Natural Deduction

Francis Jeffry Pelletier
Simon Fraser University

Allen P. Hazen
University of Melbourne

Contents

1 Introduction

Work that is called 'natural deduction' is carried out in two ways: first, as an object-language method to prove theorems and to demonstrate the validity of arguments; and secondly, as a metatheoretic investigation into the properties of these types of proofs and the use of properties of this sort to demonstrate results about other systems (such as the consistency of arithmetic or analysis). In the former realm, we turn to elementary textbooks that introduce logic to philosophy students; in the latter realm, we turn to the topic of proof theory.

2 Object Language Natural Deduction

Natural deduction for classical logic is the type of logical system that almost all philosophy departments in North America teach as their first and (often) second course in logic.[1] Since this one- or two-course sequence is all that is required by most North American philosophy departments, most non-logician philosophers educated in North America know only this about logic. Of course, there are those who take further courses, or take logic in mathematics or computer science departments. But the large majority of North American philosophers have experienced only natural deduction as taught using one or another of the myriad elementary logic textbooks (or professor's classroom notes, which are usually just "better ways to explain" what one or another of the textbooks put forth as logic).

But when a student finishes these one- or two-semester courses, he or she is often unable to understand a different elementary logic textbook, even though it and the textbook from the course are both "natural deduction". Part of the reason for this — besides the students' not yet having an understanding of what logic is — concerns the fact that many different ideas have gone into these different books, and from the point of view of an elementary student, there can seem to be very little in common in these books. There is, of course, the basic issue of the specific language under consideration: will it have upper- or lower-case letters? will it use arrows or horseshoes? will it have free variables in premises? etc. But these issues of the choice of language are not what we are pointing towards here; instead, we think of the differing ways that different textbooks represent proofs, differing rules of inference, and so on. How can all these differences be accommodated under the single term 'natural deduction'? What are the essential properties that make a proof system be 'natural deduction'? What is 'natural' about them all?

Our view is that there are quite a number of characteristics that contribute to a proof system's being called 'natural deduction', but a system need not have them all in order to be a natural deduction system. We think that the current connotation of the term functions rather like a prototype: there is/are some exemplar(s) that the term most clearly

[1] This is still true, despite the encroachment of semantic tableaux methods that are now often taught alongside natural deduction.

applies to and which manifest(s) a number of characteristics. But there are other proof systems that differ from this prototypical natural deduction system and are nevertheless correctly characterized as being natural deduction. It is not clear just how many of the properties that the prototype exemplifies can be omitted and still have a system that is correctly characterized as a natural deduction system, and we will not try to give an answer. Instead we focus on a number of features that are manifested to different degrees by the various natural deduction systems. The picture is that if a system ranks "low" on one of these features, it can "make up for it" by ranking high on different features. And it is somehow an overall rating of the total amount of conformity to the entire range of these different features that determines whether any specific logical system will be called a natural deduction system. Some of these features stem from the initial introduction of natural deduction in 1934 (Jaśkowski, 1934; Gentzen, 1934); but even more strongly, in our opinion, is the effect that elementary textbooks from the 1950s had.

This mixture of features lends itself to identifying both a wide and a narrow notion of 'natural deduction'. The narrow one comes from the formal characterization of proof systems by (especially) Gentzen, and to some extent from the elementary textbook authors. The wide notion comes by following up some informal remarks that Gentzen and Jaśkowski made and which have been repeated in the elementary textbooks. Also thrown into this are remarks that have been made by researchers in related disciplines. . . such as in mathematics and computer science. . . when they want to distinguish natural deduction from their own, different, logical proof systems.

2.1 The Wider Notion of Natural Deduction

Before moving onto distinctions among types of proof systems, we mention features that are, by some, associated with natural deduction. As we have said, we do not think of these features as *defining* natural deduction, but rather as contributing to the general mixture of properties that are commonly invoked when one thinks of natural deduction.

One meaning of 'natural deduction'—especially in the writings from computer science and mathematics, where there is often a restriction to a small set of connectives or to a normal form—focuses on the notion that systems employing it will retain the 'natural form' of first-order logic and will not restrict itself to any subset of the connectives nor any normal form representation. Although this is clearly a feature of the modern textbooks, we can easily see that such a definition is neither necessary nor sufficient for a logical system's being a natural deduction system. For, surely we can give natural deduction accounts for logics that have restricted sets of connectives, so it is not necessary. And we can have non-natural-deduction systems (e.g., axiomatic systems) that contain all the usual connectives, so it is not sufficient.

Another feature in the minds of many is that the inference rules are "natural" or "pretheoretically accepted." To show how widely accepted this feature is, here is what five elementary natural deduction textbooks across a fifty year span have to say. (Suppes, 1957, p.

viii) says: "The system of inference... has been designed to correspond as closely as possible to the author's conception of the most natural techniques of informal proof." (Kalish and Montague, 1964, p. 38) says that these systems "are said to employ natural deduction and, as this designation indicates, are intended to reflect intuitive forms of reasoning." (Bonevac, 1987, p. 89) says: "we'll develop a system designed to simulate people's construction of arguments.... it is natural in the sense that it approaches... the way people argue." (Chellas, 1997, p. 134) says "Because the rules of inference closely resemble patterns of reasoning found in natural language discourse, the deductive system is of a kind called natural deduction." And (Goldfarb, 2003, p. 181) says "What we shall present is a system for *deductions*, sometimes called a system of *natural deduction*, because to a certain extent it mimics certain natural ways we reason informally." These authors are echoing (Gentzen, 1934, p. 74) "We wish to set up a formalism that reflects as accurately as possible the actual logical reasoning involved in mathematical proofs."

But this also is neither necessary nor sufficient. An axiom system with only modus ponens as a rule of inference obeys the restriction that all the rules of inference are "natural", yet no one wants to call such a system 'natural deduction', so it is not a sufficient condition. And we can invent rules of inference that we would happily call natural deduction even when they do not correspond to particularly normal modes of thought (such as could be done if the basis had unusual connectives like the Sheffer stroke[2], and is often done in modal logics, many-valued logics, relevant logics, and other non-standard logics).

As we have said, the notion of a rule of inference "being natural" or "pretheoretically accepted" is often connected with formal systems of natural deduction; but as we also said, the two notions are not synonymous or even co-extensive. This means that there is an interesting area of research open to those who wish to investigate what "natural reasoning" is in ordinary, non-trained people, and its relationship with what logicians call 'natural deduction'. This sort of investigation is being carried out by a group of cognitive scientists, but their results are far from universally accepted (Rips, 1994; Johnson-Laird and Byrne, 1991). (See also the papers in Adler and Rips, 2008, esp. §3.)

2.2 Different Proof Systems

Another way to distinguish natural deduction from other methods is to compare what these competing proof systems offer or require, compared with natural deduction. Proof systems can be characterized by the way proofs of theorems and arguments proceed. The syllogistic, for example, is characterized by the basic proofs consisting of two premises and a conclusion, and which obeyed certain constraints on the items that made up these sentences. And all extended proofs consist of concatenations of basic proofs, so long as they obey certain other constraints. An axiomatic system contains a set of "foundational" statements (axioms) and some rules that characterize how these axioms can be transformed into other statements. A tableaux system consists of a number of rules that describe how to "decompose" a formula

[2]See Price (1961) for a nice, but unusual, trio of rules for the Sheffer stroke.

to be proved, and does so with the intent of discovering whether or not there can be a proof of the formula.

The foregoing descriptions are merely rough-and-ready. For one thing, they do not always distinguish one type of proof system from another, as we will see. For another thing, all of them are rather too informal to clearly demarcate a unique set of proof systems. Part of the explanation of these shortcomings is that the different types of systems in fact *do* overlap, and that (with a bit of latitude) one can see some systems as manifesting more than one type of proof theory.

Nonetheless, we think that each type has a central core — a prototypical manifestation — and that the resulting cores are in fact different from one another. We think that there are at least five[3] different types of modern systems of logic: axiomatic, resolution, tableaux, sequent calculus, and natural deduction. We do not intend to spend much time describing the first two types, merely enough to identify them to the reader who is already familiar with some different types of systems. Later on (§3.2) we will discuss the relation between sequent calculus and natural deduction, and also describe the way tableaux methods are related to both sequent calculus and natural deduction. Our main goal is to describe natural deduction, and it is to that end that our accounts of the other types of systems are directed: as signposts that describe what is *not* considered a natural deduction system.

Axiomatic Systems of logic[4] are characterized by having axioms — formulas that are not proved other than by a citation of the formula itself — although this is not a sufficient characterization because natural deduction systems also often have axioms, as we will see below. Prototypically, axiom systems are characterized by having only a small number of rules, often merely modus ponens (detachment; conditional elimination) plus a rule of substitution (or schema formulation of the axioms). And then a proof of formula φ is a sequence of lines, each one of which is either an axiom or follows from preceding lines by one of the rules of inference, and whose last line is φ. Premises can also be accommodated, usually by allowing them to be entered at any point in the proof, but not allowing the rule of substitution to apply to them.

Resolution Systems of logic[5] are refutation-based in that they start by assuming that the to-be-proved is false, that is, starting with its negation (in classical logic). It also

[3] There are others, no doubt, such as "the inverse method". But these lesser-known systems won't be mentioned further. (The inverse method is generally credited to Sergey Maslov (1964; 1969). Nice introductions to the inverse method are Lifschitz (1989); Degtyarev and Voronkov (2001).) There are also algebraic systems, where formulas are taken as terms and which use substitution of equals for equals as rules in an equation calculus. And there is also the interesting case of Peirce's "existential graphs."

[4] Often called Frege or Hilbert systems, although the idea of an axiomatic system seems to go back to Euclid. The name 'Hilbert system' is perhaps due to Gentzen (1934), who mentioned "einem dem Hilbertschen Formalismus angeglichenen Kalkül"; and its use was solidified by Kleene in his (1952, §15) when he called them "Hilbert-type systems". If one wanted to use a name more modern than 'Euclid', it would be historically more accurate to call them "Frege systems" (as some logicians and historians in fact do).

[5] Initially described by Robinson (1965), these are the most commonly-taught methods of logic in computer science departments, and are embedded into the declarative programming language Prolog.

employs a normal form: this negation, plus all the premises (if any), are converted to clausal normal form. In the propositional logic clausal normal form is just conjunctive normal form (a conjunction of disjunctions of literals [which are atomic sentences or their negations]). In the predicate logic one first converts to a prenex normal form where all quantifiers have scope over the remainder of a formula, and then existential quantifiers are eliminated by use of skolem functions. The remaining universal quantifiers are dropped, and variables are assumed to be universally quantified. These formulas are then converted to conjunctive normal form. Each conjunct is called a clause, and is considered as a separate formula. The resulting set of clauses then has two rules applied to them (and to the clauses that result from the application of these rules): resolution and unification. Resolution is the propositional rule:

Definition 1. Resolution:
From $(p_1 \vee p_2 \vee \cdots \vee r \vee \cdots \vee p_n)$ and $(q_1 \vee q_2 \vee \cdots \vee \neg r \vee \cdots \vee q_m)$
infer $(p_1 \vee p_2 \vee \cdots \vee p_n \vee q_1 \vee q_2 \vee \cdots \vee q_m)$. [With no occurrence of r or $\neg r$].

In the predicate logic case, if one of r and $\neg r$ contains a free variable (hence universally quantified) x while the other contains some term τ in that position, then infer the clause in accordance with the rule of resolution but where all occurrences of x in the resolvent clause are replaced by the most general unifier of x and τ. The goal of a resolution proof is to generate the null clause, which happens when (the most general unifiers of) the two parent clauses are singletons (disjunctions with only one disjunct) that are negations of each other. If the empty clause is generated, then the initial set of formulas (with its unnegated-conclusion) is a valid argument.

Tableaux Systems of logic[6] are characterized by being decompositional in nature, and often are constructed so as to mimic the semantic evaluation one would employ in determining whether some given formula is true. (Because of this one often sees the term 'semantic tableaux method', indicating to some that these methods are not really proof procedures but rather are ways to semantically evaluate formulas. We take the view here that the system can also be viewed as purely formal methods to manipulate formulas, just as much as axiomatic systems seem to do. And therefore they count as proof systems.) Generally speaking, these systems start with a formula asserting that the formula to be proved is not a designated formula. This can be done in a number of different ways, such as claiming that it is false or that its negation is true (depending on whether the system is bivalent and whether negation works "normally"). The decomposition rules are constructed so as to "preserve truth": "if the formula to be decomposed were true, then at least one of the following would be true." At the end, if the initial formula were provable, then there could be no formula that would be true. (Various adjustments to this characterization are necessary if the system is not bivalent or not 'extensional'.) Most developments of tableaux

[6] Usually these are credited to Beth (1955) and Hintikka (1953, 1955a,b), although the ideas can be teased out of Herbrand (1930) and Gentzen (1934), as done in Craig (1957). See also Anellis (1990) for a historical overview.

systems express proofs in terms of decomposition trees, and when a branch of the tree turns out to be impossible, then it is marked 'closed'; if all branches are closed, then the initial argument is valid. We return to tableaux systems in §3.7.

Sequent Calculus was invented by Gerhard Gentzen (Gentzen, 1934), who used it as a stepping-stone in his characterization of natural deduction, as we will outline in some detail in §3.2. It is a very general characterization of a proof; the basic notation being $\varphi_1, \ldots, \varphi_n \vdash \psi_1, \ldots, \psi_m$, which means that it is a consequence of the premises $\varphi_1, \ldots, \varphi_n$ that at least one of $\psi_1, \ldots, \psi_m$ holds.[7] If Γ and Σ are sets of formulas, then $\Gamma \vdash \Sigma$ means that it is a consequence of all the formulas of Γ that at least one of the formulas in Σ holds. Sequent systems take basic sequents such as $\varphi \vdash \varphi$ as axiomatic, and then a set of rules that allow one to modify proofs, or combine proofs. The modification rules are normally stated in pairs, 'x on the left' and 'x on the right': how to do something to the premise set Γ and how to do it to the conclusion set Σ. So we can understand the rules as saying "if there is a consequence of such-and-so form, then there is also a consequence of thus-and-so form". These rules can be seen as being of two types: structural rules that characterize the notion of a proof, and logical rules that characterize the behavior of connectives. For example, the rule that from $\Gamma \vdash \Sigma$ one can infer $\Gamma, \varphi \vdash \Sigma$ ("thinning on left") characterizes the notion of a proof (in classical logic), while the rule that from $\Gamma, \varphi \vdash \Sigma$ one can infer $\Gamma, (\varphi \wedge \psi) \vdash \Sigma$ ("$\wedge$-Introduction on left") characterizes (part of) the behavior of the logical connective $\wedge$ when it is a premise. We expand considerably on this preliminary characterization of sequent calculi in §3.2.

2.3 The Beginnings of Natural Deduction: Jaśkowski and Gentzen (and Suppes) on Representing Natural Deduction Proofs

Gentzen (1934) coined the term 'natural deduction', or rather the German *das natürliche Schließen*.[8] But Jaśkowski (1934) might have a better claim to have been the first to invent a system that embodies what we now recognize as natural deduction, calling it a "method of suppositions".

According to Jaśkowski (1934), Jan Łukasiewicz had raised the issue in his 1926 seminars that mathematicians do not construct their proofs by means of an axiomatic theory (the systems of logic that had been developed at the time) but rather made use of other reasoning methods; especially they allow themselves to make "arbitrary assumptions" and see where they lead. Łukasiewicz wondered whether there could be a logical theory that embodied this insight but which yielded the same set of theorems as the axiomatic systems

[7] One can also read this as asserting the existence of proofs: the just-mentioned sequent would then be understood as saying that there is a proof from the premises $\varphi_1, \ldots, \varphi_n$ to at least one of $\psi_1, \ldots, \psi_m$. Basic sequents, described just below, would then be understood as asserting the existence of "basic proofs". In such an interpretation of sequents, the inference rules would then be seen as asserting that "if there is a proof of such-and-such type, then there is also a proof of so-and-so type."

[8] For more on Gentzen, and how natural deduction fit into his broader concerns about the consistency of arithmetic, etc., see von Plato (2008b,a).

then in existence. Again according to Jaśkowski (1934), he (Jaśkowski) developed such a system and presented it to the First Polish Mathematical Congress in 1927 at Lvov, and it was mentioned in their published proceedings of 1929. There seems to be no copies of Jaśkowski's original paper in circulation, and our knowledge of the system derives from a lengthy footnote in Jaśkowski (1934). (This is also where he said that it was presented and an abstract published in the Proceedings. Jan Woleński, in personal communication, tells us that in his copy of the Proceedings, Jaśkowski's work (Jaśkowski, 1929) was reported by title in the Proceedings.) Although the footnote describes the earlier use of a graphical method to represent these proofs, the main method described in Jaśkowski (1934) is rather different—what we below call a bookkeeping method. Cellucci (1995) recounts Quine's visit to Warsaw in 1933, and his meeting with Jaśkowski, and he suggests that the change in representational method might be due to a suggestion of Quine (who also used a version of this bookkeeping method in his own later system, Quine, 1950a).

This earlier graphical method consists in drawing boxes or rectangles around portions of a proof; the other method amounts to tracking the assumptions and their consequences by means of a bookkeeping annotation alongside the sequence of formulas that constitutes a proof. In both methods the restrictions on completion of subproofs (as we now call them) are enforced by restrictions on how the boxes or bookkeeping annotations can be drawn. We would now say that Jaśkowski's system had two subproof methods: conditional-proof (conditional-introduction)[9] and reductio ad absurdum (negation-elimination). It also had rules for the direct manipulation of formulas (e.g., Modus Ponens). After formulating his set of rules, Jaśkowski remarks (p.238) that the system "has the peculiarity of requiring no axioms" but that he can prove it equivalent to the established axiomatic systems of the time. (He shows this for various axiom systems of Łukasiewicz, Frege, and Hilbert). He also remarks (p.258) that his system is "more suited to the purposes of formalizing practical [mathematical] proofs" than were the then-accepted systems, which are "so burdensome that [they are] avoided even by the authors of logical [axiomatic] systems." Furthermore, "in even more complicated theories the use of [the axiomatic method] would be completely unproductive." Given all this, one could say that Jaśkowski was the inventor of natural deduction as a complete logical theory.

Working independently of Łukasiewicz and Jaśkowski, Gerhard Gentzen published an amazingly general and amazingly modern-sounding two-part paper in (1934/35). The opening remarks of Gentzen (1934) are

[9] Obviously, this rule of Conditional-Introduction is closely related to the deduction theorem, that from the fact that $\Gamma, \varphi \vdash \psi$ it follows that $\Gamma \vdash (\varphi \to \psi)$. The difference is primarily that Conditional-Introduction is a rule of inference in the object language, whereas the deduction theorem is a metalinguistic theorem that guarantees that proofs of one sort could be converted into proofs of the other sort. According to (Kleene, 1967, p.39fn33), "The deduction theorem as an informal theorem proved about particular systems like the propositional calculus and the predicate calculus... first appears explicitly in Herbrand (1930) (and without proof in Herbrand, 1928); and as a general methodological principle for axiomatic-deductive systems in Tarski (1930b). According to (Tarski, 1956, p.32fn), it was known and applied by Tarski since 1921."

> My starting point was this: The formalization of logical deduction, especially as it has been developed by Frege, Russell, and Hilbert, is rather far removed from the forms of deduction used in practice in mathematical proofs. Considerable formal advantages are achieved in return.
>
> In contrast, I intended first to set up a formal system which comes as close as possible to actual reasoning. The result was a 'calculus of natural deduction' ('NJ' for intuitionist, 'NK' for classical predicate logic). ...

Like Jaśkowski, Gentzen sees the notion of making an assumption to be the leading idea of his natural deduction systems:

> ... the essential difference between NJ-derivations and derivations in the systems of Russell, Hilbert, and Heyting is the following: In the latter systems true formulae are derived from a sequence of 'basic logical formulae' by means of a few forms of inference. Natural deduction, however, does not, in general, start from basic logical propositions, but rather from assumptions to which logical deductions are applied. By means of a later inference the result is then again made independent of the assumption.

These two founding fathers of natural deduction were faced with the question of how this method of "making an arbitrary assumption and seeing where it leads" could be represented. As remarked above, Jaśkowski gave two methods; Gentzen also contributed a method. There is one further method that was introduced some 20 years later in Suppes (1957). All of the representational methods used in today's natural deduction systems are variants on one of these four.

The differences and similarities among these methods of representing natural deduction proofs are easiest to see in an example of a proof done in the different methods, with the proof's ebb and flow of assumptions and cancellation of assumptions. Of course, our writers had different rules of inference in mind as they described their systems, but we need not pause over that at the moment. We will use names in common use for these rules. (And also, since Jaśkowski did not have $\wedge$ in his language, we are employing a certain latitude. But it is clear how his systems would use it.) We consider the way a proof of the theorem $(((p \supset q) \wedge (\neg r \supset \neg q)) \supset (p \supset r))$ is represented in the four different formats.

Since the main connective is a conditional, the most likely strategy will be to prove it by a rule of conditional introduction. But to apply this rule one must have a subproof that assumes the conditional's antecedent and ends with the conditional's consequent. All the methods will follow this strategy; the differences among them concern only how to represent the strategy.

In Jaśkowski's graphical method, each time an assumption is made it starts a new portion of the proof which is to be enclosed with a rectangle (a "subproof"). The first line of this subproof is the assumption... in the case of trying to apply conditional introduction, the assumption will be the antecedent of the conditional to be proved and the remainder of this subproof will be an attempt to generate the consequent of that conditional. If this can be done, then Jaśkowski's rule conditionalization says that the conditional can be asserted as proved *in the subproof level of the box that surrounds the one just completed.* So the present proof will assume the antecedent, $((p \supset q) \wedge (\neg r \supset \neg q))$, thereby starting a subproof trying to generate the consequent, $(p \supset r)$. But this consequent itself has a conditional as main connective, and so it too should be proved by conditionalization with a yet-further-embedded subproof that assumes its antecedent, p, and tries to generate its consequent, r. As it turns out, this subproof calls for a yet further embedded subproof using Jaśkowski's reductio ad absurdum. Here is how this proof would be represented in his graphical method.

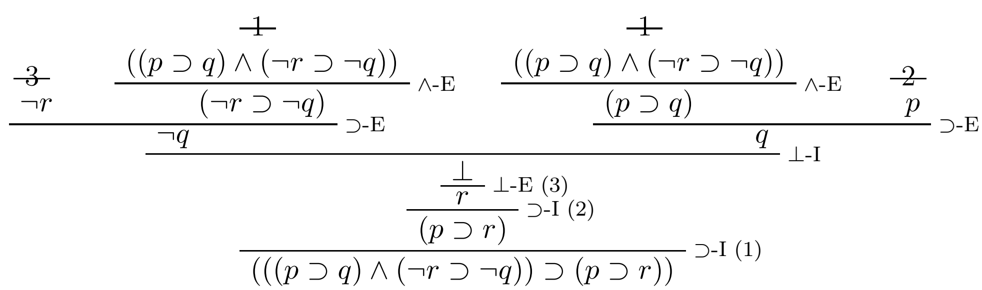
1.	$((P \supset Q)\&(\sim R \supset\sim Q))$	Supposition
2.	P	Supposition
3.	$((P \supset Q)\&(\sim R \supset\sim Q))$	1, Repeat
4.	$(P \supset Q)$	3, Simplification
5.	Q	2,4 Modus Ponens
6.	$(\sim R \supset\sim Q)$	3, Simplification
7.	$\sim R$	Supposition
8.	$(\sim R \supset\sim Q)$	6, Repeat
9.	$\sim Q$	7,8 Modus Ponens
10.	Q	5, Repeat
11.	R	7-10 Reductio ad Absurdum
12.	$P \supset R$	2-11 Conditionalization
13.	$(((P \supset Q)\&(\sim R \supset\sim Q)) \supset (P \supset R))$	1-12 Conditionalization

Jaśkowski's second method (which he had hit upon later than the graphical method, and was the main method of Jaśkowski, 1934) was to make a numerical annotation on the left-side of the formulas in a proof. Again, this is best seen by example; and so we re-present the previous proof. In this new method, Jaśkowski changed the statements of various of the rules and he gave them new names: Rule I is now the name for making a supposition, Rule II is the name for conditionalization, Rule III is the name for modus ponens, and Rule IV is the name for reductio ad absurdum. (Rules V, VI, and VII have to do with quantifier elimination and introduction). Some of the details of these changes to the rules are such that it is no longer required that all the preconditions for the applicability of a rule of inference must be in the same "scope level" (in the new method this means being in the same depth of numerical annotation), and hence there is no longer any requirement for a rule of Repetition. To indicate that a formula is a supposition, Jaśkowski now prefixes it with 'S'.

1.1.	S $((p \supset q) \wedge (\neg r \supset \neg q))$	I
2.1.	$(p \supset q)$	&E 1
3.1.	$(\neg r \supset \neg q)$	&E 1
4.1.1.	S p	I
5.1.1.	q	III 4,2
6.1.1.1.	S $\neg r$	I
7.1.1.1.	$\neg q$	III 6,3
8.1.1.	r	IV 5,7,6
9.1.	$(p \supset r)$	II 4,8
10.	$(((p \supset q) \wedge (\neg r \supset \neg q)) \supset (p \supset r))$	II 1,9

It can be seen that this second method is very closely related to the method of rectangles. (And much easier to typeset!) Its only real drawback concerns whether we can distinguish different subproofs which are at the same level of embedding. A confusion can arise when one subproof is completed and then another started, both at the same level of embedding. In the graphical method there will be a closing of one rectangle and the beginning of another; but here it could get confused. Jaśkowski's solution is to mark the second such subproof as having '2' as its rightmost numerical prefix. This makes numerals be superior to using other symbols in this role, such as an asterisk. As we will see in §2.4, this representational method was adopted by Quine (1950a), who used asterisks rather than numerals thus leading to the shortcoming just noted.

A third method was given by Gentzen. Proofs in the N calculi (the natural deduction calculi) are given in a tree format with formulas appearing as nodes of the tree. The root of the tree is the formula to be proved, and the "suppositions" are at the leaves of the tree. The following is a tree corresponding to the example we have been looking at, although it should be mentioned that Gentzen's main rule for indirect proofs first generated $\perp$ ("the absurd proposition") from the two parts of a contradiction, and then generated the negation of the relevant assumption.[10]



[10]He also considered a double negation rule.

The lines indicate a transition from the upper formula(s) to the one just beneath the line, using the rule of inference indicated on the right edge of the line. (We might replace these horizontal lines with vertical or splitting lines to more clearly indicate tree-branches, and label these branches with the rule of inference responsible, and the result would look even more tree-like). Gentzen uses the numerals on the leaves as a way to keep track of subproofs. Here the main antecedent of the conditional to be proved is entered (twice, since there are two separate things to do with it) with the numeral '1', the antecedent of the consequent of the main theorem is entered with numeral '2', and the formula $\neg r$ (to be used in the reductio part of the proof) is entered with numeral '3'. When the relevant "scope changing" rule is applied (indicated by citing the numeral of that branch as part of the citation of the rule of inference, in parentheses) this numeral gets "crossed out", indicating that this subproof is finished.

One reason that Jaśkowski's (and Quine's) bookkeeping method did not become more common is that Suppes (1957) introduced a method using the line numbers of the assumptions which any given line in the proof depended upon, rather than asterisks or arbitrary numerals. The method retained the ease of typesetting that the bookkeeping method enjoyed over the graphical and tree representations, but was much clearer in its view of how new subproofs were started. In this fourth method, when an assumption is made its line number is put in set braces to the left of the line (its "dependency set"). The application of "ordinary rules" such as &E and Modus Ponens make the resulting formula inherit the union of the dependencies of the lines to which they are applied, whereas the "scope changing" rules like $\supset$I and Reductio delete the relevant assumption's line number from the dependencies. In this way, the "scope" of an assumption is not the continuous sequence of lines that occurs until the assumption is discharged by a $\supset$I or $\neg$I rule, but rather consists of just those (possibly non-contiguous) lines that "depend upon" the assumption. But the fact that the lines in a given subproof are no longer necessarily contiguous marks a break from the idea that we are "making an assumption and seeing where it leads" — or at least, one might argue this idea has been lost. In fact, if one views the numbers in a dependency set as just an abbreviation for the actual formula that occurs on the line with that number, and the line number itself is just to be replaced with '$\vdash$', it can appear that the system is actually representing a sequent calculus proof (except for the fact that there is always just one formula on the right side). This is explored in detail in §3.3.

Without using Suppes's specific rules, we can get the flavor of this style of representation by presenting the above theorem as proved in a Suppes-like manner.

{1}	1.	$((p \supset q) \wedge (\neg r \supset \neg q))$	
{1}	2.	$(p \supset q)$	&E 1
{1}	3.	$(\neg r \supset \neg q)$	&E 1
{4}	4.	p	
{1,4}	5.	q	⊃E 4,2
{6}	6.	$\neg r$	
{1,6}	7.	$\neg q$	⊃E 6,3
{1,4}	8.	r	Reductio 5,7,6
{1}	9.	$(p \supset r)$	⊃I 4,8
∅	10.	$(((p \supset q) \wedge (\neg r \supset \neg q)) \supset (p \supset r))$	⊃I 1,9

These four methods of representing natural deduction proofs continue to characterize natural deduction, although it should be remarked that neither the bookkeeping method nor the Gentzen method became very popular. (Quine, 1950a used the bookkeeping method, and because of Quine's stature, some few other textbooks retained it, but this became very rare. Gentzen's method finds its place in more technical discussions of natural deduction, and in a few intermediate-level textbooks. It is quite rare in beginning textbooks, although it is used van Dalen (1980); Tennant (1978) if these are seen as elementary.)

We turn now to the way natural deduction became the dominant logical method that was taught to generations of (mostly philosophy) undergraduates (in North America).

2.4 Natural Deduction in Elementary Textbooks

In §2.3 we looked at the four ways that have been used in proofs to visually display the notion of "making an assumption and seeing where it leads". One reason elementary students find the generic notion of natural deduction confusing is because they find it difficult to understand how systems that have such diverse representational methods can really be "the same type of logical system".

These four ways have, each to a greater or lesser extent, been adopted by the elementary logic textbooks that started to appear in the early 1950s. Quine (1950a) was the first of these textbooks, and as we have said, it used the Jaśkowski bookkeeping method. As we also said, this way of representing natural deduction proofs was not very popular in the textbook genre, being followed by only four further textbooks in the following 60 years.[11] Gentzen's style of proof was also not very popular in the textbooks, finding its way into only five of the 50 textbooks, and as mentioned above, these are mostly claimed to be "intermediate logic books" or "intended for graduate students in philosophy".

[11] This number, as well as the ones to come, are derived from (Pelletier, 2000, p. 135), and augmented with another 17 textbooks. As is remarked there, the figures are not "scientific" in any clear sense, and represent only the 33 (at that time, and 50 now) elementary logic textbooks using natural deduction that Pelletier "had on his bookshelf". Nonetheless, they certainly seem to give the correct flavor of the way that elementary logic textbooks dealt with natural deduction in the 1950–2010 timeframe under consideration. More details of the contents of these books is given in Table I below.

Despite the fact that Jaśkowski apparently preferred his bookkeeping to his earlier graphical method, it was the graphical method that caught on. The second elementary logic textbook to use natural deduction was Fitch (1952), and the popularity of his version of the graphical method has made this method be called "the Fitch method", despite the fact that it is really Jaśkowski's method. (Fitch [p. vii] says merely "The method of subordinate proofs was suggested by techniques due to Gentzen and Jaśkowski.") The main difference between Jaśkowski's original method and Fitch's is that Fitch does not completely draw the whole rectangle around the embedded subproof (but only the left side of the rectangle), and he underlines the assumption. The same proof displayed above using Jaśkowski's graphical method is done like the following in Fitch's representation (with a little laxness on identifying the exact rules Fitch employs).

1	$\mid\ \underline{((p \supset q) \wedge (\neg r \supset \neg q))}$	
2	$\mid\ \mid\ \underline{p}$	
3	$\mid\ \mid\ ((p \supset q) \wedge (\neg r \supset \neg q))$	1, R
4	$\mid\ \mid\ (p \supset q)$	3,$\wedge$E
5	$\mid\ \mid\ q$	2,4 $\supset$E
6	$\mid\ \mid\ (\neg r \supset \neg q))$	3,$\wedge$E
7	$\mid\ \mid\ \mid\ \underline{\neg r}$	
8	$\mid\ \mid\ \mid\ (\neg r \supset \neg q))$	6,R
9	$\mid\ \mid\ \mid\ \neg q$	7,8 $\supset$E
10	$\mid\ \mid\ \mid\ q$	5,R
11	$\mid\ \mid\ r$	7–10, $\neg$E
12	$\mid\ (p \supset r)$	2–11, $\supset$I
13	$(((p \supset q) \wedge (\neg r \supset \neg q)))$	1–12,$\supset$I

The graphical method, in one form or another, is in use in very many elementary textbooks. Some authors do not make an explicit indication of the assumption, other than that it is the first formula in the scope level; and some authors do not even use the vertical lines, but rather merely employ indentation (e.g., Hurley, 1982). Other authors, e.g., Kalish and Montague (1964); Kalish et al. (1980); Bonevac (1987) keep the rectangles, but put the conclusion at the beginning (both as a way to guide what a legitimate assumption is and also to help in the statement of restrictions on existential elimination and universal introduction rules). And there are authors who retain still different parts of the rectangle, such as Copi (1954), who keeps the bottom and left side parts, and ends the top part with an arrowhead pointing at the assumption. (Similar variations are in Gamut, 1991; Harrison, 1992; Bessie and Glennan, 2000; Arthur, 2010). But clearly, all these deviations from Jaśkowski's original boxes are minor, and all these should be seen as embracing the graphical method.

Table I below reports that 60% of the elementary logic textbooks that were surveyed used some form of this graphical method, making it by far the most common representation of natural deduction proofs. The bookkeeping and Gentzen-tree methods accounted for 10% each, and the remaining 20% employed the Suppes method. (Some might think this latter percentage must be higher, but this perception is probably due to the immense popularity of Lemmon, 1965 and Mates, 1965, which both used the Suppes-method.)

2.5 More Features of the Prototype of Natural Deduction

In §2.1 we looked at two aspects of systems of natural deduction that are in the minds of many researchers. . . especially those from computer science and mathematics. . . when they think about natural deduction. Although as we said there, these do not define natural deduction, it is still the case that they form a part of the prototypical natural deduction systems. In this section we look at some more features of this sort, but where these are more closely tied to the remarks that Gentzen and Jaśkowski made in their initial discussions.

A third feature of natural deduction systems, at least in the minds of some, is that they will have two rules for each connective: an introduction rule and an elimination rule. But again this can't be necessary, because there are many systems we happily call natural deduction which do not have rules organized in this manner. The data in Table I below report that only 23 of the 50 texts surveyed even pretended to organize their natural deduction rules as matching int-elim rules (and some of these 23 acknowledged that they were "not quite pure" int-elim). Furthermore, since they are textbooks for classical (as opposed to intuitionistic) logic, these systems all have to have extra features that lead to classical logic, such as some "axioms" (see below). And anyway, even if we concocted an axiomatic system that did have rules of this nature, this would not make such a system become a natural deduction system. So it is not sufficient either.

A fourth idea in the minds of many, especially when they consider the difference between natural deduction and axiomatic systems, is that natural deduction does not have axioms. (Recall Jaśkowski's remark that his system has no requirement of axioms.) But despite the fact that Jaśkowski found no need for axioms, Gentzen *did* have them in his NK, the natural deduction system for classical logic; it was only his NJ, the intuitionistic logic, that did not have them. And many of the authors of more modern textbooks endorse methods that are difficult to distinguish from having axioms. For example, as a primitive rule many authors have a set of 'tautologies' that can be entered anywhere in a proof. This is surely the same as having axioms. Other authors have such a set of tautological implications together with a rule that allows a line in a proof to be replaced by a formula which it implies according to a member of this set of implications. So, if φ is a line in a proof, and one of the tautological implications in the antecedently-given list is $\varphi \rightarrow \psi$, then ψ can be entered as a line (with whatever dependencies the φ line had). And in the world of "having axioms" it is but a short step from here to add to the primitive formulation of the system a set of 'equivalences' and the algebraically-inspired rule that one side of the equivalence that can be substituted for a *subpart* of an existing line that matches the other side of the equivalence, as many authors do. A highly generalized form of this method is adopted by Quine (1950a), Suppes (1957), and others, which have a rule TF ("truth functional inference") that allows one to infer "any schema which is truth-functionally implied by the given line(s)".[12] Although one can detect certain differences amongst all these variants just mentioned here, they seem all to be ways of adopting axioms.[13] Table I lists 22 of the 50 textbooks (44%) as allowing either a TF inference or else tautologies or replacement from a list of equivalences in the primitive portion of their system. . . in effect, containing axioms.

A fifth idea held by some is that "real" natural deduction will have introduction and elimination rules for every connective, and there will be no further rules. But even Gentzen, to whom this int-elim ideal is usually traced, thought that he had a natural deduction system for classical logic. And as we remarked above, a difference between the natural deduction system for intuitionistic logic and that for classical logic was that the latter had every instance of $\varphi \vee \neg\varphi$ as axioms. So, not even Gentzen followed this dictum.[14] We can get around the requirement of axioms for classical logic if we have a bit more laxity in what counts as a "pure" introduction/elimination rule—for example, allowing double negation rules to be int-elim. But even with this laxity, the number of elementary textbooks that even try to be int-elim is nowhere near universal: Table I shows less than half to be of this nature, and many of them have *very* lax interpretations of what an int-elim pair is. Other textbooks happily pair modus ponens with modus tollens, pair biconditional modus ponens with two-conditionals-gives-a-biconditional, pair modus tollendo ponens (= unit resolution, disjunctive syllogism) and separation of cases with *or*-introduction, and so forth. In fact, one elegant version of natural deduction has two *pairs* of rules per connective. Fitch (1952) supplements standard rules of, e.g., Conjunction Introduction and Elimination with rules of Negative Conjunction Introduction and Elimination embodying the principle that a negated conjunction is equivalent to a disjunction. In Fitch's book, these are stated as rules making the negated conjunction inter-inferable with the equiva-

[12]Their TF rules allow one to infer anything that follows from the conjunction of lines already in the proof.

[13]One might separate systems with the TF rule from these other "axiomatic" systems in that the former do not have any list of tautological implications to employ, and instead this is formulated as a rule. Note that, in the propositional logic, there would be no real need for any other rules, including the rule of conditionalization. For, everything can be proved by the rule TF. (Any propositional theorem follows from the null set of formulas by the rule TF).

[14]However, many theorists believe that this is a reason to think that intuitionistic logic is the "real" logic and that classical logic is somehow defective. Perhaps Gentzen thought so too, many people claim. See §§4.3–4.4 below for some thoughts on this topic.

lent disjunction: Negative Conjunction Introduction licenses the inference from $\neg\phi \vee \neg\psi$ to $\neg(\phi \wedge \psi)$, Negative Conjunction Elimination licenses the converse inference. It may be that this formulation is easier for beginning students to memorize, but, at least for theoretical purposes, the explicit disjunctive formula is redundant: a more streamlined version of the system would make these rules formally parallel to the disjunction rules, so Negative Conjunction Introduction would license the inference of $\neg(\phi \wedge \psi)$ from either $\neg\phi$ or $\neg\psi$, and Negative Conjunction Elimination would allow a conclusion θ to be inferred from three items, a negated conjunction $\neg(\phi\wedge\psi)$, a subproof deriving θ from the hypothesis $\neg\phi$, and a subproof deriving θ from the hypothesis $\neg\psi$. Similarly, rules of Negative Disjunction Introduction and Negative Disjunction Elimination would parallel the standard ("positive") rules for conjunction; one could (though Fitch, due to the peculiarities of the type-free theory of classes his book presents, does not) add conjunction-like rules of Negative Implication Introduction and Negative Implication Elimination. Since there are dualities in quantificational as well as propositional logic, we can have rules of Negative Universal Quantifier Introduction and Elimination (parallel to the positive rules for the Existential Quantifier) and Negative Existential Quantifier Introduction and Elimination (parallel to the positive rules for the Universal Quantifier). All these negative rules are, of course, classically valid, and would be derivable rules in any complete natural deduction system for classical logic. In application, though, they are useful enough in shortening classical derivations to be worth making explicit. In a natural deduction formulation of any of a number of three- and four-valued logics (the Strong 3-valued logic of Kleene, 1952, the Logic of Paradox of Priest, 1979, the logic of First Degree Entailment of Anderson and Belnap, 1975), or of Intuitionistic Logic with the constructible negation (also called strong negation) introduced in Nelson (1949), it would be appropriate to take both positive and negative rules as primitive. We will refer to them below in describing a number of systems related to natural deduction.

A sixth idea, perhaps the one that most will gravitate to when the other ideas are shown not to distinguish natural deduction from other frameworks, is that the concept of a subproof is unique to natural deduction. Of course, one needs to take some care here: an axiomatic or resolution (etc.) proof could have a contiguous portion that can be seen as a subproof. . . in fact, it is pretty clear that one can find pairs of axiomatic and natural deduction proofs such that, if one were to lay the proofs side-by-side, the subproof portions of the natural deduction proof would correspond to contiguous sections of the axiomatic proof. Still, there is a feeling that the concept of "make an assumption and see where it leads, and discharge the assumption at a latter point in the proof" is a distinguishing characteristic of natural deduction, for, although *some* axiomatic/resolution/etc. proofs might have this property, it might be said that *all* natural deduction proofs do.

Actually, though, *not* all natural deduction proofs have this property. First, not all particular natural deduction proofs even make assumptions[15], and second, the Suppes rep-

[15] At least, if we don't think of the premises of an argument as assumptions. Some books treat them as assumptions—e.g., those using the Fitch style of proof, where they are all listed at the beginning of the outermost scope line; but others don't, for example, those books that simply allow a premise to be entered

resentational method does *not* follow the dictum of "make an assumption and see where it leads", because the contiguous lines of a proof need not have the same dependency set of assumptions. So it is not quite clear that this can be used as a defining characteristic, as opposed to being merely a typical characteristic of natural deduction. (There is also the issue that some textbooks that self-describe as natural deduction do not have *any* mechanism for making assumptions, such as Barker (1965); Kilgore (1968); Robison (1969).[16] Gustason and Ulrich (1973) have a system where their only subproof-requiring rule, $\supset I$, 'is dispensable' (they have $(\varphi \vee \neg\varphi)$ as an axiom scheme).[17]

Jaśkowski and Gentzen both had (what we now call) $\supset I$, and all of the textbooks in Table I have such a rule (even if it is said to be "dispensable"). Most natural deduction textbooks have further subproof-requiring rules (Table I lists 82% of the books as having some subproof-requiring rule besides $\supset I$). Gentzen for example used a subproof for his $\vee E$ rule, as do some 44% of our textbook authors. This latter was not necessary, since in the presence of $\supset I$, the $\vee E$ rule could simply have been Separation of Cases:

$$\frac{\phi \vee \psi \qquad \phi \supset \theta \qquad \psi \supset \theta}{\theta}\ \text{(SC)}$$

which it is in about 42% of the elementary textbooks surveyed in Table I; these books usually also have Disjunctive Syllogism (= unit resolution or *modus tollendo ponens*: the rule from $\phi \vee \psi, \neg\phi$ to infer ψ) together with SC. The remaining either had Disjunctive Syllogism alone or else used equivalences/tautologies for their reasoning concerning how to eliminate disjunctions.

We have seen above Gentzen's rules for negation, which used $\bot$. This subproof-using method is still in use in some six textbooks, but most elementary textbooks (64%) have this instead as their subproof-using $\neg I$ rule:

$$\begin{array}{c} [\phi] \\ \vdots \\ \psi \\ \vdots \\ \dfrac{\neg\psi}{\neg\phi}\ (\neg I) \end{array}$$

(or some version where $(\psi \wedge \neg\psi)$ is on one line). And 24% of the texts have a $\neg E$ rule of the same sort, except that the assumed formula is a negation to be eliminated. 20% have an $\equiv I$-rule that requires subproofs, the rest just directly infer $(\varphi \equiv \psi)$ from $(\varphi \supset \psi)$ and

anywhere in a proof, as in Kalish and Montague (1964).

[16] We chose not to count Barker (1965); Kilgore (1968); Robison (1969) as natural deduction despite those authors' claims.

[17] We chose to count Gustason and Ulrich (1973) as natural deduction.

$(\psi \supset \varphi)$. Two books have other subproof-requiring rules: Bostock (1999); Goodstein (1957) have, respectively,

$$\frac{\begin{array}{cc}[\phi] & [\neg\phi] \\ \vdots & \vdots \\ \psi & \psi\end{array}}{\psi} \qquad \frac{\begin{array}{c}[\phi] \\ \vdots \\ \psi\end{array}}{\neg\psi \supset \neg\phi}$$

There are further important differences among our textbook authors, having to do with the ways they handle quantification. But even at the level of propositional logic, we can see that there are no clear and unequivocal criteria according to which we can always label a logical system as natural deduction. There is instead a set of characteristics that a system can have to a greater or lesser degree, and the more of them that a system has, the more happily it wears the title of 'natural deduction'.

2.6 Natural Deduction Quantificational Rules

The story of elementary natural deduction and textbooks is not complete without a discussion of the difficulties involved with the existential quantifier (and related issues with the universal quantifier) in natural deduction systems. If a smart student who has not yet looked at the quantification chapter of his or her textbook were asked what s/he thinks ought to be the elim-rule for existential sentences like $\exists xFx$, s/he will doubtless answer that it should be eliminated in favor of Fa — "so long as a is arbitrary". That is, the rule should be

$$\frac{\exists xFx}{Fa}\ \text{(EI); } a \text{ is arbitrary}$$

This is the rule Existential Instantiation (by analogy with Universal Instantiation, but here with the condition of arbitrariness). Since it was used by Quine (1950a), sometimes the resulting systems are called Quine-systems. (Quine, 1950b, fn.3) says that Gentzen "had a more devious rule in place of EI". This "more devious" rule, which we call $\exists E$, is

$$\frac{\exists xFx \qquad \begin{array}{c}[Fa] \\ \vdots \\ \phi\end{array}}{\phi}\ (\exists E)\text{; with restrictions on } a, \phi, \text{ and all "active assumptions"}$$

Although it may be devious, the Gentzen $\exists E$ has the property that if all the assumptions on all branches that dominate a given formula are true, then that formula must be true also; and this certainly makes a soundness proof be straightforward. (In the linear forms—the bookkeeping or graphical methods—this becomes: if all the active assumptions above a given formula are true then so must that formula be. In the Suppes-style form this becomes: if all the formulas mentioned in the dependency set are true, then so is the given formula.) The reason for this is that the "arbitrary instance" of the existentially quantified formula becomes yet another assumption, and the restrictions on this instance ensure that it cannot occur outside the subproof. The same cannot be said about the Quine EI rule where the arbitrary instance is in the same subproof as the existentially quantified formula, and Quine goes to prodigious efforts to have the result come out right: in his earliest version [first edition], he introduces total ordering on the variables and decrees that multiple EIs must use this ordering[18], and he introduces the notion of an "unfinished proof"—where all the rules are applied correctly and yet the proof still might not be valid—in which some postprocessing needs to take place to determine validity.

Quine's method, complex though it was, did not torture students and professors as much as the third published natural deduction textbook, Copi (1954), which also had a rule of EI. Quine's restrictions actually separated valid arguments from invalid ones, but Copi's allowed invalid arguments to be generated by the proof theory. There were many articles written about this, and the proper way to fix it, but it remained an enormously popular textbook that was the most commonly-used textbook of the 1950s and 1960s, despite its unsoundness. The conditions on the rules changed even within the different first edition printings and again in the second edition, but the difficulties did not get resolved until the 1965 third edition, where he adopted the Gentzen $\exists E$ rule.[19] One might think that with all the difficulties surrounding EI—particularly the specification of the restrictions on the instantiated variable and the difficulties that arise from having an inferred line that is not guaranteed to be true even if its premise is true—almost all textbooks would employ $\exists E$. But Table I shows that, of the 49 books that had a rule for eliminating existential quantifiers[20], the percentage was only 61% using $\exists E$: 30 used $\exists E$ while 19 used EI.

Related to the choice of having a subproof-introducing $\exists E$ rule vs. non-subproof EI rule, is the choice of requiring a subproof for universal quantifier introduction. Table I lists only eleven of 49 texts[21] as having a subproof-requiring $\forall I$; the other 38 had generalization on the same proof-level (UG: universal generalization), with further restrictions on the variable or name being generalized, so as to ensure "arbitrariness". Of the eleven that required a

[18]The second edition changed this to "there is a total ordering on the variables that are used in the proof". See the discussion surrounding (Pelletier, 1999, p.16–17fn18,20). Other interesting aspects of this can be gathered from Cellucci (1995) and Anellis (1991).

[19]However, the "proof" later in the book of the soundness of the system was not changed to accommodate this new rule. It continued to have the incorrect proof that wrongly dealt with EI. It was not until the 1972 fourth edition that this was corrected—some 18 years of irritated professors and confused students, who nonetheless continued using the book! The circumstances surrounding all this are amusingly related in Anellis (1991). See also (Pelletier, 1999, §5) for a documentation of the various changes in the different printings and editions.

[20]One of the 50 textbooks did not have any rule for eliminating existential quantifiers, other than by appeal to "axioms".

[21]One of the 50 books did not have any generalization rules at all, and thus no universal introduction.

$\forall I$ subproof, six had $\exists E$ subproofs and four had EI non-subproofs. (The remaining one had no existential elimination at all.) So, there seems to be no particular tie between $\exists E$ subproofs and $\forall I$ subproofs… nor between EI and $\forall I$ subproofs, nor between any other combination, in elementary logic textbooks.

One further note should be added to the discussion of the quantifier rules in natural deduction, and that concerns the description of EI, $\exists E$, and $\forall I$ as requiring "arbitrary objects" for their values. In most textbooks, at least when the proof techniques are being taught, this is merely some sort of *façon de parler*. It plays no actual role in proofs; it is just a way of emphasizing to students that there are conditions on the choice of singular term that is being used in these rules. But sometimes in these textbooks, in the chapters that have some metatheoretical proofs, the notion of an arbitrary object is also appealed to. Here one might find difficulties, since some student may ask embarrassing questions concerning the nature of these objects ("Is this arbitrary person a male? A female? Or neither, since it wouldn't be arbitrary otherwise? But aren't all people either male or female?" Or perhaps: "Our domain doesn't contain this object… it only has this-and-that sort of thing as members… There are no abstract objects in it." Or perhaps from a more sophisticated student: "Objects aren't arbitrary! It is the way that a specific object is *chosen* that is arbitrary!"[22]

Some of our textbook authors were advocates of "the substitution interpretation" of quantification, and so they naturally thought that the singular term that was used in finding "the arbitrary instance" was actually a proper name of some object in the domain. For them, the arbitrariness came into play in the way this name was chosen—it had to be "foreign" to the proof as thus far constructed. And they then had a proof in their metatheory chapter to showed that this restriction was sufficient: if the conclusion could be drawn from such a substitution instance then it could be drawn from the existentially quantified formula, because these formulas were true (in their substitutional interpretation) just in case some instance with a name being substituted in the formula made it true. The "objectual interpretation" version of quantification couldn't make use of this ploy, and they naturally gravitated toward using free variables in their statements of $\forall I$. A formula with a free variable, occurring in a proof, is semantically interpreted as a universally-quantified formula—unless the free variable is due to an EI rule or is introduced in an assumption or

[22]Echoing Frege (1904):

> But are there indefinite numbers? Are numbers to be divided into definite and indefinite numbers? Are there indefinite men? Must not every object be definite? But is not the number n indefinite? I do not know the number n. 'n' is not the proper name of some number, definite or indefinite. And yet one sometimes says 'the number n'. How is that possible? Such an expression must be considered in context. … One writes the letter 'n', in order to achieve generality….
>
> Naturally one may speak of indefinite here; but here the word 'indefinite' is not an adjective of 'number', but an adverb of, for example, to 'signify'.

is "dependent on" some EI-introduced variable. There are (to judge from the textbooks) a myriad of ways of saying this that have the effect that only the variables that are in fact able to be interpreted as universal are allowed to show up in the precondition to the rule. However, the informal talk in their metatheoretical discussions continued, in many of these authors, to make use of the notion of 'arbitrary object', ignoring the philosophical objections.[23]

But there is one more group: those who use a special vocabulary category of *parameter*. These are not names, but they are not variables either. In fact, the metatheory in the textbooks that use this tends to ignore parameters altogether. In some books the rules are that existential quantifiers can never introduce them, but universal instantiation can, since a universal can instantiate to *anything*; and a legitimate $\forall I$ will require that the item being generalized be one of these parameters. Other books take it that the existential instantiation or elimination rule will make use of these parameters and that the universal generalization rule cannot generalize on such a term. It can be seen that there are two notions here of "arbitrary object" being employed by these two different versions—the former notion of an arbitrary object is "any object at all, it doesn't matter which", while the latter is "a particular object, but you can't know anything about which one". It is not obvious that either conception of 'arbitrary object' makes much sense, except as a *façon de parler*.

2.7 A Summary of Elementary Natural Deduction Textbooks

In this section we will compare how some 50 elementary logic textbooks deal with natural deduction proofs. Every textbook author surveyed here (and reported in Table I) describes their book as being an instance of natural deduction, and so to some extent contributes to the image that philosophers have of what "real" natural deduction is. Many textbooks have gone through many editions over the years, and sometimes this has resulted in differences to the sort of features we are considering here. Whenever we could find the first edition of the book, we used the statements of the rules in that edition.

Here is some background that is relevant to understanding Table I. First, column headings are the names of the first author of textbook, and the textbooks can be found in the special bibliographic listing of textbooks after the regular bibliography, below. Second, the features described belong to the *primitive basis* of the textbook. Most textbooks, no matter what their primitive basis, provide a list of "equivalences" or "derived rules." However, these features are not considered in the Table (unless indicated otherwise). We are instead interested in the basis of the natural deduction system.

The Table identifies eight different general areas where the textbooks might differ. And within some of these areas, there are sub-points along which the differences might manifest. The explanation of the eight general areas is the following.

[23] But Fine (1985) set about to make this notion have a more philosophically satisfactory interpretation. See Hazen (1987) for further discussion.

I. Proof Style This area describes which of the four different styles of representation—which we discussed above in §2.3—is employed by the textbook.

II. "Axiomatic" Describes whether, in the primitive basis of the system, there are unproved formulas that are taken for granted and allowed to be entered or used in the proof without further justification. This is subdivided into (a) having a list of tautologies that are allowed to be entered anywhere in a proof, for instance allowing the formulas $\neg\phi \supset (\phi \supset \psi)$ or $\neg\neg\phi \supset \phi$ to be entered anywhere in a proof; (b) having a list of equivalences that justify the replacement of a subformula of some formula in the proof with what it is equivalent to according to the list, for example, having $\neg(\phi \wedge \psi) \equiv (\neg\phi \vee \neg\psi)$ on the list and allowing formulas in a proof that contain as a subformula (a substitution instance of) the left side to be inferred by the same formula but where the subformula is replaced by (the same substitution instance of) the right side, as in algebraic systems of logic; (c) having a general rule of "truth functionally follows", TF, that allows the inference of a new line in a proof if it follows truth functionally from some or all the antecedent lines of the proof.

III. All Int-Elim This category is supposed to pick out those systems which self-consciously attempt to make their rules of inference come in pairs of 'introduce a connective' and 'eliminate that connective'. We have allowed quite a bit of laxity in asserting that the rules are int-elim in nature, mostly following whether the authors *think* they are giving int-elim rules. As discussed above in §2.5, some authors call rule-sets 'int-elim' which would not be thus considered by other logicians.

IV. Sub-Proofs This tracks the propositional rules of inference which are required to have a subproof as part of their preconditions. Again, keep in mind that this is for the *primitive* portion of the systems: many authors introduce derived rules that might (or not) require subproofs, but this information is not recorded in Table I.

V. One Quantifier This answers the question of whether the system has only one quantifier in its primitive basis (and introduces the other by means of some definition).

VI. Arbitrary Names The question being answered is whether the system makes use of an orthographically distinct set of terms ("parameters" or whatever) being employed in one or both of Universal Quantifier Introduction and Existential Quantifier Elimination. As we said above, most authors use for this purpose either the "names" or the "free variables" that are employed elsewhere in the proof system; but some authors have a distinct set of symbols for this purpose. Surprisingly, it can be difficult to discern whether a system does or doesn't employ orthographically different items for this—witness (Teller, 1989, p.75), who uses letters with circumflex accents for this purpose: "A name with a hat on it is not a new kind of name. A name is a name is a name, and two occurrences of the same name, one with and one without a hat, are two occurrences of the same name. A hat on a name is a kind of flag to remind us that at that point the name is occurring arbitrarily." We have chosen to classify this as using an orthographically different set of terms for use in $\forall I$ and $\exists E$, despite the informal interpretation that Teller puts on it.

VII. Existential Rule The issue here is whether the system employs the Gentzen-like Existential Elimination $\exists E$ rule that requires a subproof, or uses the Quine-like Existential

Instantiation (EI) rule which stays at the same proof level.

VIII. Universal Rule Here we track whether the Universal Introduction rule $\forall I$ requires a subproof or not.

Footnotes to Table I

(a) Reporting "the derived system". The main system is axiomatic.

(b) Hurley uses indentation only, without any other "scope marking".

(c) Meaning "allows at least one type of 'axiom'."

(d) Although equivalences are used throughout the book, apparently the intent is that they are derived, even though there is no place where they were derived. Authors claim that the system is int-elim.

(e) We allow as int-elim for a rule of Repetition, the "impure" DN rules, the "impure" negation-of-a-binary connective rules, and others—including various ways to convert intuitionistic systems to classical ones.

(f) Tennant considers a number of alternatives for converting the intuitionistic system to classical, including axioms, rules such as $[\phi]\cdots\psi, [\neg\phi]\cdots\psi \vdash \psi$, etc.

(g) Has $\phi, \neg\phi \vdash \psi$ as well as $[\phi]\cdots\psi, [\neg\phi]\cdots\psi \vdash \psi$, so the rules are not quite int-elim, although we are counting them so.

(h) Also has "impure" $\neg\wedge$ and $\neg\forall$ rules (etc.).

(i) Not strictly int-elim, because the two $\bot$ rules are both elim-rules. ($\neg$ is defined in terms of $\supset$ and $\bot$.)

(j) Meaning "has some subproof-requiring propositional rule besides $\supset I$

(k) Says the $\supset I$ rule is theoretically dispensable.

(l) Has $[\phi]\cdots\psi \vdash (\neg\psi \supset \neg\phi)$.

(m) A basic tautology is $\forall x(Fx \supset P) \supset (\exists xFx \supset P)$, if P does not contain x. Rather than an $\exists$-rule, if $\exists xFx$ is in the proof, Pollock recommends proving the antecedent of this conditional by $\forall I$, then appeal to this tautology to infer the consequent by $\supset E$, and then use $\supset E$ again with the $\exists xFx$ and thereby infer P. We count this as neither $\exists E$ nor as EI; hence totals are out of 49.

(n) The actual rule is $\exists xFx, (F\alpha/x \supset \psi) \vdash \psi$. Since there will be no way to get this conditional other than by $\supset I$, and thereby assuming the antecedent, this rule is essentially $\exists E$.

(o) After introducing the defined symbol, $\exists$

(p) The basic system is $\exists E$ and no subproof for $\forall$, but gives alternative system with EI and a subproof for $\forall I$. We are counting the basic system here.

(q) Wilson's system has only UI, EI and QN—no generalization rules. (So, totals are out of 49.)

TABLE I: Characterization of 50 Elementary Natural Deduction Textbooks

	Anderson	Arthur	Barwise	Bergmann	Bessie	Bonevac	Bostock	Byerly	Carter	Cauman	Chellas	Copi	Curry	DeHaven	Fitch	Forbes	Gamut	Georgacarakos	Goldfarb	Goodstein	Gustason	Guttenplan	Harrison	Hurley	Iseminger	Jacquette	Jennings	Kalish	Klenk	Kozy	Leblanc	Lemmon	Machina
I. Proof Style																																	
bookkeeping										x								x															
graphical	x	x	x	x	x	x		x	x			x		x	x		x				x		x	x[2]	x	x		x	x	x	x		x
Gentzen							x						x							x													
Suppes											x					x			x			x					x					x	
II. "Axiomatic"?																																	
tautologies									x			x									x		x										
Equivalences		x	[d]		x			x	x			x						x			x		x	x					x			x	
"TF inference"																			x														
III. All Int-Elim?[e]	x		x	x		x	x[g]			x			x		x[h]	x	x[g]			x		x			x	x	x				x		x
IV. Sub-Proof for:																																	
⊃I	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x[k]	x	x	x	x	x	x	x	x	x	x	x	x
vE	x		x	x			x			x			x		x	x	x			x		x			x	x	x			x	x	x	
~I	x	x	x	x		x			x	x		x	x	x		x	x	x		x		x	x	x	x	x	x	x	x	x	x	x	x
~E				x	x				x		x			x												x		x					x
≡I			x	x											x										x	x					x		
other							x[g]													x[l]													
V. One Quantifier?													x							x													
VI. Arbitrary Names?		x								x											x						x					x	
VII. Existential Rule																																	
∃E	x		x	x	x		x				x		x	x	x	x	x[n]	x	x			x				x	x			x	x	x	x
EI		x				x		x	x	x		x								x	x		x	x	x			x	x		[p]		
VIII. Universal Rule																																	
∀I subproof?			x			x									x										x			x	x		[p]		

TABLE I: Characterization of 50 Elementary Natural Deduction Textbooks

	Martin	Massey	Mates	Myro	Pollock	Purtill	Quine	Resnick	Simco[a]	Simpson	Suppes	Tapscott	Teller	Tennant	Thomason	van Dalen	Wilson	Total (of 50)	Percent
I. Proof Style																			
bookkeeping		x				x	x											5	10%
graphical	x							x	x	x		x	x		x		x	30	60%
Gentzen														x		x		5	10%
Suppes			x	x	x						x							10	20%
II. "Axiomatic"?																			22[c] of 50=
tautologies					x				x			x		x				8	44%
Equivalences					x	x			x		x	x					x	17	
"TF inference"		x	x				x				x							5	
III. All Int-Elim?[e]	x									x			x	x[f]	x	x[i]		23	46%
IV. Sub-Proof for:																			41[j]=82%
⊃I	x	x	x	x	x	x	x	x	x[k]	x	x	x	x	x	x	x	x	50	100%
vE	x			x						x				x	x			22	44%
~I	x			x				x	x	x		x	x	x	x		x	36	72%
~E	x								x							x	x	12	24%
≡I	x			x						x					x			10	20%
other																		2	4%
V. One Quantifier?			x													x		4	8%
VI. Arbitrary Names?			x			x				x	x		x		x			11	22%
VII. Existential Rule					x[m]														of 49
∃E	x		x[o]	x				x		x		x	x	x	x	x[o]		30	61%
EI		x				x	x		x		x						x	19	39%
VIII. Universal Rule																	[q]		of 49
∀I subproof?					x			x		x			x		x			11	22%

2.8 Exploiting Natural Deduction Techniques: Modal Logic

The modern study of modal logic was initiated by C. I. Lewis (1918) in an effort to distinguish two concepts of "implication" which he thought had been confused in *Principia Mathematica*: the relation between propositions expressed by the material implication connective and the relation of logical consequence. It can be argued that natural deduction formulations of logic, in which logical *axioms* formulated as material implications are replaced by logical *rules*, go a long way toward clarifying the relation between these concepts; but natural deduction formulations were first published in 1934, almost two decades after Lewis began his study of his systems of "strict implication." In any event, he presented modal logic in a series of basically *axiomatic* systems (differing, however, from the usual axiomatic formulations of non-modal logic by including an algebraically inspired, primitive rule of substitution of proved equivalents). The modal logics he formulated include some that are still central to the study of modality (S_4 and S_5), but also others (which he thought better motivated by his philosophical project) now regarded as having at best artificial interpretations, and he failed to discover yet others (like the basic alethic logic T) that now seem obvious and natural.[24] It seems plausible that the logical technology he was familiar with is what led him to prefer logics like S_2 and S_1, and arguable that this preference was a philosophical mistake: these logics do not count even logical theorems as necessarily necessary, but his rule of substitution allows proven equivalents to be substituted for each other even in positions in the scope of multiple modal operators. This combination of features seems unlikely for a well-motivated necessity concept.

Around 1950, several logicians, among them the most notable perhaps being H.B. Curry and F.B. Fitch[25], had the idea of formulating modal logic[26] in terms of Introduction and Elimination rules for modal operators. Possibility Introduction and Necessity Elimination are simple and obvious: $\Diamond\phi$ may be inferred from ϕ, and ϕ may be inferred from $\Box\phi$. For the other two rules, natural deduction's idea of subproofs—that in some rules a conclusion is inferred, not from one or more statements, but from the presentation of a derivation of some specified kind—seems to have provided useful inspiration. Now, the subproofs used in standard rules for non-modal connectives differ from the main proof in that within them we can appeal to an additional assumption, the hypothesis of the subproof, but for modal logic a more important feature is that reasoning in a subproof can be subject to special

[24] The term 'normal modal logic' was originally introduced by Kripke (1963) for a class of logics extending T, and only later extended to include such weaker systems as K when it came to be appreciated how naturally Kripke's model-theoretic approach extended to them.

[25] Curry (1950) gave a formulation of S_4; Curry (1963) gives a more satisfactory discussion, although he does not seem to have given much consideration to any other system. Nonetheless, the changes to his rules needed for a formulation of T are quite simple. Fitch's earlier contributions are in a series of articles, summarized in his textbook (Fitch, 1952, Ch. 3).

[26] Alethic modal logic, that is. The idea that deontic, etc., concepts could be represented by "modal" operators seems not to have received significant attention until several years later, despite the publication of von Wright (1951a).

restrictions, as is the case in some quantificational rules. $\forall$-Introduction allows us to assert $\forall x Fx$ if we have presented a derivation of $F\alpha$, with the restriction that no (undischarged) assumptions containing the parameter α have been used. Fitch's rule of $\Box$-Introduction allows us, similarly, to assert $\Box\phi$ if we have been able to give a proof of ϕ in which only (what we have already accepted to be) necessary truths have been assumed. More formally: $\Box\phi$ can be asserted after a subproof, with no hypothesis, in which ϕ occurs as a line, under the restriction that a statement, ψ, from outside the subproof, can only be appealed to within it if (not just ψ itself, but) $\Box\psi$ occurs in the main proof before the beginning of the subproof. In the following proof, the modal subproof comprises lines 3–5, in which q is deduced from the premises $(p \supset q)$ and p, premises which are asserted to be necessary truths by the hypotheses of the non-modal, $\supset I$, subproofs.

1	│ $\Box(p \supset q)$	
2	│ │ $\Box p$	
3	│ │ $\Box$│ $(p \supset q)$	1, $\Box R$
4	│ │ │ p	1,$\Box R$
5	│ │ │ q	3,4 $\supset$E
6	│ │ $\Box q$	3–5, $\Box I$
7	│ $(\Box p \supset \Box q)$	2–6, $\supset I$
8	$\Box(p \supset q) \supset (\Box p \supset \Box q)$	1–7, $\supset I$

This rule, together with $\Box$-Elimination, gives a formulation of the basic alethic logic T (with necessity as sole modal primitive).[27] Varying the condition under which a formula is allowed to occur as a "reiterated" premise in a modal subproof yields formulations of several other important logics. S_4, for example, with its characteristic principle that any necessary proposition is necessarily necessary, is obtained by allowing $\Box\psi$ to occur in a modal subproof if $\Box\psi$ occurs above it, and S_5 by allowing $\neg\Box\psi$ to occur in a modal subproof if it occurs above the subproof (or alternatively, if $\Diamond$ is in the language, to allow $\Diamond p$ to appear in a modal subproof if it appears in an encompassing one).

Fitch also gave a corresponding $\Diamond$-Elimination rule, using subproofs with both restrictions and hypotheses: $\Diamond\phi$ may be inferred from $\Diamond\psi$ together with a subproof in which $\Diamond\phi$ is derived from the hypothesis ψ, the subproof having the same restrictions as to reiterated premises as the subproofs used in $\Box I$. This rule is, however, somewhat anomalous: obviously, in that the operator, $\Diamond$, supposedly being eliminated occurs in the conclusion of the

[27] T was first presented in Feys (1937), but is often attributed to Feys and von Wright. Given that von Wright didn't publish the system until 1951b, it seems a bit churlish not to credit Fitch as well, since Fitch presents the natural deduction formulation in his 1952 treatise/textbook. Unlike Feys and von Wright, Fitch didn't give the system a name, saying only that it was "similar to" S_2.

rule, less obviously in being deductively weak. The possibility rules are derivable from the necessity rules if we define $\Diamond\phi$ as $\neg\Box\neg\phi$, but not conversely: the necessity rules cannot be derived from the possibility rules and the definition of $\Box\psi$ as $\neg\Diamond\neg\psi$. Again, if both modal operators are taken as primitive, the definitions of $\Diamond$ in terms of $\Box$ and of $\Box$ in terms of $\Diamond$ are not derivable from the four modal rules and the standard classical negation rules. Fitch compensates for these weaknesses by imposing Negative introduction and elimination rules for the modal operators, by which $\neg\Box\phi$ is interderivable with $\Diamond\neg\phi$ and $\neg\Diamond\psi$ with $\Box\neg\psi$.[28]

Fitch's rules for the necessity and possibility operators have a more than passing formal resemblance to the rules for the Universal and Existential quantifiers: depending on one's point of view, the parallels between the logical behaviors of the modal operators and the quantifiers can be seen as explained by or as justifying the Leibniz-Kripke paraphrases of modal statements in terms of quantification over possible worlds. Since S_5 has the simplest possible worlds semantics (no accessibility relation is needed), it is to be expected that it can be given a natural deduction formulation in which its modal rules mirror the quantifier rules particularly well. Think of ordinary, non-modal, formulas as containing an invisible term referring to the actual world. If we think of modalized formulas as disguised quantifications of possible worlds, this term is changed, in $\Box I$ and $\Diamond E$ subproofs, into the proper parameter of a quantificational subproof: hence the restriction on reiteration. The condition for legitimate reiteration is that the invisible world-term should have no free occurrences in the reiterated formula, so: a formula ϕ occurring above a modal subproof may be reiterated in it if every occurrence sentence letter in ϕ is inside the scope of some modal operator. Call such an ϕ modally closed. We can now give the $\Diamond E$ rule a formulation closer to the standard $\exists E$ rule: any modally closed formula ϕ may be inferred from $\Diamond\psi$ together with a modally restricted subproof in which ϕ is derived from the hypothesis ψ. In S_5 with these rules, the Negative modal rules are derivable from the positive ones and standard negation rules, and the necessity rules are derivable from the possibility rules when $\Box\phi$ is defined as $\neg\Diamond\neg\phi$.[29]

Not all modal logics have nice natural deduction formulations, but those that do include many of the most important for applications. Fitch (1966) gives natural deduction formulations for a number of alethic and deontic logics and systems with both alethic and deontic operators. Fitting (2002) describes natural deduction formulations of a large number of logics, in addition to tableau versions. Several introductory logic texts (for example, Anderson and Johnstone, 1962; Iseminger, 1968; Purtill, 1971; Bonevac, 1987; Gamut, 1991; Forbes, 1994; Bessie and Glennan, 2000) include natural deduction formulations of one or more systems of modal logic, and Garson (2006), intended for students who have already

[28] Fitch also notes that, if the rules for the modal operators are added to Intuitionistic rather than Classical logic, the inferability of $(\Diamond\phi \vee \Diamond\psi)$ from $\Diamond(\phi \vee \psi)$ has to be postulated specially.

[29] Fitch (1952) treats modal operators before quantifiers. Despite the reverse relative importance of the two topics, this may make pedagogical sense, as it accustoms students to the use of subproofs with restrictions on reiteration before they have to cope with the technicalities of free and bound occurrences of variables.

done a bit of logic, introduces many modal logics, both propositional and quantified, with both natural deduction and tableau formulations.

3 The Metatheory of Natural Deduction

3.1 Normalizing Natural Deduction

Gentzen's rules of his system of natural deduction come in pairs of an introduction and an elimination rule. The rules in such a pair tend to be—in an intuitively obvious way whose precise definition isn't easy to formulate—*inverses* of each other. This is clearest with the rules for conjunction: conjunction introduction infers a conjunction from its conjuncts, conjunction elimination infers the conjuncts from the conjunction. In other cases the inverse relationship is not as direct: disjunction introduction infers a disjunction from one of its disjuncts, but disjunction elimination can't infer a disjunct from the disjunction (that would be unsound, as neither disjunct is in general entailed by the disjunction). The effect of the disjunction elimination rule is rather to reduce the problem of inferring a conclusion from a disjunction to that of deriving it from the disjuncts. With other operators the inversion can take yet other forms, and negation tends to be a special case.

What this inverse relationship between rules suggests (and in particular, what it suggested to Gentzen when he was working on his thesis, which formed the basis of Gentzen (1934)) is that if, in the course of a proof, you infer a conclusion by an introduction rule and then use that conclusion as the major[30] premise for an inference by the corresponding elimination rule, you have made an avoidable detour: you have inserted a bit of conceptual complexity (represented by the operator introduced by the introduction rule) into your proof, but pointlessly, since after the elimination inference your proof uses as premises only things you had reached before the detour. Leading to the conjecture—since proven as a theorem for many systems—that if you can prove something in a system of natural deduction, you can prove it in a way that avoids all such detours. Or, as Gentzen put it in the published version of his thesis (Gentzen, 1934, p.289)

> ... every purely logical proof can be reduced to a determinate, though not unique, normal form. Perhaps we may express the essential properties of such a normal proof by saying 'it is not roundabout.' No concepts enter into the proof other than those contained in its final result, and their use was therefore essential to the achievement of that result.

Let us be a bit more formal. Define a proof in a natural deduction system as *normal* if no formula in it is both the conclusion of an inference by an introduction rule and the (major)

[30] In an elimination rule needing more than one premise, the major is the one containing the operator mentioned in the name of the rule: e.g., when B is inferred from A and $(A \to B)$ by $\to$ elimination, the conditional is the major premise.

premise of an inference by an elimination rule.[31] Normal proofs avoid the useless conceptual complexity involved in introduction-elimination detours: they proceed by first applying elimination rules to premises or hypotheses (reducing their complexity by stripping off the main logical operator at each step) and then monotonically building up the complexity of the final conclusion by introduction rules. They are thus in a sense conceptually minimal: no piece of logical construction occurs in them unless it is already contained in the original premises or the final conclusion. In technical terms, a normal proof has the *subformula property*: every formula occurring in it is a subformula[32] of the conclusion or of some premise. This feature of normal proofs has proven valuable for consistency proofs, proof-search algorithms, and decision procedures, among other properties.

We can now formulate stronger and weaker claims that might be made about a system of natural deduction. A *normal form theorem* would say that for every proof in the system there is a corresponding normal proof with the same conclusion and (a subset of) the same premises. A stronger *normalization theorem* would say that there is a (sensible[33]) procedure which, applied to a given proof, will convert it into a corresponding normal proof. Both theorems hold for nice systems.

Gentzen saw this, at least for intuitionistic logic, and in a preliminary draft of his thesis that remained generally unknown until recently (von Plato (2008b)), he proved the normalization theorem for his system of intuitionistic logic. Gentzen's proof is essentially the same as the one published thirty years later in Prawitz (1965), and the normalization procedure is simple and intuitive: look for a "detour" and excise it (this is called a reduction); repeat until done[34]. Thus, for example, suppose that in the given proof $A \vee B$ is inferred from A by disjunction introduction and C then inferred from $A \vee B$ by disjunction elimination. One of the subproofs (to use our Fitch-inspired terminology) of the disjunction elimination will contain a derivation of C from the hypothesis A. So delete the occurrence of $A \vee B$ and the second subproof, and use the steps of the A-subproof to derive C from the occurrence of A in the main proof. That's basically it, although there's a technical glitch. Suppose, in the proof to be normalized, you used an introduction rule inside the subproof of an application

[31] This is a rough preliminary definition. A technical refinement is added below in connection with "permutative reductions," and for some systems a variety of subtly different notions of normality have been studied.

[32] This is literally true for standard systems of intuitionistic logic. For classical systems it is often necessary to allow the negations of subformulas of a given formula to count as honorary subformulas of it. And in, e.g., a system of modal logic one might want to allow formulas obtained by adding some limited number of modal operators to a real subformula.

[33] Not every algorithm for converting proofs qualifies. The normal form theorem would imply the existence of an idiot's algorithm: grind out all the possible normal proofs in some systematic order and stop when you get one with the right premises and conclusion. Logicians are interested in normalization procedures in which each step in modifying the given proof into a normal one intuitively leaves the proof "idea" the same or is a simplification.

[34] Gentzen's proof that the procedure works, like that in Prawitz (1965), depends on attacking the different detours in a particular order. But the order doesn't actually matter, as was proved by the strong normalization theorem of Prawitz (1971).

of existential quantifier elimination (or inside one of the subproofs of an application of disjunction elimination) to obtain the final formula of the subproof, wrote another copy of the same formula as the conclusion of the existential quantifier elimination, and then applied the corresponding elimination rule in the main proof. Under a literal reading of the definition of a normal proof given above, this would not qualify as a detour (no one occurrence of the formula is both the conclusion of the introduction rule and premise of the elimination rule), and it won't be excised by the reduction steps described. Clearly, however, it would be cheating to allow things like this in a "normal" proof (and allowing it would mean that "normal" proofs might not have the subformula property). An honest normalization procedure will eliminate this sort of "two stage" detours. To do this Gentzen and Prawitz included a second sort of step in the procedure that Prawitz called a permutative reduction: when something like what is described above occurs, rearrange the steps so the elimination rule is applied within the subproof, reducing the two-stage detour to an ordinary detour in the subproof. So the full normalization procedure is: look for a detour or a two-stage detour, excise it by an ordinary or permutative reduction, repeat until done.

Gentzen despaired, however, of proving a normalization theorem for classical logic: in the published version of his thesis, (Gentzen, 1934, p.289), he says:

> For, though [his natural deduction system] already contains the properties essential to the validity of the *Hauptsatz*, it does so only with respect to its intuitionist form, in view of the fact that the law of excluded middle ... occupies a special position in relation to these properties.

He therefore defined a new sort of system, his L-systems or sequent calculi (described below), for which he could prove an analogue of normalization for both intuitionistic and classical systems, and left the details of normal forms and normalization of natural deduction out of the published version of the thesis.

Since classical logic is in some ways simpler than intuitionistic logic (think of truth tables!), it may seem surprising that proof-theoretic properties like the normalization theorem are harder to establish for it, and details will differ with different formulations of classical logic. Prawitz (and Gentzen in the early draft of his thesis) worked with a natural deduction system obtained from that for intuitionistic logic by adding the classicizing rule of *indirect proof*: ϕ may be asserted if a contradiction is derivable from the hypothesis $\neg\phi$. Note that this rule can be broken down into two steps: first, given the derivation of a contradiction from $\neg\phi$, assert $\neg\neg\phi$ by the intuitionistically valid form of *reductio* (that is, by the standard negation introduction rule of natural deduction systems for intuitionistic logic), and then infer ϕ by (another possible choice for a classicizing rule) *double negation elimination*. Thought of this way, there is an immediate problem: the characteristic inference pattern that allows for classical proofs that are not intuitionistically valid seems to be essentially a "detour", namely, an introduction rule is used to infer a double negation which is then used as a premise for an elimination rule, with the result that on this naïve construal of 'normal', no intuitionistically invalid classical proof can possibly be normalized!

But there is another problem, much deeper than this rather legalistic objection. Recall that the interest of normal proofs lies largely in their conceptual simplicity: in the fact that they possess the subformula property. If we ignored the legalistic objection and simply allowed indirect proof (as, perhaps, a special rule outside the pattern of introduction and elimination rules) we could produce proofs that failed this simplicity criterion very badly: working from simple premises we could build up (perhaps without any obvious detours) to an indirect proof of some immensely long and complex formula, and then (again by normal-looking means) derive a simple conclusion from it. In other words, whether we think of the classicizing rule *de jure* normal or not, an interesting notion of normal proof for classical logic must put some restriction on its use.

Prawitz (and Gentzen before him) chose an obvious restriction which would give normal proofs the subformula property[35]: in a normal proof, indirect proof is used only to establish atomic formulas.[36] With such a definition of a normal proof, a normal form theorem will hold only if the system with its classicizing rule restricted to the atomic case is complete for classical logic; that is, if, by using the classicizing rule in the restricted form we can prove everything we could by using it without restriction. Now, in some cases it is possible to "reduce" non-atomic uses of indirect proof (e.g.) to simpler cases. Suppose we used indirect proof to establish a conjunction, $\phi \wedge \psi$. $\neg(\phi \wedge \psi)$ is implied (intuitionistically) by each of $\neg\phi$ and $\neg\psi$, so the contradiction derived from the hypothesis $\neg(\phi \wedge \psi)$ could have been derived in turn from each of the two hypotheses $\neg\phi$ and $\neg\psi$. So we could have used indirect proof to get the simpler formulas ϕ and ψ instead, and then inferred $\phi \wedge \psi$ by conjunction introduction.

In other cases, however, this is not possible. $\forall x(\exists y \neg Fy \vee Fx)$, where the variable x does not occurring free in Fy, classically (but not intuitionistically) implies $(\exists y \neg Fy \vee \forall x Fx)$: this is an inference that a natural deduction system for classical logic must allow. Assuming classical logic (so allowing indirect proof or excluded middle) for formulas only of the complexity of Fx, will not help. (For, excluded middle is provable for quantifier-free formulas of intuitionistic (Heyting) arithmetic, so if we could go from having excluded middle for atomic instances to excluded middle for quantifications, then Heyting Arithmetic would collapse into classical Peano Arithmetic.) So to get full classical first-order logic *some* classicizing rule must be allowed to work on complex formulas.

On closer examination, however, it turned out that this difficulty only arises in connection with formulas containing disjunction operators or existential quantifiers. Prawitz was able to prove a normalization theorem (the definition of "normal" including the proviso that indirect proof be used only for atomic formulas) for the fragment of classical logic having only the universal quantifier and only the conjunction, conditional and negation connec-

[35]Well, "weak subformula" property: negations of genuine subformulas allowed.

[36]If a different classicizing rule is chosen, it will also have to be restricted somehow to get an "honest" normalization theorem. Results about a system with a restricted form of one rule don't carry over automatically to systems with comparably restricted forms of other rules. Some, otherwise attractive, natural deduction systems for classical logic don't seem to have good normalization theorems.

tives. For some purposes this is sufficient: disjunction and existential quantification are classically definable in terms of the operators included in the fragment, so to every proof in full classical logic there corresponds a proof, with trivially equivalent premises and a trivially equivalent conclusion, in the fragment Prawitz showed to be normalizable. Thus, for example, if one were to want to use the normalization theorem and the subformula property of normal proofs as part of a consistency proof for some classical mathematical theory, Prawitz's result would suffice even though it doesn't extend to classical logic with a full complement of operators.

This leaves open the question of whether, with a more carefully worded definition of normal proofs, a reasonable normalization theorem would be provable for full classical logic. Various results have been obtained; the current state of the art is perhaps represented by Stålmarck (1991)'s normalization theorem: here normal proofs, in addition to avoiding explicit detours, are not allowed to use indirect proof to obtain complex formulas which are then used as major premises for elimination rules.

The normalization theorems described so far have been proven by methods which are finitistic in the sense of Hilbert's program: Gentzen's hope in studying natural deduction systems was that normalization theorems would contribute to that program's goal of finding a finitistic consistency proof for arithmetic. In contrast, stronger set-theoretic methods have also been used in studying normalization. These have yielded strong normalization theorems (that any reduction process, applying normalizing transformations in any order to a proof, will terminate in a finite number of steps with a normal proof) and confluence theorems (that the normal proofs produced by different reduction processes to a given proof will, up to trivialities like relettering variables, be identical). Prawitz (1971)'s strong normalization theorem for intuitionistic logic is an early example; Stålmarck (1991) gives a set-theoretic proof of a strong normalization theorem as well as an elementary proof of normalization for classical logic.

3.2 Natural Deduction and Sequent Calculus

Gentzen's own response to the problem of normalization for classical logic was to define a new kind of formal system, closely related to natural deduction, for which he was able to prove something similar to normalization for full classical logic. This involved ideas (and useful notations!) which have proved to be of independent interest. As a first step (in fact applying only to intuitionistic logic), consider a formal system in which the lines of a proof are thought of not as formulas but as *sequents*. For the moment (it will get a bit more complex later when we actually get to classical logic), a sequent may be thought of as an ordered pair of a finite sequence of formulas (the antecedent formulas of the sequent) and a formula (its succedent formula). This is usually written

$$\phi_1, \phi_2, \ldots, \phi_n \vdash \psi$$

with commas separating the antecedent formulas, and different authors choosing for the symbol separating them from the succedent a turnstile (as here) or a colon or an arrow (if this is not used for the conditional connective within formulas). In schematic presentations of rules for manipulating sequents it is standard to use lower-case Greek letters as schematic for individual formulas and capital Greek letters for arbitrary (possibly empty) finite sequences of formulas. A sequent is interpreted as signifying that its antecedent formulas collectively imply its succedent: if the antecedent formulas are all true, so is the succedent. (So, if one wants, one can think of a sequent as an alternative notation for the theoremhood of a conditional with a conjunctive antecedent.) There are special cases: a sequent with a null sequence of antecedents is interpreted as asserting the succedent, a sequent with no succedent formula is interpreted as meaning that the antecedents imply a contradiction, and an empty sequent, with neither antecedent nor succedent formulas, is taken to be a contradiction: consistency, for formal systems based on sequents, is often defined as the unprovability of the empty sequent.

3.3 Sequent Natural Deduction

In an axiomatic system of logic each formula occurring as a line of a proof is asserted as a logical truth: it is either an axiom or follows from the axioms. The significance of a line in a natural deduction proof is less obvious: the formula may not be valid, for it may be asserted only as following from some hypothesis or hypotheses which, depending on the geometry of the proof presentation, may be hard to find! One way of avoiding this difficulty would be to reformulate a natural deduction system so that lines of a proof are sequents instead of formulas: sequents with the formula occurring at the corresponding line of a conventional proof as succedent and a list of the hypotheses on which it depends as antecedent. (Such a reformulation might seem to risk writer's cramp, since the hypotheses have to be copied down on each new line until they are discharged, but abbreviations are possible: Suppes (1957), by numbering formulas and writing only the numerals instead of the formulas in the antecedent, made a usable text-book system of sequent-based natural deduction that was followed by many others, as we have seen above.) It is straightforward to define such a system of *sequent natural deduction* corresponding to a given conventional natural deduction system:

- Corresponding to the use of hypotheses in natural deduction, and to the premises of proofs from premises, sequent natural deduction will start proofs with identity sequents, sequents of the form $\phi \vdash \phi$, which can be thought of as logical axioms.

- Corresponding to each rule of the conventional system there will be a rule for inferring sequents from sequents. For example, the reformulated rule of conjunction introduction will allow the inference of $\Gamma \vdash (\phi_1 \wedge \phi_2)$ from the two sequents $\Gamma \vdash \phi_1$ and $\Gamma \vdash \phi_2$; conjunction elimination will allow the inference of either of the latter sequents from the former. Corresponding to a rule of the conventional system that discharges a hypothesis there will be a rule deleting a formula from the antecedent of a sequent: conditional introduction will allow the inference of $\Gamma \vdash (\phi_1 \rightarrow \phi_2)$ from $\Gamma, \phi_1 \vdash \phi_2$.

- There will be a few book-keeping, or structural rules. One, which Gentzen called thinning[37], allows extra formulas to be inserted in a sequent. To see the use of this, note that as stated above the conjunction introduction rule requires the same antecedent in its two premises. To derive $(\phi_1 \wedge \phi_2)$ from the two hypotheses ϕ_1 and ϕ_2, then, one would start with the identity sequents $\phi_1 \vdash \phi_1$ and $\phi_2 \vdash \phi_2$, use thinning to get $\phi_1, \phi_2 \vdash \phi_1$ and $\phi_1, \phi_2 \vdash \phi_2$, and then apply conjunction introduction to these sequents. Two other structural rules would be demoted to notational conventions if one was willing (as Gentzen was not[38]) to regard the antecedents of sequents as unstructured *sets* of formulas. *Permutation* allows the order of the antecedent formulas to be changed. *Contraction*[39] allows multiple occurrences of a single formula in the antecedent to be reduced to a single occurrence.

As an example of a proof in sequent natural deduction, consider the commutativity of conjunction:

$\phi_1 \wedge \phi_2$	$\vdash \phi_1 \wedge \phi_2$	Axiom
$\phi_1 \wedge \phi_2$	$\vdash \phi_1$	by conjunction elimination
$\phi_1 \wedge \phi_2$	$\vdash \phi_2$	also by conjunction elimination
$\phi_1 \wedge \phi_2$	$\vdash \phi_2 \wedge \phi_1$	by conjunction introduction

3.4 From Sequent Natural Deduction to Sequent Calculus

The action, so to speak, in a sequent natural deduction proof is all in the succedents of the sequents: whole formulas can be added or subtracted from the antecedents, or rearranged within them, by the structural rules, but the rules for the connectives only add or subtract operators from succedent formulas. (Not surprisingly, since it is the succedent formulas that correspond to the formulas occurring in the conventional variant of the natural deduction proof!) What Gentzen saw was that, though this is natural enough for the introduction rules, the elimination rules could also be seen as corresponding to manipulations of antecedent formulas: in particular, just as the conclusion of the sequent form of an introduction rule is a sequent whose succedent formula has one more operator than the succedent

[37]German *Verdünnung*. This is thinning in the sense in which one thins paint: some logicians translate it as *dilution*.

[38]The logics he considered—classical and intuitionistic—can both be described by sequents with sets of antecedent formulas. Sequents with more structured antecedents are useful in formulating things like *relevance logics* or Girard's *linear logic*: systems which, precisely because their formulation omits or restricts some of Gentzen's structural rules, are often called substructural logics. See Restall (2000).

[39]Following Curry, many logicians refer to contraction as rule W. Unfortunately other logicians call Thinning *Weakening*, which also abbreviates to W.

formulas of the premise sequents, analogues of the elimination rules can be given in which the conclusion is a sequent containing an antecedent formula with an extra operator. Since, for example, the effect of the conjunction elimination rule is to allow the derivation from a conjunction of any formula derivable from either of its conjuncts, a sequent natural deduction analogue (henceforth called conjunction left; the analogue of conjunction introduction stated above will be called conjunction right) will tell us that $\Gamma, \phi_1 \wedge \phi_2 \vdash \psi$ can be inferred from either $\Gamma, \phi_1 \vdash \psi$ or $\Gamma, \phi_2 \vdash \psi$. As an illustration of how this works, let us see the commutativity of conjunction in this framework:

ϕ_1	$\vdash \phi_1$	Axiom
$\phi_1 \wedge \phi_2$	$\vdash \phi_1$	By conjunction left
ϕ_2	$\vdash \phi_2$	Axiom
$\phi_1 \wedge \phi_2$	$\vdash \phi_2$	By conjunction left (second axiom)
$\phi_1 \wedge \phi_2$	$\vdash \phi_2 \wedge \phi_1$	By conjunction right (2nd and 4th lines)[40]

Our trivial example about the commutativity of conjunction has involved reformulations of a *normal* natural deduction proof: conjuncts are inferred from an assumed conjunction by an elimination rule and a new conjunction is then inferred from them by an introduction rule. What if we *wanted* to make a detour, to infer something by an introduction rule and then use it as the major premise for an application of the corresponding elimination rule? Sequent natural deduction amounts to ordinary natural deduction with hypothesis-recording antecedents decorating the formulas, so there is no difficulty here. In Gentzen's *sequent calculi* (or *L-systems*: he named his sequent formulations of intuitionistic and classical logic LJ and LK respectively) things aren't as straightforward. Suppose, for example, that we can derive both ϕ and ψ from some list of hypotheses (so: start with the sequents $\Gamma \vdash \phi$ and $\Gamma \vdash \psi$) and we decide to infer $\phi \wedge \psi$ from them by conjunction introduction, and then (having somehow forgotten how we got $\phi \wedge \psi$) decide to infer ϕ from it by conjunction elimination. The introduction step is easy: the sequent

(i) $\Gamma \vdash \phi \wedge \psi$

follows immediately from the two given sequents by conjunction right. And we can prove something corresponding to the elimination step,

(ii) $\Gamma, \phi \wedge \psi \vdash \phi$

almost as immediately:

ϕ	$\vdash \phi$	Axiom
$\phi \wedge \psi$	$\vdash \phi$	By conjunction left
$\Gamma, \phi \wedge \psi$	$\vdash \phi$	By as many thinnings as there are formulas in Γ

But how do we put them together and infer (forgetting we already have it!)

(iii) $\Gamma \vdash \phi$

[40]Obviously the identity sequent $\phi_1 \wedge \phi_2 \vdash \phi_1 \wedge \phi_2$ can be derived in a similar fashion. A useful exercise for newcomers to Gentzen's rules is to show that all identity sequents are derivable when only the identities for atomic formulas are assumed as axioms.

from (i) and (ii)? The short answer is that we can't: sequent (iii) is shorter than (i) or (ii) and (except for the irrelevant Contraction) none of the rules stated so far allow us to infer a shorter sequent from a longer. To allow the inference, Gentzen appeals to a fourth structural rule, *Cut*. Cut says that, given a sequent with a formula θ in its antecedent and another sequent with θ as succedent formula, we may infer a sequent with the same succedent as the first and an antecedent containing the formulas other than θ from the antecedent of the first sequent along with the formulas from the antecedent of the second sequent.[41] Schematically, from $\Gamma, \theta \vdash \delta$ and $\Delta \vdash \theta$ we may infer $\Gamma, \Delta \vdash \delta$. The inference from (ii) and (i) above to (iii) is an instance of Cut.[42]

Interpreting sequents as statements of what is implied by various hypotheses, it is clear that Cut is a valid rule: it amounts to a generalization of the principle of the transitivity of implication. On the other hand it plays a somewhat anomalous role in the system. Normal natural deduction proofs can be translated into L-system proofs in which the Cut rule is never used, and—our example is typical—any "detour" in a non-normal proof corresponds to an application of Cut.[43] It is also analogous to detours in its effect on conceptual complexity in proofs: in a proof in which Cut is not used, every formula occurring in any sequent of the proof is a subformula of some formula in the final sequent. Cut is the only rule which allows material to be removed from a sequent: Thinning adds whole formulas, the left and right rules for the different logical operators all add an operator to what is present in their premise sequents, Permutation just rearranges things, and even Contraction only deletes extra copies. Possible L-system proofs of a sequent not using Cut, in other words, are constrained in the same way that normal natural deduction derivations of the sequent's succedent from its antecedent formulas are, and Cut relaxes the constraints in the same way that allowing detours does.

In the published version of his thesis, then, Gentzen, after presenting his natural deduction systems NJ and NK for intutitionistic and classical logic, sets them aside in favor of the sequent calculi LJ and LK. In place of the normalization theorem for the natural deduction systems he proves his *Hauptsatz* (literally: 'principal theorem', but now used as a proper name for this particular result) or *Cut-elimination theorem*: any sequent provable in the L-systems is provable in them without use of Cut, and indeed there is a procedure, similar to the reduction procedure for normalizing natural deduction proofs, which allows

[41] Technically, if the first sequent has multiple copies of θ in its antecedent, Gentzen's cut only removes one. This complicates the details of his proof of the Hauptsatz, but can be overlooked in an informal survey: the version stated here is closer to his rule *Mix*, which he shows equivalent to Cut in the presence of the other structural rules.

[42] For convenience we have shown proofs in sequent calculi as simple columns of sequents. Gentzen displayed such proofs as trees of sequents (this is probably the most perspicuous representation for theoretical purposes.). If we were to use LJ or LK in practice for writing out proofs, it would be inconvenient because the same sequent will often have to have multiple occurrences in an arboriform proof, and converting a columnar proof into a tree can lead to a superpolynomial increase in proof size.

[43] Though not every application of Cut corresponds to a normality-destroying manoeuvre in a natural deduction proof.

any L-system proof to be converted into a Cut-free proof of the same final sequent.[44]

3.5 The Sequent Calculus LK

So far we have mentioned the classical sequent calculus LK, but not described it. What Gentzen saw was that a very simple modification of the intuitionistic sequent calculus LJ—one which, unlike the classical system of natural deduction, did not involve adding a new rule for some logical operator—produced a formulation of classical logic. Recall that the difficulties in defining a normal form and proving a normalization theorem for classical logic centered on the treatment of disjunction (and the existential quantifier). The classically but not intuitionistically valid inference from $\forall x(P \vee Fx)$ (x not occurring free in P) to $(P \vee \forall xFx)$ is typical. For any particular object in the domain we can (intuitionistically) infer that either it is F or else P is true: Universal Quantifier Elimination applied to the premise gives us $(P \vee Fa)$. Applying Disjunction Elimination, we adopt the two disjuncts in succession as hypotheses and see what they give us. P is helpful: Disjunction Introduction gives us the conclusion, $(P \vee \forall xFx)$. Fa, however, isn't: we want to be able to infer $\forall xFx$, but being told that one object in the domain is F doesn't tell us they all are! What we would like to do, classically, is to consider the two disjuncts, but somehow to still be generalizing when we think about Fa.[45] This is, of course, objectionable from the point of view of intuitionistic philosophy (it assumes that the entire, possibly infinite, domain exists as a given whole, whereas the premise only tells us that, whatever individuals there are in the domain, a certain disjunction holds of each one). There is also no obvious way of fitting this kind of thinking into the framework of natural deduction: introduction rules install new main operators in formulas, so we can only "generalize"—formally, use Universal Quantifier Introduction to insert a new universal quantifier—if we are in a position to assert an (arbitrary) instance of the quantification as an independent formula, but in this situation we are only able present it as an alternative: as one disjunct of a disjunction.

Gentzen's solution was to generalize the notion of a sequent to allow one sequent to have multiple succedent formulas, interpreted disjunctively, and to allow the analogues of natural deduction's Introduction rules—the rules that add an operator in the succedent—to act on one succedent formula while ignoring the rest. Formally, a sequent is redefined as an ordered pair of sequences of formulas (again allowing the special cases of one or the other or both being the null sequence),

$$\Gamma \vdash \Delta$$

[44]Gentzen defined his L-systems with the Cut rule, so this is his formulation of the result. Some subsequent writers have defined their sequent calculi as *not* containing Cut. They thus rephrase the Hauptsatz as: adding Cut as an additional rule would not enable us to prove any sequents we couldn't prove in their official systems. This yields another name for Gentzen's result: *Admissibility of Cut.*

[45]One possibility would be to use something like Quine's rule of U.G., letting a be an "arbitrary." Unless serious restrictions are imposed, use of rules like E.I. and U.G. in combination with intuitionistic propositional rules will yield a superintuitionistic quantified logic.

As before, the interpretation is a sort of implication: this time, that *if* all the antecedent formulas are true, *then* at least one of the succedent formulas is. (Thus it can be thought of as a notational variant of a conditional with a conjunctive antecedent and a disjunctive consequent.) But, now that we are looking in more detail about how quantificational inferences are handled in the L systems, it is worth being more precise about the interpretation. The formulas in a sequent may, after all, contain free variables (= parameters), and they are to be given a *generality* interpretation: the sequent says that, for any values assigned to the free variables (assigning values uniformly, so if the same variable occurs free in two or more formulas of the sequent they will all be taken to be about the same object), if all the antecedent formulas are true for those values, then at least one of the sucedent formulas will be true for the same values. (So, if we are to think of a sequent as a notational variant of some sentence, it will have to be, not just an implication, but a generalized implication: variables free in the sequent are to be thought of as implicitly bound by universal quantifiers having the entire implication in their scope.)

The rules for the classical sequent calculus LK are virtually identical to those of the intuitionistic LJ, but now applied to sequents with multiple succedent formulas. In particular, the structural rules of Permutation and Contraction now apply to succedents ("Permutation right," etc) as well as to antecedents, and Thinning can add additional formulas to non-null succedents. (Thinning was allowed on the right in LJ, but only if the premise had a null succedent: the inference from $\Gamma \vdash$ to $\Gamma \vdash \phi$ is the formalization in LJ of the *ex falso quodlibet* principle.) Cut takes a more general form: from $\Gamma, \theta \vdash \Delta$ and $\Phi \vdash \Psi, \theta$ to infer $\Gamma, \Phi \vdash \Psi, \Delta$[46] The rules for the logical operators are unchanged, but now the right rules as well as the left ones tolerate neighbors to the formulas they act on. We illustrate the workings of the system by showing how to derive some standard examples of intuitionistically invalid classical validities.

The negation rules in the L-systems switch a formula from one side to the other, adding a negation to it in the process. (If the validity of any of these rules, on the interpretation described above for sequents, is not obvious, it is a useful exercise to think it through until it becomes obvious: none are hard.) In the intuitionistic case, LJ allows us to prove

ϕ	$\vdash \phi$	Axiom
$\phi, \neg\phi$	$\vdash$	Negation left

(contradictions are repugnant to intuitionists as well as to classical logicians). Since LJ allows at most a single succedent formula, the move in the opposite direction is impossible, but in LK we have

ϕ	$\vdash \phi$	Axiom
	$\vdash \neg\phi, \phi$	Negation right

[46]The classical validity of this rule is easily seen. Suppose there is an interpretation (and assignment to free variables) falsifying the conclusion: making all the formulas in Γ and Φ true and all those in Ψ and Δ false. If formula θ is true on this interpretation, the first premise sequent is falsified, and if it is false the second is.

which, on the disjunctive interpretation given to multiple succedents, amounts to the Law of Excluded Middle. Since the rule for Disjunction on the right corresponds to the Disjunction Introduction rule of natural deduction, we can continue this derivation to get a more explicit statement of the Law:

$\vdash \neg\phi, (\phi \vee \neg\phi)$	Disjunction right
$\vdash (\phi \vee \neg\phi), (\phi \vee \neg\phi)$	Disjunction right
$\vdash (\phi \vee \neg\phi)$	Contraction right[47]

As an even simpler example,

ϕ	$\vdash \phi$	Axiom
	$\vdash \neg\phi, \phi$	Negation right
$\neg\neg\phi$	$\vdash \phi$	Negation left.

And, finally, our example of restricting a universal quantification to one disjunct:

ϕ	$\vdash \phi$	Axiom
ϕ	$\vdash \phi, Fa$	Thinning right
Fa	$\vdash Fa$	Axiom
Fa	$\vdash \phi, Fa$	Thinning right
$(\phi \vee Fa)$	$\vdash \phi, Fa$	Disjunction left, from 2nd and 4th sequents

Note here an interesting general feature of LK: we can prove a sequent saying that an explicit disjunction implies the pair of its disjuncts (interpreted disjunctively!). The dual principle, that a pair of conjuncts implies their conjunction, is expressed by a sequent with a single succedent formula, provable in both LJ and LK. Continuing our example,

$\forall x(P \vee Fx) \vdash P, Fa$ Universal quantifier left.

The condition on the Universal quantifier right rule—corresponding to the condition on Universal Quantifier Introduction, that the free variable to be replaced by a universally quantified one may not occur free in assumptions on which the instance to be generalized depends—is that the free variable cannot occur in any other formula, left or right, of the premise sequent: but we now have a sequent satisfying this, so we may continue with

$\forall x(P \vee Fx) \vdash P, \forall x Fx$ Universal quantifier right,

after which a couple of Disjunction rights and a contraction will give us the desired

$$\forall x(P \vee Fx) \vdash (P \vee \forall x Fx).$$

What gets pulled out of the hat probably got put into it earlier. The rules which allowed us to prove a sequent expressing the inference from $\forall x(P \vee Fx)$ to $(P \vee \forall x Fx)$ can be seen as a generalization of that very inference. As mentioned above, a sequent can be thought

[47]Gentzen actually formulated his Disjunction right rule as applying only to the last formula of the succedent, so technically there should have been a Permutation right inference between the two Disjunction rights: a complication we will ignore in examples.

of as a notational variant of a universally quantified implication, with a disjunction as its consequent (= the succedent of the sequent). Viewed this way, the rule of Universal Quantifier Right amounts to taking a universal quantifier from the front of a sentence, binding a variable that occurs in only one disjunct of the matrix, and moving it in to that disjunct! So what, exactly, has been gained? The problem with classical logic, when it came to normalizing natural deduction proofs, was that classicizing rules had to be applied to logically complex formulas. (It's easy to derive $(P \vee \forall xFx)$ from $\forall x(P \vee Fx)$ by indirect proof, deriving a contradiction from the premise and the hypothesis $\neg(P \vee \forall xFx)$!) In order to get around this problem, the sequent calculus has rules that in effect amount to rules for manipulating formulas (viz., quantified conditionals with disjunctions as consequents) more complex than the formulas explicitly displayed. What keeps this from being a trivial cheat is the fact that the additional complexity is strictly limited. The formulas occurring in the antecedents and succedents of the sequents in an LK or LJ proof are the ones we are really interested in. The sequents themselves are equivalent to more complex formulas built up from them, but not arbitrarily more complex: only a single layer each, so to speak, of conjunctive (in the antecedent), disjunctive (in the succedent), implicational and universal-quantificational structure is added beyond the formulas we are interested in. This is the justification for using different notation rather than writing sequents as quantified implications: in building up formulas with normal connectives and quantifiers we can iterate ad infinitem, but in writing commas and turnstiles instead, we are noting the strictly limited amount of additional structure required by Gentzen's rules. And it is because of the limitation on this additional structure that Gentzen's Hauptsatz and the subformula property of Cut-free proofs are "honest": yes, there is a bit more structure than is contained in the formulas whose implicational relations are demonstrated in the proof, but it is strictly constrained. We have what we really need from a subformula property: a limit on the complexity of structure that can appear in the proof of a given logical validity.

3.6 Variants of Sequent Calculi

Following Gentzen's initial introduction of sequent calculi, many variants have been been defined. Gentzen himself noted that alternative formulations of the rules for the logical operators are possible. For example, Gentzen's Disjunction left rule requires that the premise sequents be identical except for containing the left and right disjuncts in their antecedents:

From $\phi, \Gamma \vdash \Theta$ and $\psi, \Gamma \vdash \Theta$ to infer $(\phi \vee \psi), \Gamma \vdash \Theta$

and his Disjunction right rule adds a disjunct to a single formula in the succedent:

From $\Gamma \vdash \Theta, \phi$ to infer $\Gamma \vdash \Theta, (\phi \vee \psi)$ (or $\Gamma \vdash \Theta, (\psi \vee \phi)$

One could equally well have a Disjunction left rule that combined the extra material in the two premises:

From $\delta, \Gamma \vdash \Delta$ and $\theta, \Phi \vdash \Psi$ to infer $(\delta \vee \theta), \Gamma, \Phi \vdash \Delta, \Psi$

and a Disjunction right rule that joined two distinct formulas in the succedent into a disjunction:

From $\Gamma \vdash \Delta, \phi, \psi$ to infer $\Gamma \vdash \Delta, (\phi \vee \psi)$.

Which to choose? In the presence of the structural rules, the two pairs of rules are equivalent. Without them, or with only some of them, the two pairs of rules can be seen as defining different connectives: what the Relevance Logic tradition calls *extensional* (Gentzen's rules) and *intensional* (the alternative pair) disjunction and the tradition stemming from Girard's work on Linear Logic calls *additive* and *multiplicative* disjunction. The conjunction rules are exact duals of these: Gentzen's version defining, when structural rules are restricted or omitted, what the two traditions call extensional or additive conjunction, the alternative pair intensional or multiplicative conjunction.[48] Gentzen's rules for the conditional are more nearly parallel to the alternative (intensional/multiplicative) versions of the conjunction and disjunction rules than to Gentzen's own:

From $\Gamma \vdash \Theta, \phi$ and $\psi, \Delta \vdash \Phi$ to infer $(\phi \rightarrow \psi), \Gamma, \Delta \vdash \Theta, \Phi$

on the left, and

From $\phi, \Gamma \vdash \Theta, \psi$ to infer $\Gamma \vdash \Theta, (\phi \rightarrow \psi)$

on the right. Here again, an alternative, "extensional", pair of rules is possible:

From $\phi, \Gamma \vdash \Delta$ to infer $\Gamma \vdash \Delta, (\phi \rightarrow \psi)$

and

From $\Gamma \vdash \Delta, \psi$ to infer $\Gamma \vdash \Delta, (\phi \rightarrow \psi)$

on the right, and

From $\Gamma \vdash \Delta, \phi$ and $\psi, \Gamma \vdash \Delta$ to infer $(\phi \rightarrow \psi), \Gamma \vdash \Delta$

on the left. Once again, in the presence of all the structural rules, the two rule-pairs are equivalent.

Other choices and combinations of rules can "absorb" some of the functions of the structural rules, making possible variant sequent calculi in which the use of some structural rules is largely or completely avoided. Thus, for example, if the axioms are allowed to contain extraneous formulas (that is, to take the form $\phi, \Gamma \vdash \Delta, \phi$ instead of just $\phi \vdash \phi$), applications of Thinning can be avoided. In what is called the "Ketonen" version of LK (Ketonen, 1944), these generalized axioms are combined with what we have called above the "extensional" forms of rules with two premises and the "intensional" forms of the rules with a single premise. In the propositional case, this permits proofs of classically valid sequents containing no applications of Thinning or Contraction: the system, therefore, makes

[48]The Relevance Logic tradition also calls intensional disjunction and conjunction *fission* and *fusion*, respectively.

searching for proofs comparatively easy. (Curry, 1963, §5C–5E) gives detailed coverage of this and some other variant forms of sequent calculus, including systems for classical logic with single formulas in the succedent and for intuitionistic logic with multiple succedent formulas: an additional rule is needed in one case, restrictions on the rules in the other.

Other modifications lead to a simplification of the structure of sequents. Among Gentzen's rules, only those for negation and the conditional move material from one side of the sequent to the other.[49] By changing the rules for these connectives we can have a variant of LK in which everything stays on its own side of the fence. In detail, this involves four changes to Gentzen's system. First, keeping his rules for operators other than negation and the conditional, we add negative rules, analogous to the negative introduction and elimination rules of Fitch's version of natural deduction. Thus, for example, Negated Disjunction left,

From either $\neg\phi, \Gamma \vdash \Delta$ or $\neg\psi, \Gamma \vdash \Delta$ to infer $\neg(\phi \vee \psi), \Gamma \vdash \Delta$

and Negated Disjunction right,

From the two premises $\Gamma \vdash \Delta, \neg\phi$ and $\Gamma \vdash \Delta, \neg\psi$ to infer $\Gamma \vdash \Delta, \neg(\phi \vee \psi)$,

treat negated disjunctions like the conjunctions to which they are classically equivalent, and similarly for Negated Conjunction, Negated Existential Quantifier and Negated Universal Quantifier rules. Second, add Negated Negation (double negation) rules:

From $\phi, \Gamma \vdash \Delta$ to infer $\neg\neg\phi, \Gamma \vdash \Delta$

and

From $\Gamma \vdash \Delta, \phi$ to infer $\Gamma \vdash \Delta, \neg\neg\phi$.

Third, add "extensional" rules, both positive and negative, for the conditional, treating $(\phi \rightarrow \psi)$ like the classically equivalent $(\neg\phi \vee \psi)$. Finally, take as axioms (all with α atomic: corresponding forms with complex formulas are derivable) not just Gentzen's original $\alpha \vdash \alpha$ but also (n) $\neg\alpha \vdash \neg\alpha$, (em) $\vdash \alpha, \neg\alpha$ and (efq) $\alpha, \neg\alpha \vdash$.[50] The new rules are of forms parallel

[49] Which makes it easy to see that Cut-free proofs have a certain strengthening of the subformula property. Define *positive* and *negative* subformulas by simultaneous induction: Every formula is a positive subformula of itself, if it is a quantification, conjunction or disjunction its positive (negative) subformulas also include the positive (negative) subformulas of its instances, conjuncts and disjuncts respectively, if it is a negation its positive (negative) subformulas include the negative (positive) subformulas of the negated formula, and if it is a conditional its positive (negative) subformulas include the positive (negative) subformulas of its consequent and the negative (positive) subformulas of its antecedent. Then, in a Cut-free proof, every formula occurring in the antecedent of some sequent of the proof is either a positive subformula of a formula in the antecedent of the final sequent or a negative subformula of a formula in the succedent of the final sequent, and every formula in the succedent of a sequent of the proof is either a positive subformula of a succedent formula of the final sequent or a negative subformula of an antecedent formula of the final sequent.

[50] Leave out the axioms of form (em) and you get an L-system for Kleene's (strong) three-valued logic (Kleene (1952)). Leave out the axioms of form (efq) and you get a formulation of Graham Priest's three-valued "Logic of Paradox" (Priest (1979)). Leave out axioms of both these forms and you get a formulation of Anderson and Belnap's First Degree Entailment (Anderson and Belnap, 1975).

to inferences already available in LK, so essentially Gentzen's original reasoning proves the eliminability of Cut: not just standard Cut, but also "Right Handed Cut" (from $\Gamma \vdash \Delta, \phi$ and $\Phi \vdash \Psi, \neg\phi$ to infer $\Gamma, \Phi \vdash \Delta, \Psi$ and "Left Handed Cut" (from $\phi, \Gamma \vdash \Delta$ and $\neg\phi, \Phi \vdash \Psi$ to infer $\Gamma, \Phi \vdash \Delta, \Psi$) [51] Cut free proofs have an appropriately weakened subformula property.

Once the rules have been modified so that nothing ever changes side, it is possible to have proofs in which every sequent has one side empty: a logical theorem can be given a proof in which no sequent has any antecedent formulas, and a refutation—showing that some collection of formulas is contradictory—can be formulated as a proof with no succedent formulas. Both of these possibilities have been exploited in defining interesting and widely-studied systems. The deductive apparatus considered in Schütte (1960) , and in the large volume of work influenced by Schütte, is a right-handed sequent calculus (with its succedent-only sequents written as disjunctions).

3.7 Sequent Calculi and Tableaux

Left-handed (and so refutational) sequent calculi—in somewhat disguised notation and under another name—have proven very popular: they may now be the deductive systems most often taught to undergraduates! Beth (1955) developed his systems of semantic tableaux as a simplification of sequent calculi.[52] Actually writing out proofs in Gentzen's L systems is laborious, in part because the "inactive" formulas in sequents (the ones represented by the Greek capitals in the rule schemas) have to be copied from one line of the proof to the next, with the result that a proof is likely to contain many copies of such formulas. One way of looking at Beth's systems is that they optimize Gentzen's LK (which, after all, was originally introduced only for theoretical purposes) for use by incorporating notational conventions that minimize this repetition. This is perhaps easiest to explain using a modern version of tableaux.[53] A tableau, in this formulation, is a tree of formulas, constructed in accordance with certain rules, and constitutes a refutation of a finite set of formulas. To construct a propositional tableau, start with the set of formulas to be refuted, arranged in an initial, linear, part of the tree. One then extends the tree by applying, in some order, branch extension rules to formulas occurring on the tree, "checking off" (notationally, writing a check beside) a formula when a rule is applied to it. A branch of the tree (i.e., a

[51] Right Handed Cut is a derivable rule of both LJ and LK: use standard Cut and provable sequents of the form $\phi, \neg\phi \vdash$. Left Handed Cut is similarly derivable in LK. Addition of Left Handed Cut as a new rule to LJ would yield a formulation of Classical logic.

[52] Beth describes a tableau as a "natural deduction" proof, but we find the analogy with sequent calculi clearer, although given the close relationship between the two kinds of system, the question may be moot. (Recall from our discussion above in §2 that many logicians use the term 'natural deduction' in a broad sense, thereby including sequent calculi as well as the systems we are calling natural deduction.)

[53] Beth's original notation for tableaux is a bit clumsy. A variety of improvements were defined by several logicians in the 1960s: the best and most influential seems to be due to Smullyan (1968) and Jeffery (1967). For more on the history of tableaux, mentioning especially the role Jaakko Hintikka (1953; 1955a; 1955b) might have played in developing it, see Anellis (1990)

maximal part of the tree linearly ordered by the tree partial ordering) is closed if it contains some formula and its negation; otherwise it is open. A tableau is closed, and constitutes a refutation, just in case all its branches are closed. The branch extension rules are analogous to the (positive and negative) elimination rules of natural deduction:

- if a conjunction-like formula (i.e. a conjunction, negated disjunction, or negated conditional) occurs in the tree and is not yet checked off, extend each open branch of the tree passing through it by adding its conjuncts (or the negations of the disjuncts, or the antecedent and the negation of the consequent, as the case may be) and check the formula off
- if a disjunction-like formula (i.e. a disjunction, conditional, or negated conjunction) occurs unchecked, split each open branch passing through it in two, adding one disjunct to each of the new branches so formed (or: adding the consequent to one and the negation of the antecedent to the other, or: adding the negation of one conjunct to each) and check the formula off
- if a double negation occurs unchecked, add the un-doubly negated formula to each open branch passing through it and check the double negation off.

The finished product, a closed tableau, is a much "skinnier" array of symbols than the corresponding LK proof would be, but the logic is essentially the same. To see this, consider a record of the stages of the construction of a closed tableau:

- for the starting point, write out the formulas the tableau starts with horizontally, separated by commas,
- for any stage of construction involving a non-splitting rule, write, above the top of the representation of any branch of the tableau affected, the list of the formulas remaining unchecked on that branch, horizontally and separated by commas,
- for any stage of construction involving a splitting rule, write side-by-side, above the top of the representation of any branch affected, the lists of unchecked formulas on each of the branches introduced by the splitting rule.

The result (ignoring Permutation) will be a proof in a left-handed sequent calculus, using Ketonen forms of the rules and axioms with extra formulas! The tree structure of the two proofs is nearly the same (the tableau's tree has extra nodes if it starts with more than one formula and in places where a conjunction-like rule has added two formulas to a branch), but where the tableau has a single formula at a node of the tree, the sequent proof has a (left-handed) sequent made up of the formulas remaining unchecked on the branch when that node was reached in the process of tableau construction.

Tableaux with quantifiers can be construed as similarly abbreviated presentations of sequent calculus proofs, but with an additional modification to Gentzen's rules. The branch extension rule for existential and negated universal quantifications

- if $\exists x Fx$ (or $\neg\forall x Fx$) occurs unchecked, extend each branch through it by adding Fa (or $\neg Fa$), where a is a parameter (=free variable, instantial term, dummy name) that does not yet occur on the branch and check the quantified formula off,

presents no problems: it corresponds exactly to the usual sequent calculus rule for introducing an existential quantifier (negated universal quantifier) on the left, with the proviso that the parameter be new corresponding to the condition on variables in this rule. The branch extension rule for universal and negated existential quantifications, on the other hand, introduces something new:

- if $\forall x Fx$ (or $\neg\exists x Fx$) occurs, then, for any term t occurring free in a formula on the branch through it (and for an arbitrarily chosen parameter if there are none) extend any branch through it by adding Ft (or $\neg Ft$), and do not check the quantified formula off.

The quantified formula, in other words, "stays active": at a later stage in the construction of the tableau (after, perhaps, new terms have appeared through the use of the existential/negated universal rule) we can re-apply the rule to add a new instance of the same quantified formula. Thinking of the finished tableau as an abbreviated presentation of a sequent proof with the sequents containing the formulas active at a given stage of tableau construction, this corresponds to a new rule:

From $\Gamma, Ft, \forall x Fx \vdash$ to infer $\Gamma, \forall x Fx \vdash$
(and from $\Gamma, \neg Ft, \neg\exists x Fx \vdash$ to infer $\Gamma, \neg\exists x Fx \vdash$)

and the "extra formulas" in the axiom sequents at the top of the sequent proof will include all the universal (and negated existential) quantifications occurring in sequents below them. This variant form of the quantifier rules, however, is like the Ketonen variant of the propositional rules in "absorbing" some of the functions of structural rules into rules for the logical operators, and so simplifying the search for proofs. Often, in proving a sequent containing quantified formulas in versions of sequent calculus with rules more like Gentzen's it is necessary to apply a quantifier rule (universal quantifier on the left, existential quantifier on the right) more than once, generalizing on different free variables, to produce duplicate copies of the same quantified formula, with the duplication then eliminated by use of Contraction.[54] The variant quantifier rules given in this paragraph get the effect of this sort of use of Contraction—contracting, as it were, the newly inferred quantified formulas into the spare copies already present—and allow proofs in which Contraction is not otherwise used.

That comparison with (single-sided) sequent calculi seems to us the most informative way of comparing tableaux to other systems, but—illustrating again the close relationship

[54]This reflects a deep, essential, property of First-Order Logic that shows up in different ways in different deductive systems: it is responsible for the undecidability of First-Order Logic, in the sense that provability without use of Contraction in standard LK is a decidable property. To see a simple example, write out a proof of $\forall x(Fx \wedge \neg\forall y Fy) \vdash$.

between natural deduction and sequent calculi—a tableau can also be seen as a special sort of natural deduction derivation. Textbook versions of tableaux often include the instruction that some special symbol be written at the bottom of each closed branch to mark its closed status. We can pretend that this symbol is a propositional *Falsum* constant $\bot$, and then interpret a closed tableau (which is, after all, a *refutation* of the tableau's initial formula(s)) as a derivation of $\bot$ from the set of formulas the tableau started with. Some of the branch extension rules for tableau construction are virtually identical to standard natural deduction rules: the rules for conjunctions and Universal Quantifications simply extend branches by adding formulas derivable by Conjunction Elimination and Universal Quantifier Elimination. The, formally parallel, branch extension rules for negated disjunctions, negated conditionals and negated Existential Quantifications correspond in the same way to natural deduction inferences by the Negative Disjunction (Conditional, Existential Quantifier) Elimination rules. The branch extension rule for disjunction splits the branch: the two new sub-branches formed amount to the two subproofs of an application of Disjunction Elimination, the two formulas heading the new sub-branches being the hypotheses of the two subproofs: requiring that all branches be closed (contain $\bot$) corresponds to the requirement, for Disjunction Elimination, that the conclusion be derived in each subproof. The splitting branch extension rule for negated conjunctions corresponds to the parallel rule of Negative Conjunction Elimination, and that for conditionals to a non-standard Conditional Elimination rule treating material implications like the disjunctions to which they are semantically equivalent. Finally, the branch extension rules for Existential (and Negated Universal) Quantifications can be seen as initiating the subproofs for inferences by Existential (and Negative Universal) Quantifier Elimination: the requirement, in the branch extension rule, that the instantial term (free variable, parameter, dummy name. . .) in the "hypothesis" of the subproof be new to the branch is precisely what is needed to guarantee that any formula occurring earlier on the branch can be appealed to within the subproof. Writing in the "branch closed" symbol after a pair of contradictory formulas has appeared on a branch can be seen as a use of Negation Elimination (i.e., *ex falso quodlibet*) to infer $\bot$. In other words, allowing for the graphical presentation (which is not very explicit in showing which lines depend on hypotheses), a closed tableau is a deduction of $\bot$ from its initial formulas in a variant of natural deduction with Negative rules and a non-standard (but classically valid) rule for the conditional. A deduction, moreover, which satisfies strict normality conditions: only elimination rules are used, and Negation Elimination is used *only* to infer $\bot$.

Variant forms of tableaux are, as one should expect, known, and tableaux formulations can be given for a wide range of logics, including intuitionistic and modal logics. Fitting (2002) has given an encyclopedic survey of such systems. Toledo (1975) has used tableaux, rather than Schütte (1960)'s own sort of sequent calculus, in expounding much of the content of Schütte (1960).[55]

[55]Toledo's book is based on the first edition of Schütte's, and so includes material—for instance, on Ramified Type Theory and on the Type-free logic of Ackermann—that was dropped from the second

The popularity of tableaux in recent undergraduate textbooks is, in the opinion of the present writers, unfortunate. By casting all proofs into the form of refutations and then putting refutations into the specialized form of tree diagrams they obscure the relation between formal derivations and naturally occurring, informal, intuitive deductive argument. This surely detracts from the philosophical value of an introductory logic course! It might be forgivable if tableau systems had outstanding practical virtues, but in fact they (like any system permitting only normal or Cut-free derivations) are, in application to complex examples, very inefficient in terms of proof size. Their academic advantage seems to be in connection with assessment: proof-search with tableaux is routine enough, and tableaux for very simple examples small enough, that a reasonable proportion of undergraduates can be taught to produce them on an examination.

3.8 Natural Deduction with Sequences

It is also possible to have a sort of hybrid system, using disjunctively-interpreted sequences of formulas as the lines of a proof, with structural rules to manipulate them, as in a right-handed sequent calculus, but using both introduction and elimination rules, as in standard natural deduction. In such a system, Disjunction Elimination literally breaks a disjunction down into its component disjuncts: the rule licenses the inference from something like $\Gamma, (\phi \vee \psi), \Delta$ to $\Gamma, \phi, \psi, \Delta$. Such systems have been proposed by Boričić (1985) and Cellucci (1987, 1992). These systems are typically called systems of *sequent natural deduction*, but they are radically unlike the systems described in §3.3. The systems of that section are simply reformulations of standard natural deduction systems, using the antecedent formulas of a sequent to record what hypotheses the single succedent formula depends on. In contrast, the systems described here use sequents with no antecedent formulas (dependence on hypotheses is represented in Gentzen's way, by writing proofs in tree form), but they exploit the disjunctive interpretation of multiple succedent formulas in the way LK does. Thus, to use a familiar illustrative example, start with the hypothesis $\forall x(P \vee Fx)$, thought of now as a single-formula sequent. Infer $(P \vee Fa)$, another single-formula sequent, by Universal Quantifier Elimination. Now the novel step: Disjunction Elimination allows us to infer the two-formula sequent P, Fa. The parameter a occurs only in the second formula, and does not occur in the hypothesis from which this sequent is derived, so Universal Quantifier Introduction allows us to infer the $P, \forall xFx$, and Disjunction Introduction (twice, with a Contraction) yields $(P \vee \forall xFx)$ as a final, single-formula, sequent.

The system shares with sequent calculi the laborious-seeming feature that "inactive" formulas have to be copied repeatedly from one line to the next, but Cellucci's system is very efficient: on many test problems (though there are exceptions) a proof in Cellucci's system is smaller (contains fewer symbols) than one in, for example, Fitch's version of natural deduction. This is partly due to careful choice of rules: Cellucci's Conjunction Introduction, for example, unlike Gentzen's Conjunction Right rule for LK (but like the

edition and the English translation based on the second edition.

alternative rule for "intensional" conjunction), allows different sets of inactive formulas in the two premises; the total rule-set is perhaps comparable to Ketonen's (1944) version of sequent calculus. (Boričić's system, with a different choice of rule formulations, seems to be much less efficient.) For further efficiency, Cellucci considers quantifier rules like the EI described in §2.6. The whole package is perhaps the best available version of natural deduction for anyone intending actually to write out large numbers of serious complex derivations (particularly since even minimal editing computer software can automate the task of recopying inactive formulas)!

For theoretical purposes, system of this sort can be as well-behaved as any natural deduction or sequent calculi: Cellucci has proved normalization theorems for them.

3.9 Size of Normal Proofs

Normalizing a natural deduction proof (or eliminating the applications of Cut from a sequent calculus proof) *simplifies* it in an obvious intuitive sense: it gets rid of detours. Why, one might wonder, would one ever want to make use of a *non-normal* proof? Why prove something in a more complicated manner than necessary? The surprising answer is that detours can sometimes be shortcuts! Getting rid of a detour sometimes makes a proof smaller (in terms of number of lines, or number of symbols), but not always. If a detour involves proving a Universal Quantification, for example, which is then used several times to give a series of its instances by Universal Quantifier Elimination, normalizing will yield a proof in which the steps used in deriving the universal formula will be repeated several times, once for each instance. The resulting proof will still be a simplification—the repeated bits will be near duplicates, differing only in the term occurring in them, and all will be copies of the Universal Quantifier Introduction subproof with new terms substituted for its proper parameter—but it may be larger, and indeed much larger.[56] A simple example illustrates the problem. Consider a sequence of arguments, Argument$_0$, Argument$_1$, Argument$_2 \cdots$. For each n, Argument$_n$ is formulated in a First-Order language with one monadic predicate, F, and $n+1$ dyadic predicates, $R_0, \cdots, R_n$, and $n+1$ premises. All share the common premise

$$\forall x \forall y((R_0(xy) \wedge Fx) \to Fy).$$

For each k, Argument$_{k+1}$ adds the premise

$$\forall x \forall y(R_{k+1}(xy) \to \exists s \exists t(R_k(xs) \wedge R_k(st) \wedge R_k(ty)))$$

to the premises of the preceding argument. The conclusion of Argument$_n$ is[57]

$$\forall x \forall y((R_n(xy) \wedge Fx) \to Fy).$$

[56] Toledo (1975) discusses this and shows that Cut elimination can be thought of as "repetition introduction."

[57] So Argument$_0$ is a trivial argument, with the conclusion identical to the premise.

(To put it in English, the first premise says that the property of F-ness is hereditary along the R_0 relation and the rest say that any two objects related by one of the "higher" relations are connected by a chain of objects related by "lower" relations: note that the length of this chain roughly triples for each succeeding relation. The conclusion is that F-ness is hereditary along R_n.)

A normal proof of the conclusion of Argument$_n$ from its premises will have an Existential Quantifier Elimination subproof for each object in the chain of R_0-related objects connecting a pair of R_n-related objects, applying the common premise in each one to establish that each object in the chain is an F-object. Since the length of the chain increases exponentially as one moves to higher relations[58], and the number of lines in the normal proof increases exponentially as one goes to later arguments in the series. In contrast, if we allow "detours" the length of proofs is much more civilized. To prove the conclusion of Argument$_n$ for any large n, one should first use the first two premises (the premises, that is, of Argument$_1$) and prove (using two Existential Quantifier Elimination subproofs) the conclusion of Argument$_1$:

$$\forall x \forall y((R_1(xy) \wedge Fx) \rightarrow Fy)$$

This "lemma" has the same syntactic form as the common premise, but it is a distinct formula, and deriving it is a detour and a conceptual complication in the proof: it is not a subformula of any premises or of the conclusion of Argument$_n$ (for $n > 1$). Since it is of the same form as the common premise, however, an entirely similar series of steps leads from it and the premise

$$\forall x \forall y(R_2(xy) \rightarrow \exists s \exists t(R_1(xs) \wedge R_1(st) \wedge R_1(ty)))$$

(i.e., the new premise first added in Argument$_2$) to obtain the conclusion of Argument$_2$,

$$\forall x \forall y((R_2(xy) \wedge Fx) \rightarrow Fy)$$

And we continue in this way until we reach the conclusion of Argument$_n$. The number of lines of the proofs obtained in this way for the different Arguments$_k$ grows only linearly in k. Very quickly[59] we reach arguments that can be shown valid by formal deduction if we use lemmas but would be humanly impossible if normal proofs were required.[60]

Since the 1970s there has been much theoretical investigation of the comparative efficiency of different kinds of logical proof-procedure,[61] and the topic is linked to one of

[58]Looking at the $(n+1)^{st}$ case, $R_n(xy)$, we can see that the chain of R_0-related objects starting with x and ending with y is $(3^n) + 1$ objects long.

[59]The normal proof is longer than the lemma-using proof from around Argument$_3$ on. For Argument$_3$, each method—with reasonably-sized handwriting—requires around a half-dozen pages. After that, each succeeding argument requires another two pages with "detours" and about three times as many with normal proofs.

[60]The example is adapted from Hazen (1999); it was inspired by the example in Boolos (1984), a very readable paper strongly recommended to anyone wanting further discussion of the issue.

[61]For an introduction, see the review article Urquhart (1995).

the most famous open questions of contemporary mathematics and theoretical computer science: if there is a proof procedure in which any valid formula of classical propositional logic containing n symbols has a proof containing less than $f(n)$ symbols, where f is a polynomial function, then NP=co-NP. In this rarified theoretical context, two systems are counted as tied in efficiency if each can p-simulate the other: if, that is, there is a polynomial function g such that, for any proof containing n symbols of a valid formula in one there is a proof of the same formula in the other containing less than $g(n)$ symbols. By this standard, all typical natural deduction systems are equivalent to each other and to axiomatic systems.[62] On the other hand, as the example above suggests, systems restricted to *normal* proofs—tableaux, sequent calculi without cut—have been proven *not* to p-simulate natural deduction.

The differences in efficiency between systems that can p-simulate each other, though irrelevant to abstract complexity theory, can be practically important. Quine's proof of the soundness of his formulation of First Order Logic Quine (1950a), using an Existential Instantiation rule, suggests an algorithm for translating proofs in his system into proofs in a system with something like Gentzen's Existential Quantifier Elimination with a theoretically trivial increase in the number of lines: multiplication by a constant factor. Since the factor is > 2, however, the difference in efficiency can be of considerable practical relevance.

4 Problems and Projects

Our investigations have in passing brought up a number of topics that deserve further study, and in this section we discuss them with an eye to presenting enough information about each so that the motivated reader would be equipped to pursue it.

4.1 The Concept of Natural Deduction, Some Further Informal Thoughts

In §§2.1 and 2.5 we discussed "the wider notion of natural deduction." We discussed the sorts of features that are associated—or thought to be associated—with natural deduction, and we argued that none of these informal features were either necessary or sufficient for a natural deduction system—even though they are generally found in the elementary logic textbooks.

One aspect that is often thought to be central to natural deduction is that it doesn't have axioms from which one reasons towards a conclusion. Rather, the thought goes, natural deduction systems are comprised of sets of *rules*. Now, we have pointed out that very many systems that we are happy to call natural deduction in fact *do* have axioms.

[62]This result—that the proof procedures for propositional logic of Frege and Gentzen and the myriad variants of each are, up to p-simulation, of equivalent efficiency—seems to be due to Robert Reckhow (see Cook and Reckhow, 1979). The equivalence is not obvious: although the algorithm for converting natural deduction proofs into axiomatic ones suggested by textbook proofs of the deduction system seems to lead to exponential increases in proof size, more efficient algorithms are available.

But even setting that thought aside, the mere fact that a logical system contains only rules does not on its own confer the title of "natural deduction". And it is here, we think, that some scholars of logic in antiquity have been misled. Łukasiewicz (1951) reconstructed the Aristotelian syllogistic by means of an axiomatic system, but many researchers have thought that the particular axioms Łukasiewicz employed were alien to Aristotle. This has led other scholars to try a reconstruction along some other lines, and many of them claim to discern "natural deduction" in the syllogism... on the grounds that it is merely a set of rules. (See Corcoran, 1972, 1973, 1974; Martin, 1997; Andrade and Becerra, 2008.) However, these reconstructions do not contain any way to "make an assumption and see where it leads"; instead, they just apply the set of rules to premises. Those of us who see the making of assumptions and then their discharge as being crucial to natural deduction will not wish to have Aristotelian syllogistic be categorized as natural deduction. On the other hand, there is the metatheoretic aspect of Aristotelian logic, where he shows that all the other valid syllogisms can be "reduced" to those in the first figure. In doing this, Aristotle makes use of both the method of *ecthesis*—which seems to be a kind of $\exists E$, with its special use of arbitrary variables (see Smith, 1982)—and Reductio ad Absurdum. There is a reading of Aristotle where his use of these rules seems to involve assumptions from which conclusions are drawn, leading to new conclusions based on what seems to be an embedded subproof. So, the background metatheory maybe could be thought to be, or presuppose, some sort of natural deduction framework.

We now turn our attention to some other general features of various systems of logic (or rather, features often claimed to belong to one or another type of system). For example as we noted in §2.1, natural deduction systems of logic are said to be "natural" *because* the rules "are intended to reflect intuitive forms of reasoning" and *because* they "mimic certain natural ways we reason informally." In fact, though, research into the psychology of reasoning (e.g., Evans et al., 1993; Manktelow, 1999; Manktelow et al., 2010) has uniformly shown that people do not in general reason in accordance with the rules of natural deduction. The only ones of these rules that are clearly accepted are MP (from $\phi, \phi \supset \psi$ infer ψ), Biconditional MP (from $\phi, \phi \equiv \psi$ infer ψ, and symmetrically the reverse), $\wedge I$, and $\wedge E$.

The majority of "ordinary people" will deny $\vee I$. Now, it is probably correct that their denial is due to "pragmatic" or "conversational" reasons rather than to "logical" ones. But this just confirms that $\vee I$ is *not* "pretheoretically accepted." Study after study has shown that "ordinary people" do not find MT (from $\neg\psi, \phi \supset \psi$ infer $\neg\phi$) any more logically convincing than denying the antecedent or affirming the consequent ($\neg\phi, \phi \supset \psi$ infer $\neg\psi$; $\psi, \phi \supset \psi$ infer ϕ). Similar remarks hold for $\vee E$, Disjunctive Syllogism (from $\neg\phi, \phi \vee \psi$ infer ψ), and Reduction ad Absurdum (in any of its forms).

So, maybe the more accurate assessment would be what Jaśkowski and Gentzen originally claimed: that natural deduction (as developed in elementary textbooks) is a formalization of certain informal methods of reasoning employed by *mathematicians*... and not of "ordinary people."

Another claim that used to be more common than now is that natural deduction is the easiest logical method to teach to unsophisticated students. This usually was interpreted as saying that natural deduction was easier than "the logistic method" (axiomatic logic). Nowadays it is somewhat more common to hear that tableaux methods are the easiest logical method to teach to unsophisticated students. Now, Tableaux are *intrinsically* simpler, as shown in §3.7 where they are presented as a special case of natural deduction derivations, so it would not be surprising if they were more easily learned. But so far as we are aware, there has been no empirical study of this, although anecdotal evidence does seem to point in that direction. (It is also not so clear how to test this claim; but it would be quite a pedagogically useful thing to know, even though ease of learning is not the only consideration relevant to the choice of what to cover in a logic course. Should we stop teaching natural deduction in favor of tableaux methods, as seems to be a trend in the most recent textbooks?)

And again, axiomatic systems are usually alleged to be more suited to metatheoretic analysis than are natural deduction systems (and also tableaux systems?). Gentzen agreed with this in his opening remarks of Gentzen (1934):

> ... The formalization of logical deduction, especially as it has been developed by Frege, Russell, and Hilbert, is rather far removed from the forms of deduction used in practice in mathematical proofs. Considerable formal advantages are achieved in return.

Even in some textbooks that teach natural deduction we will find a shift to an axiomatic system in the proof of soundness and completeness (for instance, in Kalish and Montague, 1964; Thomason, 1970; Kalish et al., 1980, where reference is made to "equivalent axiomatic systems"). But we might in fact wonder just how much easier it really is to prove that (say) propositional *Principia Mathematica* with its five axioms and two rules of inference is complete than is a propositional natural deduction system with nine rules of inference. (And if we wanted really to have a "fair" comparison, the two systems ought to have the same connectives.) Or for that matter, is it really easier to show that five axioms are universally true and that $\supset E$ plus Substitution are truth-preserving than it is to show that some nine rules "preserve truth, if all active assumptions are true"? It would seem that before such claims are asserted by teachers of elementary logic on the basis of their intuitions, there should be some controlled psychological/educational studies to gather serious empirical data relevant to the claim. Again, so far as we are aware, this has not been done.

4.2 Natural Deduction and Computers

As mentioned in §2.2, resolution is the standard proof method employed in computer-oriented automated proof systems. There have been some automated theorem proving systems that employ natural deduction techniques; however, some of the systems, especially the early ones, that called themselves natural deduction (e.g., Bledsoe, 1971, 1977; Nevins, 1974; Kerber and Präcklein, 1996) were really various different sorts of "front end programs" that had the effect of classifying problems as being of one or another type, breaking them down into simpler problems based on this classification, and then feeding the various simpler problems to standard resolution provers. Still, there have been some more clearly natural deduction automated theorem provers, such as Li (1992); Pelletier (1998); Pollock (1992); Sieg and Byrnes (1998); Pastre (2002).

In the mid-1990s, a contest pitting automated theorem proving systems against one another was inaugurated.[63] Although various of the natural deduction theorem proving systems have been entered into the contests over the last 15 years, none of them has come close to the performance of the highest-rated resolution-based provers. In the most recent contest in which a natural deduction system was entered (2008)[64], it came in second-to-last in a field of 13 for the competition in which it would be expected to show its strength best: the first-order format (FOF) category which has no normal forms that might give resolution systems an edge. Even though all the resolution-based systems had to first convert the problems to clause form before starting their proofs, they managed to win handily over MUSCADET. For example, of the 200 problems to be solved within the 300 second time window, the winning system solved 169 whereas MUSCADET solved only 38. Despite the fact that MUSCADET solved some problems that none of the other systems could solve, it seems that the lesson learned from computerized chess also applies to theorem proving: systems that follow human patterns of heuristic reasoning about solving problems cannot successfully compete against brute force algorithms that employ massive amounts of memory and extremely fast deep and broad search mechanisms. (See Slate and Atkin, 1977, for a typical account of this view in the realm of chess). It just may be that the goal set by the founders of natural deduction, and its many followers, of presenting logical proofs "in the way that ordinary mathematicians construct their proofs" is not really the most effective way to prove logic problems, even if it is the way that mathematicians proceed when they are proving mathematical problems. (A different rationale for the relatively poor showing of natural deduction systems over the last 15 years that direct comparisons have been made has pointed to the amount of effort that has been poured into resolution theorem proving since the mid-1960s as compared to natural deduction. See for instance Pelletier, 1998, p.33.)

Of course, there have all along been computer programs that were designed to help elementary logic students learn natural deduction. Probably the earliest of these were the suite developed under the guidance of Patrick Suppes and designed to help school students learn "the new math" of the 1960s, partially summarized in Goldberg and Suppes (1972) for learning natural deduction. In more recent times, it has become more or less of a requirement that a beginning natural deduction logic textbook have some computer assistance, and this includes a "proof assistant" that will help the student in the construction

[63] CASC: the CADe ATP System Competition—CADe stands for "Conference on Automated Deduction" and ATP stands for "automated theorem proving". CASC in general is described in Sutcliffe and Suttner (2006); Pelletier et al. (2002); the most recent contest (2009) is reported in Sutcliffe (2009). See also the website `http://www.cs.miami.edu/~tptp/CASC`.

[64] The system is called MUSCADET; see Pastre (2002).

of proofs in the chosen natural deduction system. Some of the more long-standing systems that are associated with textbooks are BERTIE3, DAEMON, PLATO, SYMLOG, PANDORA, FITCH, INFERENCE ENGINE, JAPE (see Bergmann et al., 2008; Allen and Hand, 2001; Bonevac, 1987; Portoraro, 1994; Broda et al., 1994; Barwise and Etchemendy, 2002; Bessie and Glennan, 2000; Bornat, 2005 respectively)[65]. The student constructs a proof as far as s/he can, and the assistant will help by giving a plausible next step (or explain some mistake the student has made). Most of these systems also have an "automatic mode" that will generate a complete proof, rather than merely suggesting a next step or correcting previous steps, and in this mode they can be considered automated theorem proving systems along the lines of the ones just considered in the preceding paragraphs. (However, since they were not carefully and especially tuned for this role, it is easy to understand why they can be "stumped" by problems that the earlier-mentioned systems can handle). There are a number of further systems on the web which are not associated with any particular textbook: Saetti (2010); McGuire (2010); Kaliszyk (2010); Andrews (2010); Christensen (2010); Gottschall (2010); Frické (2010); von Sydow (2010).

One of the initial motivations for programming a natural deduction automated theorem proving system was to assist mathematicians in the quest for a proof of some new theorems. (Such a view is stated clearly in Benzmüller, 2006; Pastre, 1978, 1993; Siekmann et al., 2006; Autexier et al., 2008; and the history of this motivation from the point of view of one project is in Matuszewski and Rudnicki, 2005). The rationale for using the natural deduction framework for this purpose was (channeling Jaśkowski and Gentzen) that natural deduction was the way that "ordinary mathematicians" reasoned and gave their proofs. Thus, the sort of help desired was of a natural deduction type, and the relevant kinds of hints and strategic recommendations that a mathematician might give to the automated assistant would be of a nature to construct a natural deduction proof. A related motivation has been the search for "real proofs" of accepted mathematical theorems, as opposed to the "informal proofs" that mathematicians give. (In this latter regard, see the QED project, outlined in Anonymous (1994), and the earlier MIZAR project at both `http://mizar.org/project` and `http://webdocs.cs.ualberta.ca/~piotr/Mizar/` which is now seen as one of the leading forces within the QED initiative. Many further aspects are considered by Freek Wiedijk on his webpage `http://www.cs.ru.nl/~freek`, from which there are many links to initiatives that are involved in the task of automating and "formalizing" mathematics—for instance, `http://www.cs.ru.nl/~freek/digimath/index.html`). A very nice summary of the goals, obstacles and the issues still outstanding is in Wiedijk (2007), in which Wiedijk ranks the systems in terms of how many well-known theorems each has proved. The three systems

[65]The Association for Symbolic Logic maintains a website for educational logic software, listing some 46 programs (not all of which are natural deduction). See `http://www.ucalgary.ca/aslcle/logic-courseware`.

judged to be "state of the art" are HOL-systems[66], the COQ system[67], and MIZAR. Mizar history is recounted in Matuszewski and Rudnicki (2005). The crucial Mizar Mathematical Library (MML) contains the axioms of the theory (set-theoretic axioms) added to the basic underlying natural deduction proving engine, and also a number of works written using the system. These works undergo a verification of their results, extracting facts and definitions for the Library that can then be used by new submissions to the Library.

A study of the successes of giving a complete formalization of mathematics by means of derivations from first principles of logic and axioms of set theory will show quite slow progress. A number of writers have thought that the entire QED goal to be unreachable for a variety of reasons, such as disagreement over the underlying logical language, the unreadability of machine-generated proofs, the implausibility of creating a suitably large background of previously-proved theorems with the ability to know which should be used and when, and the general shortcomings of automated theorem proving systems on difficult problems. Wiedijk (2007) evaluates many of these types of problems but nonetheless ends on a positive note that the QED system "will happen earlier than we now expect…in a reasonable time".

4.3 Natural Deduction and Semantics

Natural deduction, at least for standard logics with standard operators, is, well, natural and in use in most informal mathematical reasoning. The rules are intuitive: most people, or at least, most mathematicians, after a bit of reflective thought, can come to see them as obviously valid. It is therefore plausible to think that the rules are, somehow and in some sense, closely tied to our understanding of the logical operators. Almost from the start the study of natural deduction has been accompanied by the search for some philosophical payoff from this connection. In one way or another, many philosophers have thought that the natural deduction rules for the logical operators are semantically informative: they reveal or determine the meanings of the operators. Since the rules, thought of naïvely, govern a practice—reasoning, or arguing, or proof presentation—it has seemed that a success in grounding the semantics of the operators on the rules would be a success in a more general philosophical project: that of seeking (an explanation of) meanings (of at least these bits of language) in *use*. The first published suggestion along these lines is in Gentzen (1934), where, discussing the natural deduction calculus after presenting it, Gentzen writes:

> The introductions represent, as it were, the 'definitions' of the symbols concerned, and the eliminations are no more, in the final analysis, than the con-

[66]This Higher-Order Logic family consists of three different research efforts: HOL-LIGHT, ProofPower, Isabelle-HOL. As Wiedijk points out, of these three only Isabelle-HOL employs a clearly natural deduction system. HOL-LIGHT is described in Harrison (2007); ProofPower at `http://www.lemma-one.com/ProofPower`; and Isabelle-HOL is described in Nipkow et al. (2002) and with the natural deduction language ISAR in Wenzel (2002).

[67]See `http://pauillac.inria.fr/coq/doc/main.html`.

> sequences of these definitions. This fact may be expressed as follows: In eliminating a symbol, the formula... may be used 'only in the sense afforded it by the introduction of that symbol'.[68]

Other writers have thought of the rules as *collectively* giving the meanings of the operators. Perhaps the first to bring such a proposal to the attention of a wide philosophical audience was Karl Popper, in Popper (1947a,b).

There are problems with these suggestions, but they are attractive and the problems have, at least in part, been overcome. One problem is that not every imaginable collection of rules can be taken to define a connective. The classic statement of this problem (in an era when philosophers were less verbose than now) is Prior's (1960). Prior considers an alleged connective, **tonk**, with an introduction rule allowing (A **tonk** B) to be inferred from A and an elimination rule licensing the inference of B from (A **tonk** B). Together the two rules allow any proposition to be inferred from any other: since, whatever propositions may be, some are true and others false, there can be no operation on them satisfying these rules. Prior suggests that rules cannot define a connective, but rather must be responsive to its prior meaning. The equally classic reply is Belnap (1962). Definitions, no less than assertions, are subject to logical restriction, as shown by the familiar fallacious "proofs" that smuggle in an illicit assumption by concealing it in the definition of an unusual arithmetic operator. The fundamental restriction, violated by Prior's **tonk**, is that a definition should be non-creative: should not allow us to deduce conclusions which could not have been derived without using the defined expression. In the particular case of definitions of logical operators by introduction and elimination rules, Belnap gives a precise formulation of this restriction: assuming a language with a notion of deduction (involving logical operators already in the language and/or non-logical inference rules) satisfying the structural rules of Gentzen's L-calculus, the result of adding the new operator with its rules must be a conservative extension: it should yield no inference, whose premises and conclusion are stated in the original language, which could not be derived in the original system. A sufficient condition for this is that Gentzen's rule Cut should be admissible in the extended calculus.[69]

If Prior's objection is that rules can do too much, other challenges claim they cannot do enough: even the best pairs of introduction and elimination rules can fail to determine the meaning of the operator they supposedly define. This can happen in different ways, illustrated by two examples.

> **EXAMPLE A.** Two different notions of necessity—say, logical necessity and physical necessity—might have the same formal logic—perhaps S5. The in-

[68]This resembles Heyting's account of the meaning of intuitionistic logical operators closely enough that some writers speak of the *Gentzen-Heyting* account of the logical operators.

[69]Prior's objection applies to proposals on which the introduction and elimination rules jointly define the new operator. Gentzen's proposal, on which it is the introduction rules which are definitional and the elimination rules are required to respect the meaning so defined, seems to avoid it: Belnap's discussion can be seen as a precise working out of the details of Gentzen's sketchily-presented suggestion.

troduction and elimination rules of that logic's necessity operator will then be neutral between the two interpretations, and so cannot by themselves determine a unique meaning for the $\Box$.

EXAMPLE B. We ordinarily think of disjunction as truth-functional: a disjunction is true if and only if at least one of its disjuncts is. As Rudolf Carnap (1944) observed, however, the rules of classical logic do not require this interpretation: the entire deductive apparatus of classical logic will also be validated by an interpretation on which sentences can take values in a Boolean Algebra with more than two elements, with only those taking the top value counted as *true*. On such an interpretation a disjunction can be true even if neither of its disjuncts is: in particular, assuming negation is interpreted by Boolean complement, all instances of the Law of Excluded Middle will be true, but most will not have true disjuncts.[70]

Carnap's own response to the problem in the second example was to consider an enriched logical framework. The problem arises when a logic is understood as defining a consequence relation in the sense of Tarski (1930a,b, 1935): a relation holding between a finite set of premises and a given conclusion if and only if the inference from the premises to the conclusion is valid. It disappears if a logic is thought of as providing a multiple conclusion consequence relation: a relation holding between two finite sets of sentences just in case the truth of all the members of the first set implies that at least one of the second is true. These abstract relations relate in an obvious way to syntactic notions we have seen: an ordinary, Tarskian, consequence relation is (if we identify relations with sets of ordered pairs) a set of the sort of single-succedent sequents used in Gentzen's LJ, and the generalized consequence relation Carnap moved to is a set of generalized sequents of the sort introduced for LK.[71] If classical logic is defined as a generalized consequence relation, then it excludes non-truth-functional interpretations of the sort Carnap noted: the validity of the sequent $\phi \vee \psi \vdash \phi, \psi$ means that a disjunction can only be true if at least one of its disjuncts is. Critics, notably Church (1944), have thought this was cheating: the move to multiple-succedent sequents amounts to introducing a second notation for disjunction, *stipulating* that it is to be given a truth-functional interpretation, and then defining the interpretation of the ordinary disjunction connective in terms of it.

So far we have discussed whether the consequence relation induced by the logical rules can determine the interpretation of the logical operators. Natural deduction, with rules allowing inferences from subproofs and not merely from (n-tuples of) formulas has a richer

[70] Readers familiar with van Fraassen's notion of *supervaluations* will recognize here another consequence of the same fact about Boolean algebras.

[71] The correspondence between multiple conclusion consequence relations and semantic interpretations of logics has been widely studied since the 1970s. Such relations are often called *Scott consequence relations* in reference to Scott (1974).

structure, and we can ask whether attention to this structure can close the gap.[72] In fact it seems to come tantalizingly close. The disjunction elimination rule allows a conclusion to be inferred from a disjunction together with two subproofs, one deriving the conclusion from one disjunct and the other from the other, but we have to be careful in specifying the conditions the subproofs must satisfy. Define a subproof to be *de facto truth preserving* iff it satisfies the condition that either the conclusion derived in it is true or the hypothesis (or one of the reiterated premises appealed to in the subproof) is not true. If disjunction elimination is postulated for all de facto truth preserving subproofs, then its validity forces the truth-functional interpretation of disjunction! (The rule, with this kind of subproof allowed, would not be sound on a non-truth-functional interpretation. To see this, let $\phi \vee \psi$ be a true disjunction with two non-true disjuncts, and θ some untruth. θ is implied by $\phi \vee \psi$ on this version of the rule, since the two degenerate subproofs in which θ is directly inferred from ϕ and ψ are, trivially, de facto truth preserving: since their hypotheses are not true, they have, so to speak, no truth *to* preserve.) If, however, we require that the subproofs embody *formally* valid reasoning, disjunction elimination doesn't do any more toward ruling out a non-truth-functional interpretation of disjunction than imposing the consequence relation of classical logic does. (By the Deduction Theorem, if there are formally valid subproofs from ϕ to θ and from ψ to θ, we would have proofs of $(\phi \supset \theta)$ and $(\psi \supset \theta)$, and the rule of disjunction elimination tells us no more than that the consequence relation includes $(\phi \vee \psi), (\phi \supset \theta), (\psi \supset \theta) \vdash \theta$.) The philosophical significance of this contrast depends, obviously, on whether or not there is a principled reason for preferring one or the other class of admissible subproofs. If we see the project of defining logical operators by their rules as a contribution to a more general explication of meaning in terms of "use", however, it would seem that only rules which could be adopted or learned by reasoners are relevant. From this point of view it would seem that the restriction to formally valid subproofs is appropriate: recognizing de facto truth preserving subproofs is not something a (non-omniscient) being could learn in the way it can learn to reason in accordance with formal rules.

Philosophical discussion of the significance of Example B, and in particular of the legitimacy of Carnap's appeal to a "multiple conclusion" consequence relation, continues: cf., e.g., Restall (2005); Rumfitt (2008). At least for those who share Church's intuitions about "multiple conclusion" consequence, however, Carnap's observation would appear to set limits to what we can hope for. If you insist that semantics must include something like a model-theoretic interpretation, assigning truth values (or other values) to formulas, then their logical rules cannot fully determine the semantics of the logical operators.[73] It

[72]Readers familiar with supervaluations will have seen other examples where these rules introduce novelties. Supervaluational interpretations (of, e.g., a language with self-reference and a truth predicate) validate the same consequence relation as truth-functional, but do not validate all the subproof-using rules of classical logic.

[73]It follows from the completeness theorem that the possible interpretations of a classically consistent theory will include standard model theoretic ones, in which the connectives are truth- functional. Whether,

remains possible, however, that the logical rules determine some more general kind of meaning, or some aspect of the meaning, of the logical operators. Here, for a change, there is an impressive positive result! Think of the meanings of logical operators abstractly, without requiring any particular connection between their meanings and the assignment of truth values (or similar) to complex sentences. Say that two operators are semantically *equivalent* if any sentence formed using one is logically equivalent to the sentence formed in the same way from the same constituents using the other.[74] Then the standard introduction and elimination rules determine the meanings of the operators at least up to equivalence. Thus, for example, suppose we had two operators, $\wedge_1$ and $\wedge_2$, each governed by "copies" of the $\wedge I$ and $\wedge E$ rules. Then we could derive $\phi \wedge_1 \psi$ from $\phi \wedge_2 \psi$ and conversely. (Left to right: use $\wedge_1 E$ to get ϕ and ψ, then combine them by $\wedge_2 I$. Right to left similarly.) Parallel results hold for the other standard connectives $\vee, \supset, \neg$, in both Classical and Intuitionistic logic. Parallel results hold for the quantifiers in both logics, though here there is an added subtlety: after all, in Higher Order logics (as in other many-sorted logics), the quantifier rules are (or at least can be formulated so as to be) the same for different orders of quantification, but nobody wants to say that First and Second Order universal quantifiers are equivalent! The key here is that the rules for different quantifiers are differentiated, not by formal differences in the rule schemas, but by the classes of terms substitutable for their bound variables: in order to show two universal quantifiers to be equivalent, we must be able to instantiate each, by its $\forall E$ rule, to the parameters (free variables, arbitrary names,...) used in the other's $\forall I$ rule. Leading us, at last, back to **Example A**. If we think of modal operators as quantifiers over possible worlds, it would seem that the case of distinct modalities with formally identical worlds *ought* to be explained in a way analogous to our treatment of many-sorted quantificational logic...but the "terms" for possible worlds are "invisible." Looking back to our discussion of natural deduction formulations of modal logics (§2.8), the idea of "invisible" terms for possible worlds was used to motivate the restrictions on reiteration into modal subproofs. So we can say that for two necessity operators to have "the same" introduction and elimination rules, something more than identity of the schematic presentation of the rules is needed: the two versions of $\Box I$ must have the same restrictions on what can be reiterated into a modal subproof. And, in fact, two necessity operators governed by the $\Box I$ and $\Box E$ rules of, say, the logic T will be equivalent if formulas governed by each can be reiterated into the modal subproofs of the other.

These results—showing that Introduction and Elimination rules characterize the meanings of the logical operators up to "equivalence"—are impressive, and it is hard not to see

in all cases, any of the models mathematically possible deserve to be thought of as possible semantic interpretations, in a strong sense of semantic, is a debated philosophical question: certainly the Löwenheim-Skolem theorem and Non-Standard models of arithmetic suggest that not all models are semantically relevant.

[74]We will not discuss the question of whether equivalent operators are synonymous here. Certainly many philosophers—notably Bertrand Russell in his logical writings from the first decade of the 20th Century, e.g., (Russell, 1906, p.201)—have wanted to allow equivalence to be properly weaker than synonymy.

them as relevant to semantics, in some sense of 'semantics'. They depend on what is often referred to as *harmony* between the Introduction and Elimination rules. Avoiding Prior's problem with **tonk** required that the elimination rules for an operator not be too strong: in Gentzen's words, they must use the operator only in the sense afforded it by its introduction rules. Conversely, the uniqueness up to equivalence results require elimination rules that are strong enough: they must make full use of the sense afforded the operator by the introduction rules. Working in the context of sequent calculus, Belnap proposed a formal test of this harmony: cut must be admissible if the elimination rules are not too strong, and the identity sequents for compound formulas must be derivable when only identity sequents for atoms are taken as axiomatic if they are not too weak.

It is worth noting possible limitation on what can be achieved along these lines. Even if we restrict our attention to operators broadly analogous to standard logical ones, it is not clear that all operators can be characterized by their logical rules. The vocabulary of an intuitionistic theory of real numbers, for example, can be enriched with two different *strong (constructible) negation* operators with the same logical rules. This is possible because the purely schematic rules for Nelson's negation (Nelson, 1949), though determining how complex formulas containing negations relate to each other, do not afford a sense to the negation of an atomic formula: there is no negation introduction rule allowing the proof of a negated atom. In adding two strong negations to an intuitionistic theory, then, we would have to postulate rules giving the content of negated atoms in terms of non-logical, specifically mathematical, ideas: allowing us to infer the negation (in the sense of one strong negation) of an identity if its terms are non-identical (in the sense of ordinary Heyting negation) but to infer its negation (in the sense of the other strong negation) only if they are separated.[75] Natural deduction was first introduced for Intuitionistic and Classical logics with their standard operators, and works best on its "native turf": natural deduction techniques can be extended beyond this, but it cannot be assumed that the extension will have all thenice properties of the original. One nice property is that the natural deduction rules can, to a degree, be seen as defining the logical operators, and what we have seen is that this does not always extend, even to non-standard negation operators. The success, such as it is, of the idea that logical operators are defined by their rules has helped to inspire larger programs of *inferential semantics*[76] in the general philosophy of language. The case of strong negation can perhaps be taken as an illustration of how difficult it may be to extend this sort of treatment beyond the strictly logical domain.

The possibility of characterizing a logical operator in terms of its Introduction and Elimination rules has made possible a precise formulation of an interesting question. One of the properties of classical logic that elementary students are often told about is *functional*

[75] Formally, the negation operator of classical logic has rules extending those of both Heyting's negation and Nelson's. It is usually taken to be a logical operator generalizing Heyting's negation, but Bertrand Russell's discussion of negative facts, in his Russell (1918), suggests treating it as less than purely logical, on the analogy of Nelson's negation.

[76] Cf. Brandom (1994) and the literature it has stimulated, for example Peregrin (2008).

completeness: every possible truth-functional connective (of any arity) is explicitly definable in terms of the standard ones. The question *should* present itself of whether there is any comparable result for intuitionistic logic, but this can't be addressed until we have some definite idea of what counts as a possible intuitionistic connective. We now have a proposal: a possible intuitionistic connective is one that can be added (conservatively) to a formulation of intuitionistic logic by giving an introduction rule (and an appropriately matched, not too strong and not too weak) elimination rule for it. Appealing to this concept of a possible connective, Zucker and Tragesser (1978) prove a kind of functional completeness theorem. They give a general format for stating introduction rules, and show that any operator that can be added to intuitionistic logic by a rule fitting this format can be defined in terms of the usual intuitionistic operators. Unexpectedly, the converse seems not to hold: there are operators, explicitly definable from standard intuitionistic ones, which do *not* have natural deduction rules of the usual sort. For a simple example, consider the connective $\ddot{\vee}$ defined in intuitionistic logic by the equivalence:[77]

$$(\phi \ddot{\vee} \psi) =_{df} ((\phi \supset \psi) \supset \psi).$$

(In classical logic, this equivalence is a well-known possible definition for disjunction, but intuitionistically $(\phi \ddot{\vee} \psi)$ is much weaker than $(\phi \vee \psi)$.) The introduction and elimination rules for the standard operators of intuitionistic logic are pure, in the sense that no operator other than the one the rules are for appears in the schematic presentation of the rule, and $\ddot{\vee}$ has no pure introduction and elimination rules. (Trivially, it has impure rules: an introduction rule allowing $(\phi \ddot{\vee} \psi)$ to be inferred from its definiens and a converse elimination rule.) To get around this problem, Schroeder-Heister (1984b) introduces a generalization of natural deduction: subproofs may have inferences instead of (or in addition to) formulas as hypotheses.[78] In this framework we can have rules of $\ddot{\vee}I$ allowing the inference of $(\phi \ddot{\vee} \psi)$ from a subproof in which ψ is derived on the hypothesis that $\phi \vdash \psi$ is valid, and $\ddot{\vee}E$ allowing ψ to be inferred from $(\phi \ddot{\vee} \psi)$ and a subproof in which ψ is derived on the hypothesis ϕ. Schroeder-Heister proves that any connective characterized by introduction and elimination rules of this generalized sort is definable in terms of the standard intuitionistic connectives and vice versa. Schroeder-Heister (1984a) proves a similar result for intuitionistic logic with quantifiers.

4.4 The One True Logic? Some Philosophical Reflections

Here we very briefly survey some points at which natural deduction's format for proofs has been employed to argue for some more philosophical-logic conclusions, or at least to lend support or clarity to certain doctrines. Although we do not follow up on the literature surrounding these topics, we commend the issues to further study by those interested in natural deduction.

[77]This connective was suggested to Allen Hazen by Lloyd Humberstone.

[78]Fitch (1966) defines a similarly generalized sort of subproof, but uses it only to abbreviate proofs.

As we have mentioned at various places, the straightforward int-elim rules for natural deduction as chosen by Gentzen generate intuitionistic logic. Further additions[79] are needed to extend the rules to describe classical logic. Some might—indeed, some have—taken the simplicity and aesthetic qualities of the intuitionistic rules over the kinds of rules identified in the preceding footnote to be an argument in favor of intuitionistic logic over classical logic. (One argument among many, perhaps.[80] But some think it is quite strong.) In this sort of view, the other logics are generated for special purposes by adding some axiom or new rule of inference to show how some connective will act in the new, "unnatural" environment where it is being employed. Intuitionism becomes the Pure Logic of Inference to which special additions might be appended. Views like this are often associated with Michael Dummett (for example, Dummett, 1978), although Dummett also has further views concerning verification that make intuitionism a natural choice for a logic that correctly describes his view.

Looking more closely at the possible ways of strengthening a set of intuitionistic rules, we see that classical logic can be obtained by adding rules for negation or the conditional, and a superintuitionistic First Order logic (the logic of "constant domains") can be obtained by adding a rule for distributing the universal quantifier over disjunction. On the other hand, there is a sense in which the standard intuitionistic rules for conjunction, disjunction and the existential quantifier already imply all the classically valid rules for these operators.[81] To see this, note that Gentzen's intuitionistic sequent calculus LJ can, without strengthening intuitionistic logic, be modified to allow multiple succedent formulas as long

[79] As we said above, Gentzen added (all instances of) $(\phi \vee \neg\phi)$ as "axioms", yielding proof structures that look like the *Law of the Excluded Middle*, LEM, left-most of the following five rules. But other writers would prefer to use $\neg\neg E$ as in the second-left rule, or perhaps $\neg E$ as shown in the middle of the five rules. Or perhaps a rule that embodies "Pierce's Law", or a contraposition law in the form given rightmost:

$$\frac{\begin{array}{cc}[\phi] & [\neg\phi]\\ \vdots & \vdots\\ \psi & \psi\end{array}}{\psi}\,(\text{LEM}) \qquad \frac{\neg\neg\phi}{\phi}\,(\neg\neg E) \qquad \frac{\begin{array}{c}[\neg\phi]\\ \vdots\\ \psi\\ \vdots\\ \neg\psi\end{array}}{\phi}\,(\neg E) \qquad \frac{\begin{array}{c}[\phi\supset\psi]\\ \vdots\\ \phi\end{array}}{\phi}\,(\text{Pierce}) \qquad \frac{\begin{array}{c}[\neg\phi]\\ \vdots\\ \neg\psi\end{array}}{\psi\supset\phi}\,(\text{Contrapose})$$

and various other rules, to effect the extension to classical logic.

[80] In some theorists' minds this aesthetic argument is just a "follow-on argument" that started with the considerations we brought out in §4.3. That argument starts with the a view on how to define the "meaning" of logical connectives and culminates in the conclusion that the meanings of the natural language 'and', 'or', 'not', 'if–then–', 'if and only if' are precisely given by the int-elim rules. It is yet another step to infer from that the conclusion that intuitionism is The One True Logic. But one could. And in the course of doing so the present aesthetic consideration might be raised.

[81] Cf. Belnap and Thomason (1963); Belnap et al. (1963). The result seems at odds with the common assumption that intuitionism incorporates a distinctive, strong, "sense" of disjunction: see the discussion in Hazen (1990).

as the rules for negation, conditional and universal quantifier on the right are applied only to premise-sequents with at most a single succedent formula: the rules for conjunction, disjunction and existential quantification in the modified system are identical to those for the classical LK. This suggests that the two sets of operators are somehow different in status. Perhaps we could say that there is a *basic logic*[82] with only conjunction, disjunction and existential quantification as logical operators, and that the distinction between classical and intuitionistic logics only applies to systems extending the basic logic with other operators.

Gentzen separated the rules he described into two sorts: the logical rules and the structural rules. The former give the ways to employ the logical connectives, while the latter characterize the form that proofs may take. From the point of view of natural deduction, the way to characterize the differences between some pairs of logics—e.g., between classical logic and some particular modal logic—is to point to the existence of some new logical operators and the (int-elim) rules that govern them. The way to characterize the differences between other pairs of logics—e.g., between classical logic and intuitionistic logic—is to talk about the differences between their respective int-elim rules for the same logical operators. This brings into the foreground that there could be other pairs of logics—e.g., intuitionistic logic and some relevant logic—that differ instead in their structural rules governing proofs. And in fact, that is a fruitful way to look at a whole group of logics that were initially developed axiomatically: the various relevant logics. (See Restall, 2000). This is arguably a clearer viewpoint from which to evaluate their properties than the axiomatic versions (in, e.g., Anderson and Belnap, 1975, or the semantic characterizations (in, e.g., Routley and Meyer, 1972, 1973), or the algebraic perspectives (as lattices with residuated families of operators in, e.g., Dunn, 1993), and the existence of such a fruitful viewpoint gives some reason to appreciate natural deduction (and sequent calculi) as an approach to logic generally.

Logical pluralists (see Beall and Restall, 2006) take the view that there are many differing notions of validity of arguments. In turn, this leads them to interpret the view just outlined about the structural versus logical rules of a proof theory in such a way that the *language of logic*, including the connectives and the rules governing them, stays constant in the different views of logic, but the *fine structure of proofs* as described by the structural rules is allowed to vary from one application to another. It is this variability of the structural rules that allows for the distinctively different features of the consequence relations in the differing logics. They are all "legal" views about consequence, just employed for different purposes. Again, it is the viewpoint offered by natural deduction (and sequent calculi) that make this approach viable.

Finally, the history of thought is replete with confrontations between, on the one hand,

[82]Terminology due to Fitch: conjunction, disjunction and existential quantification are the logical operators of the systems considered in Fitch (1942) and subsequent papers. First Order logic restricted to these three basic operators has an interesting recursion-theoretic property that might make it useful in expositions of recursion theory: if the primitive predicates of such a language all express recursively enumerable sets or relations, so do all its formulas.

those thinkers who wish to explain complex phenomena in terms of the parts of the complex and the ways these parts interact with one another, and on the other hand, those thinkers who seek to explain parts of a complex system in terms of the rules that the parts play within the system. Pelletier (forthcoming), reflecting what is common language in philosophy, calls the former mode of explanation "atomism" and the latter "holism."

When applied to the topic of proof theory and model theory, some thinkers have called the atomistic method of explanation "the analytic mode" and the holistic method "the synthetic mode" (see Belnap, 1962). Applying the distinction to the logical connectives, the idea is that the analytic mode would wish to define or explain a connective—as well as correct inferences/deductions that involve that connective—in terms of antecedently given features, such as truth tables, or preservation of some property (such as truth, or truth-in-a-model, or truth-in-a-possible-world, and so on) that is thought to be antecedently known or understood. The synthetic mode, which Belnap (1962) favored over the analytic mode of Prior (1960), takes the notion of a good type of derivation as the given and defines the properties of the connectives in terms of how they contribute to these types of derivation. (See our discussion in §4.3 for Belnap's considerations.)

There has not been much interaction between the philosophical writings on the atomism/holism debate and the debate within the philosophy of logic about the analytic/synthetic modes of explanation. We think that a clearer understanding of the issue, prompted by consideration of natural deduction rules, could open the door to some fruitful interchange. It might also open some interesting interaction with the position known as structuralism in logic (e.g., Koslow, 1992), which currently seems quite out of the mainstream discussions in the philosophy of logic.

The precision possible in logical metatheory has made it an attractive laboratory for the philosophy of language: success in sustaining the claim that logical operators are defined by their rules doesn't necessarily imply that inferential semantics will succeed elsewhere, but it is the work in this area, stemming from Gentzen's suggestions, that inspires the hope that the inferentialist project can be carried out rigorously and in detail.

Acknowledgments

We are grateful for discussions, assistance, and advice from Bernard Linsky, Jack MacIntosh, Greg Restall, Piotr Rudnicki, Jane Spurr, Geoff Sutcliffe, and Alasdair Urquhart. Pelletier also acknowledges the help of (Canadian) NSERC grant A5525.

References

Adler, Jonathan and Lance Rips. 2008. *Reasoning: Studies of Human Inference and Its Foundations*. New York: Cambridge University Press.

Allen, Colin and Michael Hand. 2001. *Logic Primer*. Cambridge: MIT Press.

Anderson, Alan and Nuel Belnap. 1975. *Entailment: Vol. 1*. Princeton, NJ: Princeton University Press.

Anderson, John and Henry Johnstone. 1962. *Natural Deduction: The Logical Basis of Axiom Systems*. Belmont, CA: Wadsworth Pub. Co.

Andrade, Edgar and Edward Becerra. 2008. Establishing connections between Aristotle's natural deduction and first-order logic. *History and Philosophy of Logic* 29:309–325.

Andrews, Peter. 2010. (educational) theorem proving system. http://gtps.math.cmu.edu/tps.html.

Anellis, Irving. 1990. From semantic tableaux to Smullyan trees: A history of the development of the falsifiability tree method. *Modern Logic* 1:36–69. See also the errata in *Modern Logic* **1:** 263 and **2:** 219.

Anellis, Irving. 1991. Forty years of "unnatural" natural deduction and quantification: A history of first-order systems of natural deduction from Gentzen to Copi. *Modern Logic* 2:113–152. See also the "Correction" in *Modern Logic* (1992) **3:** 98.

Anonymous. 1994. The QED manifesto. In A. Bundy, ed., *Automated Deduction — CADE 12*, pages 238–251. Berlin: Springer-Verlag. Volume **814** of *Lecture Notes in Artificial Intelligence*.

Arthur, Richard. 2010. *Natural Deduction: An Introduction to Logic*. Belmont, CA: Wadsworth Pub. Co.

Autexier, Serge, Christoph Benzmüller, Dominik Dietrich, and Marc Wagner. 2008. Organisation, transformation, and propagation of mathematical knowledge in ΩMEGA. *Mathematics in Computer Science* 2:253–277.

Barker, Stephen. 1965. *Elements of Logic*. NY: McGraw-Hill.

Barwise, Jon and John Etchemendy. 2002. *Language, Proof and Logic*. Stanford: CSLI Press.

Beall, J.C and Greg Restall. 2006. *Logical Pluralism*. Oxford: Oxford University Press.

Belnap, Nuel. 1962. Tonk, plonk and plink. *Analysis* 22:130–134.

Belnap, Nuel, Hugues Leblanc, and Richmond Thomason. 1963. On not strengthening intuitionistic logic. *Notre Dame Journal of Formal Logic* 4:313–320.

Belnap, Nuel and Richmond Thomason. 1963. A rule-completeness theorem. *Notre Dame Journal of Formal Logic* 4:39–43.

Benzmüller, Christoph. 2006. Towards computer aided mathematics. *Journal of Applied Mathematics* 4:359–365.

Bergmann, Merrie, James Moor, and Jack Nelson. 2008. *The Logic Book, Fifth Edition*. New York: Random House.

Bessie, Joseph and Stuart Glennan. 2000. *Elements of Deductive Inference*. Belmont, CA: Wadsworth Pub. Co.

Beth, Evert Willem. 1955. Semantic entailment and formal derivability. *Mededlingen van den Koninklijke Nederlandse Akademie van Wetenschappen* 18:309–342. Reprinted in J. Hintikka (ed.) (1969) *The Philosophy of Mathematics*, Oxford UP, pp. 9–41.

Bledsoe, Woody. 1971. Splitting and reduction heuristics in automatic theorem proving. *Artificial Intelligence* 2:55–78.

Bledsoe, Woody. 1977. Non-resolution theorem proving. *Artificial Intelligence* 9:1–35.

Bonevac, Daniel. 1987. *Deduction*. Mountain View, CA: Mayfield Press.

Boolos, George. 1984. Don't eliminate cut. *Journal of Philosophical Logic* 13:373–378.

Boričić, Branislav. 1985. On sequence-conclusion natural deduction systems. *Journal of Philosophical Logic* 14:359–377.

Bornat, Richard. 2005. *Proof and Disproof in Formal Logic*. Oxford: Oxford Univ. Press.

Bostock, David. 1999. *Intermediate Logic*. Oxford: Oxford UP.

Brandom, Robert. 1994. *Making it Explicit*. Cambridge, MA: Harvard University Press.

Broda, Krysia, Susan Eisenbach, Hessam Khoshnevisan, and Steve Vickers. 1994. *Reasoned Programming*. Englewood Cliffs, NJ: Prentice-Hall.

Carnap, Rudolf. 1944. *The Formalization of Logic*. Cambridge: Harvard University Press.

Cellucci, Carlo. 1987. Efficient natural deduction. In C. Cellucci and G. Sambin, eds., *Atti del Congresso Temi e Prospettive della Logica e della Filosofia della Scienza Contemporanee, Volume I—Logica*, pages 29–57. Bologna: Cooperativa Libraria Universitaria Editrice Bologna.

Cellucci, Carlo. 1992. Existential instantiation and normalization in sequent natural deduction. *Annals of Pure and Applied Logic* 58:111–148.

Cellucci, Carlo. 1995. On Quine's approach to natural deduction. In P. Leonardi and M. Santambrogio, eds., *On Quine: New Essays*, pages 314–335. Cambridge: Cambridge UP.

Chellas, Brian. 1997. *Elementary Formal Logic*. Calgary: Penny Lane Press.

Christensen, Dan. 2010. DC Proof. http://www.dcproof.com.

Church, Alonzo. 1944. Review of Carnap (1944). *Philosophical Review* 53:493–498.

Cook, Stephen and Robert Reckhow. 1979. The relative efficiency of propositional proof systems. *Journal of Symbolic Logic* 44:36–50.

Copi, Irving. 1954. *Symbolic Logic*. NY: Macmillan Co.

Corcoran, John. 1972. Completeness of an ancient logic. *Journal of Symbolic Logic* 37:696–702.

Corcoran, John. 1973. A mathematical model of Aristotle's syllogistic. *Archiv für Geschichte der Philosophie* 55:191–219.

Corcoran, John. 1974. Aristotle's natural deduction system. In J. Corcoran, ed., *Ancient Logic and Its Modern Interpretations*, pages 85–131. Dordrecht: Reidel.

Craig, William. 1957. Review of E.W. Beth, 'Remarks on natural deduction', 'Semantic entailment and formal derivability'; K.J.J. Hintikka, 'A new approach to sentential logic'. *Journal of Symbolic Logic* 22:360–363.

Curry, Haskell. 1950. *A Theory of Formal Deducibility*. Notre Dame, IN: Univ. Notre Dame Press.

Curry, Haskell. 1963. *Foundations of Mathematical Logic*. New York: McGraw-Hill.

van Dalen, Dirk. 1980. *Logic and Structure*. Berlin: Springer-Verlag.

Degtyarev, Anatoli and Andrei Voronkov. 2001. The inverse method. In A. Robinson and A. Voronkov, eds., *Handbook of Automated Reasoning*, pages 178–272. Amsterdam: Elsevier.

Dummett, Michael. 1978. The philosophical basis of intuitionistic logic. In *Truth and Other Enigmas*, pages 215–247. London: Duckworth.

Dunn, Michael. 1993. Partial-gaggles applied to logics with restricted structural rules. In P. Schroeder-Heister and K. Dosen, eds., *Substructural Logics*, pages 63–108. Oxford: Oxford UP.

Evans, Jonathan St.B. T., Stephen Newstead, and Ruth Byrne. 1993. *Human Reasoning: The Psychology of Deduction*. Hove, East Sussex: Lawrence Erlbaum Associates.

Feys, Robert. 1937. Les logiques nouvelles des modalités. *Revue Néoscholastique de Philosophie* 40:517–553.

Fine, Kit. 1985. *Reasoning with Arbitrary Objects*. Oxford: Blackwell.

Fitch, Fredric. 1942. A basic logic. *Journal of Symbolic Logic* 7:105–114.

Fitch, Fredric. 1952. *Symbolic Logic*. NY: Roland Press.

Fitch, Fredric. 1966. Natural deduction rules for obligation. *American Philosophical Quarterly* 3:27–28.

Fitting, Melvin. 2002. *Types, Tableaus, and Gödel's God*. Dordrecht: Kluwer.

Forbes, Graeme. 1994. *Modern Logic*. Oxford: Oxford UP.

Frege, Gottlob. 1904. What is a function? In S. Meyer, ed., *Festschrift Ludwig Boltzmann gewidmet zum sechzigsten Geburtstage*, pages 656–666. Leipzig: Barth. Translation in P. Geach & M. Black (1952) *Translations from the Philosophical Writings of Gottlob Frege* [Blackwell: Oxford], pp. 107–116.

Frické, Martin. 2010. SoftOption. `http://www.softoption.us`.

Gamut, L.T.F. 1991. *Logic, Language and Meaning, Vol. I*. Chicago: University of Chicago Press.

Garson, James. 2006. *Modal Logic for Philosophers*. Cambridge: Cambridge Univ. Press.

Gentzen, Gerhard. 1934. Untersuchungen über das logische Schließen, I and II. *Mathematische Zeitschrift* 39:176–210, 405–431. English translation "Investigations into Logical Deduction" published in *American Philosophical Quarterly*, **1:** 288–306 (1964), and **2:** 204–218 (1965). Reprinted in M. E. Szabo (ed.) (1969) *The Collected Papers of Gerhard Gentzen*, North-Holland, Amsterdam, pp. 68–131. Page references to the original for German and the *APQ* version for English.

Goldberg, Adele and Patrick Suppes. 1972. A computer-assisted instruction program for exercises on finding axioms. *Educational Studies in Mathematics* 4:429–449.

Goldfarb, Warren. 2003. *Deductive Logic*. Indianapolis: Hackett Pub. Co.

Goodstein, Rueben. 1957. *Mathematical Logic*. Leicester, UK: Leicester UP.

Gottschall, Christian. 2010. Gateway to logic. `http://logik.phl.univie.ac.at/~chris/gateway/formular-uk.html`.

Gustason, William and Dolph Ulrich. 1973. *Elementary Symbolic Logic*. NY: Holt, Rinehart & Winston.

Harrison, Frank. 1992. *Logic and Rational Thought*. St. Paul: West Pub. Co.

Harrison, John. 2007. *Introduction to Logic and Automated Theorem Proving*. Cambridge: Cambridge University Press.

Hazen, Allen. 1987. Natural deduction and Hilbert's ϵ-operator. *Journal of Philosophical Logic* 16:411–421.

Hazen, Allen. 1990. The myth of the intuitionistic 'or'. In J. Dunn and A. Gupta, eds., *Truth or Consequences: Essays in Honor of Nuel Belnap*, pages 177–195. Dordrecht: Kluwer.

Hazen, Allen. 1999. Logic and analyticity. In A. Varzi, ed., *European Review of Philosophy, Vol. 4: The Nature of Logic*, pages 79–110. Stanford: CSLI Press.

Herbrand, Jacques. 1928. Sur la théorie de la démonstration. *Comptes rendus hebdomadaires des séances de l'Académie des Sciences* (Paris) 186:1274–1276.

Herbrand, Jacques. 1930. *Recherches sur la théorie de la démonstration*. Ph.D. thesis, University of Paris. Reprinted in Warren Goldfarb (ed. & trans.) (1971) *Logical Writings*, D. Reidel, Dordrecht.

Hintikka, K. Jaakko. 1953. A new approach to sentential logic. *Societas Scientiarum Fennica. Commentationes Physico-Mathematicae* 17:1–14.

Hintikka, K. Jaakko. 1955a. Form and content in quantification theory. *Acta Philosophica Fennica* 8:7–55.

Hintikka, K. Jaakko. 1955b. Notes on quantification theory. *Societas Scientiarum Fennica: Commentationes Physico-Mathematicae* 17:1–13.

Hurley, Patrick. 1982. *A Concise Introduction to Logic*. Belmont, CA: Wadsworth.

Iseminger, Gary. 1968. *Introduction to Deductive Logic*. New York: Appleton-Century-Crofts.

Jaśkowski, Stanisław. 1929. Teoria dedukcji oparta na regułach założeniowych (Theory of deduction based on suppositional rules). In *Księga pamiątkowa pierwszego polskiego zjazdu matematycznego* (Proceedings of the First Polish Mathematical Congress)*, 1927*, page 36. Krakow: Polish Mathematical Society.

Jaśkowski, Stanisław. 1934. On the rules of suppositions in formal logic. *Studia Logica* 1:5–32. Reprinted in S. McCall (1967) *Polish Logic 1920–1939* Oxford UP, pp. 232–258.

Jeffery, Richard. 1967. *Formal Logic: Its Scope and Limits*. New York: McGraw-Hill.

Johnson-Laird, Philip and Ruth Byrne. 1991. *Deduction*. East Sussex: Lawrence Erlbaum Assoc.

Kalish, Donald and Richard Montague. 1964. *Logic: Techniques of Formal Reasoning*. NY: Harcourt, Brace, World.

Kalish, Donald, Richard Montague, and Gary Mar. 1980. *Logic: Techniques of Formal Reasoning, Second Edition*. NY: Harcourt, Brace, World.

Kaliszyk, Cezary. 2010. Proofweb. `http://proofweb.cs.ru.nl/login.php`.

Kerber, Manfred and Axel Präcklein. 1996. Using tactics to reformulate formulae for resolution theorem proving. *Annals of Mathematics and Artificial Intelligence* 18:221–241.

Ketonen, Oiva. 1944. Untersuchungen zum Prädikatenkalkül. *Annales Academiae Scientiarum Fennicae, Series A* 1(23):1–77.

Kilgore, William. 1968. *An Introductory Logic*. NY: Holt, Rinehart & Winston.

Kleene, Stephen. 1952. *Introduction to Metamathematics*. NY: van Nostrand.

Kleene, Stephen. 1967. *Elementary Logic*. NY: Wiley.

Koslow, Arnold. 1992. *A Structuralist Theory of Logic*. Cambridge: Cambridge Univ. Press.

Kripke, Saul. 1963. Semantical analysis of modal logic I: Normal propositional calculi. *Zeitschrift für Mathematische Logik und Grundlagen der Mathematik* 9:67–96.

Lemmon, E. John. 1965. *Beginning Logic*. London: Nelson.

Lewis, Clarence. 1918. *A Survey of Symbolic Logic*. Berkeley: University of California Press.

Li, Dafa. 1992. A natural deduction automated theorem proving system. In *Automated Deduction—CADe–11*, pages 668–672. Berlin: Springer. Volume **607/1992** of *Lecture Notes in Computer Science*.

Lifschitz, Vladimir. 1989. What is the inverse method? *Journal of Automated Reasoning* 5:1–23.

Łukasiewicz, Jan. 1951. *Aristotle's Syllogistic from the Standpoint of Modern Formal Logic*. Oxford: Clarendon Press.

Manktelow, Ken. 1999. *Thinking and Reasoning*. Hove, East Sussex: Psychology Press.

Manktelow, Ken, David Over, and Shira Elqayam. 2010. *The Science of Reason*. Hove, East Sussex: Psychology Press.

Martin, John. 1997. Aristotle's natural deduction reconsidered. *History and Philosophy of Logic* 18:1–15.

Maslov, Sergey Yu. 1964. An inverse method for establishing deducibility in classical predicate calculus. *Doklady AN SSSR* 159:1420–1424. In Russian. English translation in J. Siekmann & G. Wrightson (1983) *The Automation of Reasoning, Vol. 2* (Springer: Berlin), pp. 48–54.

Maslov, Sergey Yu. 1969. Relationship between tactics of the inverse method and the resolution method. In A. Slisenko, ed., *Zapiski Nauchnykh Seminarov LOMI v.16*, pages 139–146. In Russian. English translation in J. Siekmann & G. Wrightson (1983) *The Automation of Reasoning, Vol. 2* (Springer: Berlin), pp. 264–272.

Mates, Benson. 1965. *Elementary Logic*. NY: Oxford UP.

Matuszewski, Roman and Piotr Rudnicki. 2005. Mizar: The first 30 years. *Mechanized Mathematics and Its Applications* 4:3–24.

McGuire, Hugh. 2010. Proofbuilder. `http://www.cis.gvsu.edu/~mcguire/ProofBuilder/`.

Nelson, David. 1949. Constructible falsity. *Journal of Symbolic Logic* 14:16–26.

Nevins, Arthur. 1974. A human oriented logic for automatic theorem-proving. *Journal of the ACM* 21:606–621.

Nipkow, Tobias, Larry Paulson, and Markus Wenzel. 2002. Isabelle/hol—*A Proof Assistant for Higher-Order Logic*. Berlin: Springer. Volume **2283** of *Lecture Notes in Computer Science*.

Pastre, Dominique. 1978. Automated theorem proving in set theory. *Artificial Intelligence* 10:1–27.

Pastre, Dominique. 1993. Automated theorem proving in mathematics. *Annals of Mathematics and Artificial Intelligence* 8:425–447.

Pastre, Dominique. 2002. Strong and weak points of the MUSCADET theorem prover—examples from CASC-JC. *AI Communications* 15:147–160.

Pelletier, Francis Jeffry. 1998. Natural deduction theorem proving in THINKER. *Studia Logica* 60:3–43.

Pelletier, Francis Jeffry. 1999. A brief history of natural deduction. *History and Philosophy of Logic* 20:1–31.

Pelletier, Francis Jeffry. 2000. A history of natural deduction and elementary logic textbooks. In J. Woods and B. Brown, eds., *Logical Consequence: Rival Approaches, Vol. 1*, pages 105–138. Oxford: Hermes Science Pubs.

Pelletier, Francis Jeffry. forthcoming. Holism and compositionality. In W. Hinzen, E. Machery, and M. Werning, eds., *Handbook of Compositionality*. Oxford: Oxford Univ. Press. Forthcoming in 2010.

Pelletier, Francis Jeffry, Geoff Sutcliffe, and Christian Suttner. 2002. The Development of CASC. *AI Communications* 15(2-3):79–90.

Peregrin, Jaroslav. 2008. What is *the* logic of inference? *Studia Logica* 88:263–294.

von Plato, Jan. 2008a. Gentzen's logic. In D. Gabbay and J. Woods, eds., *Handbook of the History and Philosophy of Logic, Vol. 5: From Russell to Church*, pages 607–661. Amsterdam: Elsevier.

von Plato, Jan. 2008b. Gentzen's proof of normalization for natural deduction. *Bulletin of Symbolic Logic* 14:240–257.

Pollock, John. 1992. Interest-driven suppositional reasoning. *Journal of Automated Reasoning* 6:419–462.

Popper, Karl. 1947a. Logic without assumptions. *Proceedings of the Aristotelian Society* 47:251–292.

Popper, Karl. 1947b. New foundations for logic. *Mind* 56:193–235. Corrigenda in Vol. 57 (1948, pp. 69-70).

Portoraro, Frederic. 1994. *Logic with Symlog*. Englewood Cliffs, NJ: Prentice-Hall.

Prawitz, Dag. 1965. *Natural Deduction: A Proof-theoretical Study*. Stockholm: Almqvist & Wicksell.

Prawitz, Dag. 1971. Ideas and results in proof theory. In J. Fenstad, ed., *Proceedings of the Second Scandinavian Logic Symposium*, pages 235–307. Amsterdam: North-Holland.

Price, Richard. 1961. The stroke function and natural deduction. *Zeitschrift für mathematische Logik und Grundlagen der Mathematik* 7:117–123.

Priest, Graham. 1979. The logic of paradox. *Journal of Philosophical Logic* 8:219–241.

Prior, Arthur. 1960. The runabout inference ticket. *Analysis* 21:38–39.

Purtill, Richard. 1971. *Logic for Philosophers*. New York: Harper & Row.

Quine, Willard V. 1950a. *Methods of Logic*. New York: Henry Holt & Co.

Quine, Willard V. 1950b. On natural deduction. *Journal of Symbolic Logic* 15:93–102.

Restall, Greg. 2000. *An Introduction to Substructural Logics*. London and New York: Routledge.

Restall, Greg. 2005. Multiple conclusions. In P. Hájek, L. Valdes-Villanueva, and D. W. hl, eds., *Logic, Methodology and Philosophy of Science: Proceedings of the Twelfth International Congress*, pages 189–205. London: Kings' College Publications.

Rips, Lance. 1994. *The Psychology of Proof: Deduction in Human Thinking*. Cambridge: MIT Press.

Robinson, J. Alan. 1965. A machine-oriented logic based on the resolution principle. *Journal of the ACM* 12:23–41.

Robison, Gerson. 1969. *An Introduction to Mathematical Logic*. Englewood Cliffs, NJ: Prentice-Hall.

Routley, Richard and Robert Meyer. 1972. The semantics of entailment – II. *Journal of Philosophical Logic* 1:53–73.

Routley, Richard and Robert Meyer. 1973. The semantics of entailment. In H. Leblanc, ed., *Truth, Syntax and Modality*, pages 194–243. Amsterdam: North Holland.

Rumfitt, I. 2008. Knowledge by deduction. *Grazer philosophische Studien* 77:61–84.

Russell, Bertrand. 1906. The theory of implication. *American Journal of Mathematics* 28:159–202.

Russell, Bertrand. 1918. Lectures on logical atomism. *The Monist* 28:495–527. Continues in Volume 29: pp. 32–63, 190–222, 345–380.

Saetti, John. 2010. Logic toolbox. `http://philosophy.lander.edu/~jsaetti/Welcome.html`.

Schroeder-Heister, Peter. 1984a. Generalized rules for quantifiers and the completeness of the intuitionistic operators $\wedge, \vee, \supset, \bot, \forall, \exists$. In M. Richter, E. Börger, W. Oberschelp, B. Schinzel, and W. Thomas, eds., *Computation and Proof Theory: Proceedings of the Logic Colloquium Held in Aachen, July 18–23, 1983, Part II*, pages 399–426. Berlin: Springer-Verlag. Volume **1104** of *Lecture Notes in Mathematics*.

Schroeder-Heister, Peter. 1984b. A natural extension of natural deduction. *Journal of Symbolic Logic* 49:1284–1300.

Schütte, Kurt. 1960. *Beweistheorie*. Berlin: Springer-Verlag. English translation *Proof Theory*, Springer-Verlag: New York, 1977.

Scott, Dana. 1974. Completeness and axiomatizability in many-valued logic. In L. Henken, J. Addison, W. Craig, C. Chang, D. Scott, and R. Vaught, eds., *Proceedings of the Tarski Symposium: Proceedings of Symposia in Pure Mathematics* **25**, pages 411–435. Providence: American Mathematical Society.

Sieg, Wilfried and John Byrnes. 1998. Normal natural deduction proofs (in classical logic). *Studia Logica* 60:67–106.

Siekmann, Jörg, Christoph Benzmüller, and Serge Autexier. 2006. Computer supported mathematics with ΩMEGA. *Journal of Applied Logic* 4:533–559.

Slate, David and Larry Atkin. 1977. The northwestern university chess program. In P. Frey, ed., *Chess Skill in Man and Machine*, pages 82–118. Springer.

Smith, Robin. 1982. What is Aristotelian echthesis? *History and Philosophy of Logic* 3:113–127.

Smullyan, Raymond. 1968. *First Order Logic*. New York: Springer-Verlag.

Stålmarck, Gunnar. 1991. Normalization theorems for full first-order classical natural deduction. *Journal of Symbolic Logic* 56:129–149.

Suppes, Patrick. 1957. *Introduction to Logic*. Princeton: Van Nostrand/Reinhold Press.

Sutcliffe, Geoff. 2009. The 4th IJCAR Automated Theorem Proving Competition. *AI Communications* 22(1):59–72.

Sutcliffe, Geoff and Christian Suttner. 2006. The State of CASC. *AI Communications* 19(1):35–48.

von Sydow, Björn. 2010. ALFIE. `http://www.cs.chalmers.se/~sydow/alfie`.

Tarski, Alfred. 1930a. Fundamentale Begriffe der Methodologie der deduktiven Wissenschaften I. *Monatshefte für Mathematiik und Physik* 37:361–404. English translation "Fundamental Concepts of the Methodology of the Deductive Sciences" in (Tarski, 1956, pp.60–109).

Tarski, Alfred. 1930b. Über einige fundamentalen Begriffe der Metamathematik. *Comptes rendus des séances de la Société des Sciences et Lettres de Varsovie* (Classe III) 23:22–29. English translation "On Some Fundamental Concepts of Metamathematics" in (Tarski, 1956, pp.30–37).

Tarski, Alfred. 1935. Grundzüge der Systemenkalküls. *Fundamenta Mathematicae* 25:503–526. Part II of article in Volume 26, pp.283–301. English translation "Foundations of the Calculus of Systems" in (Tarski, 1956, pp.342–383).

Tarski, Alfred. 1956. *Logic, Semantics, Metamathematics*. Oxford: Clarendon.

Teller, Paul. 1989. *A Modern Formal Logic Primer*. Englewood Cliffs, NJ: Prentice-Hall.

Tennant, Neil. 1978. *Natural Logic*. Edinburgh: Edinburgh Univ. Press.

Thomason, Richmond. 1970. *Symbolic Logic*. New York: Macmillan.

Toledo, Sue. 1975. *Tableau Systems for First Order Number Theory and Certain Higher Order Theories*. Berlin: Springer-Verlag. Volume **447** of *Lecture Notes in Mathematics.*

Urquhart, Alasdair. 1995. The complexity of propositional proofs. *Bulletin of Symbolic Logic* 1:425–467.

Wenzel, Markus. 2002. The isabelle/isar reference manual. `http://isabelle.in.tum.de/doc/isar-ref.pdf`.

Wiedijk, Freek. 2007. The qed manifesto revisited. In R. Matuszwski and A. Zalewska, eds., *From Insight to Proof: Festschrift in Honour of Andrzej Trybulec*, pages 121–133. Białystok, Poland: University of Białystok. Volume 10(23) of the series *Studies in Logic, Grammar and Rhetoric*, available online at `http://logika.uwb.edu.pl/studies/index.html`.

von Wright, Georg. 1951a. Deontic logic. *Mind* 60:1–15.

von Wright, Georg. 1951b. *An Essay in Modal Logic*. Amsterdam: North-Holland.

Zucker, Jeffery and Robert Tragesser. 1978. The adequacy problem for inferential logic. *Journal of Philosophical Logic* 7:501–516.

Appendix: Elementary Logic Textbooks Described in Table I

Anderson, J. and Johnstone, H. (1962). *Natural Deduction: The Logical Basis of Axiom Systems*. Wadsworth Pub. Co., Belmont, CA.

Arthur, R. (2010). *Natural Deduction: An Introduction to Logic*. Wadsworth Pub. Co., Belmont, CA.

Barwise, J. and Etchemendy, J. (2002). *Language, Proof and Logic*. CSLI Press, Stanford.

Bergmann, M., Moor, J., and Nelson, J. (1980). *The Logic Book*. Random House, New York.

Bessie, J. and Glennan, S. (2000). *Elements of Deductive Inference*. Wadsworth Pub. Co., Belmont, CA.

Bonevac, D. (1987). *Deduction*. Mayfield Press, Mountain View, CA.

Bostock, D. (1999). *Intermediate Logic*. Oxford UP, Oxford.

Byerly, H. (1973). *A Primer of Logic*. Harper & Row, New York.

Carter, C. (2005). *A First Course in Logic*. Pearson/Longman, New York.

Cauman, L. (1998). *First-Order Logic: An Introduction*. Walter de Gruyter, Hawthorne, NY.

Chellas, B. (1997). *Elementary Formal Logic*. Penny Lane Press, Calgary.

Copi, I. (1954). *Symbolic Logic*. Macmillan Co., NY.

Curry, H. (1963). *Foundations of Mathematical Logic*. McGraw Hill, New York.

DeHaven, S. (1996). *The Logic Course*. Broadview Press, Peterburough, Ontario.

Fitch, F. (1952). *Symbolic Logic*. Roland Press, NY.

Forbes, G. (1994). *Modern Logic*. Oxford Univ. Press, Oxford.

Gamut, L. (1991). *Logic, Language and Meaning, Vol. I, II*. University of Chicago Press, Chicago.

Georgacarakos, G. and Smith, R. (1979). *Elementary Formal Logic*. McGraw Hill, New York.

Goldfarb, W. (2003). *Deductive Logic*. Hackett Pub. Co., Indianapolis.

Goodstein, R. (1957). *Mathematical Logic*. Leicester UP, Leicester, UK.

Gustason, W. and Ulrich, D. (1973). *Elementary Symbolic Logic*. Holt, Rinehart & Winston, NY.

Guttenplan, S. (1997). *The Languages of Logic*. Blackwells, Oxford.

Harrison, F. (1992). *Logic and Rational Thought*. West Pub. Co., St. Paul.

Hurley, P. (1982). *A Concise Introduction to Logic*. Wadsworth, Belmont, CA.

Iseminger, G. (1968). *Introduction to Deductive Logic*. Appleton-Century-Crofts, New York.

Jacquette, D. (2001). *Symbolic Logic*. Wadsworth, Belmont, CA.

Jennings, R. and Friedrich, N. (2006). *Truth and Consequence*. Broadview, Peterborough, Ont.

Kalish, D. and Montague, R. (1964). *Logic: Techniques of Formal Reasoning*. Harcourt, Brace, World, NY.

Klenk, V. (1983). *Understanding Symbolic Logic*. Prentice-Hall, Englewood Cliffs, NJ.

Kozy, J. (1974). *Understanding Natural Deduction*. Dickenson Pub. Co., Encino, CA.

Leblanc, H. and Wisdom, W. (1965). *Deductive Logic*. Allyn & Bacon, Boston.

Lemmon, E. J. (1965). *Beginning Logic*. Nelson, London.

Machina, K. (1978). *Basic Applied Logic*. Scott-Foreman, Glenview, IL.

Martin, R. (2004). *Symbolic Logic*. Broadview Press, Peterborough, Ontario.

Massey, G. (1970). *Understanding Symbolic Logic*. Harper & Row, New York.

Mates, B. (1965). *Elementary Logic*. Oxford UP, NY.

Myro, G., Bedau, M., and Monroe, T. (1987). *Rudiments of Logic*. Prentice-Hall, Englewood Cliffs, NJ.

Pollock, J. (1969). *Introduction to Symbolic Logic*. Holt, Rinehart & Winston, New York.

Purtill, R. (1971). *Logic for Philosophers*. Harper & Row, New York.

Quine, W. V. (1950). *Methods of Logic*. Henry Holt & Co., New York.

Resnick, M. (1970). *Elementary Logic*. McGraw-Hill, New York.

Simco, N. and James, G. (1976). *Elementary Logic*. Dickenson, Encino, CA.

Simpson, R. (1987). *Essentials of Symbolic Logic*. Broadview Press, Peterborough, Ontario.

Suppes, P. (1957). *Introduction to Logic*. Van Nostrand/Reinhold Press, Princeton.

Tapscott, B. (1976). *Elementary Applied Symbolic Logic*. Prentice-Hall, Englewood Cliffs, NJ.

Teller, P. (1989). *A Modern Formal Logic Primer*. Prentice-Hall, Englewood Cliffs, NJ.

Tennant, N. (1978). *Natural Logic*. Edinburgh Univ. Press, Edinburgh.

Thomason, R. (1970). *Symbolic Logic*. Macmillan, New York.

van Dalen, D. (1980). *Logic and Structure*. Springer-Verlag, Berlin.

Wilson, J. (1992). *Introductory Symbolic Logic*. Wadsworth Pub. Co., Belmont, CA.