t.

By the hypothesis $c(\Pi_1), c(\Pi_2) \leq |\forall x B|$, it is clear that for the new proof Π we have $c(\Pi) \leq |\forall x B|$, since $c(\Pi) = max(\{|B[t/x]| + 1, c(\pi_1[t/y]), c(\pi_2)\})$, $c(\pi_1[t/y]) = c(\pi_1)$, $c(\pi_1) \leq c(\Pi_1)$, and $c(\pi_2) \leq c(\Pi_2)$. It is also easy to establish the upper bound on $l(\Pi)$.

Cross-cuts version (Some obvious simplifications apply when either $m = 0$ or $n = 0$).

$$\dfrac{\dfrac{\begin{array}{c}\pi_1\\ \Gamma \vdash \Delta, m(\forall x B), B[y/x]\end{array}}{\Gamma \vdash \Delta, (m+1)(\forall x B)} \qquad \dfrac{\begin{array}{c}\pi_2\\ B[t/x], n(\forall x B), \Lambda \vdash \Theta\end{array}}{(n+1)(\forall x B), \Lambda \vdash \Theta}}{\Gamma, \Lambda \vdash \Delta, \Theta}$$

Let Π'_1 be the proof tree obtained by applying the induction hypothesis to

$$\dfrac{\begin{array}{c}\pi_1\\ \Gamma \vdash \Delta, m(\forall x B), B[y/x]\end{array} \qquad \dfrac{\begin{array}{c}\pi_2\\ B[t/x], n(\forall x B), \Lambda \vdash \Theta\end{array}}{(n+1)(\forall x B), \Lambda \vdash \Theta}}{\Gamma, \Lambda \vdash \Delta, \Theta, B[y/x]}$$

and Π'_2 the proof tree obtained by applying the induction hypothesis to

$$\dfrac{\dfrac{\begin{array}{c}\pi_1\\ \Gamma \vdash \Delta, m(\forall x B), B[y/x]\end{array}}{\Gamma \vdash \Delta, (m+1)(\forall x B)} \qquad \begin{array}{c}\pi_2\\ B[t/x], n(\forall x B), \Lambda \vdash \Theta\end{array}}{B[t/x], \Gamma, \Lambda \vdash \Delta, \Theta}$$

and finally let Π be

$$\dfrac{\dfrac{\begin{array}{c}\Pi'_1[t/y]\\ \Gamma, \Lambda \vdash \Delta, \Theta, B[t/x]\end{array} \qquad \begin{array}{c}\Pi'_2\\ B[t/x], \Gamma, \Lambda \vdash \Delta, \Theta\end{array}}{\Gamma, \Gamma, \Lambda, \Lambda \vdash \Delta, \Delta, \Theta, \Theta}}{\overline{\Gamma, \Lambda \vdash \Delta, \Theta}}$$

In the above, it may be necessary to rename some eigenvariables in Π'_1 so that they are distinct from all the variables in t.

Since $c(\pi_1) \leq c(\Pi_1)$, and $c(\pi_2) \leq c(\Pi_2)$, by the induction hypothesis, $c(\Pi'_1), c(\Pi'_2) \leq |\forall x B|$, and it is clear that for the new proof Π we have $c(\Pi) \leq |\forall x B|$, since $c(\Pi) = max(\{|B[t/x]| + 1, c(\Pi'_1[t/y]), c(\Pi'_2)\})$ and $c(\Pi'_1[t/y]) = c(\Pi'_1)$. It is also easy to establish the upper bound on $l(\Pi)$.

(vi) ($\exists$: *right*) and ($\exists$: *left*)

$$\dfrac{\dfrac{\begin{array}{c}\pi_1\\ \Gamma \vdash \Delta, B[t/x]\end{array}}{\Gamma \vdash \Delta, \exists x B} \qquad \dfrac{\begin{array}{c}\pi_2\\ B[y/x], \Lambda \vdash \Theta\end{array}}{\exists x B, \Lambda \vdash \Theta}}{\Gamma, \Lambda \vdash \Delta, \Theta}$$

$$\Longrightarrow$$

$$\dfrac{\begin{array}{c}\pi_1\\ \Gamma \vdash \Delta, B[t/x]\end{array} \qquad \begin{array}{c}\pi_2[t/y]\\ B[t/x], \Lambda \vdash \Theta\end{array}}{\Gamma, \Lambda \vdash \Delta, \Theta}$$

In the above, it may be necessary to rename some eigenvariables in π_2 so that they are distinct from all the variables in t.

By the hypothesis, $c(\Pi_1), c(\Pi_2) \leq |\exists x B|$, It is clear that for the new proof Π we have $c(\Pi) \leq |\exists x B|$, since $c(\Pi) = max(\{|B[t/x]| + 1, c(\pi_1), c(\pi_2[t/y])\})$, $c(\pi_2[t/y]) = c(\pi_2)$, $c(\pi_1) \leq c(\Pi_1)$, and $c(\pi_2) \leq c(\Pi_2)$. It is also easy to establish the upper bound on $l(\Pi)$.

Cross-cuts version (Some obvious simplifications apply when either $m = 0$ or $n = 0$). Similar to (v) and left as an exercise.

(2) Either the root of Π_1 or the root of Π_2 is the conclusion of some logical rule, the cut rule, or some structural rule having some occurrence of a formula $X \neq A$ as principal formula. We say that A is *passive*.

We only show the transformations corresponding to the case where A is passive on the left, the case in which it is passive on the right being symmetric. For this case (where A is passive on the left), we only show the transformation where the last inference applied to the left subtree is a right-rule, the others being similar.

(i) ($\vee$: *right*)

$$\dfrac{\dfrac{\begin{array}{c}\pi_1\\ \Gamma \vdash \Delta, mA, B\end{array}}{\Gamma \vdash \Delta, mA, B \vee C} \qquad \begin{array}{c}\pi_2\\ nA, \Lambda \vdash \Theta\end{array}}{\Gamma, \Lambda \vdash \Delta, \Theta, B \vee C}$$

$$\Longrightarrow$$

$$\dfrac{\dfrac{\begin{array}{c}\pi_1\\ \Gamma \vdash \Delta, mA, B\end{array} \qquad \begin{array}{c}\pi_2\\ nA, \Lambda \vdash \Theta\end{array}}{\Gamma, \Lambda \vdash \Delta, \Theta, B}}{\Gamma, \Lambda \vdash \Delta, \Theta, B \vee C}$$

Note that $c(\pi_1) \leq c(\Pi_1)$ and $c(\pi_2) \leq c(\Pi_2)$. We conclude by applying the induction hypothesis to the subtree rooted with $\Gamma, \Lambda \vdash \Delta, \Theta, B$. It is also easy to establish the upper bound on $l(\Pi)$.

(ii) ($\wedge$: *right*)

$$\dfrac{\dfrac{\begin{array}{c}\pi_1\\ \Gamma \vdash \Delta, mA, B\end{array} \qquad \begin{array}{c}\pi_2\\ \Gamma \vdash \Delta, mA, C\end{array}}{\Gamma \vdash \Delta, mA, B \wedge C} \qquad \begin{array}{c}\pi_3\\ nA, \Lambda \vdash \Theta\end{array}}{\Gamma, \Lambda \vdash \Delta, \Theta, B \wedge C}$$

$$\Longrightarrow$$

$$\dfrac{\dfrac{\overset{\pi_1}{\Gamma \vdash \Delta, mA, B} \quad \overset{\pi_3}{nA, \Lambda \vdash \Theta}}{\Gamma, \Lambda \vdash \Delta, \Theta, B} \qquad \dfrac{\overset{\pi_2}{\Gamma \vdash \Delta, mA, C} \quad \overset{\pi_3}{nA, \Lambda \vdash \Theta}}{\Gamma, \Lambda \vdash \Delta, \Theta, C}}{\Gamma, \Lambda \vdash \Delta, \Theta, B \wedge C}$$

Note that $c(\pi_1) \leq c(\Pi_1)$, $c(\pi_2) \leq c(\Pi_1)$, and $c(\pi_3) \leq c(\Pi_2)$. We conclude by applying the induction hypothesis to the subtrees rooted with $\Gamma, \Lambda \vdash \Delta, \Theta, B$ and $\Gamma, \Lambda \vdash \Delta, \Theta, C$. It is also easy to establish the upper bound on $l(\Pi)$.

(iii) ($\supset$: *right*)

Left as an exercise.

(iv) ($\neg$: *right*)

$$\dfrac{\dfrac{\overset{\pi_1}{B, \Gamma \vdash \Delta, mA}}{\Gamma \vdash \Delta, mA, \neg B} \qquad \overset{\pi_2}{nA, \Lambda \vdash \Theta}}{\Gamma, \Lambda \vdash \Delta, \Theta, \neg B}$$

$$\Longrightarrow$$

$$\dfrac{\dfrac{\overset{\pi_1}{B, \Gamma \vdash \Delta, mA} \qquad \overset{\pi_2}{nA, \Lambda \vdash \Theta}}{B, \Gamma, \Lambda \vdash \Delta, \Theta}}{\Gamma, \Lambda \vdash \Delta, \Theta, \neg B}$$

Note that $c(\pi_1) \leq c(\Pi_1)$, and $c(\pi_2) \leq c(\Pi_2)$. We conclude by applying the induction hypothesis to the subtree rooted with $B, \Gamma, \Lambda \vdash \Delta, \Theta$. It is also easy to establish the upper bound on $l(\Pi)$.

(v) ($\forall$: *right*)

$$\dfrac{\dfrac{\overset{\pi_1}{\Gamma \vdash \Delta, mA, B[y/x]}}{\Gamma \vdash \Delta, mA, \forall x B} \qquad \overset{\pi_2}{nA, \Lambda \vdash \Theta}}{\Gamma, \Lambda \vdash \Delta, \Theta, \forall x B}$$

$$\Longrightarrow$$

$$\dfrac{\dfrac{\overset{\pi_1[z/y]}{\Gamma \vdash \Delta, mA, B[z/x]} \qquad \overset{\pi_2}{nA, \Lambda \vdash \Theta}}{\Gamma, \Lambda \vdash \Delta, \Theta, B[z/x]}}{\Gamma, \Lambda \vdash \Delta, \Theta, \forall x B}$$

In the above, some renaming may be necessary to ensure the eigenvariable condition.

Note that $c(\pi_1[z/y]) = c(\pi_1)$, $c(\pi_1) \leq c(\Pi_1)$, and $c(\pi_2) \leq c(\Pi_2)$. We conclude by applying the induction hypothesis to the subtree rooted with $\Gamma, \Lambda \vdash \Delta, \Theta, B[z/x]$. It is also easy to establish the upper bound on $l(\Pi)$.

(vi) ($\exists$: *right*)

$$\dfrac{\dfrac{\begin{matrix}\pi_1\\ \Gamma \vdash \Delta, mA, B[t/x]\end{matrix}}{\Gamma \vdash \Delta, mA, \exists x B} \qquad \begin{matrix}\pi_2\\ nA, \Lambda \vdash \Theta\end{matrix}}{\Gamma, \Lambda \vdash \Delta, \Theta, \exists x B}$$

$$\Longrightarrow$$

$$\dfrac{\dfrac{\begin{matrix}\pi_1\\ \Gamma \vdash \Delta, mA, B[t/x]\end{matrix} \qquad \begin{matrix}\pi_2\\ nA, \Lambda \vdash \Theta\end{matrix}}{\Gamma, \Lambda \vdash \Delta, \Theta, B[t/x]}}{\Gamma, \Lambda \vdash \Delta, \Theta, \exists x B}$$

Note that $c(\pi_1) \leq c(\Pi_1)$, and $c(\pi_2) \leq c(\Pi_2)$. We conclude by applying the induction hypothesis to the subtree rooted with $\Gamma, \Lambda \vdash \Delta, \Theta, B[t/x]$. It is also easy to establish the upper bound on $l(\Pi)$.

(vii) (*cut*)

$$\dfrac{\dfrac{\begin{matrix}\pi_1\\ \Gamma_1 \vdash \Delta_1, m_1 A, pB\end{matrix} \qquad \begin{matrix}\pi_2\\ qB, \Lambda_1 \vdash \Theta_1, m_2 A\end{matrix}}{\Gamma \vdash \Delta, mA} \qquad \begin{matrix}\pi_3\\ nA, \Lambda \vdash \Theta\end{matrix}}{\Gamma, \Lambda \vdash \Delta, \Theta}$$

where in the above proof π, $m_1 + m_2 = m$, $\Gamma = \Gamma_1, \Lambda_1$, and $\Delta = \Delta_1, \Theta_1$. Since by the hypothesis, $c(\Pi_1), c(\Pi_2) \leq |A|$, and $c(\Pi_1) = max(\{|B| + 1, c(\pi_1), c(\pi_2)\})$, we must have $|B| < |A|$, $c(\pi_1) \leq |A|$, $c(\pi_2) \leq |A|$, and $c(\pi_3) \leq |A|$. Thus in particular, $B \neq A$. We show the transformation in the case where $m_1 > 0$ and $m_2 > 0$, the cases where either $m_1 = 0$ or $m_2 = 0$ being special cases.

Let Π_1' be the result of applying the induction hypothesis to

$$\dfrac{\begin{matrix}\pi_1\\ \Gamma_1 \vdash \Delta_1, pB, m_1 A\end{matrix} \qquad \begin{matrix}\pi_3\\ nA, \Lambda \vdash \Theta\end{matrix}}{\Gamma_1, \Lambda \vdash \Delta_1, \Theta, pB}$$

let Π_2' be the result of applying the induction hypothesis to

$$\dfrac{\begin{matrix}\pi_2\\ qB, \Lambda_1 \vdash \Theta_1, m_2 A\end{matrix} \qquad \begin{matrix}\pi_3\\ nA, \Lambda \vdash \Theta\end{matrix}}{qB, \Lambda_1, \Lambda \vdash \Theta_1, \Theta}$$

and let Π be the proof

$$\dfrac{\begin{matrix}\Pi_1'\\ \Gamma_1, \Lambda \vdash \Delta_1, \Theta, pB\end{matrix} \qquad \begin{matrix}\Pi_2'\\ qB, \Lambda_1, \Lambda \vdash \Theta_1, \Theta\end{matrix}}{\dfrac{\Gamma, \Lambda, \Lambda \vdash \Delta, \Theta, \Theta}{\overline{\Gamma, \Lambda \vdash \Delta, \Theta}}}$$

Since by the induction hypothesis, $c(\Pi'_1), c(\Pi'_2) \leq |A|$, and since $|B| < |A|$, we have $c(\Pi) \leq |A|$. It is also easy to establish the upper bound on $l(\Pi)$.

(viii) (*contrac*: *right*)

$$\dfrac{\dfrac{\overset{\pi_1}{\Gamma \vdash \Delta, B, B, mA}}{\Gamma \vdash \Delta, B, mA} \qquad \overset{\pi_2}{nA, \Lambda \vdash \Theta}}{\Gamma, \Lambda \vdash \Delta, \Theta, B}$$

$$\Longrightarrow$$

$$\dfrac{\dfrac{\overset{\pi_1}{\Gamma \vdash \Delta, B, B, mA} \qquad \overset{\pi_2}{nA, \Lambda \vdash \Theta}}{\Gamma, \Lambda \vdash \Delta, \Theta, B, B}}{\Gamma, \Lambda \vdash \Delta, \Theta, B}$$

Note that $c(\pi_1) \leq c(\Pi_1)$, and $c(\pi_2) \leq c(\Pi_2)$. We conclude by applying the induction hypothesis to the subtree rooted with $\Gamma, \Lambda \vdash \Delta, \Theta, B, B$. It is also easy to establish the upper bound on $l(\Pi)$.

(ix) (*weakening*: *right*)

$$\dfrac{\dfrac{\overset{\pi_1}{\Gamma \vdash \Delta, mA}}{\Gamma \vdash \Delta, mA, B} \qquad \overset{\pi_2}{nA, \Lambda \vdash \Theta}}{\Gamma, \Lambda \vdash \Delta, \Theta, B}$$

$$\Longrightarrow$$

$$\dfrac{\dfrac{\overset{\pi_1}{\Gamma \vdash \Delta, mA} \qquad \overset{\pi_2}{nA, \Lambda \vdash \Theta}}{\Gamma, \Lambda \vdash \Delta, \Theta}}{\Gamma, \Lambda \vdash \Delta, \Theta, B}$$

Note that $c(\pi_1) \leq c(\Pi_1)$, and $c(\pi_2) \leq c(\Pi_2)$. We conclude by applying the induction hypothesis to the subtrees rooted with $\Gamma, \Lambda \vdash \Delta, \Theta$. It is also easy to establish the upper bound on $l(\Pi)$.

(3) Either Π_1 or Π_2 is an axiom. We consider the case in which the left subtree is an axiom, the other case being symmetric.

$$\dfrac{A \vdash A \qquad \overset{\pi_2}{nA, \Lambda \vdash \Theta}}{A, \Lambda \vdash \Theta}$$

$$\Longrightarrow$$

$$\dfrac{\overset{\pi_2}{nA, \Lambda \vdash \Theta}}{\overline{A, \Lambda \vdash \Theta}}$$

Note that $c(\pi_2) \leq c(\Pi_2)$. Since by hypothesis $c(\Pi_1), c(\Pi_2) \leq |A|$, it is clear that $c(\Pi) \leq |A|$.

(4) Either the root of Π_1 or the root of Π_2 is the conclusion of some thinning or contraction resulting in an occurrence of the cut formula A. We consider the case in which this happens in the succedent of the left subtree, the other case being symmetric.

(i) (*weakening*: *right*)

$$\dfrac{\dfrac{\overset{\pi_1}{\Gamma \vdash \Delta}}{\Gamma \vdash \Delta, A} \qquad \overset{\pi_2}{A, \Lambda \vdash \Theta}}{\Gamma, \Lambda \vdash \Delta, \Theta}$$

$$\Longrightarrow$$

$$\dfrac{\overset{\pi_1}{\Gamma \vdash \Delta}}{\Gamma, \Lambda \vdash \Delta, \Theta}$$

and when $m, n > 0$,

$$\dfrac{\dfrac{\overset{\pi_1}{\Gamma \vdash \Delta, mA}}{\Gamma \vdash \Delta, (m+1)A} \qquad \overset{\pi_2}{nA, \Lambda \vdash \Theta}}{\Gamma, \Lambda \vdash \Delta, \Theta}$$

$$\Longrightarrow$$

$$\dfrac{\overset{\pi_1}{\Gamma \vdash \Delta, mA} \qquad \overset{\pi_2}{nA, \Lambda \vdash \Theta}}{\Gamma, \Lambda \vdash \Delta, \Theta}$$

Since by the hypothesis we have $c(\Pi_1), c(\Pi_2) \leq |A|$, it is clear that $c(\Pi) \leq |A|$ in the first case. In the second case, since $c(\pi_1) \leq c(\Pi_1)$ and $c(\pi_2) \leq c(\Pi_2)$, we conclude by applying the induction hypothesis.

(ii) (*contrac*: *right*)

$$\dfrac{\dfrac{\overset{\pi_1}{\Gamma \vdash \Delta, (m-1)A, A, A}}{\Gamma \vdash \Delta, mA} \qquad \overset{\pi_2}{nA, \Lambda \vdash \Theta}}{\Gamma, \Lambda \vdash \Delta, \Theta}$$

$$\Longrightarrow$$

$$\dfrac{\overset{\pi_1}{\Gamma \vdash \Delta, (m+1)A} \qquad \overset{\pi_2}{nA, \Lambda \vdash \Theta}}{\Gamma, \Lambda \vdash \Delta, \Theta}$$

Since by the hypothesis we have $c(\Pi_1), c(\Pi_2) \leq |A|$, and we have $c(\pi_1) \leq c(\Pi_1)$ and $c(\pi_2) \leq c(\Pi_2)$, we conclude by applying the induction hypothesis. $\square$

We can now prove the following major result (essentially due to Tait [29], 1968), showing not only that every proof can be transformed into a cut-free proof, but also giving an upper bound on the size of the resulting cut-free proof.

Theorem 12.4 *Let Π be a proof with cut-rank $c(\Pi)$ of a sequent $\Gamma \vdash \Delta$. A cut-free proof Π^* for $\Gamma \vdash \Delta$ can be constructed such that $l(\Pi^*) \leq exp(4, c(\Pi), l(\Pi))$.*

Proof. We prove the following claim by induction on the depth of proof trees.

Claim: Let Π be a proof with cut-rank $c(\Pi)$ for a sequent $\Gamma \vdash \Delta$. If $c(\Pi) > 0$ then we can construct a proof Π' for $\Gamma \vdash \Delta$ such that

$$c(\Pi') < c(\Pi) \quad \text{and} \quad l(\Pi') \leq 4^{l(\Pi)}.$$

Proof of Claim: If either the last inference of Π is not a cut, or it is a cut and $c(\Pi) > |A| + 1$, we apply the induction hypothesis to the immediate subtrees Π_1 or Π_2 (or Π_1) of Π. We are left with the case in which the last inference is a cut and $c(\Pi) = |A| + 1$. The proof is of the form

$$\frac{\begin{array}{c}\Pi_1\\ \Gamma \vdash \Delta, mA\end{array} \qquad \begin{array}{c}\Pi_2\\ nA, \Lambda \vdash \Theta\end{array}}{\Gamma, \Lambda \vdash \Delta, \Theta}$$

By the induction hypothesis, we can construct a proof Π'_1 for $\Gamma \vdash \Delta, mA$ and a proof Π'_2 for $nA, \Lambda \vdash \Theta$, such that $c(\Pi'_i) \leq |A|$ and $l(\Pi'_i) \leq 4^{l(\Pi_i)}$, for $i = 1, 2$. Applying the reduction lemma (Lemma 12.3), we obtain a proof Π' such that, $c(\Pi') \leq |A|$ and $l(\Pi') \leq 2(l(\Pi'_1) + l(\Pi'_2))$. But

$$2(l(\Pi'_1) + l(\Pi'_2)) \leq 2(4^{l(\Pi_1)} + 4^{l(\Pi_2)}) \leq 4^{max(l(\Pi_1), l(\Pi_2))+1} = 4^{l(\Pi)}.$$

$\square$

The proof of Theorem 12.4 follows by induction on $c(\Pi)$, and by the definition of $exp(4, m, n)$.
$\square$

It is easily verified that the above argument also goes through for the system $\mathcal{LJ}$. Thus, we obtain Gentzen's original cut elimination theorem.

Theorem 12.5 [Cut Elimination Theorem, Gentzen (1935)] *There is an algorithm which, given any proof Π in $\mathcal{LK}$ produces a cut-free proof Π' in $\mathcal{LK}$. There is an algorithm which, given any proof Π in $\mathcal{LJ}$ produces a cut-free proof Π' in $\mathcal{LJ}$.*

It is instructive to see exactly where (*contraction*) and (*weakening*) are actually used in the transformations. Contraction is used in the cross-cuts, in the case of axioms (case (3)), and of course when it is the last inference of a proof (in case (2) and (4)). Weakening is only used in case (2) and (4), and this because we have chosen the minimal axioms $A \vdash A$. If we use the version of the cut rule (the *additive* version) in which the contexts are merged rather than concatenated

$$\frac{\Gamma \vdash \Delta, A \quad A, \Gamma \vdash \Delta}{\Gamma \vdash \Delta} \quad (cut)$$

and the corresponding extended cut rule

$$\frac{\Gamma \vdash \Delta, mA \quad nA, \Gamma \vdash \Delta}{\Gamma \vdash \Delta} \quad (cut)$$

we can verify that contraction is no longer needed in the cross-cuts, but it is still needed for the axioms, and of course when it is the last inference of a proof (in case (2) and (4)). If in addition we use the "fat" axioms $A, \Gamma \vdash \Delta, A$, it seems that the weakening rules are no longer needed, but this is in fact erroneous. It is true that weakening is not needed to handle the case of axioms (only contraction is needed), but in case (2) for (*contrac*: *left*) and (*contrac*: *right*), weakening is needed! (see case (2)(viii) for example). This is somewhat ennoying since the transformation rules for weakening discard one of the two subproofs appearing as premises of a cut. Indeed, as observed by Yves Lafont, it is this property which causes any two classical proofs to be identified under the equivalence relation induced by the reduction rules for cut elimination. The proof

$$\dfrac{\dfrac{\begin{array}{c}\pi_1 \\ \Gamma \vdash \Delta\end{array}}{\Gamma \vdash \Delta, A} \qquad \dfrac{\begin{array}{c}\pi_2 \\ \Gamma \vdash \Delta\end{array}}{A, \Gamma \vdash \Delta}}{\Gamma \vdash \Delta}$$

reduces both to π_1 and π_2, showing that the equivalence of classical proofs induced by cut-elimination is trivial. However, note that if we allow the new transformations which take a proof π of a sequent $\Gamma \vdash \Delta$ and create (the obvious) proof $\Lambda + \pi$ of $\Lambda, \Gamma \vdash \Delta$ and $\pi + \Lambda$ of $\Gamma \vdash \Delta, \Lambda$, then weakening can be dispended with, and I conjecture that the equivalence of proofs is nontrivial. Still, note that the proofs of the axioms $A \vdash A$ and $B \vdash B$ (with $A \neq B$) will be equivalent since they are both equivalent to the proof of $A, B \vdash A, B$, but this does not seem to be problematic. The intuitionistic calculus does not suffer from the same problem: the equivalence of proofs induced by cut elimination is nontrivial (this can be seen by mapping sequential proofs to natural deduction proofs and using the normal form property of λ-terms).

From theorem 12.4, we see that the complexity of cut-elimination is horrendous, since the upper bound on the size of a cut-free proof is $exp(4, c(\Pi), l(\Pi))$, a super exponential. Careful inspection of the proof of lemma 12.3 shows that $l(\Pi) \leq l(\Pi_1) + l(\Pi_2)$, except for case (1)(iii). This is the reason for the upper bound $exp(4, c(\Pi),\ l(\Pi))$, which applies even in the propositional case. However, Hudelmaier [17] has shown that for the system $\mathcal{LJT}^{\supset,\wedge,\vee,\perp}$, this upper bound can be reduced to a quadruple exponential, and in the classical case for a system similar to $\mathcal{GK}^{\supset,\wedge,\vee,\perp}$ to a double exponential. But nothing is obtained for free! Indeed, the lower complexity of cut-elimination in the system $\mathcal{LJT}^{\supset,\wedge,\vee,\perp}$ is attained at the cost of the computational expressive power of the system. In the system of natural deduction $\mathcal{N}^{\supset,\wedge,\vee,\perp}$, the proposition $(P \supset P) \supset (P \supset P)$ has infinitely many inequivalent proofs (in fact, the Church numerals for P). On the other hand, in $\mathcal{LJT}^{\supset,\wedge,\vee,\perp}$, $(P \supset P) \supset (P \supset P)$ only has finitely many proofs, and in fact only two, corresponding to the Church numerals "zero" and "one". This is regrettable, but a cynic would not be surprised: there is no free lunch!

The above considerations lead naturally to the following question: what is the exact relationship between cut-elimination and (strong) normalization in intuitionistic logic? In [37], Jeff Zucker makes an in-depth study of this relationship. Although Zucker obtains some very nice results in this formidable paper, he does not provide a complete answer to the problem. Intuitively, the reason is that the reduction relation induced by cut-elimination transformations is finer than β-reduction (as alluded to in section 10). Thus, the exact relationship between cut-elimination and (strong) normalization remains a challenging open problem. For the latest results, see Ungar [33].

A few more remarks about the role of contraction and weakening will be useful as a motivation for linear logic. We already noticed with the cut rule that contexts (the Γ, Δ occurring in the premise(s) of inference rules) can be treated in two different ways: (1) either they are merged (which implies that they are identical), or (2) they are concatenated.

In order to search for proof backwards, it is more convenient to treat contexts in mode (1), but this hides some subtleties. For example, the ($\wedge$: *right*) rule can be written either as

$$\frac{\Gamma \vdash \Delta, A \quad \Gamma \vdash \Delta, B}{\Gamma \vdash \Delta, A \wedge B}$$

where the contexts are merged, or as

$$\frac{\Gamma \vdash \Delta, A \quad \Lambda \vdash \Theta, B}{\Gamma, \Lambda \vdash \Delta, \Theta, A \wedge B}$$

where the contexts are just concatenated but not merged. Following Girard, let's call the first version *additive*, and the second version *multiplicative*. Under contraction and weakening, the two versions are equivalent: the first rule can be simulated by the second rule using contractions:

$$\begin{array}{c} \dfrac{\Gamma \vdash \Delta, A \qquad\qquad \Gamma \vdash \Delta, B}{\Gamma, \Gamma \vdash \Delta, \Delta, A \wedge B} \\ \hline\hline \Gamma \vdash \Delta, A \wedge B \end{array}$$

and the second rule can be simulated by the first rule using weakenings:

$$\dfrac{\begin{array}{c} \Gamma \vdash \Delta, A \\ \hline\hline \Gamma, \Lambda \vdash \Delta, \Theta, A \end{array} \qquad\qquad \begin{array}{c} \Lambda \vdash \Theta, B \\ \hline\hline \Gamma, \Lambda \vdash \Delta, \Theta, B \end{array}}{\Gamma, \Lambda \vdash \Delta, \Theta, A \wedge B}$$

Similarly, the ($\wedge$: *left*) rules can be written either as

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

or as

$$\frac{A, B, \Gamma \vdash \Delta}{A \wedge B, \Gamma \vdash \Delta}$$

Again, let's call the first version *additive*, and the second version *multiplicative*. These versions are equivalent under contraction and weakening. The first version can be simulated by the second rule using weakening:

$$\dfrac{\dfrac{A, \Gamma \vdash \Delta}{A, B, \Gamma \vdash \Delta} \quad (\textit{weakening: left})}{A \wedge B, \Gamma \vdash \Delta}$$

and the second version can be simulated by the first rule and contraction:

$$\dfrac{\dfrac{\dfrac{A, B, \Gamma \vdash \Delta}{A \wedge B, B, \Gamma \vdash \Delta}}{A \wedge B, A \wedge B, \Gamma \vdash \Delta}}{A \wedge B, \Gamma \vdash \Delta} \quad (\textit{contrac: left})$$

If we take away contraction and weakening, the additive and multiplicative versions are no longer equivalent. This suggests, and this path was followed by Girard, to split the connectives $\wedge$ and $\vee$ into two versions: the *multiplicative* version of $\wedge$ and $\vee$, denoted as $\otimes$ and $\wp$, and the *additive* version of $\wedge$ and $\vee$, denoted as & and $\oplus$. In *linear logic*, due to Girard [12], the connectives $\wedge$ and $\vee$ are split into multiplicative and additive versions, contraction and weakening are dropped, negation denoted $A^{\perp}$ is involutive, and in order to regain the loss of expressiveness due to the absence of contraction and weakening, some new connectives (the exponentials ! and ?) are introduced. The main role of these connectives is to have better control over contraction and weakening. Thus, at the heart of linear logic lies the notion that resources are taken into account.

13 Reductions of Classical Logic to Intuitionistic Logic

Although there exist formulae that are provable classically but not intuitionistically, there are several ways of embedding classical logic into intuitionistic logic. More specifically, there are functions * from formulae to formulae such that for every formula A, its translation A^* is equivalent to A classically, and A is provable classically iff A^* is provable intuitionistically. Stronger results can be obtained in the propositional case. Since $\neg\neg A \supset A$ is provable classically but not intuitionistically, whereas $A \supset \neg\neg A$ is provable both classically and intuitionistically, we can expect that double-negation will play a crucial role, and this is indeed the case. One of the crucial properties is that triple negation is equivalent to a single negation. This is easily shown as follows (in $\mathcal{G}_i^{\supset,\wedge,\vee,\neg}$):

$$\dfrac{\dfrac{\dfrac{\dfrac{A \vdash A}{A, \neg A \vdash}}{A \vdash \neg\neg A}}{A, \neg\neg\neg A \vdash}}{\neg\neg\neg A \vdash \neg A}$$

Since we also have the following proof (in $\mathcal{G}_i^{\supset,\wedge,\vee,\neg}$)

$$\dfrac{\dfrac{\neg A \vdash \neg A}{\neg A, \neg\neg A \vdash}}{\neg A \vdash \neg\neg\neg A}$$

it is clear that $\neg\neg\neg A \equiv \neg A$ is provable intuitionistically.

The possibility of embedding classical logic into intuitionistic logic is due to four crucial facts which we show step by step:

(1) $\neg\neg\neg A \equiv \neg A$ is provable intuitionistically;

(2) If a formula A is provable classically without using the ($\forall$: *right*)-rule, then $\neg\neg A$ is provable intuitionistically;

(3) For a class of formulae for which $\neg\neg A \vdash A$ is provable intuitionistically, (2) holds unrestricted. This means that if a formula A in this class is provable classically then $\neg\neg A$ is provable intuitionistically;

(4) For every formula A built only from $\supset, \wedge, \neg$ and $\forall$, if $\Delta = \neg\neg P_1 \supset P_1, \ldots, \neg\neg P_k \supset P_k$ where $P_1, \ldots, P_k$ are all the atoms occurring in A, then $\Delta, \neg\neg A \vdash A$ is provable intuitionistically.

The "trick" of the double-negation translation (often attributed to Gödel (1933), although it was introduced independently by Kolmogorov (1925) and Gentzen (1933)) is that if we consider a formula A only built from $\supset, \wedge, \neg, \forall$, and replace every atomic subformula P by $\neg\neg P$ obtaining $A^\dagger$, we get a subclass of formulae for which (4) holds without the Δ, and thus (3) also holds. For this class, A is provable classically iff $A^\dagger$ is provable intuitionistically.

Our first result will concern propositions. Given $\Gamma = A_1, \ldots, A_m$, let $\neg\neg\Gamma = \neg\neg A_1, \ldots, \neg\neg A_m$.

Lemma 13.1 *Given a sequent* $\Gamma \vdash B_1, \ldots, B_n$ *of propositions, if* $\Gamma \vdash B_1, \ldots, B_n$ *is provable in* $\mathcal{G}_c^{\supset,\wedge,\vee,\neg}$, *then* $\neg\neg\Gamma \vdash \neg(\neg B_1 \wedge \ldots \wedge \neg B_n)$ *is provable in* $\mathcal{G}_i^{\supset,\wedge,\vee,\neg}$.

Proof. We proceed by induction on proofs. In fact, it is easier to work in $\mathcal{G}_i^{\supset,\wedge,\vee,\neg,cut}$ and use cut elimination. It is necessary to prove that a number of propositions are provable intuitionistically. First, observe that if $A_1, \ldots, A_m \vdash B$ is provable in $\mathcal{G}_i^{\supset,\wedge,\vee,\neg,cut}$, then $\neg\neg A_1, \ldots, \neg\neg A_m \vdash \neg\neg B$ is also provable in $\mathcal{G}_i^{\supset,\wedge,\vee,\neg,cut}$. The following sequents are provable in $\mathcal{G}_i^{\supset,\wedge,\vee,\neg,cut}$:

$$\begin{aligned}
\neg\neg A &\vdash \neg(\neg A \wedge D),\\
\neg(\neg A \wedge D), \neg(\neg B \wedge D) &\vdash \neg(\neg(A \wedge B) \wedge D),\\
\neg\neg(A \wedge B) &\vdash \neg\neg A \wedge \neg\neg B,\\
\neg(\neg A \wedge \neg B \wedge D) &\vdash \neg(\neg(A \vee B) \wedge D),\\
\neg\neg(A \vee B) &\vdash \neg(\neg A \wedge \neg B),\\
(\neg\neg A \supset \neg(\neg B \wedge D)) &\vdash \neg(\neg(A \supset B) \wedge D),\\
\neg\neg(A \supset B), \neg(\neg A \wedge D), (\neg\neg B \supset \neg D) &\vdash \neg D,\\
(\neg\neg A \supset \neg D) &\vdash \neg(\neg\neg A \wedge D),\\
\neg(\neg A \wedge D), \neg\neg\neg A &\vdash \neg D,\\
\neg(\neg A \wedge \neg A \wedge D) &\vdash \neg(\neg A \wedge D).
\end{aligned}$$

Given $\Delta = D_1, \ldots, D_m$, we let $D = \neg D_1 \wedge \ldots \wedge \neg D_m$. This way, observe that $\neg(\neg A \wedge \neg D_1 \wedge \ldots \wedge \neg D_m) = \neg(\neg A \wedge D)$. The reader should be aware that the case where $m = 0$ in $\Delta = D_1, \ldots, D_m$ can arise and requires special treatment. In this case, D can be considered to be $\neg\perp$. Actually, if we add $\perp$ and the axioms $\perp, \Gamma \vdash \Delta$ where Δ is either empty or a single formula to $\mathcal{G}_i^{\supset,\wedge,\vee,\neg,cut}$, then

$\neg\neg\perp\equiv\perp$ and $A\wedge\neg\perp\equiv A$ are provable, and we just have to simplify the above sequents accordingly. We proceed assuming that $m \geq 1$, leaving the case $m = 0$ to the reader. Now, consider the axioms and each inference rule. An axiom $\Gamma, A \vdash A, \Delta$ becomes $\neg\neg\Gamma, \neg\neg A \vdash \neg(\neg A \wedge D)$, which is provable in $\mathcal{G}_i^{\supset,\wedge,\vee,\neg,cut}$ since $\neg\neg A \vdash \neg(\neg A \wedge D)$ is. Let us also consider the case of the ($\supset$: *right*)-rule, leaving the others as exercises.

$$\frac{\Gamma, A \vdash B, \Delta}{\Gamma \vdash A \supset B, \Delta}$$

By the induction hypothesis, $\neg\neg\Gamma, \neg\neg A \vdash \neg(\neg B \wedge D)$ is provable in $\mathcal{G}_i^{\supset,\wedge,\vee,\neg,cut}$, and so is

$$\neg\neg\Gamma \vdash (\neg\neg A \supset \neg(\neg B \wedge D)).$$

Since

$$(\neg\neg A \supset \neg(\neg B \wedge D)) \vdash \neg(\neg(A \supset B) \wedge D)$$

is also provable in $\mathcal{G}_i^{\supset,\wedge,\vee,\neg,cut}$, by a cut, we obtain that

$$\neg\neg\Gamma \vdash \neg(\neg(A \supset B) \wedge D)$$

is provable in $\mathcal{G}_i^{\supset,\wedge,\vee,\neg,cut}$, as desired. $\square$

In order to appreciate the value of Lemma 13.1, the reader should find a direct proof of $\neg\neg(\neg\neg P \supset P)$ in $\mathcal{G}_i^{\supset,\wedge,\vee,\neg}$.

Since $\neg\neg\neg A \equiv \neg A$ is provable intuitionistically, we obtain the following lemma known as Glivenko's Lemma.

Lemma 13.2 [Glivenko, 1929] *Given a sequent $\neg\Gamma, \Delta \vdash \neg B_1, \ldots, \neg B_n$ made of propositions, if $\neg\Gamma, \Delta \vdash \neg B_1, \ldots, \neg B_n$ is provable in $\mathcal{G}_c^{\supset,\wedge,\vee,\neg}$, then $\neg\Gamma, \neg\neg\Delta \vdash \neg(B_1 \wedge \ldots \wedge B_n)$ is provable in $\mathcal{G}_i^{\supset,\wedge,\vee,\neg}$. In particular, if $\neg\Gamma \vdash \neg B$ is a propositional sequent provable in $\mathcal{G}_c^{\supset,\wedge,\vee,\neg}$, then it is also provable in $\mathcal{G}_i^{\supset,\wedge,\vee,\neg}$.*

Proof. By Lemma 13.1, using the fact that $\neg\neg\neg A \equiv \neg A$ is provable intuitionistically, and that the sequent

$$\neg(\neg\neg B_1 \wedge \ldots \wedge \neg\neg B_n) \vdash \neg(B_1 \wedge \ldots \wedge B_n)$$

is provable in $\mathcal{G}_i^{\supset,\wedge,\vee,\neg}$. $\square$

As a consequence of Lemma 13.1, if a proposition A is provable classically, then $\neg\neg A$ is provable intuitionistically, and as a consequence of Lemma 13.2, if a proposition $\neg A$ is provable classically, then it is also provable intuitionistically. It should be noted that Lemma 13.1 *fails* for quantified formulae. For example, $\forall x(P(x) \vee \neg P(x))$ is provable classically, but we can show that $\neg\neg\forall x(P(x) \vee \neg P(x))$ is not provable intuitionistically, for instance using the system of Lemma 8.9. Similarly, $\forall x \neg\neg P(x) \supset \neg\neg\forall x P(x)$ is provable classically, but it is not provable intuitionistically, and neither is $\neg\neg(\forall x \neg\neg P(x) \supset \neg\neg\forall x P(x))$. As observed by Gödel, Lemma 13.2 has the following remarkable corollary.

Lemma 13.3 [Gödel, 1933] *For every proposition A built only from $\wedge$ and $\neg$, if A is provable classically, then A is also provable intuitionistically.*

Proof. By induction on A. If $A = \neg B$, then this follows by Glivenko's Lemma. Otherwise, it must be possible to write $A = B_1 \wedge \ldots \wedge B_n$ where each B_i is not a conjunct and where each B_i is provable classically. Thus, each B_i must be of the form $\neg C_i$, since if B_i is an atom it is not provable. Again, each B_i is provable intuitionistically by Glivenko's Lemma, and thus so is A. $\square$

Lemma 13.1 confirms that double-negation plays an important role in linking classical logic to intuitionistic logic. The following lemma shows that double-negation distributes over the connectives $\wedge$ and $\supset$.

Lemma 13.4 *The following formulae are provable in* $\mathcal{G}_i^{\supset,\wedge,\vee,\neg}$:

$$\begin{aligned} \neg\neg(A \wedge B) &\equiv \neg\neg A \wedge \neg\neg B, \\ \neg\neg(A \supset B) &\equiv \neg\neg A \supset \neg\neg B. \end{aligned}$$

Proof. We give proofs for

$$\neg\neg(A \supset B) \vdash \neg\neg A \supset \neg\neg B$$

and

$$\neg\neg A \supset \neg\neg B \vdash \neg\neg(A \supset B),$$

leaving the others as exercises.

$$\begin{prooftree}
\AxiomC{$A, \neg B \vdash A$}
\AxiomC{$A, B \vdash B$}
\UnaryInfC{$A, B, \neg B \vdash$}
\BinaryInfC{$A, A \supset B, \neg B \vdash$}
\UnaryInfC{$A, \neg B \vdash \neg(A \supset B)$}
\UnaryInfC{$A, \neg\neg(A \supset B), \neg B \vdash$}
\UnaryInfC{$\neg\neg(A \supset B), \neg B \vdash \neg A$}
\UnaryInfC{$\neg\neg(A \supset B), \neg\neg A, \neg B \vdash$}
\UnaryInfC{$\neg\neg(A \supset B), \neg\neg A \vdash \neg\neg B$}
\UnaryInfC{$\neg\neg(A \supset B) \vdash \neg\neg A \supset \neg\neg B$}
\end{prooftree}$$

$$\begin{prooftree}
\AxiomC{$A \vdash A$}
\UnaryInfC{$\neg A, A \vdash$}
\UnaryInfC{$\neg A, A \vdash B$}
\UnaryInfC{$\neg A \vdash A \supset B$}
\UnaryInfC{$\neg(A \supset B), \neg A \vdash$}
\UnaryInfC{$\neg(A \supset B) \vdash \neg\neg A$}
\AxiomC{$B, A \vdash B$}
\UnaryInfC{$B \vdash A \supset B$}
\UnaryInfC{$B, \neg(A \supset B) \vdash$}
\UnaryInfC{$\neg(A \supset B) \vdash \neg B$}
\UnaryInfC{$\neg\neg B, \neg(A \supset B) \vdash$}
\BinaryInfC{$\neg\neg A \supset \neg\neg B, \neg(A \supset B) \vdash$}
\UnaryInfC{$\neg\neg A \supset \neg\neg B \vdash \neg\neg(A \supset B)$}
\end{prooftree}$$

$\square$

Lemma 13.4 *fails* for disjunctions. For example,

$$\neg\neg(P \vee \neg P) \vdash (\neg\neg P \vee \neg\neg\neg P)$$

is *not* provable in $\mathcal{G}_i^{\supset,\wedge,\vee,\neg}$, since $\neg\neg(P \vee \neg P)$ is provable but $(\neg\neg P \vee \neg\neg\neg P)$ is *not* provable in $\mathcal{G}_i^{\supset,\wedge,\vee,\neg}$ (this is easily shown using the system $\mathcal{GK}_i$). Lemma 13.4 also *fails* for the quantifiers. For example, using the system of Lemma 8.9, we can show that $\forall x\neg\neg P(x) \supset \neg\neg\forall x P(x)$ and $\neg\neg\exists x P(x) \supset \exists x\neg\neg P(x)$ are not provable intuitionistically.

Even though Lemma 13.1 fails in general, in particular for universal formulae, Kleene has made the remarkable observation that the culprit is precisely the ($\forall$: *right*)-rule [18] (see Theorem 59, page 492). Indeed, the lemma still holds for arbitrary sequents $\Gamma \vdash B_1, \ldots, B_n$, provided that their proofs in $\mathcal{G}_c^{\supset,\wedge,\vee,\neg,\forall,\exists}$ do *not* use the rule ($\forall$: *right*).

Lemma 13.5 *Given a first-order sequent $\Gamma \vdash B_1, \ldots, B_n$, if $\Gamma \vdash B_1, \ldots, B_n$ is provable in $\mathcal{G}_c^{\supset,\wedge,\vee,\neg,\forall,\exists}$ without using the rule ($\forall$: right), then $\neg\neg\Gamma \vdash \neg(\neg B_1 \wedge \ldots \wedge \neg B_n)$ is provable in $\mathcal{G}_i^{\supset,\wedge,\vee,\neg,\forall,\exists}$.*

Proof. As in the proof of Lemma 13.1, we proceed by induction on proofs. It is necessary to prove that the following sequents are provable in $\mathcal{G}_i^{\supset,\wedge,\vee,\neg,\forall,\exists}$:

$$\begin{aligned}
\neg(\neg A[t/x] \wedge D) &\vdash \neg(\neg\exists x A \wedge D),\\
\forall x(\neg\neg A \supset \neg D), \neg\neg\exists x A &\vdash \neg D,\\
(\neg\neg A[t/x] \supset \neg D), \neg\neg\forall x A &\vdash \neg D.
\end{aligned}$$

where x does not occur in D in the second sequent. Proofs for the above sequents follow:

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{A[t/x], D \vdash A[t/x]}{A[t/x], D \vdash \exists x A}}{A[t/x], \neg\exists x A, D \vdash}}{\neg\exists x A, D \vdash \neg A[t/x]} \qquad \neg\exists x A, D \vdash D}{\neg\exists x A, D \vdash \neg A[t/x] \wedge D}}{\neg(\neg A[t/x] \wedge D), \neg\exists x A, D \vdash}}{\neg(\neg A[t/x] \wedge D), \neg\exists x A \wedge D \vdash}}{\neg(\neg A[t/x] \wedge D) \vdash \neg(\neg\exists x A \wedge D)}$$

$$
\dfrac{\dfrac{\dfrac{D, A[y/x] \vdash A[y/x]}{D, A[y/x], \neg A[y/x] \vdash}}{D, A[y/x] \vdash \neg\neg A[y/x]} \qquad \dfrac{D, A[y/x] \vdash D}{\neg D, D, A[y/x] \vdash}}{\dfrac{\neg\neg A[y/x] \supset \neg D, D, A[y/x] \vdash}{\dfrac{\forall x(\neg\neg A \supset \neg D), D, A[y/x] \vdash}{\dfrac{\forall x(\neg\neg A \supset \neg D), D, \exists x A \vdash}{\dfrac{\forall x(\neg\neg A \supset \neg D), D \vdash \neg\exists x A}{\dfrac{\forall x(\neg\neg A \supset \neg D), \neg\neg\exists x A, D \vdash}{\forall x(\neg\neg A \supset \neg D), \neg\neg\exists x A \vdash \neg D}}}}}}
$$

where x does not occur in D, and y is a new variable.

$$
\dfrac{\dfrac{\dfrac{D, A[t/x] \vdash A[t/x]}{D, A[t/x], \neg A[t/x] \vdash}}{D, A[t/x] \vdash \neg\neg A[t/x]} \qquad \dfrac{D, A[t/x] \vdash D}{\neg D, D, A[t/x] \vdash}}{\dfrac{(\neg\neg A[t/x] \supset \neg D), D, A[t/x] \vdash}{\dfrac{(\neg\neg A[t/x] \supset \neg D), D, \forall x A \vdash}{\dfrac{(\neg\neg A[t/x] \supset \neg D), D \vdash \neg\forall x A}{\dfrac{(\neg\neg A[t/x] \supset \neg D), \neg\neg\forall x A, D \vdash}{(\neg\neg A[t/x] \supset \neg D), \neg\neg\forall x A \vdash \neg D}}}}}
$$

We now have to consider the cases where the last inference is one of ($\forall$: *left*), ($\exists$: *left*), or ($\exists$: *right*). We treat the case of the rule ($\exists$: *right*), leaving the others as exercises.

$$
\frac{\Gamma \vdash A[t/x], \Delta}{\Gamma \vdash \exists x A, \Delta}
$$

Given $\Delta = D_1, \ldots, D_m$, we let $D = \neg D_1 \wedge \ldots \wedge \neg D_m$. The reader should be aware that the case where $m = 0$ in $\Delta = D_1, \ldots, D_m$ can arise and requires special treatment (as in the proof of Lemma 13.1). We proceed with the case where $m \geq 1$, leaving the case $m = 0$ to the reader. By the induction hypothesis, $\neg\neg\Gamma \vdash \neg(\neg A[t/x] \wedge D)$ is provable in $\mathcal{G}_i^{\supset,\wedge,\vee,\neg,\forall,\exists}$. On the other hand, since the sequent

$$
\neg(\neg A[t/x] \wedge D) \vdash \neg(\neg\exists x A \wedge D)
$$

is provable in $\mathcal{G}_i^{\supset,\wedge,\vee,\neg,\forall,\exists}$, using a cut, we obtain that the sequent

$$
\neg\neg\Gamma \vdash \neg(\neg\exists x A \wedge D)
$$

is provable in $\mathcal{G}_i^{\supset,\wedge,\vee,\neg,\forall,\exists}$, as desired. $\square$

Technically, the problem with Lemma 13.5, is that the sequent

$$\forall x \neg(\neg A \wedge D) \vdash \neg(\neg \forall x A \wedge D)$$

(where x does not occur in D) is not provable in $\mathcal{G}_i^{\supset,\wedge,\vee,\neg,\forall,\exists}$. In order to see where the problem really lies, we attempt to construct a proof of this sequent.

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\neg(\neg A[y/x] \wedge D), D \vdash A[y/x]}{\forall x \neg(\neg A \wedge D), D \vdash A[y/x]}}{\forall x \neg(\neg A \wedge D), D \vdash \forall x A}}{\forall x \neg(\neg A \wedge D), \neg \forall x A, D \vdash}}{\forall x \neg(\neg A \wedge D), \neg \forall x A \wedge D \vdash}}{\forall x \neg(\neg A \wedge D) \vdash \neg(\neg \forall x A \wedge D)}$$

where x does not occur in D, and y is a new variable. The problem is that we cannot apply the ($\forall$: *left*)-rule before $\neg \forall x A$ has been transferred to the righthand side of the sequent (as $\forall x A$) and before the ($\forall$: *right*)-rule has been applied to $\forall x A$, since this would violate the eigenvariable condition. Unfortunately, we are stuck with the sequent $\neg(\neg A[y/x] \wedge D), D \vdash A[y/x]$ which is unprovable in $\mathcal{G}_i^{\supset,\wedge,\vee,\neg,\forall,\exists}$. However, note that the sequent $\neg(\neg A[y/x] \wedge D), D \vdash \neg\neg A[y/x]$ in which $A[y/x]$ has been replaced with $\neg\neg A[y/x]$ is provable in $\mathcal{G}_i^{\supset,\wedge,\vee,\neg,\forall,\exists}$:

$$\dfrac{\dfrac{\dfrac{D, \neg A[y/x] \vdash \neg A[y/x] \qquad D, \neg A[y/x] \vdash D}{D, \neg A[y/x] \vdash \neg A[y/x] \wedge D}}{\neg(\neg A[y/x] \wedge D), D, \neg A[y/x] \vdash}}{\neg(\neg A[y/x] \wedge D), D \vdash \neg\neg A[y/x]}$$

Thus, if the sequent $\neg\neg A \vdash A$ was provable in $\mathcal{G}_i^{\supset,\wedge,\vee,\neg,\forall,\exists}$, the sequent

$$\forall x \neg(\neg A \wedge D) \vdash \neg(\neg \forall x A \wedge D)$$

would also be provable in $\mathcal{G}_i^{\supset,\wedge,\vee,\neg,\forall,\exists}$. It is therefore important to identify a subclass of first-order formulae for which $\neg\neg A \vdash A$ is provable in $\mathcal{G}_i^{\supset,\wedge,\vee,\neg,\forall,\exists}$, since for such a class, Lemma 13.5 holds without restrictions. The following lemma showing the importance of the axiom $\neg\neg P \vdash P$ where P is atomic, leads us to such a class of formulae. It is at the heart of the many so-called "double-negation translations".

Lemma 13.6 *For every formula A built only from $\supset, \wedge, \neg, \forall$, the sequent $\neg\neg A \vdash A$ is provable in the system $\mathcal{G}_{i+}^{\supset,\wedge,\neg,\forall}$ obtained from $\mathcal{G}_i^{\supset,\wedge,\neg,\forall}$ by adding all sequents of the form $\neg\neg P, \Gamma \vdash P$ where P is atomic as axioms. In the propositional case, if $\Delta = \neg\neg P_1 \supset P_1, \ldots, \neg\neg P_k \supset P_k$ where $P_1, \ldots, P_k$ are all the atoms occurring in A, then $\Delta, \neg\neg A \vdash A$ is provable in $\mathcal{G}_i^{\supset,\wedge,\neg}$.*

Proof. It proceeds by induction on the structure of A. If A is an atom P, this is obvious since $\neg\neg P \vdash P$ is an axiom. If $A = B \wedge C$, by the induction hypothesis, both $\neg\neg B \vdash B$ and

$\neg\neg C \vdash C$ are provable in $\mathcal{G}_{i+}^{\supset,\wedge,\neg,\forall}$, and so is $\neg\neg B \wedge \neg\neg C \vdash B \wedge C$. We just have to prove that $\neg\neg(B \wedge C) \vdash \neg\neg B \wedge \neg\neg C$ is provable in $\mathcal{G}_{i}^{\supset,\wedge,\neg,\forall}$, which is easily done. If $A = \neg B$, since we have shown that $\neg\neg\neg B \vdash \neg B$ is provable in $\mathcal{G}_{i}^{\supset,\wedge,\neg,\forall}$, so is $\neg\neg A \vdash A$. If $A = B \supset C$, then by the induction hypothesis, $\neg\neg C \vdash C$ is provable in $\mathcal{G}_{i+}^{\supset,\wedge,\neg,\forall}$ (and so is $\neg\neg B \vdash B$, but we won't need it). Observe that the sequent $\neg\neg C \supset C, \neg\neg(B \supset C) \vdash B \supset C$ is provable in $\mathcal{G}_{i}^{\supset,\wedge,\neg,\forall}$:

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{B \vdash B \qquad B, C \vdash C}{B, (B \supset C) \vdash C}}{B, \neg C, (B \supset C) \vdash}}{B, \neg C \vdash \neg(B \supset C)}}{\neg\neg(B \supset C), B, \neg C \vdash}}{\neg\neg(B \supset C), B \vdash \neg\neg C} \qquad \neg\neg(B \supset C), B, C \vdash C}{\neg\neg C \supset C, \neg\neg(B \supset C), B \vdash C}}{\neg\neg C \supset C, \neg\neg(B \supset C) \vdash B \supset C}$$

Using the fact that $\neg\neg C \vdash C$ is provable in $\mathcal{G}_{i+}^{\supset,\wedge,\neg,\forall}$ and a suitable cut, $\neg\neg(B \supset C) \vdash B \supset C$ is provable in $\mathcal{G}_{i+}^{\supset,\wedge,\neg,\forall}$. If $A = \forall x B$, we can show easily that $\neg\neg\forall x B \vdash \neg\neg B[t/x]$ is provable in $\mathcal{G}_{i}^{\supset,\wedge,\neg,\forall}$. Since by the induction hypothesis, $\neg\neg B \vdash B$ is provable in $\mathcal{G}_{i+}^{\supset,\wedge,\neg,\forall}$, for any new variable y, $\neg\neg B[y/x] \vdash B[y/x]$ is also provable in $\mathcal{G}_{i+}^{\supset,\wedge,\neg,\forall}$, and thus by choosing $t = y$, the sequent $\neg\neg\forall x B \vdash B[y/x]$ is provable where y is new, so that $\neg\neg\forall x B \vdash \forall x B$ is provable in $\mathcal{G}_{i+}^{\supset,\wedge,\neg,\forall}$. $\square$

Unfortunately, Lemma 13.6 *fails* for disjunctions and existential quantifiers. For example,

$$\neg\neg P \supset P \vdash \neg\neg(P \vee \neg P) \supset (P \vee \neg P)$$

is not provable in $\mathcal{G}_{i}^{\supset,\wedge,\vee,\neg,\forall,\exists}$. This can be shown as follows. Since $P \vee \neg P$ is provable in $\mathcal{G}_{c}^{\supset,\wedge,\vee,\neg}$, by Lemma 13.1, $\neg\neg(P \vee \neg P)$ is provable in $\mathcal{G}_{i}^{\supset,\wedge,\vee,\neg}$. Thus, $\neg\neg P \supset P \vdash (P \vee \neg P)$ would be provable in $\mathcal{G}_{i}^{\supset,\wedge,\vee,\neg,\forall,\exists}$, but we can show using the system of Lemma 8.9 that this is not so.

The sequent

$$\neg\neg P \supset P \vdash \neg\neg(\neg\neg\exists x P(x) \supset \exists x \neg\neg P(x)) \supset (\neg\neg\exists x P(x) \supset \exists x \neg\neg P(x))$$

is also not provable in $\mathcal{G}_{i}^{\supset,\wedge,\vee,\neg,\forall,\exists}$. This is because $(\neg\neg\exists x P(x) \supset \exists x \neg\neg P(x))$ is provable in $\mathcal{G}_{c}^{\supset,\wedge,\vee,\neg,\forall,\exists}$ without using the ($\forall$: *right*)-rule, and so, by Lemma 13.5, $\neg\neg(\neg\neg\exists x P(x) \supset \exists x \neg\neg P(x))$ is provable in $\mathcal{G}_{i}^{\supset,\wedge,\vee,\neg,\forall,\exists}$. Then,

$$\neg\neg P \supset P \vdash (\neg\neg\exists x P(x) \supset \exists x \neg\neg P(x))$$

would be provable in $\mathcal{G}_{i}^{\supset,\wedge,\vee,\neg,\forall,\exists}$, but we can show using the system of Lemma 8.9 that this is not so.

Since the sequent $A \vee \neg A \vdash \neg\neg A \supset A$ is easily shown to be provable in $\mathcal{G}_{i}^{\supset,\wedge,\vee,\neg,\forall,\exists}$, Lemma 13.6 also holds with the axioms $\Gamma \vdash P \vee \neg P$ substituted for $\neg\neg P, \Gamma \vdash P$ (for all atoms P). In fact, with such axioms, we can even show that Lemma 13.6 holds for disjunctions (but not for existential formulae).

In view of Lemma 13.6 we can define the following function † on formulae built from $\supset, \wedge, \neg, \forall$:

$$\begin{aligned} A^\dagger &= \neg\neg A, \quad \text{if } A \text{ is atomic,} \\ (\neg A)^\dagger &= \neg A^\dagger, \\ (A * B)^\dagger &= (A^\dagger * B^\dagger), \quad \text{if } * \in \{\supset, \wedge\}, \\ (\forall x A)^\dagger &= \forall x A^\dagger. \end{aligned}$$

Given a formula built only from $\supset, \wedge, \neg, \forall$, the function † simply replaces every atom P by $\neg\neg P$. It is easy to show that A and $A^\dagger$ are classically equivalent. The following lemma shows the significance of this function.

Lemma 13.7 *For every formula A built only from $\supset, \wedge, \neg, \forall$, the sequent $\neg\neg A^\dagger \vdash A^\dagger$ is provable in the system $\mathcal{G}_i^{\supset,\wedge,\neg,\forall}$.*

Proof. Since $\neg\neg\neg A \equiv \neg A$ is provable in $\mathcal{G}_i^{\supset,\wedge,\neg,\forall}$, the sequent $\neg\neg\neg\neg P \equiv \neg\neg P$ is provable in $\mathcal{G}_i^{\supset,\wedge,\neg,\forall}$ for every atom P, and thus the result follows from the definition of $A^\dagger$ and Lemma 13.6. $\square$

Actually, we can state a slightly more general version of Lemma 13.7, based on the observation that $\neg\neg\neg A \equiv \neg A$ is provable in $\mathcal{G}_i^{\supset,\wedge,\neg,\forall}$.

Lemma 13.8 *For every formula A built only from $\supset, \wedge, \neg, \forall$ and where every atomic subformula occurs negated (except $\perp$), the sequent $\neg\neg A \vdash A$ is provable in the system $\mathcal{G}_i^{\supset,\wedge,\neg,\forall}$.*

The formulae of the kind mentioned in Lemma 13.8 are called *negative* formulae. The following lemma shows that if we use double-negation, then $\vee$, $\supset$, and $\exists$ are definable intuitionistically from the connectives $\wedge, \neg, \forall$.

Lemma 13.9 *The following formulae are provable in $\mathcal{G}_i^{\supset,\wedge,\neg,\forall}$:*

$$\begin{aligned} \neg\neg(A \vee B) &\equiv \neg(\neg A \wedge \neg B), \\ \neg\neg\exists x A &\equiv \neg\forall x \neg A, \\ \neg(A \wedge \neg B) &\equiv \neg\neg(A \supset B). \end{aligned}$$

Proof. We give a proof of the sequents $\neg\neg\exists x A \vdash \neg\forall x\neg A$ and $\neg(A \wedge \neg B) \vdash \neg\neg(A \supset B)$, leaving the others as exercises.

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{A[y/x] \vdash A[y/x]}{\neg A[y/x], A[y/x] \vdash}}{\forall x\neg A, A[y/x] \vdash}}{\forall x\neg A, \exists x A \vdash}}{\forall x \neg A \vdash \neg\exists x A}}{\neg\neg\exists x A, \forall x\neg A \vdash}}{\neg\neg\exists x A \vdash \neg\forall x\neg A}$$

where y is a new variable,

$$
\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\neg(A \supset B), A \vdash A \qquad \dfrac{\dfrac{\dfrac{A, B, A \vdash B}{A, B \vdash A \supset B}}{\neg(A \supset B), A, B \vdash}}{\neg(A \supset B), A \vdash \neg B}}{\neg(A \supset B), A \vdash A \wedge \neg B}}{\neg(A \wedge \neg B), \neg(A \supset B), A \vdash}}{\neg(A \wedge \neg B), \neg(A \supset B), A \vdash B}}{\neg(A \wedge \neg B), \neg(A \supset B) \vdash A \supset B}}{\neg(A \wedge \neg B), \neg(A \supset B), \neg(A \supset B) \vdash}}{\dfrac{\neg(A \wedge \neg B), \neg(A \supset B) \vdash}{\neg(A \wedge \neg B) \vdash \neg\neg(A \supset B)}}
$$

□

We are now ready to prove the main lemma about the double-negation translation. The correctness of many embeddings of classical logic into intuitionistic logic follows from this lemma, including those due to Kolmogorov, Gödel, and Gentzen.

Lemma 13.10 *Let $\Gamma \vdash B_1, \ldots, B_n$ be any first-order sequent containing formulae made only from $\supset, \wedge, \neg$, and $\forall$. If $\Gamma \vdash B_1, \ldots, B_n$ is provable in $\mathcal{G}_c^{\supset,\wedge,\vee,\neg,\forall,\exists}$ then its translation $\Gamma^\dagger \vdash \neg(\neg B_1^\dagger \wedge \ldots \wedge \neg B_n^\dagger)$ is provable in $\mathcal{G}_i^{\supset,\wedge,\neg,\forall}$. In particular, if B is provable in $\mathcal{G}_c^{\supset,\wedge,\vee,\neg,\forall,\exists}$, then $B^\dagger$ is provable in $\mathcal{G}_i^{\supset,\wedge,\neg,\forall}$.*

Proof. First, we prove that if $\Gamma \vdash B_1, \ldots, B_n$ is provable in $\mathcal{G}_c^{\supset,\wedge,\vee,\neg,\forall,\exists}$ then $\Gamma^\dagger \vdash B_1^\dagger, \ldots, B_n^\dagger$ is also provable in $\mathcal{G}_c^{\supset,\wedge,\vee,\neg,\forall,\exists}$. This is done by a simple induction on proofs. Next, we prove that $\neg\neg\Gamma^\dagger \vdash \neg(\neg B_1^\dagger \wedge \ldots \wedge \neg B_n^\dagger)$ is provable in $\mathcal{G}_i^{\supset,\wedge,\neg,\forall}$. The only obstacle to Lemma 13.5 is the use of the ($\forall$: *right*)-rule. However, we have seen in the discussion following Lemma 13.5 that the problem is overcome for formulae such that $\neg\neg A \vdash A$ is provable in $\mathcal{G}_i^{\supset,\wedge,\neg,\forall}$. But this is the case by Lemma 13.7 (which itself is a direct consequence of Lemma 13.6), since we are now considering formulae of the form $A^\dagger$. Since $B \vdash \neg\neg B$ is provable in $\mathcal{G}_i^{\supset,\wedge,\neg,\forall}$ for any B, using cuts on the premises in $\neg\neg\Gamma^\dagger$, we obtain a proof of $\Gamma^\dagger \vdash \neg(\neg B_1^\dagger \wedge \ldots \wedge \neg B_n^\dagger)$ in $\mathcal{G}_i^{\supset,\wedge,\neg,\forall}$. In the special case where $n = 1$ and Γ is empty, we have shown that $\vdash \neg\neg B^\dagger$ is provable in $\mathcal{G}_i^{\supset,\wedge,\neg,\forall}$, and using Lemma 13.7, we obtain that $B^\dagger$ is provable in $\mathcal{G}_i^{\supset,\wedge,\neg,\forall}$. □

It is trivial that the converse of Lemma 13.10 holds (since $\mathcal{G}_i^{\supset,\wedge,\neg,\forall}$ is a subsystem of $\mathcal{G}_c^{\supset,\wedge,\vee,\neg,\forall,\exists}$) (and using the fact that $A^\dagger$ is classically equivalent to A). As a corollary of Lemma 13.10, observe that for negative formulae (defined in Lemma 13.8), A is provable in $\mathcal{G}_c^{\supset,\wedge,\neg,\forall}$ iff A is provable in $\mathcal{G}_i^{\supset,\wedge,\neg,\forall}$. This is because for a negative formula A, all atoms appears negated, and thus $A \equiv A^\dagger$ is provable in $\mathcal{G}_i^{\supset,\wedge,\neg,\forall}$.

We now define several translations of classical logic into intuitionistic logic.

Definition 13.11 The function $^\circ$ (due to Gentzen) is defined as follows:

$$
\begin{aligned}
A^\circ &= \neg\neg A, \quad \text{if } A \text{ is atomic},\\
(\neg A)^\circ &= \neg A^\circ,\\
(A \wedge B)^\circ &= (A^\circ \wedge B^\circ),\\
(A \supset B)^\circ &= (A^\circ \supset B^\circ),\\
(A \vee B)^\circ &= \neg(\neg A^\circ \wedge \neg B^\circ),\\
(\forall x A)^\circ &= \forall x A^\circ,\\
(\exists x A)^\circ &= \neg\forall x \neg A^\circ.
\end{aligned}
$$

The function * (due to Gödel) is defined as follows:

$$
\begin{aligned}
A^* &= \neg\neg A, \quad \text{if } A \text{ is atomic},\\
(\neg A)^* &= \neg A^*,\\
(A \wedge B)^* &= (A^* \wedge B^*),\\
(A \supset B)^* &= \neg(A^* \wedge \neg B^*),\\
(A \vee B)^* &= \neg(\neg A^* \wedge \neg B^*),\\
(\forall x A)^* &= \forall x A^*,\\
(\exists x A)^* &= \neg\forall x \neg A^*.
\end{aligned}
$$

The function $^\kappa$ (due to Kolmogorov) is defined as follows:

$$
\begin{aligned}
A^\kappa &= \neg\neg A, \quad \text{if } A \text{ is atomic},\\
(\neg A)^\kappa &= \neg A^\kappa,\\
(A \wedge B)^\kappa &= \neg\neg(A^\kappa \wedge B^\kappa),\\
(A \supset B)^\kappa &= \neg\neg(A^\kappa \supset B^\kappa),\\
(A \vee B)^\kappa &= \neg\neg(A^\kappa \vee B^\kappa),\\
(\forall x A)^\kappa &= \neg\neg\forall x A^\kappa,\\
(\exists x A)^\kappa &= \neg\neg\exists x A^\kappa.
\end{aligned}
$$

Since all atoms are negated twice in A°, A^*, and A^κ, we can use Lemma 13.6 to show that $A^\circ \equiv \neg\neg A^\circ$, $A^* \equiv \neg\neg A^*$, and $A^\kappa \equiv \neg\neg A^\kappa$ are provable in $\mathcal{G}_i^{\supset,\wedge,\neg,\forall}$. Then, using this fact and Lemma 13.9, we can show by induction on formulae that $A^\circ \equiv A^*$, $A^* \equiv A^\kappa$, and $A^\kappa \equiv A^\circ$ are provable in $\mathcal{G}_i^{\supset,\wedge,\neg,\forall}$. Consequently, for any sequent $\Gamma \vdash B_1, \ldots, B_n$, the sequent

$$\Gamma^\circ \vdash \neg(\neg B_1^\circ \wedge \ldots \wedge \neg B_n^\circ)$$

is provable in $\mathcal{G}_i^{\supset,\wedge,\neg,\forall}$ iff

$$\Gamma^* \vdash \neg(\neg B_1^* \wedge \ldots \wedge \neg B_n^*)$$

is provable in $\mathcal{G}_i^{\supset,\wedge,\neg,\forall}$ iff

$$\Gamma^\kappa \vdash \neg(\neg B_1^\kappa \wedge \ldots \wedge \neg B_n^\kappa)$$

is provable. Furthermore, it is easily shown that $A \equiv A^\circ$, $A \equiv A^*$, and $A \equiv A^\kappa$, are provable classically.

Theorem 13.12 *For any sequent $\Gamma \vdash B_1, \ldots, B_n$, if $\Gamma \vdash B_1, \ldots, B_n$ is provable in $\mathcal{G}_c^{\supset,\wedge,\vee,\neg,\forall,\exists}$, then the sequents $\Gamma^\circ \vdash \neg(\neg B_1^\circ \wedge \ldots \wedge \neg B_n^\circ)$, $\Gamma^* \vdash \neg(\neg B_1^* \wedge \ldots \wedge \neg B_n^*)$, and $\Gamma^\kappa \vdash \neg(\neg B_1^\kappa \wedge \ldots \wedge \neg B_n^\kappa)$, are provable in $\mathcal{G}_i^{\supset,\wedge,\neg,\forall}$. In particular, if A is provable in $\mathcal{G}_c^{\supset,\wedge,\vee,\neg,\forall,\exists}$, then A°, A^*, and A^κ, are provable in $\mathcal{G}_i^{\supset,\wedge,\neg,\forall}$.*

Proof. We simply have to observe that the translation $^\circ$ is in fact the composition of two functions: the first one $^\bullet$ is defined as in Definition 13.11, *except* that atoms remain unchanged, and the second function is just †. This translation has the property that $A^\bullet$ only contains the connectives $\supset, \wedge, \neg$, and $\forall$. Furthermore, it is easily shown that $\Gamma \vdash B_1, \ldots, B_n$ is provable in $\mathcal{G}_c^{\supset,\wedge,\vee,\neg,\forall,\exists}$ iff $\Gamma^\bullet \vdash B_1^\bullet, \ldots, B_n^\bullet$ is (because $A \equiv A^\bullet$ is easily provable in $\mathcal{G}_c^{\supset,\wedge,\vee,\neg,\forall,\exists}$). Therefore, Lemma 13.10 applies to $\Gamma^\bullet \vdash B_1^\bullet, \ldots, B_n^\bullet$, and we obtain the desired result. $\square$

It is trivial that the converse of Theorem 13.12 holds.

We shall now discuss another translation of classical logic into intuitionistic logic due to Girard [14]. Girard has pointed out that the usual double-negation translations have some rather undesirable properties:

(1) They are not compatible with substitution. Indeed, the translation $A[B/P]^*$ of $A[B/P]$ is not equal to $A^*[B^*/P]$ in general, due to the application of double negations to atoms.

(2) Negation is not involutive. For instance, $A[B/P]^*$ and $A^*[B^*/P]$ are related through the erasing of certain double negations (passing from $\neg\neg\neg P$ to $\neg P$), but this erasing is not harmless.

(3) Disjunction is not associative. For example, if $A \vee B$ is translated as $\neg(\neg A \wedge \neg B)$, then $(A \vee B) \vee C$ is translated as $\neg(\neg(\neg(\neg A \wedge \neg B)) \wedge \neg C)$, and $A \vee (B \vee C)$ is translated as $\neg(\neg A \wedge \neg(\neg(\neg B \wedge \neg C)))$.

Girard has discovered a translation which does not suffer from these defects, and this translation also turns out to be quite economical in the number of negation signs introduced [14]. The main idea is to assign a *sign* or *polarity* ($+$ or $-$) to every formula. Roughly speaking, a positive literal P (where P is an atom) is a formula of polarity $+$, a negative literal $\neg P$ is a formula of polarity $-$, and to determine the polarity of a compound formula, we combine its polarities as if they were truth values, except that $+$ corresponds to **false**, $-$ corresponds to **true**, existential formulae are always positive, and universal formulae are always negative. Given a sequent $\Gamma \vdash \Delta$, the idea is that right-rules have to be converted to left-rules, and in order to do this we need to move formulae in Δ to the lefthand side of the sequent. The new twist is that formulae in Δ will be treated differently according to their polarity. One of the key properties of polarities is that every formula A of polarity $-$ turns out to be equivalent to a formula of the form $\neg B$ with B of polarity $+$. Then every formula $A \equiv \neg B$ in Δ of polarity $-$ will be transferred to the lefthand side as B (and not as $\neg\neg B$), and every formula A in Δ of polarity $+$ will be transferred to the lefthand side as $\neg A$. The translation is then completely determined if we add the obvious requirement that the translation of a classically provable sequent should be intuitionistically provable, and that it should be as simple as possible. Let us consider some typical cases.

Case 1. The last inference is

$$\frac{\Gamma \vdash \neg C, \neg D}{\Gamma \vdash \neg C \vee \neg D}$$

where C and D are positive. The sequent $\Gamma \vdash \neg C, \neg D$ is translated as $\Gamma, C, D \vdash$, and we have the inference

$$\frac{\Gamma, C, D \vdash}{\Gamma, C \wedge D \vdash}$$

It is thus natural to translate $\neg C \vee \neg D$ as $\neg(C \wedge D)$, since then $C \wedge D$ will be placed on the lefthand side (because $\neg(C \wedge D)$ is negative).

Case 2. The last inference is

$$\frac{\Gamma \vdash C, D}{\Gamma \vdash C \vee D}$$

where C and D are positive. The sequent $\Gamma \vdash C, D$ is translated as $\Gamma, \neg C, \neg D \vdash$, and we have the inference

$$\frac{\Gamma, \neg C, \neg D \vdash}{\Gamma, \neg C \wedge \neg D \vdash}$$

This time, the simplest thing to do is to translate $C \vee D$ as $C \vee D$ (since $C \vee D$ is positive), so that $\neg(C \vee D)$ is placed on the lefthand side of the sequent. This is indeed legitimate because $\neg(C \vee D) \equiv \neg C \wedge \neg D$ is provable intuitionistically.

Case 3. The last inference is

$$\frac{\Gamma \vdash C[y/x]}{\Gamma \vdash \forall x C}$$

where C is positive. The sequent $\Gamma \vdash C[y/x]$ is translated as $\Gamma, \neg C[y/x] \vdash$, and we have the inference

$$\frac{\Gamma, \neg C[y/x] \vdash}{\Gamma, \exists x \neg C \vdash}$$

We translate $\forall x C$ as $\neg \exists x \neg C$, so that $\exists x \neg C$ is placed on the lefthand side of the sequent.

Case 4. The last inference is

$$\frac{\Gamma \vdash \neg C[y/x]}{\Gamma \vdash \forall x \neg C}$$

where C is positive. The sequent $\Gamma \vdash \neg C[y/x]$ is translated as $\Gamma, C[y/x] \vdash$, and we have the inference

$$\frac{\Gamma, C[y/x] \vdash}{\Gamma, \exists x C \vdash}$$

We translate $\forall x \neg C$ as $\neg \exists x C$, so that $\exists x C$ is placed on the lefthand side of the sequent.

Case 5. The last inference is

$$\frac{\Gamma \vdash C[t/x]}{\Gamma \vdash \exists x C}$$

where C is positive. The sequent $\Gamma \vdash C[t/x]$ is translated as $\Gamma, \neg C[t/x] \vdash$, and we have the inference

$$\frac{\Gamma, \neg C[t/x] \vdash}{\Gamma, \forall x \neg C \vdash}$$

The simplest thing to do is to translate $\exists x C$ as $\exists x C$, so that $\neg\exists x C$ is placed on the lefthand side of the sequent. This is possible because $\neg\exists x C \equiv \forall x \neg C$ is provable intuitionistically.

Case 6. The last inference is

$$\frac{\Gamma \vdash \neg C[t/x]}{\Gamma \vdash \exists x \neg C}$$

where C is positive. The sequent $\Gamma \vdash \neg C[t/x]$ is translated as $\Gamma, C[t/x] \vdash$. The simplest thing to do is to translate $\exists x \neg C$ as $\exists x \neg C$, so that $\neg\exists x \neg C$ is placed on the lefthand side of the sequent. This is possible because $\neg\exists x \neg C \equiv \forall x \neg\neg C$ is provable intuitionistically, and we have the sequence of inferences

$$\cfrac{\cfrac{\cfrac{\Gamma, C[t/x] \vdash}{\Gamma \vdash \neg C[t/x]}}{\Gamma, \neg\neg C[t/x] \vdash}}{\Gamma, \forall x \neg\neg C[t/x] \vdash}$$

Note that it was necessary to first double-negate $C[t/x]$. This is because $\neg\exists x \neg C \equiv \forall x \neg\neg C$ is provable intuitionistically, but $\neg\exists x \neg C \equiv \forall x C$ is not.

Case 7. The last inference is

$$\frac{\Gamma \vdash C \quad \Gamma \vdash D}{\Gamma \vdash C \wedge D}$$

where C, D are positive. The sequents $\Gamma \vdash C$ and $\Gamma \vdash D$ are translated as $\Gamma, \neg C \vdash$ and $\Gamma, \neg D \vdash$. Since $C \wedge D$ is positive, the simplest thing to do is to translate $C \wedge D$ as $C \wedge D$, so that $\neg(C \wedge D)$ is placed on the lefthand side of the sequent. This is possible because $\neg(\neg\neg C \wedge \neg\neg D) \equiv \neg(C \wedge D)$ is provable intuitionistically, and we have the sequence of inferences

$$\cfrac{\cfrac{\cfrac{\Gamma, \neg C \vdash}{\Gamma \vdash \neg\neg C} \qquad \cfrac{\Gamma, \neg D \vdash}{\Gamma \vdash \neg\neg D}}{\Gamma \vdash \neg\neg C \wedge \neg\neg D}}{\Gamma, \neg(\neg\neg C \wedge \neg\neg D) \vdash}$$

Case 8. The last inference is

$$\frac{\Gamma \vdash \neg C \quad \Gamma \vdash D}{\Gamma \vdash \neg C \wedge D}$$

where C, D are positive. The sequents $\Gamma \vdash \neg C$ and $\Gamma \vdash D$ are translated as $\Gamma, C \vdash$ and $\Gamma, \neg D \vdash$. Since $\neg C \wedge D$ is positive, the simplest thing to do is to translate $\neg C \wedge D$ as $\neg C \wedge D$, so that $\neg(\neg C \wedge D)$ is placed on the lefthand side of the sequent. This is possible because $\neg(\neg C \wedge \neg\neg D) \equiv \neg(\neg C \wedge D)$ is provable intuitionistically, and we have the sequence of inferences

$$\cfrac{\cfrac{\cfrac{\Gamma, C \vdash}{\Gamma \vdash \neg C} \qquad \cfrac{\Gamma, \neg D \vdash}{\Gamma \vdash \neg\neg D}}{\Gamma \vdash \neg C \wedge \neg\neg D}}{\Gamma, \neg(\neg C \wedge \neg\neg D) \vdash}$$

Case 9. The last inference is

$$\frac{\Gamma \vdash \neg C \quad \Gamma \vdash \neg D}{\Gamma \vdash \neg C \wedge \neg D}$$

where C, D are positive. The sequents $\Gamma \vdash \neg C$ and $\Gamma \vdash \neg D$ are translated as $\Gamma, C \vdash$ and $\Gamma, D \vdash$, and we have the inference

$$\frac{\Gamma, C \vdash \quad \Gamma, D \vdash}{\Gamma, C \vee D \vdash}$$

We translate $\neg C \wedge \neg D$ as $\neg(C \vee D)$, so that $C \vee D$ is placed on the lefthand side of the sequent.

Considering all the cases, we arrive at the following tables defining the Girard-translation $\widehat{A}$ of a formula.

Definition 13.13 Given any formula A, its sign (polarity) and its Girard-translation $\widehat{A}$ are given by the following tables:

If $A = P$ where P is an atom, including the constants $\top$ (**true**) and $\perp$ (**false**), then $sign(A) = +$ and $\widehat{A} = A$, and if A is a compound formula then $\widehat{A}$ is given by the following tables:

Girard's $\neg\neg$-Translation				
A	B	$A \wedge B$	$A \vee B$	$A \supset B$
$+, C$	$+, D$	$+, C \wedge D$	$+, C \vee D$	$-, \neg(C \wedge \neg D)$
$+, C$	$-, \neg D$	$+, C \wedge \neg D$	$-, \neg(\neg C \wedge D)$	$-, \neg(C \wedge D)$
$-, \neg C$	$+, D$	$+, \neg C \wedge D$	$-, \neg(C \wedge \neg D)$	$+, C \vee D$
$-, \neg C$	$-, \neg D$	$-, \neg(C \vee D)$	$-, \neg(C \wedge D)$	$-, \neg(\neg C \wedge D)$

Girard's $\neg\neg$-Translation			
A	$\forall x A$	$\exists x A$	$\neg A$
$+, C$	$-, \neg\exists x \neg C$	$+, \exists x C$	$-, \neg C$
$-, \neg C$	$-, \neg\exists x C$	$+, \exists x \neg C$	$+, C$

In order to state the main property of the Girard-translation, we need one more definition. Given a formula A, we define its translation $\overline{A}$ as follows:

$$\overline{A} = \begin{cases} \neg\widehat{A} & \text{if } sign(A) = +, \\ B & \text{if } sign(A) = - \text{ and } \widehat{A} = \neg B. \end{cases}$$

Given $\Delta = B_1, \ldots, B_n$, we let $\overline{\Delta} = \overline{B_1}, \ldots, \overline{B_n}$. Then, a sequent $\Gamma \vdash \Delta$ is translated into the sequent $\Gamma, \overline{\Delta} \vdash$.

We have the following theorem due du Girard [14].

Theorem 13.14 *Given any sequent $\Gamma \vdash \Delta$, if $\Gamma \vdash \Delta$ is provable classically (in $\mathcal{G}_c^{\supset,\wedge,\vee,\neg,\forall,\exists}$), then its translation $\Gamma, \overline{\Delta} \vdash$ is provable intuitionistically (in $\mathcal{G}_i^{\supset,\wedge,\vee,\neg,\forall,\exists}$).*

Proof. By induction on the structure of proofs. We have already considered a number of cases in the discussion leading to the tables of Definition 13.13. As an auxiliary result, we need to show that the following formulae are provable intuitionistically (in fact, the top four follow from lemma 13.4, lemma 13.9, and the fact that $\neg\neg\neg A \equiv \neg A$ is provable intuitionistically):

$$
\begin{aligned}
\neg(C \vee D) &\equiv \neg C \wedge \neg D, \\
\neg(\neg\neg C \wedge \neg\neg D) &\equiv \neg(C \wedge D), \\
\neg(\neg C \wedge \neg\neg D) &\equiv \neg(\neg C \wedge D), \\
\neg(\neg\neg C \wedge \neg D) &\equiv \neg(C \wedge \neg D), \\
\neg\exists x C &\equiv \forall x \neg C, \\
\neg\exists x \neg C &\equiv \forall x \neg\neg C.
\end{aligned}
$$

We leave the remaining cases as an exercise. □

Observe that a formula A of any polarity can be made into an equivalent formula of polarity $+$, namely $A^+ = A \wedge \top$, or an equivalent formula of polarity $-$, namely $A^- = A \vee \neg\top$. The Girard-translation has some nice properties, and the following lemma lists some of them [14].

Lemma 13.15 *The translation $A \mapsto \widehat{A}$ given in Definition 13.13 is compatible with substitutions respecting polarities. Furthermore, it satisfies a number of remarkable identities (in the equivalences below, it is assumed that we are considering the Girard-translations of the formulae involved. For example, in (i), we really mean $\widehat{\neg\neg A} \equiv \widehat{A}$. To unclutter the notation, hats will be omitted):*

(i) *Negation is involutive: $\neg\neg A \equiv A$.*

(ii) *De Morgan identities: $\neg(A \wedge B) \equiv \neg A \vee \neg B$; $\neg(A \vee B) \equiv \neg A \wedge \neg B$; $A \supset B \equiv \neg A \vee B$; $\neg\forall x A \equiv \exists x \neg A$; $\neg\exists x A \equiv \forall x \neg A$.*

(iii) *Associativity of $\wedge$ and $\vee$; as a consequence, $(A \wedge B) \supset C \equiv A \supset (B \supset C)$, and $A \supset (B \vee C) \equiv (A \supset B) \vee C$.*

(iv) *Neutrality identities: $A \vee \perp \equiv A$; $A \wedge \neg \perp \equiv A$.*

(v) *Commutativity of $\wedge$ and $\vee$ (as a consequence, $A \supset B \equiv \neg B \supset \neg A$).*

(vi) *Distributivity identities with restriction on polarities: $A \wedge (P \vee Q) \equiv (A \wedge P) \vee (A \wedge Q)$; $A \vee (L \wedge M) \equiv (A \vee L) \wedge (A \vee M)$ (where P, Q are positive, and L, M negative).*

(vii) *Idempotency identities: $P^+ \equiv P$ where P is positive; $N^- \equiv N$ where N is negative; as a consequence, $A^{++} \equiv A^+$ and $A^{--} \equiv A^-$.*

(viii) *Quantifier isomorphisms: $A \wedge \exists x P \equiv \exists x(A \wedge P)$ if x is not free in A and P is positive; $A \vee \forall x N \equiv \forall x(A \vee N)$ if x is not free in A and N is negative.*

Proof. The proof is quite straightforward, but somewhat tedious. Because of the polarities, many cases have to be considered. Some cases are checked in Girard [14], and the others can be easily verified. □

Acknowledgement. I wish to thank Hassan Aït-Kaci, Val Breazu-Tannen, Jin Choi, Philippe de Groote, Andreas Podelski, Andre Scedrov, and Ascánder Suárez, for their comments. Special thanks to Peter Baumann, Kathleen Milsted, Marcin Skubiszewski, and Jean-Christophe Patat, for proofreading earlier versions very carefully, and to Jean Yves Girard who has been and remains a constant source of inspiration (and invented the "Girardian turnstile" $\vdash$). I am particularly grateful to Peter Baumann who was relentless in debugging this paper. Not only did he read the entire paper with the eyes of an eagle, but he suggested many improvements, discovered errors, and helped me fix them. Finally, I wish to thank the Digital Equipment Corporation for offering me an ideal environment to spend a sabbatical year in 1991-1992, and where 99% of this research was accomplished.

References

[1] S. Abramsky. Computational interpretation of linear logic. Technical Report DOC 90/20, Imperial College, London SW7 2BZ, October 1990. To appear in Theor. Comp. Sci.

[2] H.P. Barendregt. *The Lambda Calculus*, volume 103 of *Studies in Logic*. North-Holland, second edition, 1984.

[3] H.B. Curry and R. Feys. *Combinatory Logic, Vol. I.* Studies in Logic. North-Holland, third edition, 1974.

[4] A.G. Dragalin. *Mathematical intuitionism – introduction to proof theory*, volume 67 of *Translations of Mathematical Monographs*. Amer. Math. Soc., Providence (RI), 1988.

[5] R. Dyckhoff. Contraction-free sequent calculi for intuitionistic logic. *J. Symbolic Logic*, 57, 1992. To appear.

[6] Jean H. Gallier. *Logic for Computer Science*. Harper and Row, New York, 1986.

[7] Jean H. Gallier. On Girard's "candidats de reductibilités". In P. Odifreddi, editor, *Logic And Computer Science*, pages 123–203. Academic Press, London, New York, May 1990.

[8] G. Gentzen. Investigations into logical deduction. In M.E. Szabo, editor, *The Collected Papers of Gerhard Gentzen*. North-Holland, 1969.

[9] J.-Y. Girard, Y. Lafont, and P. Taylor. *Proofs and Types*, volume 7 of *Cambridge Tracts in Theoretical Computer Science*. Cambridge University Press, 1989.

[10] Jean-Yves Girard. Une extension de l'interprétation de gödel à l'analyse, et son application à l'élimination des coupures dans l'analyse et la théorie des types. In J.E. Fenstad, editor, *Proc. 2nd Scand. Log. Symp.*, pages 63–92. North-Holland, 1971.

[11] Jean-Yves Girard. *Interprétation fonctionnelle et élimination des coupures de l'arithmétique d'ordre supérieur*. PhD thesis, Université de Paris VII, June 1972. Thèse de Doctorat d'Etat.

[12] Jean-Yves Girard. Linear logic. *Theoretical Computer Science*, 50:1–102, 1987.

[13] Jean-Yves Girard. *Proof Theory and Logical Complexity*. Bibliopolis, 1987.

[14] Jean-Yves Girard. A new constructive logic: classical logic. *Math. Struc. in Comp. Sci.*, pages 255–296, 1991.

[15] J.R. Hindley and J.P. Seldin. *Introduction to Combinators and λ-Calculus*, volume 1 of *London Mathematical Society Student texts.* Cambridge University Press, 1986.

[16] W. A. Howard. The formulae-as-types notion of construction. In J. P. Seldin and J. R. Hindley, editors, *To H. B. Curry: Essays on Combinatory Logic, Lambda Calculus and Formalism*, pages 479–490. Academic Press, London, 1980. Reprint of manuscript first published in 1969.

[17] J. Hudelmaier. *Bounds for cut-elimination in intuitionistic propositional logic.* PhD thesis, University of Tübingen, 1989. Dissertation.

[18] S. Kleene. *Introduction to Metamathematics.* North-Holland, seventh edition, 1952.

[19] J.L. Krivine. *Lambda-Calcul, types et modèles.* Etudes et recherches en informatique. Masson, 1990.

[20] P. Lincoln, A. Scedrov, and N. Shankar. Linearizing intuitionistic implication. In G. Kahn, editor, *Sixth IEEE LICS Symposium*, pages 51–62, Vrije University, Amsterdam, July 1991. IEEE.

[21] P. Martin-Löf. An intuitionistic theory of types. Technical report, University of Stokholm, Stockholm, Sweden, 1972. Privately circulated manuscript.

[22] A.M. Pitts. On an interpretation of second-order quantification in first-order intuitionistic propositional logic. *J. Symbolic Logic*, 57:33–52, 1992.

[23] D. Prawitz. *Natural deduction, a proof-theoretical study.* Almquist & Wiksell, Stockholm, 1965.

[24] D. Prawitz. Ideas and results in proof theory. In J.E. Fenstad, editor, *Proc. 2nd Scand. Log. Symp.*, pages 235–307. North-Holland, 1971.

[25] P. Rozière. *Règles admissibles en calcul propositionnel intuitionniste.* PhD thesis, Université de Paris VII, May 1992. Thèse de Doctorat.

[26] H. Schwichtenberg. Proof theory: some applications of cut-elimination. In J. Barwise, editor, *Handbook of Mathematical Logic*, volume 90 of *Studies in Logic*, pages 867–895. North Holland, Amsterdam, 1977.

[27] S. Stenlund. *Combinators, Lambda Terms, and Proof Theory.* D. Reidel, Dordrecht, Holland, 1972.

[28] W.W. Tait. Intensional interpretation of functionals of finite type i. *J. Symbolic Logic*, 32:198–212, 1967.

[29] W.W. Tait. Normal derivability in classical logic. In J. Barwise, editor, *The Syntax and Semantics of Infinitary Languages*, volume 72 of *Lecture Notes in Math.*, pages 204–236. Springer Verlag, 1968.

[30] W.W. Tait. A realizability interpretation of the theory of species. In R. Parikh, editor, *Logic Colloquium*, volume 453 of *Lecture Notes in Math.*, pages 240–251. Springer Verlag, 1975.

[31] G. Takeuti. *Proof Theory*, volume 81 of *Studies in Logic.* North-Holland, 1975.

[32] A.S. Troelstra and D. van Dalen. *Constructivism in Mathematics: An Introduction, Vol. I and II*, volume 123 of *Studies in Logic.* North-Holland, 1988.

[33] A.M. Ungar. *Normalization, cut-elimination and the theory of proofs*, volume 28 of *CSLI Lecture Notes.* CSLI, 1992.

[34] D. van Dalen. *Logic and Structure.* Universitext. Springer Verlag, second edition, 1980.

[35] N.N. Vorob'ev. The derivability problem in the constructive propositional calculus with strong negation. *Doklady Akademii Nauk SSSR*, 85:689–692, 1952. in Russian.

[36] N.N. Vorob'ev. A new algorithm for derivability in the constructive propositional calculus. *Amer. Math. Soc.*, (2) 94:37–71, 1970. English Translation.

[37] Jeff Zucker. The correspondence between cut-elimination and normalization. *Annals of Mathematical Logic*, 7:1–112, 1974.