Cut as Consequence

CURTIS FRANKS

Department of Philosophy, University of Notre Dame, Indiana, USA

The papers where Gerhard Gentzen introduced natural deduction and sequent calculi suggest that his conception of logic differs substantially from now dominant views introduced by Hilbert, Gödel, Tarski, and others. Specifically, (1) the definitive features of natural deduction calculi allowed Gentzen to assert that his classical system NK is complete based purely on the sort of evidence that Hilbert called 'experimental', and (2) the structure of the sequent calculi LI and LK allowed Gentzen to conceptualize completeness as a question about the relationships among a system's individual rules (as opposed to the relationship between a system as a whole and its 'semantics'). Gentzen's conception of logic is compelling in its own right. It is also of historical interest because it allows for a better understanding of the invention of natural deduction and sequent calculi.

1. Introduction

In the opening remarks of *1930*, Kurt Gödel described the construction and use of formal axiomatic systems as found in the work of Whitehead and Russell—the procedure of 'initially taking certain evident propositions as axioms and deriving the theorems of logic and mathematics from these by means of some precisely formulated principles of inference in a purely formal way'—and remarked that

> when such a procedure is followed, the question at once arises whether the initially postulated system of axioms and principles of inference is complete, that is, whether it actually suffices for the derivation of *every* logico-mathematical proposition, or whether, perhaps, it is conceivable that there are true [*wahre*[1]] propositions ... that cannot be derived in the system under consideration. (p. 103)

Gödel proceeded to establish the completeness of one fundamental system which, following Whitehead and Russell, he called 'the restricted functional calculus', known today as first-order quantification theory. Then he proved several classical corollary results and strengthenings.

Gödel wrote as if with the hindsight of today's logicians, who, when they encounter a new formal system, reflexively ask for an interpretive scheme to make sense of the

[1] In the main body of the paper Gödel used the expression '*allgemeingültige*' which he explained was a slightly imprecise way of saying '*in jedem Individuenbereich allgemeingültig*'. Evidently, Gödel has in mind not *Wahrheit* but what in contemporary terms is distinguished from it as *logische Wahrheit*. Gödel's use of an alternate expression denoting 'universal validity' follows the preference of Hilbert and others who wished to avoid philosophical commitments involved in talk of truth.

symbolism and then ask whether the system is complete with respect to this 'semantics'. But in their historical context, the questions that Gödel posed and solved hardly arose immediately. In the writing of most logicians, the question of the semantic completeness of quantification theory did not arise at all.

This has struck some historians as remarkable because it can indeed be difficult to see how this question would escape the attention of logicians who developed sufficient apparatus to solve it. Most striking is the fact that in the work of Thoralf Skolem and Jacques Herbrand one finds in explicit detail all the reasoning needed to establish the semantic completeness of quantification theory. Yet the question did not occur even to them. Thus Gödel, in a 1967 letter to Hao Wang:

> The completeness theorem, mathematically, is indeed an almost trivial consequence of *Skolem 1923*. However, the fact is that, at that time, nobody (including Skolem himself) drew this conclusion (neither from *Skolem 1923* nor, as I did, from similar considerations of [Skolem's].) (*Wang 1974*, p. 8)

Herbrand even went so far as to say in *1930* that it is tempting to infer from his results a certain statement[2], which combined with the inferences he does draw establishes the semantic completeness of quantification theory, but that the statement is too idealistic to make real, concrete sense (p. 165).[3] Subject to an insensitive, modern re-reading, what appear in the writing of these two thinkers are something like proofs of the completeness theorem with the theorem's statement and concluding sentence erased. But there is something obviously chauvinistic about reading in this way while ignoring the hermeneutical problem of recovering the conceptual framework that allowed these logicians to skirt the completeness phenomenon without recognizing it as a phenomenon.

[2]The statement is that if, for every n, the n^{th} Herbrand expansion of a quantificational formula is truth-functionally satisfiable, then the formula is satisfiable in an infinite domain.

[3]Herbrand wrote in *1930*:

> We observe that this definition differs from the definition that would seem most natural only in that, as the number p increases, the new domain C' and the new values need not be regarded as forming an 'extension' of the previous ones. Clearly, if we know C' and the values for any given p, then for each smaller number we know a domain and values that answer to the number; but only a 'principle of choice' could lead us to take a fixed system of values in an infinite domain. (p. 165)

In this passage one finds explicitly an objection to a certain use of Zermelo's set-theoretical choice axiom. But implicitly a deeper objection is hinted at, namely, one against the entire standard definition of satisfaction. The axiom of choice is needed to construct infinite domains of interpretation out of sequences of increasingly large finite ones. Herbrand is suspicious of the meaningfulness of such a construction, or at least of its appropriateness in meta-mathematical investigations. But he is also suspicious of the meaningfulness of finite domains of interpretation, in the set-theoretical sense, for the reason that these are far removed from the immanent features of logical syntax. His own '*champs finis*' are entirely different sorts of structures. For example, the universal and existential quantifiers in a given formula range over different objects. For a detailed contrast between the standard notion of a model and Herbrand's notion of a *champ fini*, see *van Heijenoort 1985*, especially pp. 100-1, 110-3.

That problem is made difficult by the fact that Gödel's conception of logic, according to which the completeness question is central and unavoidable, proved to be such a productive conception. Because no other way of looking at things ushered in nearly so many deep results as Gödel's, logicians inevitably came to adopt that conception in the process simply of tracking the development of logic over time. In science, productive ways of thinking are also seductive ways of thinking, and rightly so.

But the hermeneutical problem appears more urgent when one takes account of more historical data. Notably, some seven years passed between the establishment of the semantic completeness of propositional logic and the first written[4] acknowledgment (in *Hilbert and Ackermann 1928*) even of the question of the completeness of quantification theory. Thus it is not possible, for example, to attribute Herbrand's failure to recognize the question solely to his adherence to intuitionistic principles, nor Skolem's to his reservations about using the axiom of choice in an 'investigation in the foundations of set theory' (*Skolem 1923*, p. 293). Plenty of logicians were perfectly comfortable with non-constructive apparatus, but attempts to pose the question of semantic completeness outside of the propositional framework are not found outside the work of Hilbert and Bernays.

Still more perplexing is the fact that a considerable amount of logical work *after* 1930 proceeded largely uninfluenced by Gödel's result. Perhaps most striking is Skolem's own continued disinterest, surmised by Burton Dreben and Jean van Heijenoort who write:

> If comment is a measure of interest, then the completeness of quantification theory held absolutely no interest for Skolem. There is not one reference to completeness in the fifty-one papers on logic, dating from 1913 through 1963, collected in *Fenstad 1970.* (*1986*, p. 54)

Not quite as statistically impressive, but perhaps equally of interest (and the principal topic of this paper), is Gerhard Gentzen's attitude. Only once in the ten papers compiled in *Szabo 1970* did Gentzen mention the completeness of quantification theory. The reference is in the *1936* paper 'Die Widerspruchsfreiheit der reinen Zahlentheorie'. Gentzen's tone is nonchalant, his explanation of the concept literally parenthetical:

> The completeness of the purely logical rules of inference, i.e., the rules belonging to the connectives $\&, \vee, \supset, \neg, \forall, \exists$, has already been proved elsewhere (completeness here means that all correct inferences of the same type [as those] represented by the stated rules [are already represented]).[5] (p. 154)

[4]In *1986* Dreben and van Heijenoort observe that Hilbert had raised the question in a lecture that same year, which was later published as *1929*. Recent scholarship has traced the question to Hilbert's lectures from the academic year 1917-8. The notes for these lectures were prepared by Paul Bernays, as detailed in *Ewald and Sieg 2010*. Indeed, *Hilbert and Ackermann 1928* appears to be largely a redaction of these lectures with little if any input from Ackermann.

[5]The brackets indicate my deviation from Szabo's translation. Gentzen wrote 'alle richtigen Schlüsse von gleicher Art bereits durch die angegebenen Schlußregeln darstellbar sind, ist durch besondere Un-

There are some oddities in this passage that prove important for my reading of Gentzen in section 3 below. But more interesting than the peculiarities of the way Gentzen discussed completeness where he did is the fact that in his other logical writings he never mentioned the result. Most crucially, in his *1934–35* dissertation 'Untersuchungen über das logische Schliessen' Gentzen developed the systems of natural deduction and sequent calculus for quantification theory and proved their deductive equivalence as well as their equivalence with 'a calculus modeled on the formalism of Hilbert'. Because the latter system was known already to be sound and complete with respect to the standard quantificational semantics, Gentzen could have immediately inferred the same properties for his classical calculi NK and LK. But he neither referenced the completeness theorem nor posed the question. Indeed, not even in his expository *1938* paper 'Die gegenwärtige Lage in der mathematischen Grundlagenforschung', did Gentzen mention the completeness of quantification theory. This, despite Gentzen's claim in a section called 'Exact foundational research in mathematics: axiomatics, metalogic, metamathematics: The theorems of Gödel and Skolem' that his purpose was to 'discuss some of the more recent findings and, in particular, some of the especially important earlier results obtained in the exact foundational research in mathematics'.

Gentzen wrote:

> A main task of metamathematics is the development of the *consistency proofs* required for the realization of *Hilbert's* programme. Other major problems are: The *decision problem*, i.e., the problem of finding a procedure for a given theory which enables us to decide of every conceivable assertion whether it is true or false; further, the question of *completeness*, i.e., the question of whether a specific system of axioms and forms of inference for a specific theory is *complete*, in other words, whether the truth or falsity of every conceivable assertion of that theory can be proved by means of these forms of inference. (p. 238)

Against this backdrop Gentzen then reviewed, in order, Gödel's so-called second incompleteness theorem about the unprovability in elementary number theory of that very theory's consistency, Gentzen's own arithmetical consistency proof using transfinite induction, Church's theorem on the undecidability of quantification theory as well as Gödel's preliminary work in this direction, Gödel's first incompleteness theorem generating, in Gentzen's words, for 'every formally delimited consistent mathematical theory ... number theoretical theorems ... which are true, but which are not provable with the techniques of that theory', Ackermann's proof of the consistency of 'general set theory' relative to the consistency of elementary number theory, the Löwenheim-Skolem theorem,

---

tersuchen bewiesen worden', which Szabo translates ambiguously as 'all correct inferences of the same type are representable by the stated rules'. That translation invites the (now standard) interpretation 'all first-order inferences' as opposed to Gentzen's intended, weaker claim about 'all $\&, \vee, \supset, \neg, \forall$, and $\exists$ inferences'. (Szabo further miscopies Gentzen's reference, which is to *Gödel 1930*, as being to *Gödel 1931*. Two anonymous referees for *History and Philosophy of Logic* steered me to these two observations.)

and Skolem's proof of the existence of nonstandard models of first-order arithmetical systems. Obviously the completeness of quantification theory is a fundamental metalogical result quite difficult to omit from such a discussion, but Gentzen never mentioned it.

To sum up, there are features of the thought of figures like Herbrand, Skolem, and Gentzen that, variously, prevented them from recognizing semantic completeness as a phenomenon or dissuaded them from acknowledging the relevance of the completeness theorem after it was proved. One might suppose that those ways of thinking would be uninteresting to a modern logician if they could be recovered today. One might despair of the possibility of recovering them anyway due to the ossification of the point of view that Gödel introduced—the point of view that inclines us to think that these 'deviant' logicians were missing something instead of being on to something. But I am more optimistic.

Against these assumptions, I suggest first, that the profundity of the logical innovations that these logicians devised speaks for the validity of their conception of their craft, and second, that in the case of Gentzen, there are ample clues in his writing from which to reconstruct his thought. What led Gentzen to ignore questions of and theorems about semantic completeness are the same details of his conception of logic that led to his invention of natural deduction and sequent calculi. The differences between how he thought about basic notions like logical consequence and how they are ordinarily understood today are not subtle but fundamental. This is not to say that there is something wrong about the now standard view of logic, only that to fully appreciate Gentzen's accomplishments (which are still central in modern logic) one must relinquish that view.

2. Syntax and semantics

A picture of the development of the standard view of logic will provide a helpful foil for the details of Gentzen's thought. I sketch such a picture here and postpone reading Gentzen until the following section. The 'standard view' I have in mind is based on a theoretical distinction between logical syntax and logical semantics. The story of its development begins with the invention of formal logical languages and precise rules for their systematic deployment.

Ambiguities in natural, spoken and written languages hinder the study of the subtlest details of the laws of thought, particularly of mathematical thought. Piecemeal attempts to disambiguate these languages according to any uniform principles largely fail. Indeed, serious attempts to disambiguate natural language ordinarily reveal deeper ambiguities than are recognizable at first reflection. According to Gottlob Frege:

> A strictly defined group of modes of inference is simply not present in [natural] language, so that on the basis of linguistic form we cannot distinguish between a 'gapless' advance [in reasoning] and an omission of connected links. We can even say that the former almost never occur in language, that it runs against the feel of language because it would involve an insufferable prolixity. In

> language, logical relations are almost always only hinted at—left to guessing, not actually expressed. (*1882*, p. 85)

Determined to overcome these hindrances, logicians devised artificial languages with precise, recursively defined grammars and emphasized that a purely mechanical procedure suffices to determine if a string of symbols is a sentence of one of these languages. In a similar fashion, they devised proof systems over these languages by replacing the usual free-wheeling progression from sentence to sentence in learned discourse with formal rules of inference. Repeated applications of these rules—depending on the particulars of the system in question, either on a privileged set of sentences called axioms, or from arbitrary hypotheses, or in some settings from no starting point at all—generate finite, ordered structures (e.g., sequences or acyclic graphs) whose elements are sentences from the formal language. Just as with the recognition of sentences of the language, a purely mechanical procedure suffices to determine whether an arbitrary sequence or tree of sentences is a proof. Frege repeatedly emphasized the calculus-like nature of logical systems. These remarks from *Grundlagen* can be compared with similar emphases on page 86 of *Frege 1882* and on page 104 of *Frege 1879*:

> The demand is not to be denied: every jump must be barred from our deductions. That it is so hard to satisfy must be set down to the tediousness of proceeding step by step. Every proof which is even a little complicated threatens to become inordinately long. ... [To this end] I invented my concept writing[6]. It is designed to produce expressions which are shorter and easier to take in, and to be operated like a calculus by means of a small number of standard moves, so that no step is permitted which does not conform to the rules which are laid down once and for all. (*1884*, §91)

Of course the details of a logical system will be set up with an eye to what its various parts mean. But the purely mechanical tests for grammaticality and derivational propriety guarantee that no attention to the interpretation of the various signs in the system ever is needed in order to deploy the system correctly. The hallmark, according to Gödel in *1930*, of a formal system is that it is possible to 'derive the theorems of logic and mathematics ... without making further use of the meaning of the symbols' (p. 103). What Gödel thereby de-emphasized is that this possibility is extremely idealistic, that attention to meaning is typically essential for a system's scientific, goal-directed deployment. But the general idea is clear: The meaning of the calculus is completely absent from its operational description. Indeed, the formal notion of logical consequence dominant in the early days of formal logic made no reference to meaning. According to this notion, the logical consequences of a set of sentences are the sentences obtainable from them by repeated applications of the various inference rules.

[6]The formal system of higher order logic found in *Frege 1879*. 'Concept writing' is Austin's literal translation of '*Begriffsschrift*', the title of this work.

According to Frege, this is a purely *perspectival* point. It is possible to view a logical system purely formally, and it is also possible to view it as a system of contentful expressions. Whereas the latter point of view is essential for the system's original design and, typically, also for making good use of it, it is crucial that the purely formal point of view suffices to verify the correctness of its expressions and proofs. But Frege thought it is a mistake to make heavy weather out of the autonomy of a system's formal features. In a reply to criticisms of E. Schröder he wrote:

> I did not wish to present an abstract logic in formulas, but to express a content through written symbols in a more precise and perspicuous way than is possible with words. In fact, I wished to produce, not a mere *calculus ratiocinator*, but a *lingua characteristica* in the Leibnizian sense. In doing so, however, I recognize that deductive calculus is a necessary part of a conceptual notation. If this was misunderstood, perhaps it is because I let the abstract logical aspect stand too much in the foreground. (*1883*, pp. 90-1)

In fact, Frege thought that the formal precision of his logical system, because 'every ambiguity is banned' from its expressions and every gap filled in its derivations, necessitates this conception. In his words, it produces 'a strict logical form from which meaning cannot escape' (*1882*, p. 86).

The fate of the *lingua characteristica* conception of formal logic is well known. In moving from the rough terrain of natural language and ordinary argumentative compulsion to the precisely defined setting of a formal logical system, one raises questions about how the consequence relation defined in the formal system is coordinated with the notion of logical consequence that underlies those laws of thought that the logician originally set out to study. These questions appear at every stage of the formalization: Which thoughts amenable to logical relationships are translatable into the language of the formal system? Restricted to these thoughts, does the system have enough rules to emulate all the intricate argument patterns of ordinary learned discourse? Or conversely, does the formal system ever err by sanctioning an inference that, under the intended interpretation of the language, we should not accept as truth-preserving? These last two questions are the original forms of the completeness and soundness questions about logical systems. More simply: 'In this pristine landscape, have we left anything out? Have we got too much?'

To address these questions one must attend not only to the formal properties of the system, but also and especially to how its language is interpreted. Interpreting a logical system is typically straightforward. Because of the recursively defined generation rules for sentences, one is able to stipulate a compositional semantics on the language: The meaning of a complex sentence is calculable from the meanings of its components, whose meanings are in turn calculable from the meanings of their components, and so on, until at the end of this process are reached simple components with fully specified meanings and so-called 'logical particles' that specify which calculations to perform on the meanings of the components they govern. The notions of component and complexity are given by correlate notions from the recursively defined grammar; the logical particles are the logical constants of that grammar.

The theory of such a compositional semantics can be fully precise, just not formal in the sense of being open to investigation independent of considerations of meaning. Take for example the standard set-theoretical semantics for first-order quantificational theory. It is a precise theory *of the meanings* of the purely formal logical system. Meanwhile, that logical system, born an interpreted language, becomes *syntax*, i.e., fully uninterpreted: The system was designed so that all of its properties can be mechanically verified, and the unambiguosness engendered by this feature was its principal *scientific merit.* Now, with meaning excised into the semantic realm, this same feature becomes its *essence.*

Only in the semantic setting is one able to the define truth and falsity of a sentence. Here one also typically defines 'validity' (Gödel's notion of '*Allgemeingültigheit (in jedem Individuenbereich)*') and a new consequence relation—semantic consequence—according to which a sentence is a consequence of a set of sentences if it is true in all those interpretations that make all sentences from the set true. Indeed, history has not abided simultaneous notions of semantic and syntactic consequence, but instead has seen the semantic definition replace its syntactic forebear. In current discussion of the syntactic phenomenon, the expression 'consequence of' is replaced by 'derivable from'.

Thus from the features of unambiguous logical languages arose a two-tiered mathematical scaffolding. The perfectly natural *perspectival* distinction between syntax and semantics evolved into a *theoretical* distinction. Frege's original distinction asks about ordinary objects whether their form or their content is under investigation. The new one considers two different sorts of objects: signs, which are immanent, and their meanings, which are transcendent, and directs one's investigations to an appropriate theory depending on which objects are being asked about.

This theoretical syntax/semantics distinction leads to reformulations of the various questions that logical system building invites. One can, for example, determine quite readily whether higher order quantification or certain modal notions are recoverable in the semantics. Typically something will be lacking, but one can simply restrict one's attention to the laws of some sharply delineated fragment of human thought and decide that one is studying the logical relationships that attain between sentences expressing thoughts in that category. More pointedly, one can take on the questions about the sufficiency and possible fallaciousness of the axioms and inference rules—the questions of the logical system's completeness and soundness—quite directly. The first question becomes 'semantic completeness': 'Are all truths derivable?' or more generally 'Are all the semantic consequences of a set of sentences formally derivable from those sentences?' The question of 'semantic soundness' is its converse: 'Do all derivations exhibit legitimate semantic consequences?' 'Does this system accommodate a proof of anything that isn't actually true?' Because of the precision of all the definitions involved—truth, interpretation, validity, and consequence on the semantic side; sentence, axiom, and proof on the syntactic side—these questions are in fact amenable to mathematical solution.[7]

Until now I have spoken only about the completeness and soundness *of a logical system.* This question began vague but rich: 'How accurately do the rules of this system capture the intuitive notion of logical consequence?' In time it became precise but bland: 'How accurately are this system's semantics traced by the purely syntactic machinery?' Precision is a sign of some headway, but one can reasonably ask where, in the move from 'completeness' to 'semantic completeness', the richness went.

Drawing the theoretical syntax/semantics distinction simply pushes both the imprecision and the richness of the original 'Is anything missing?' form of completeness elsewhere. Semantic completeness is a question about the coordination of the transcendent, semantic realm with the immanent, syntactic one that it hovers over. But the original question resurfaces as a question about the propriety of those very realms. The full force of that question remains, though outside of the two-tiered scaffolding that we erect—and the grandeur of that edifice holds our attention and makes us forget about the surrounding landscape.

Given a formal semantics for a logical system, it is possible to ask about *its* soundness and completeness. One might ask, for example, whether the usual set-theoretical semantics are a proper precisification of intuitions about first-order logical truths. This question is rarely asked today because it is not considered at all doubtful that the set-theoretical semantics are a complete triumph in this regard.[8] On this point, John Etchemendy remarks:

> Our attitude here is characteristic of our attitude toward an analysis: extensional adequacy is guaranteed on a *conceptual* level, by our close adherence to the intuitive notion we aim to characterize. It is in this sense that the model-theoretic account is treated as a genuine analysis of the intuitive notions of logical truth and logical consequence. (*1988*, p. 67)

But Etchemendy shows vividly that no arm-chair analysis of our intuitions resulted in the ossification of the set-theoretical semantics of first-order logic. Gödel, in the very *1930* paper where he proved the soundness and completeness of first-order predicate logic relative to this semantics, never considered a notion of semantic consequence and reserved the designation of 'consequence' for the syntactic notion. More surprisingly, and this is Etchemendy's main point, it is Alfred Tarski to whom we usually attribute both the turn from the syntactic to the semantic notion of consequence and the orthodox account of what the semantic notion is. Yet in his landmark *1936* paper in which he did the former, he actually advocated a *different* semantics with respect to which the ordinary logical calculus is *not* complete. Thus Tarski's own attempt at a 'close adherence to the intuitive

[7]But because of the complicated nature of the semantic setting, a relatively large amount of mathematics is needed to solve completeness problems. For the relevance of this point see the concluding remarks in section 5.

[8]An anonymous referee for *History and Philosophy of Logic* reminds me that such glib optimism is happily far less common in Continental Europe than in other geographical areas.

notion we aim to characterize' led him elsewhere than to where modern logicians take themselves to be led. In his later writing, Tarksi advocated the orthodox semantics, but without argument, and it seems reasonable to suggest with Etchemendy that it is the development of model-theory and the comparative fruitfulness of the standard notion that shaped both his and our own intuitions, rather than vice-versa.

The completeness or incompleteness of quantificational theory is then a matter of taste, according to how well one takes to the standard semantics. We might only add that rejecting completeness for the Π^1 fragment of quantification theory is a matter of *bad taste*—not because the now standard notions of first-order logical consequence and truth are *a priori* correct, but because the logical study of first-order theories made possible by those notions is so formidable as in any case to make the taste worth acquiring.

3. Experiment and proof

When Hilbert and Bernays raised the question of the semantic completeness of quantification theory (in *Hilbert and Ackermann 1928* and the lectures[9] that led to the presentation there), they suggested that it had already been answered, that it was already 'known' that the axiomatization of quantification theory they specified 'suffices for all applications'. But from their point of view, the question remained open in a crucial sense, for, they claimed, the fact that nothing is missing from that axiom system was 'only known purely empirically', i.e., from observing that all known legitimate first-order inferences could be emulated purely formally in the system. They concluded that '[w]hether the axiom system is complete in the sense that from it all logical formulas that are correct for each domain of individuals can be derived is still an unsolved question', because it had not been shown mathematically that this is true (*Hilbert and Ackermann 1928*, p. 68). Hilbert had made an analogous remark in an address at the Bologna International Congress of Mathematicians: '[T]he question of the completeness of the system of logical rules, put in general form, constitutes a problem of theoretical logic. Up till now we have come to the view that these rules suffice only through experiment [*probieren*]' (*Hilbert 1929*, p. 140).[10]

It is possible to interpret Hilbert's words in two different ways. Perhaps he felt that our 'knowledge' that nothing is missing from the axiomatization lacks the certainty of mathematical conviction so long as it is attained empirically, that mathematical proof would rule out the haunting but unlikely possibility that we have overlooked something. On the other hand, in the same breath in which they questioned the appropriateness of empirical methods in this arena, Hilbert and Bernays did claim to know that their system is fully sufficient. Perhaps, therefore, the completeness question did not remain open in the sense that they did not know its answer, but only in the sense that *as a mathematical problem* it was still unsolved.

[9]See footnote 4.

[10]The translations of passages from *Hilbert 1929* and *Hilbert and Ackermann 1928* are from quotations in *Dreben and Van Heijenoort 1986*.

My view is that the second reading has more weight behind it. Hilbert seems to have been motivated to demonstrate that questions hitherto conceived of essentially as part of empirical science or transcendental philosophy are actually better understood as part of mathematics. Elsewhere[11] I have defended an interpretation of the notorious Hilbert program driven by this theme, and I suspect that it is at work again here in Hilbert's formulation of the problem of semantic completeness. But my purpose in drawing attention to Hilbert's words, now, is not to defend any particular reading of them. They simply provide the right background against which Gentzen's views stand out.

In the opening sentence of the *Untersuchungen*, Gentzen cited *Hilbert and Ackermann 1928*, but in the only passage that follows in which he discussed a question of completeness he wrote the following:

> The fact that this formal system does actually allow us to represent the types of proof customary in informal arithmetic (as long as they do not use complete induction) cannot be *proved*, since for considerations of an informal character no precisely delimited framework exists. We can merely verify this in the case of individual informal proofs by testing them [*durch den Versuch davon*]. (p. 112)

In this passage Gentzen was concerned, not with pure quantification theory, but with a formal system of arithmetic without complete induction. His purpose was to apply his *verschärfte Hauptsatz* to show that this system is consistent. But Gentzen did not merely want to show how the *Hauptsatz* can be applied. He had independent interests in arithmetical consistency proofs. Thus before beginning the proof, he explicitly considered whether the system could rightly be considered a systematization of arithmetic, i.e., whether the system is complete.

As the passage indicates, Gentzen immediately concluded that consideration with two claims—first, that the only means available to approach the completeness question are empirical, and second, that those methods are nevertheless insufficient for a conclusive answer in this case because informal arithmetic is imprecise in a way that prevents us from attaining a definite positive answer to the question. The question, as Gentzen saw it, is whether or not his system is adequate with respect to a particular subject matter. Whereas the system is perfectly precise, the relevant subject matter—informal arithmetic—is not. An answer to the question in the form of a proof is therefore impossible, and we are left to consider examples of actual arithmetical reasoning as we encounter them and to ask whether they can be formalized in the system.

Why does Gentzen not ask a similar question about pure quantification theory? One expects scruples analogous to those Gentzen expressed about his systematization of informal arithmetic to surface in sections II and III of the *Untersuchungen*, immediately

[11] In *Franks 2009*, especially chapter 2.

after the definitions of the logical calculi and prior to the proof of the *Hauptsatz*. Are these systems adequate to their subject matter? Is anything missing from them?

On Gentzen's view, these two questions are distinct and, in the case of quantification theory, have different answers. The distinction is subtle and foreign to the modern view of logic. Moreover, it is only implicit in Gentzen's writing. Indeed, when Gentzen asked these questions explicitly, he asked them about his system of induction-free arithmetic, where they coincide. Their answers were that we don't know whether anything is missing from that system nor whether the system is fully adequate, for the single reason that the informal theory of numbers is not sufficiently precise to allow definitive answers. But the answers Gentzen would have given to the same questions, if he had thought it sensible to ask them in this arena, about his classical calculi NK and LK, are that nothing is missing from them and that they do not have a subject matter.

In fact, Gentzen's certainty that his classical calculi are complete, in the original sense that they do not lack anything, stems directly from the fact that he did not conceive of a set of first-order logical truths that his calculi were supposed to capture. In section II.1, Gentzen called the expressions $(X \vee (Y \mathbin{\&} Z)) \supset ((X \vee Y) \mathbin{\&} (X \vee Z))$, $(\exists x \forall y Fxy) \supset (\forall y \exists x Fxy)$, and $(\neg \exists x Fx) \supset (\forall y \neg Fy)$ 'true formulae', but he wrote quotation marks around this label and never explained what it meant. What he did explain was 'how to see their truth in the most natural way', i.e., how to prove them informally. Their three proofs then became the templates for his natural deduction rules. So Gentzen did not distinguish being true from being informally provable, and thus he considered truth an inherently informal, though immanent rather than transcendent notion.[12] Rather than a set of first-order logical truths, he conceived only of valid quantificational schemata determined by actual patterns of reasoning. That these patterns could be directly formalized and captured was the point of the theory of natural deduction.

In subsection 1 of the synopsis, Gentzen announced his aim 'to set up a formal system that comes as close as possible to actual reasoning' (p. 68). The advantage of natural deduction has since been thought to lie primarily in the fact that the actual implementation of the calculus *comes naturally* and feels familiar compared to the implementation of axiomatic systems: Because natural deduction systems are modeled after the flow of reasoning in mathematical proofs, formalizing known mathematical proofs will be a matter of direct translation, and discovering new ones will be facilitated. But Gentzen thought that this same feature of natural deduction affords a more serious advantage: It

[12] In the middle of a section called 'Application of the sharpened Hauptsatz to a new consistency proof for arithmetic without complete induction' Gentzen mentions a 'customary way' of explaining 'the truth or falsity of [formulas of the form] $\mathfrak{A} \mathbin{\&} \mathfrak{B}, \mathfrak{A} \vee \mathfrak{B}, \neg\mathfrak{A}$, and $\mathfrak{A} \supset \mathfrak{B}$, as functions of the truth or falsity of the subformulae' (p. 114). This passage—the only mention in the *Untersuchungen* of the conception of propositional connectives as truth functions—occurs, not as one might expect in the definition of the logical connectives, but in the stipulation of an algorithm for reducing the complexity of arithmetical proofs. This algorithm makes no analogous use of a conception of the truth of quantificational formulas: Instead quantifiers are eliminated from a given normal-form arithmetical proof through a numerical substitution determined by the proof's structure.

allows for the empirical verification that the calculus is complete to be conclusive.

To see why Gentzen thought this, it suffices to look at the way he described logic and its relationship to mathematics in the *Untersuchungen*. In the opening sentences, Gentzen wrote of 'predicate logic' that '[i]t comprises the types of inference that are continually used in all parts of mathematics', and added:

> What remains to be added to these are axioms and forms of inference that may be considered as being proper to the particular branches of mathematics, e.g., in elementary number theory the axioms of the natural numbers, of addition, multiplication, and exponentiation, as well as the inference of complete induction; in geometry the geometric axioms. (p. 68)

Gentzen should be read literally: Specific branches of mathematics have independent subject matters which we can try to formalize by adding to 'those inferences continually used in all parts of mathematics' (predicate logic) axioms and forms of inference corresponding to the methods used by the mathematicians who research those branches of mathematics. The completeness question for our formalization of those branches of mathematics will always be elusive, Gentzen thought, because the subject matter itself is vague. At any time inspired mathematicians might devise new methods to uncover further of its details. No corresponding problem arises in pure predicate logic, though, because its principles are designed to track, not any independent mathematical subject matter, but the inferences that mathematicians actually use. Something 'remains to be added' to the formalization of these principles before the system can rightly be considered a systematization of any independent subject matter.

Still, a type of completeness question could arise for predicate logic, on this view, if mathematical practice exhibited so wide a range of inference types that it proved difficult to ensure that one has incorporated them all. Alternatively, one's logical system could host principles of an entirely different sort than the informal ones used by mathematicians—different not merely in their *formality*, but in their very *form*—so that the coextensionality of the formal system and the body of informal inference patterns could only be checked globally. Gentzen's method of section II of the *Untersuchungen* was to address the second of these possibilities directly, by presenting 'a formalism that reflects as accurately as possible the actual reasoning involved in mathematical proofs' (p. 74). He had only to devise the correct grid to impose on mathematical discourse in order to recognize that single inferences could span a large swath of the text of a proof, and that this property could be preserved in a formal calculus. But this same realization led immediately, though indirectly, to an answer to the first completeness question. Gentzen discovered that under the conceptual grid he devised to isolate and categorize individual mathematical inferences, an extremely small number of inference types appear. These inference types, moreover, are systematically coordinated with the logical particles that characterize the statements involved in the inferences, so that with each logical particle are associated two rules, one rule for its introduction into mathematical discourse, and a dual, 'elimination' rule governing how to reason from statements characterized by that particle—a ‘classification’, that Gentzen remarked in *1936* ‘almost suggests itself’ (p. 148).

For example, a *conjunctive* statement of the form ‘$\mathfrak{A}$ and $\mathfrak{B}$’ can be established directly from the previous establishment of the statements ‘$\mathfrak{A}$’ and ‘$\mathfrak{B}$’, and from the statement ‘$\mathfrak{A}$ and $\mathfrak{B}$’ one can establish individually as true its two components ‘$\mathfrak{A}$’ and ‘$\mathfrak{B}$’. In Gentzen’s calculi NI and NK, these inference patterns are represented by the rules:

$$\dfrac{\mathfrak{A} \qquad \mathfrak{B}}{\mathfrak{A} \,\&\, \mathfrak{B}}\,\&I \qquad\qquad \dfrac{\mathfrak{A} \,\&\, \mathfrak{B}}{\mathfrak{A}} \qquad \dfrac{\mathfrak{A} \,\&\, \mathfrak{B}}{\mathfrak{B}}\,\&E$$

Gentzen observed that this ‘introduction/elimination’ grid brought to the fore pairs of coordinated rules for each of the logical particles ‘or’, ‘for all’, ‘there exists’, ‘if ... then’, and ‘not’ analogous to those for ‘and’:

$$\dfrac{\mathfrak{A}}{\mathfrak{A} \vee \mathfrak{B}} \qquad \dfrac{\mathfrak{B}}{\mathfrak{A} \vee \mathfrak{B}}\,\vee I \qquad\qquad \dfrac{\mathfrak{A} \vee \mathfrak{B} \qquad \begin{array}{c}[\mathfrak{A}]\\ \mathfrak{C}\end{array} \qquad \begin{array}{c}[\mathfrak{B}]\\ \mathfrak{C}\end{array}}{\mathfrak{C}}\,\vee E$$

$$\dfrac{\mathfrak{F}\mathfrak{a}}{\forall \mathfrak{x}\mathfrak{F}\mathfrak{x}}\,\forall I \qquad\qquad \dfrac{\forall \mathfrak{x}\mathfrak{F}\mathfrak{x}}{\mathfrak{F}\mathfrak{a}}\,\forall E$$

$$\dfrac{\mathfrak{F}\mathfrak{a}}{\exists \mathfrak{x}\mathfrak{F}\mathfrak{x}}\,\exists I \qquad\qquad \dfrac{\exists \mathfrak{x}\mathfrak{F}\mathfrak{x} \qquad \begin{array}{c}[\mathfrak{F}\mathfrak{a}]\\ \mathfrak{C}\end{array}}{\mathfrak{C}}\,\exists E$$

$$\dfrac{\begin{array}{c}[\mathfrak{A}]\\ \mathfrak{B}\end{array}}{\mathfrak{A} \supset \mathfrak{B}}\,\supset I \qquad\qquad \dfrac{\mathfrak{A} \qquad \mathfrak{A} \supset \mathfrak{B}}{\mathfrak{B}}\,\supset E$$

$$\dfrac{\begin{array}{c}[\mathfrak{A}]\\ \curlywedge\end{array}}{\neg\mathfrak{A}}\,\neg I \qquad\qquad \dfrac{[\mathfrak{A}] \qquad \neg\mathfrak{A}}{\curlywedge}\,\neg E$$ [13]

Let us designate the calculus comprised of the foregoing rules as ‘the minimal cal-

[13] The rules in this table have the exact form of Gentzen’s (p. 77): The flow of reasoning is downward in each diagram; each rule warrants the inference of the expression below the horizontal line from the expression(s) above that same line and may be instantiated in the presence of zero or more assumptions upon which its lower formula is said to depend; expressions in square brackets represent assumptions, and the expressions immediately beneath them are meant to be known to be deducible from them; $\curlywedge$ is Gentzen’s symbol for absurdity. The free object variable $\mathfrak{a}$ in the rules $\forall$I and $\exists E$ is called an *eigenvariable* and is subject to the familiar restrictions: In an instance of $\forall I$, $\mathfrak{a}$ can occur neither in $\forall\mathfrak{x}\mathfrak{F}\mathfrak{x}$ nor in any assumption formula upon which that formula depends; in an instance of $\exists E$, $\mathfrak{a}$ can occur neither in $\exists\mathfrak{x}\mathfrak{F}\mathfrak{x}$ nor in $\mathfrak{C}$, nor in any assumption formula upon which the upper occurrence of that formula depends apart from the displayed occurrence of the formula $\mathfrak{F}\mathfrak{a}$.

culus', as it is a formalization of I. Johansson's (*1937*) minimal logic. The coordination of its rules, Gentzen explained, consists in the fact that, given a logical particle, once it is observed how a statement governed by that particle is introduced in reasoning, the corresponding 'elimination' rule describing how to reason from that statement can be deduced:

> The introductions represent, as it were, the 'definitions' of the symbols concerned, and the eliminations are no more, in the final analysis, than the consequences of these definitions. This fact may be expressed as follows: In eliminating a symbol, we may use the formula with whose terminal symbol we are dealing only 'in the sense afforded it by the introduction of that symbol'. (p. 80)

Gentzen illustrated this method of rule-deduction with the example of the rules for the symbol $\supset$, which corresponds with the informal expression 'if ... then'. The introduction rule for that symbol tells us that statements of the form $\mathfrak{A} \supset \mathfrak{B}$ are warranted when there is known to be a derivation of $\mathfrak{B}$ from $\mathfrak{A}$. Thus Gentzen explained that '[i]f we then wished to use that formula [$\mathfrak{A} \supset \mathfrak{B}$] by eliminating the $\supset$-symbol ..., we could do this precisely by inferring $\mathfrak{B}$ directly once $\mathfrak{A}$ has been proved' (Ibid.).

Gentzen concluded this observation with two noteworthy remarks: First, regarding this specific example of rule-deduction, he pointed out that the procedure could be performed without considering what he called 'the informal sense' of the symbol $\supset$ (p. 81). By the informal sense of the symbol, Gentzen surely had in mind the semantic view of the symbol as a truth function, a view that does not figure into Gentzen's development of logic anywhere in the *Untersuchungen*.[14] This is the significance of Gentzen's claim that the introduction rules themselves represent their symbols' definitions: They are legitimate rules not because of their ability to coordinate with the 'informal sense' of those symbols—what matters is that they correspond with the way statements characterized by the logical particles those symbols represent are actually introduced in mathematical practice. Because this correspondence suffices also to determine exactly how to reason from those statements, the rules fully embody the logically salient meanings of those particles.

Second, Gentzen claimed that '[b]y making these ideas more precise it should be possible to display the E-inferences as unique functions of their corresponding I-inferences, on the basis of certain requirements' (Ibid.). In sections 10 and 11 of *1936* Gentzen pointed out that the association is 1-1. Although somewhat vague, Gentzen's point here is that the general procedure of rule-derivation is precise. Much has been made of this idea as a principle of 'logical harmony'.[15] But most obviously what one can say about Gentzen's observation is that it simplifies the experimental task of verifying that the calculus adequately encompasses all the inference types that 'are continuously used in

[14]But see footnote 12.
[15]See *Belnap 1962*.

all parts of mathematics'. Each logical particle is introduced and eliminated in only one way, and the respective ways of reasoning one's way to a statement governed by that particle and of drawing inferences from a statement governed by that particle are recoverable from one another.

Thus these rules are conveniently categorized so as to almost preclude any possibility that anything has been left out from the system in one sense: A symbol's introduction and elimination rules are a full analysis of the way that symbol is used in mathematical reasoning. All other reasoning patterns that mathematical thought follows are inevitably tied to some specific field of inquiry, save two, to which Gentzen attended closely:

$$\frac{\curlywedge}{\mathfrak{D}}\,\curlywedge E \qquad\qquad \frac{\neg\neg\mathfrak{A}}{\mathfrak{A}}\,\neg\neg E$$

Gentzen wrote: 'The schema $\frac{\curlywedge}{\mathfrak{D}}$ expresses the fact that if a false proposition holds, any proposition holds' (p. 79) and noted two pages later that this rule 'occupies a special place among the schemata: It does not belong to a logical symbol but to the propositional symbol $\curlywedge$'. NI is the system that results by adding to the minimal calculus only this first inference rule. It corresponds to intuitionistic predicate logic. To obtain 'a complete [*vollständiger*] classical calculus NK' Gentzen observed that it suffices to add to NI a single schema for 'basic formulae' of the form '$\mathfrak{A} \vee \neg\mathfrak{A}$', corresponding to the principle *tertium non datur*. Then he explained (and it is easy to verify) that '[i]t would be perfectly feasible to introduce a new inference figure schema, say $\frac{\neg\neg\mathfrak{A}}{\mathfrak{A}}$, in place of the basic formula schema $\mathfrak{A} \vee \neg\mathfrak{A}$', because the calculi that result in each case are deductively equivalent. 'However', Gentzen cautioned, 'such a schema still falls outside the framework [of introduction and elimination rules], because it represents a new elimination of the negation whose admissibility does not follow at all from our method of introducing the $\neg$-symbol by the $\neg I$' (p. 81). In *1936* he elaborated:

> This form of inference conflicts in fact quite categorically with the remaining forms of inference. In the case of the logical connectives $\forall$, &, $\exists$, $\vee$, and $\supset$ we had in each case an introduction and an elimination inference corresponding to each other in a certain way. ...Double negation [elimination, by contrast] renders possible indirect proofs of positive propositions from their denials by means of contradiction, in cases where a positive [i.e., direct] proof of the same proposition may be completely unobtainable[16] [*gar nicht zu erhalten ist*]. (p. 169)

Strictly speaking, the calculi NI and NK are slightly more closely related to the minimal calculus than this progression indicates. NI is the result of adding the rule $\curlywedge E$ to the minimal calculus. Gentzen then characterized NK as the result of adding the basic formula schema $\mathfrak{A} \vee \neg\mathfrak{A}$ to NI. But it is easy to verify that in NK the rule $\curlywedge E$ does not

[16]Szabo has 'inaccessible', somewhat misleadingly.

allow for the derivation of any expressions that cannot be derived without it—i.e., that in the presence of the rule $\neg\neg E$ (or, equivalently, of *tertium non datur*), the use of $\curlywedge E$ is eliminable from all derivations. For suppose that one has derived the expression $\curlywedge$. Then $\curlywedge E$ allows one to infer an arbitrary expression $\mathfrak{D}$ forthright. But instead one might note the validity of the scheme $\begin{matrix}[\neg\mathfrak{D}]\\ \curlywedge\end{matrix}$, which immediately extends to $\dfrac{\begin{matrix}[\neg\mathfrak{D}]\\ \curlywedge\end{matrix}}{\neg\neg\mathfrak{D}}\,\neg I$ and then to $\dfrac{\dfrac{\begin{matrix}[\neg\mathfrak{D}]\\ \curlywedge\end{matrix}}{\neg\neg\mathfrak{D}}\,\neg I}{\mathfrak{D}}\,\neg\neg E$. Thus NK extends the minimal calculus simply in its accommodation of the classical principle *tertium non datur*, where NI had only the weaker principle *ex falso quodlibet*. Gentzen's principal observation was that, apart from principles of inference contentfully associated with specific branches of mathematical inquiry (e.g., mathematical induction in number theory), ordinary mathematical practice exhibits no reasoning patterns not already present in the minimal calculus other than these. Nothing could be missing from a calculus that accommodates two rules for each logical particle so tightly coordinated as to effectively define that particle, together with the single other principle ever appealed to in general mathematical discourse.

The sufficiency of the natural deduction rules is thus guaranteed by (1) their nature together with (2) a contingent fact about mathematical proofs and (3) Gentzen's modest vision of the purpose of the predicate calculus. To sum up: (1) The rules of the calculi NI and NK take the exact form of actual mathematical inferences as they are represented in written proofs, and (2) the inference types immediately recognizable in this scheme are extremely few and categorically arranged so that one can be sure that none have been left out, because (3) if there are any inferences that have been overlooked, then for that very reason they surely fail to meet the criterion of ubiquity in mathematical practice that Gentzen imposed.

Gentzen's analysis is obviously counter to Hilbert and Bernays' suggestion (from *Hilbert and Ackermann 1928* et. al.) that empirical methods could not properly secure an answer to the completeness problem. If one has a theory of first-order truths that functions independently of a formal calculus, and if one wants assurance that all and only the truths from the semantic realm are theorems of that calculus, then however much inductive evidence one amasses that this is so, *a priori* methods surely are still welcome. But that semantic realm is itself an analysis of our intuitive notions of truth and consequence—notions that Gentzen thought mathematicians would be the first and only authority on. To analyze *their* intuitions, Gentzen simply modeled his logical calculus on the inferences they actually make[17]. Then he observed that the empirical

[17]The idea is not that natural deduction somehow models the experience of doing mathematics. Gentzen did not pursue psychological or phenomenological investigations. But he did formulate his logical rules according to the form in which the deductive part of mathematical theories is presented in texts, i.e., as proofs, and it was an empirical discovery that doing so led to calculi with unexpected systematic symmetries.

task of determining whether he had left anything out was trivial. No coordination of syntax and semantics was called for: Because Gentzen analyzed the informal notions of logical truth and consequence directly with his logical calculus, the analysis did not take the form of a 'semantic theory' and natural deduction retained its status of *lingua characteristica*.

There remains a question: why, in *1936*[18], did Gentzen cite Gödel's completeness theorem? Had he loosened his hold on his contrary conception of this problem in the year following the submission of his thesis? I think that the counter-evidence alone, in the simple fact that Gentzen mentioned the theorem nowhere else in his later writing—including several places where its omission is jarring—outweighs any suggestion that he did. But a careful look at his description of Gödel's result reveals that even as he paid homage to the theorem, Gentzen preserved key features of his own point of view.

To begin with, in his explanation of completeness, Gentzen wrote about 'correct inferences' rather than correct or valid formulas. This could not have been inspired by Gödel's own discussions, for Gödel always used 'logical consequence' to refer to the purely syntactic relation of formal derivability. But as we have seen, Gentzen did not share Gödel's transcendent notion of quantificational truth and thought of true expressions of the predicate calculus merely as formalizations of sentence schemata that are informally provable. To claim that all such expressions are provable would have sounded trivial and been confusing. Clearer is Gentzen's claim that all correct, informal inferences 'are already represented'.

Also, though it is customary to speak of the *soundness* of an individual rule—so that a system is sound if, and only if, each of its rules is—one typically poses the question of *completeness* only of systems as a whole. This certainly is the conception driving Gödel's own discussion of the completeness phenomenon. Yet Gentzen's conception of completeness in terms of the ability to capture correct informal inferences engenders the unorthodox notion of an individual rule being complete with respect to the correct informal inferences 'of its type'. Thus he cites Gödel as having proved the completeness of 'the rules belonging to the connectives $\&, \vee, \supset, \neg, \forall, \exists$', seemingly oblivious to the fact that Gödel showed that these rules not only allow for the formal regimentation of any informal inference involving their specified logical particles, but indeed for 'the derivation of *every* logico-mathematical proposition' whatever. Thus, even as Gentzen acknowledges that it is possible to prove that a system of rules is complete, he means only that with each individual logical particle are associated rules that completely specify the legitimate inferences it affords. The conviction that these logical particles are all the ones whose legitimate inferences must be so specified must still derive from empirical observation.

Finally, the usual understanding of completeness as a coordination of syntax and semantics is entirely missing from Gentzen's explanation of Gödel's result. In place of

[18] An anonymous referee for the Journal informs me that the relevant passage was written in the spring of 1935.

that image, Gentzen preserved the more traditional image of having not left out any of the essential rules, rules being fully interpreted inference patterns. The coordination that Gentzen thought was relevant, even to the extent that he used it to explain Gödel's accomplishment, is one between informal inference patterns and the rules of a formal system.

In *1936*, two sections after the reference to Gödel's theorem, Gentzen claimed that '[i]t is easily proved that the logical rules of inference, applied to [the theory of a finite domain of objects], are correct [*richtig*] in the sense that their application to "true" basic mathematical sequents leads to "true" derivable sequents' (p. 159). Here Gentzen focuses on those rules' *soundness* rather than their *completeness*, and the reason for this is clear: When applied to the basic mathematical sequents designed to axiomatize the theories that Gentzen was interested in, the rules are *not* complete. Of course, from the modern point of view, what is more fundamental is the soundness of those rules in and of themselves, apart from any application. And when their soundness is viewed in this context-less way, the corresponding question of their completeness gets an affirmative answer. But as we have seen, the conception of soundness and completeness of a logical system in terms of the truth of the various expressions is available to Gentzen only in the context of its application.

But crucially, as 'easily proved' as such a soundness result might be, Gentzen chose to omit its proof and remarked:

> A verification of this statement would mean no more than an *acknowledgment*[19] [*Bestätigung*] of the fact that we have indeed chosen our formal rules of inference in such a way that they are in harmony with the informal sense of the logical connectives. (Ibid.)

Gentzen omitted the soundness proof because he believed that it would be uninformative. His system's rules are guaranteed from the start to allow only legitimate inferences, because they are an analysis of the inferences mathematicians make. There is no corresponding completeness proof for Gentzen to omit on the same grounds: Applied to any mathematical theory, the rules of the predicate calculus will be incomplete because missing from them will be any axioms and rules unique to that theory. And in their pure, unapplied form, the expressions of the predicate calculus are, according to Gentzen, not about anything so that the question of their truth does not arise. But clearly the complicated proof of the completeness of NK would be uninteresting to Gentzen for the same reason that the simple proof of its soundness is: It would be 'no more than an *acknowledgment* of the fact that we have indeed chosen our formal rules of inference' appropriately—a fact Gentzen had already confirmed empirically.

[19] Szabo de-emphasizes Gentzen's intended contrast between '*Nachweis*' (verification) and '*Bestätigung*' by rendering the latter as 'confirmation'. Gentzen's point is that the system's soundness is guaranteed by its design and therefore need not be proved.

4. Analysis and synthesis

The principal result [*Hauptsatz*] of the *Untersuchungen* is the normalization technique for quantification theory known today as cut-elimination. In subsection 2 of the synopsis, Gentzen explained that '[t]he *Hauptsatz* says that every purely logical proof can be reduced to a definite, though not unique, normal form', and added: 'Perhaps we may express the essential properties of such a normal proof by saying: it is not roundabout [*er macht keine Umwege]*' (p. 69). Thus everything provable in predicate logic turns out in fact to have a direct proof into which '[n]o concepts enter ... other than those contained in its final result' (Ibid.). These same comments are repeated in section III.2 immediately prior to the proof of the *Hauptsatz*.

Although the first logical systems Gentzen presented in the *Untersuchungen* were the 'natural' calculi NI and NK, these present certain obstacles for the proof of the *Hauptsatz*. For that reason, in sections III and IV, Gentzen developed a different style of logical calculus, called sequent calculi, proved the normalization result for its intuitionistic and classical varieties, and exhibited several applications of that result. He explained:

> In order to be able to enunciate and prove the *Hauptsatz* in a convenient form, I had to provide a logical calculus especially suited for this purpose. For this the natural calculus proved unsuitable. For, although it already contains the properties essential to the validity of the *Hauptsatz*, it does so only with respect to its intuitionistic form, in view of the fact that the law of excluded middle ... occupies a special position in relation to these properties. (p. 69)

With this development in mind, the sequent calculi that Gentzen introduced appear at first to be a later development than the natural calculi of section II.[20] And indeed, if one views the value of the *Hauptsatz* purely in terms of its vast applicability, the sequent calculi appear to have merely instrumental value, in so far as they allow for such applications. Gentzen's claim that '[t]heir form is largely determined ... by considerations connected with the "*Hauptsatz*"' has led some readers to think that the calculi were designed strictly in order to exhibit the sort of combinatorial features that facilitate consistency proofs and the like (p. 83).

But it is essential to acknowledge that the basic framework of the sequent calculus predated the *Untersuchungen* in the work[21] of Paul Hertz and in Gentzen's own *1932*

[20]Nor in one crucial sense is this appearance misleading: The calculi LK and LI emerged from the details of Gentzen's original translation scheme from proofs of the natural deduction calculus N2 to proofs of the axiomatic calculus of *Hilbert and Ackermann 1928*. This translation appears prominently in the unpublished, handwritten draft of the *Untersuchungen*. Jan von Plato mentions the relevance of this translation for the discovery of the first sequent calculus, a contraction-free system (NL3), on page 242 of his *2008* and illustrates the emergence of the rules for the calculus N2 in §8 of his *2009*.

[21]See especially *Hertz 1929*. For more discussion of Hertz' contributions, see *Bernays 1965* and *Schroeder-Heister 2002*.

paper 'Über die Existenz unabhängiger Axiomsysteme zu unendlichen Satzsysteme'. In the *Untersuchungen*, Gentzen simply extended the purely atomic system of *1932* with rules for the logical particles. 'What we want to do', he claimed there, 'is to formulate a deductive calculus (for predicate logic)' modeled after his earlier system

> on the one hand, i.e., in which the derivations do not, as in the calculus NI, contain assumption formulae, but which, on the other hand, takes over from the calculus NI the division of the forms of inference into introductions and eliminations of the various logical symbols. (p. 82)[22]

What prompted Gentzen to think to devise a calculus with just these properties? It is tempting to answer that he foresaw that these were the properties that would allow for the troublesome aspects of *tertium non datur* to vanish and remove all obstacles to a proof of the *Hauptsatz* for classical predicate logic. But in comments about the sequent calculus in *1938a* Gentzen emphasized that he had no such foresight:

> All of the connectives &, $\vee$, $\forall$, $\exists$, $\neg$ [23] have, to a large extent, equal status in the system; no connective ranks notably above any other connective. The special position of the negation, in particular, which constituted a troublesome exception in the natural calculus has been completely removed in a seemingly magical way. The manner in which this observation is expressed is undoubtedly justified since I myself was completely surprised by this property of the 'LK-calculus' when first formulating that calculus. (p. 259)

The special position of the negation that Gentzen mentioned is the fact that the classically valid rule for $\frac{\neg\neg\mathfrak{A}}{\mathfrak{A}}$ does not figure into the introduction and elimination classification for logical symbols. Because this rule is equivalent to the axiom scheme for *tertium non datur*, one observes that it is possible to reason classically from statements characterized by negation in a way other than what is licensed by $\neg E$. Thus negation stands out as an exception to the general principle that all logical connectives are on a par with one another, and this exception proves troublesome because it creates an obstacle for the proof of the *Hauptsatz*. But the suggestion that his initial motivation in creating the sequent calculus could have been to facilitate a proof of the *Hauptsatz* is implausible when one considers Gentzen's claim that the suppression of these features in the sequent calculus came to him as a surprise.

Gentzen's dissatisfaction with natural deduction that motivated his development of a new calculus was not the failure of NK proofs to normalize. He sought a logical calculus that both (1) is 'logistic' like the calculus of *1932* and (2) 'takes over from the

[22] See §8 of von Plato's *2009* for an illustration of how, in Gentzen's words, NL3's 'logistic form of proof arises from the natural one' from the details of the translation scheme of Gentzen's handwritten thesis.

[23] For technical reasons related to the consistency proof of this paper, Gentzen eliminated the symbol $\supset$ and its rules from the calculus.

calculus NI the division of the forms of inference into introductions and eliminations of the various logical symbols', because he was unhappy with certain aspects of the analysis of the meanings of logical particles into introduction and elimination rules in the natural calculi. That analysis had proved adequate to ensure that the calculus is complete, because it tracked exactly the inferences that mathematicians make. But it failed to be analytic in an important sense, because it entangled the meanings of logical particles with the synthetic notion of logical consequence. Thus Gentzen designed the sequent calculus to separate the analytic and synthetic components of proofs, thereby allowing the independent contribution to the meanings of our expressions that logical particles make to stand out more distinctly.

In order to understand how Gentzen was led to this view, it is essential to review the relatively unknown work *Gentzen 1932*. In that earlier paper, Gentzen presented a 'formal definition of provability' that he proved to be sound and complete with respect to the 'informal' notion of logical consequence. From the modern point of view, these soundness and completeness proofs are squarely in the paradigm of syntax/semantics coordination—the same paradigm so conspicuously absent from *Gentzen 1934–35* and onward.

Conceived as a formal system, Gentzen's formal definition of provability consists of 'sentences' of the form $M \longrightarrow v$, where v is an 'element' and M is a 'complex' (a non-empty set of finitely many elements.) Gentzen's typographic convention is that concatenation of complex letters represents their set-theoretical union. Sentences can also be written with the elements of a complex displayed, thus: $u_1, u_2, \ldots u_n \longrightarrow v$. Because complexes are sets, the same element cannot appear multiple times in the same complex, and the order in which the elements of a complex are listed is immaterial. Gentzen referred to the complex of a sentence as its antecedent and to the lone element on the right of the arrow symbol as the succedent. He defined tautologies to be those sentences whose antecedent is the singleton set containing the same element that appears in the sentence's succedent.

Gentzen specified two inference rules for his system, which he called 'thinning' and 'cut':

$$\frac{L \longrightarrow v}{ML \longrightarrow v}\,\textit{thinning} \qquad\qquad \frac{L \longrightarrow u \qquad Mu \longrightarrow v}{LM \longrightarrow v}\,\textit{cut}$$

Then he defined a 'proof' of a sentence $\mathfrak{q}$ from the sentences $\mathfrak{p}_1, \ldots, \mathfrak{p}_\mathfrak{v}$ to be 'an ordered succession of inferences (i.e., thinnings and cuts) arranged in such a way that the conclusion of the last inference is $\mathfrak{q}$ and that its premises are either premises of the $\mathfrak{p}$'s or tautologies' (p. 31). In contemporary terminology, this system corresponds to a purely atomic single conclusion sequent-calculus (The structural rules for contraction and interchange are implicit because of Gentzen's use of set-theoretical complexes instead of sequences); if we conceive of the comma as logical conjunction and the sequent arrow as a material conditional, then the proof system is also equivalent to SLD-resolution.[24]

In section 4, Gentzen wrote:

> Our formal definition of provability, and, more generally, our choice of the forms of inference will seem appropriate only if it is certain that a sentence $\mathfrak{q}$ is 'provable' from the sentences $\mathfrak{p}_1, \ldots, \mathfrak{p}_\mathfrak{v}$ if and only if it represents informally a consequence of the $\mathfrak{p}$'s. We shall be able to show that this is indeed so as soon as we have fixed the meaning of the still somewhat vague notion of 'consequence', in accordance with a particular informal interpretation of our 'sentences' ... (p. 33).

Gentzen distinguished in this question of the appropriateness of a formal proof system the two familiar components, the system's 'informal correctness' and its 'informal completeness'. Then, quite contrary to the claim in the *Untersuchungen* that 'for considerations of an informal character no precisely delimited framework exists', he specified a precise notion of 'informal consequence':

> We say that a complex of elements *satisfies* a given sentence if it either does not contain all antecedent elements of the sentence, or alternatively, contains all of them and also the succedent of that sentence. ... We now look at the complex K of all (finitely many) elements of $\mathfrak{p}_1, \ldots, \mathfrak{p}_\mathfrak{v}$ and $\mathfrak{q}$ and call $\mathfrak{q}$ a *consequence* of $\mathfrak{p}_1, \ldots, \mathfrak{p}_\mathfrak{v}$ if (and only if) every subcomplex of K which satisfies the sentences $\mathfrak{p}_1, \ldots, \mathfrak{p}_\mathfrak{v}$ also satisfies $\mathfrak{q}$. (p. 33)

Theorem I of *Gentzen 1932* states that the proof system is 'correct': 'if a sentence $\mathfrak{q}$ is "provable" from the sentences $\mathfrak{p}_1, \ldots, \mathfrak{p}_\mathfrak{v}$ then it is a "consequence" of them'. Thus Theorem I is a soundness theorem. It is 'informal', not because of any lack of precision in its methods, but in so far as the result coordinates Gentzen's formal sentence system with the informal, intuitive notion of consequence. As one would expect, Gentzen's statement of 'informal completeness' is the converse of this result: 'If a sentence $\mathfrak{q}$ is a "consequence" of the sentences $\mathfrak{p}_1, \ldots, \mathfrak{p}_\mathfrak{v}$, then it is also "provable" from them'. Gentzen established the informal completeness of his sentence system in Theorem II, where he in fact showed that proofs of a specific 'normal form' alone suffice to exhibit all the consequences among sentences in his system.

The comfort with which Gentzen works with these notions of soundness and completeness makes all the more striking his distance from them in *1935* and thereafter. But this shift in methodology need not be puzzling. In *1932* Gentzen dealt with the pure notion of logical consequence. The elements of his formal calculus are logically unrelated to one another, and the 'sentences' comprised of these elements are themselves interpretable as statements of logical consequence. Gentzen considered the exemplar sentence $u_1, \ldots u_n \longrightarrow v$ and suggested several interpretations of it ranging from event causation to property containment. Two crucial suggestions stand out:

[24]On this point, see *Schroeder-Heister 2002*.

> The 'sentence' $[u_1, \ldots u_n \longrightarrow v]$ may also be understood thus: A domain of elements containing the elements $u_1, \ldots u_n$ also contains the element v.
>
> ...
>
> Or we might imagine the elements to stand for 'propositions', in the sense of the propositional calculus, and the 'sentence' then reads: If the propositions $u_1, \ldots u_n$ are true, then proposition v is also true. (p. 30)

Obviously, these two interpretations correspond exactly with the informal relation of logical consequence that Gentzen defined over the sentences themselves (according to whether one construes elements generally or as propositions).

Gentzen's point was that the pure notion of logical consequence is at once simple, uncontroversial, and easily enough specified to be captured formally by a logical calculus, and the soundness and completeness results of *1932* are the proof that his 'formal definition of provability' does just that. In other words, the system of *1932* based on 'cut' and 'thinning' is a fully adequate systematization of the pure notion of logical consequence. Gentzen described this result, not as a coordination of the semantic notion of consequence and the syntactic notion of derivability, but as the formalization of the informal notion. Thus his correctness and completeness theorems serve again as an analysis of an informal notion—this time, the notion of pure logical consequence—directly into a logical system.

Crucially, logical consequence so conceived is synthetic in the Aristotelian sense. If one knows only that a sentence $u_1, \ldots u_n \longrightarrow v$ was obtained from an application of 'cut', it is not possible to determine what sentences were used as premises for that inference, because the 'cut element' vanishes in the course of the inference. Conversely, however, given a collection of truths, presented as sentences in the style of *Gentzen 1932*, from some field of inquiry, it is possible to attempt various pairings of sentences from this collection as premises of a cut inference in order to obtain new sentences, thereby expanding the size of the collection. What Gentzen took himself to have proved is that *all* purely logical reasoning that does not turn on a specific understanding of the meanings of any components of individual expressions can be recovered in just this way. Because the rule for thinning is doing very little real work in this system, it makes sense to say that Gentzen proved that the cut rule is a fully adequate formalization of the pure notion of logical consequence.

By contrast, the expressions of predicate logic and even of propositional logic are logically interrelated because of the contribution that their constituent logical particles make to their meanings. Articulating a detailed account of the consequence relation among these expressions requires a semantic theory. Such a semantic theory will inevitably be less simple, more controversial, and more difficult to specify than the pure consequence relation, for it requires a successful analysis of the various logical particles.

It was this sort of semantic analysis that Gentzen was skeptical of. The very idea of such a semantic theory presupposes a notion of quantificational truth that transcends human reasoning. In place of such a theory about the meanings of the logical particles of predicate logic, Gentzen opted for a formal emulation of the use mathematicians make of statements characterized by those particles—effectively an analysis of those particles directly into the structure of his logical system. This analysis disclosed that the use that mathematicians make of the expression 'if ... then' involves a form of 'cut'. (The rule $\supset E$ is the culprit.) This is not surprising, given the close relationship between material conditionals and statements of logical consequence. But Gentzen hoped to tease these notions apart: He wanted a logical system in which the contribution to the meaning of a sentence that the logical particles make (including the $\supset$-symbol) is perspicuous and separate from the synthetic method of generating new theorems from lemmas.

Thus the idea of the sequent calculus was to map the expressions of pure predicate logic onto the 'elements' of the *1932* formal definition of provability. To the rules 'cut' and 'thinning' Gentzen added new rules for the analysis of the logical symbols that appear within individual elements. Now that the synthetic notion of logical consequence was already a native structural rule of the system, the 'introduction' and 'elimination' rules for the logical symbols did not have to reproduce its effects.

It is not difficult to appreciate Gentzen's report that the outcome of this endeavor struck him as magical. Consider the extended passage from section III.1 of the *Untersuchungen* where the intuitionistic sequent calculus LI takes its form:

> The most obvious method of converting an NI-derivation into a logistic one is this: We replace a [derivation]-formula[25] $\mathfrak{B}$, which depends on the assumption formulae $\mathfrak{A}_1, \ldots, \mathfrak{A}_\mu$, by the new formula $(\mathfrak{A}_1 \mathbin{\&} \ldots \mathbin{\&} \mathfrak{A}_\mu) \supset \mathfrak{B}$. This we do with all [derivation]-formulae.
>
> We thus obtain formulae which are already true *in themselves*, i.e., whose truth is no longer *conditional* on the truth of certain assumption formulae. This procedure, however, introduces new [occurrences of the] logical symbols & and $\supset$, necessitating additional inference figures for & and $\supset$, and thus upsets the systematic character of our method of introducing and eliminating symbols. (p. 82)

The simple typographical solution that Gentzen proposes to address this conceptual difficulty appears at first like a joke:

> Instead of a formula $(\mathfrak{A}_1 \mathbin{\&} \ldots \mathbin{\&} \mathfrak{A}_\mu) \supset \mathfrak{B}$, we therefore write the sequent $\mathfrak{A}_1, \ldots, \mathfrak{A}_\mu \longrightarrow \mathfrak{B}$.
>
> The informal meaning of this sequent is no different from that of the above formula; the expressions differ merely in their formal structure. (Ibid.)

How does this solve the difficulty? Gentzen confessed that '[e]ven now new inference figures are required that cannot be integrated into our system of introductions and

[25]Gentzen's notion '*Herleitungsformel*': a token occurrence of a formula in a derivation tree.

eliminations'. He had in mind the inference figures for 'cut', 'thinning', and—because he now used sequences in place of sets—'contraction' and 'interchange'. 'But', he continued, 'we have the advantage of being able to reserve them special places within our system, since they no longer refer to logical symbols, but merely to the structure of the sequents' (Ibid.)

Informally, these structural rules do refer to logical particles: Gentzen himself explained that a sequent's sequent arrow and antecedent commas have the same informal sense as the symbols $\supset$ and &.[26] But the typographical distinction of the sequent calculus allows the rules governing the use of these logical particles to split, so that the aspect of informal mathematical reasoning used to synthesize lemmas into new theorems is relegated to formal structural rules, and only the remaining aspect of that reasoning occurs in the rules governing the logical symbols. The fact that the structural rules do not figure into the scheme of introductions and eliminations is immaterial, because only the rules used to *analyze* the logical particles were meant to fit that paradigm.

The extension of this construction to a classical sequent calculus is then made possible by allowing sequents to have multiple formulas in their succedents, again separated by commas, with the informal sense of these succedent commas being the same as for the symbol $\vee$. The calculus LK has the following rules:

Structural Rules

$$\frac{\Gamma \longrightarrow \Theta}{\mathfrak{D}, \Gamma \longrightarrow \Theta}\ \textit{thinning(L)} \qquad \frac{\Gamma \longrightarrow \Theta}{\Gamma \longrightarrow \Theta, \mathfrak{D}}\ \textit{thinning(R)}$$

$$\frac{\mathfrak{D}, \mathfrak{D}, \Gamma \longrightarrow \Theta}{\mathfrak{D}, \Gamma \longrightarrow \Theta}\ \textit{contraction(L)} \qquad \frac{\Gamma \longrightarrow \Theta, \mathfrak{D}, \mathfrak{D}}{\Gamma \longrightarrow \Theta, \mathfrak{D}}\ \textit{contraction(R)}$$

$$\frac{\Delta, \mathfrak{D}, \mathfrak{E}, \Gamma \longrightarrow \Theta}{\Delta, \mathfrak{E}, \mathfrak{D}, \Gamma \longrightarrow \Theta}\ \textit{interchange(L)} \qquad \frac{\Gamma \longrightarrow \Theta, \mathfrak{E}, \mathfrak{D}, \Delta}{\Gamma \longrightarrow \Theta, \mathfrak{D}, \mathfrak{E}, \Delta}\ \textit{interchange(R)}$$

$$\frac{\Gamma \longrightarrow \Theta, \mathfrak{D} \qquad \mathfrak{D}, \Delta \longrightarrow \Lambda}{\Gamma, \Delta \longrightarrow \Theta, \Lambda}\ \textit{cut}$$

Operational Rules

[26] Often, today one thinks of the sequent calculus as a 'meta-calculus' for some other, typically natural-deduction calculus. Under this interpretation, the sequent arrow represents the derivability relation, the antecedent of a sequent represent assumptions, and the succedent of a sequent represents a conclusion (or, if the succedent contains more than one formula, a list of cases). Thus each sequent in a sequent calculus proof represents a complete judgment that a certain object level proof in some other calculus exists (or has been obtained), and one reasons 'at the meta-level' from the existence of some such proofs to the existence of others. Crucially, Gentzen does not advocate any such meta-level/object-level distinction in his thesis.

$$\frac{\Gamma \longrightarrow \Theta, \mathfrak{A} \qquad \Gamma \longrightarrow \Theta, \mathfrak{B}}{\Gamma \longrightarrow \Theta, \mathfrak{A} \,\&\, \mathfrak{B}}\,\&(R) \qquad\qquad \frac{\mathfrak{A}, \Gamma \longrightarrow \Theta \qquad \mathfrak{B}, \Gamma \longrightarrow \Theta}{\mathfrak{A} \vee \mathfrak{B}, \Gamma \longrightarrow \Theta}\,\vee(L)$$

$$\frac{\mathfrak{A}, \Gamma \longrightarrow \Delta}{\mathfrak{A} \,\&\, \mathfrak{B}, \Gamma \longrightarrow \Delta}\;\frac{\mathfrak{B}, \Gamma \longrightarrow \Delta}{\mathfrak{A} \,\&\, \mathfrak{B}, \Gamma \longrightarrow \Delta}\,\&(L) \quad \frac{\Gamma \longrightarrow \Delta, \mathfrak{A}}{\Gamma \longrightarrow \Delta, \mathfrak{A} \vee \mathfrak{B}}\;\frac{\Gamma \longrightarrow \Delta, \mathfrak{B}}{\Gamma \longrightarrow \Delta, \mathfrak{A} \vee \mathfrak{B}}\,\vee(R)$$

$$\frac{\Gamma \longrightarrow \Theta, \mathfrak{F}\mathfrak{a}}{\Gamma \longrightarrow \Theta, \forall\mathfrak{x}\mathfrak{F}\mathfrak{x}}\,\forall(R) \qquad\qquad \frac{\mathfrak{F}\mathfrak{a}, \Gamma \longrightarrow \Theta}{\exists\mathfrak{x}\mathfrak{F}\mathfrak{x}, \Gamma \longrightarrow \Theta}\,\exists(L)$$

$$\frac{\mathfrak{F}\mathfrak{a}, \Gamma \longrightarrow \Theta}{\forall\mathfrak{x}\mathfrak{F}\mathfrak{x}, \Gamma \longrightarrow \Theta}\,\forall(L) \qquad\qquad \frac{\Gamma \longrightarrow \Theta, \mathfrak{F}\mathfrak{a}}{\Gamma \longrightarrow \Theta, \exists\mathfrak{x}\mathfrak{F}\mathfrak{x}}\,\exists(R)$$

$$\frac{\mathfrak{A}, \Gamma \longrightarrow \Theta}{\Gamma \longrightarrow \Theta, \neg\mathfrak{A}}\,\neg(R) \qquad\qquad \frac{\Gamma \longrightarrow \Theta, \mathfrak{A}}{\neg\mathfrak{A}, \Gamma \longrightarrow \Theta}\,\neg(L)$$

$$\frac{\Gamma \longrightarrow \Theta, \mathfrak{A} \qquad \mathfrak{B}, \Gamma \longrightarrow \Theta}{\mathfrak{A} \supset \mathfrak{B}, \Gamma \longrightarrow \Theta}\,\supset(L) \qquad\qquad \frac{\mathfrak{A}, \Gamma \longrightarrow \Theta, \mathfrak{B}}{\Gamma \longrightarrow \Theta, \mathfrak{A} \supset \mathfrak{B}}\,\supset(R)$$
[27]

The first question Gentzen considered after presenting the calculus LK is whether it contains any redundancies. This is a natural question: Because some of the logical connectives appear in duplicate in this calculus, one suspects that some of the rules might be redundant and therefore eliminable in the same way that $\curlywedge I$ was seen to be eliminable from NK. According to Gentzen, these suspicions are justified: 'The schemata are not at all mutually independent, i.e., certain schemata could be eliminated with the help of the remaining ones' (p. 85). But, he claimed, the *Hauptsatz* would not be valid for the resulting calculus. He continued: 'In general, we could *simplify* the calculi in various respects if we attached no importance to the *Hauptsatz*' (Ibid.).

These remarks deserve some explanation.[28] After all, the *Hauptsatz* itself is nothing more than the elimination of the cut rule. Presumably, in its absence other eliminations that were available before would no longer be valid. Why does Gentzen attach more importance to the eliminability of 'cut' than he does to the eliminability of the other rules?

The obvious answer to this question is that the *Hauptsatz* has important applications that distinguish it from ordinary rule-eliminability results. But a deeper explanation is available. According to Gentzen, the cut rule is a complete formalization of the pure

[27]German letters represent formulas of predicate logic; Greek letters represent finite sequences of such formulas. Here again, the rules $\forall(R)$ and $\exists(L)$ are subject to eigenvariable conditions, but the condition is easier to state: In both schemata, the variable $\mathfrak{a}$ must not appear in the lower sequent. The schematic display follows pp. 192–3 of Gentzen's original manuscript, in which various dualities are visually recognizable, and not Szabo's translation, which does not preserve this feature.

[28]On page 688 of *2009* von Plato suggests that Gentzen is referring to the consequences of treating one or more of the logical connectives as an abbreviation of some complex expression, e.g., defining $\vee$ in terms of & and $\neg$ via DeMorgan's equivalence, and observes that the relevant equivalences 'can be put to use' after such measures 'only by using them as cut formulas'.

notion of logical consequence. It is, moreover, by far the most widely emulated pattern of reasoning used in mathematics, so that mathematical thought is predominantly synthetic in nature. But the *Hauptsatz* shows that all mathematical thought that does not rest on the principles of any specific mathematical theory can be simulated with purely analytical reasoning. This is the conceptual significance of the subformula property, the fact that '[i]n an LI- or LK-derivation without cuts, all [formula occurrences in a sequent that occurs in the derivation] are *subformulae* of the [formula-occurrences] that occur in [its] endsequent' (pp. 87-8). Thus the *Hauptsatz* shows that everything provable in the predicate calculus can in fact be proved with a derivation exhibiting this subformula property. Gentzen added (Ibid.) that '[t]he final result [of such a derivation] is, as it were, gradually built up from its constituent elements'—i.e., that the derivation is an analysis, in the Aristotelian sense, of the truth of the derived sequent.

The synthetic nature of 'cut' as opposed to the analytic nature of the operational rules is reflected in Gentzen's formulation of LK. It is well-known today that classical logic can be presented equivalently in context-sharing and context-free sequent calculi. In the context-sharing presentation, rules with multiple upper sequents have the same side formulas Γ, Θ, etc., while the same rules are presented in context-free calculi with different side formulas in each upper sequent. Thus for example in the context-free version of $\supset$*(L)*,

$$\frac{\Gamma \longrightarrow \Theta, \mathfrak{A} \qquad \mathfrak{B}, \Delta \longrightarrow \Lambda}{\mathfrak{A} \supset \mathfrak{B}, \Gamma, \Delta \longrightarrow \Theta, \Lambda} \supset\!\textit{(L)},$$

the contexts of the upper sequents are joined together into larger contexts in the lower sequent. It is obvious that, in the presence of the weak structural rules for thinning, contraction, and interchange, this rule is equivalent to Gentzen's context-sharing version: To emulate an instance of a context-sharing $\supset$*(L)* in the context-free calculus, simply perform a context-free $\supset$*(L)* and then apply interchange and contraction rules to its lower sequent until all redundancies are erased; to emulate an instance of a context-free $\supset$*(L)* in the context-sharing calculus, apply thinning and interchange rules to the upper sequents until their contexts agree and perform a context-sharing $\supset$*(L)* on the resulting sequents. Similar observations confirm that analogous presentations of $\vee$*(L)*, &*(R)*, and 'cut' are equivalent. This equivalence notwithstanding, how one chooses to treat contexts can simplify or complicate the proofs of meta-logical theorems involving proof-transformations. Therefore the preference for context-sharing or context-free presentations is often determined by how one intends to apply the calculus.

In *1932* Gentzen had already presented a context-free version of 'cut', so one might expect to find fully context-free calculi in the *Untersuchungen*. Oddly though, Gentzen did not treat context consistently in his presentation of LK: He gave $\vee$*(L)*, &*(R)*, and $\supset$*(L)* context-sharing presentations alongside a context-free presentation of 'cut'. This lack of uniformity does not evidently simplify his proof of the *Hauptsatz* and demands explanation. If one distinguishes, as Aristotle did, between the two methods of logical discovery $\sigma\acute{\upsilon}\nu\theta\varepsilon\sigma\iota\varsigma$ and $\grave{\alpha}\nu\acute{\alpha}\lambda\upsilon\sigma\iota\varsigma$, then the explanation is forthcoming. In synthesis, one generates new theorems by methodically combining previously established results that may be obtained from disparate sources—i.e., they may have different contexts. But in analysis, context is determined in advance and can only be narrowed: One begins with a claim and successively breaks it down to components in order either to refute the claim through the discovery that it rests on some untenable premise or to uncover the elementary facts that attest to the claim's truth. Thus sequent calculus rules that are supposed to be analytic should be read upwards from the bottom sequent to its analyzing upper sequent(s). Clearly this reading necessitates a context-sharing presentation. But logical synthesis is more naturally understood 'downward from the top': The premises of a 'cut' are typically drawn from distant quarters, and the inference generates new information about how their contexts are related simultaneously with the dispensation of the cut formula. Only a context-free presentation brings out this reading.

Gentzen's conception of 'cut' as consequence thus leads to a new formulation of the original completeness question, a conception of cut-elimination as completeness. Semantic completeness is about the coordination of the syntactic and semantic realms. The former, immanent realm hosts the logical system, and the latter, transcendent realm is home to the notion of logical consequence. In this framework, the question of the completeness of a logical system asks whether all logical consequences can be demonstrated in the logical system. But on Gentzen's view, the question of the completeness of the calculi NK and LK is trivial: NK is engineered to ensure its adequacy with respect to the inferences that mathematicians make, and LK is deductively equivalent to it (section V of *Gentzen 1934–35*). So it is uninteresting to ask whether all valid inferences can be recovered in these systems. Nevertheless, Gentzen conceived of logical consequence as the ubiquitous activity of logical synthesis, represented in LK by the cut rule, and he asked whether the activity of logical analysis suffices to demonstrate the truth of all logical consequences (so conceived) of previously analyzed truths. Thus LK invites a highly non-trivial completeness question, about the adequacy of its purely analytic, cut-free fragment: Is this fragment closed under the synthetic operation of logical consequence (cut)? Far from being a question about the correspondence of realms, this question is about the coordination of methods—analysis and synthesis. Logical consequence is on the synthetic side of this *methodological* divide, but both analytic and synthetic styles of reasoning live side by side in the immanent features of the proof system.

The dispensability of the 'cut' rule follows immediately from two straightforward observations one can make from the point of view of syntax and semantics coordination: First, prove that with respect to the usual quantificational semantics the cut-free fragment of LK is complete. Then verify the full calculus's soundness with respect to that same semantics. The possibility that a formula could be provable in the full calculus but not in its cut-free fragment is thereby ruled out. This 'semantic proof of Gentzen's Hauptsatz' was discovered by Stig Kanger and included in his *1957* dissertation 'Provability in logic'.

In *Buss 1998*, the dispensability of 'cut' is initially established in just this manner, and Buss describes the method as 'slick' (p. 16). He cautions, though, that this method affords no insight into how an arbitrary LK proof can be effectively transformed into an analytic, cut-free one and is for that reason not of interest in proof-theoretical research. Gentzen's proof of cut-elimination in the *Untersuchungen* takes a completely different form and provides maximal insight by specifying an algorithm for the step-wise transformation of a given LK proof of a sequent into a purely analytic derivation of that same sequent. This constructive approach highlights many of the quantitative aspects of cut-elimination that are crucial in ordinal analyses and other applications, and it is fair to say that these aspects of the proof have been more influential than the fact of cut-elimination itself.

Reflecting on this situation, historians have conjectured that Gentzen subscribed to a constructivist ideology, which paid off in the end because the content of his theorems has been rivaled in importance by the structure of his proofs. In his *1976* essay on the origins of set-theoretical semantics, van Heijenoort suggests more modestly that the route to the *Hauptsatz* via semantic completeness and soundness was certainly open to Gentzen, but that he deliberately avoided it because he thought that it 'was a bit like cheating' compared to the transformation algorithm (p. 47). The present view of Gentzen's work sheds more light on the line of thought that led to the constructive proof of the *Hauptsatz* and casts suspicion on these conjectures.

The third theorem of Gentzen's *1932* paper states: 'If a nontrivial sentence $\mathfrak{q}$ is provable from the sentences $\mathfrak{p}_1, \ldots, \mathfrak{p}_\mathfrak{v}$, then there exists a normal proof for $\mathfrak{q}$ from $\mathfrak{p}_1, \ldots, \mathfrak{p}_\mathfrak{v}$'. It is illuminating to contrast Gentzen's comments about this normalization result with his approach in the *Untersuchungen*. He wrote:

> This follows at once from theorems I [informal soundness] and II [informal completeness of normal-form proofs] together. The theorem can also be obtained directly without reference to the notion of consequence by taking an arbitrary proof and transforming it step by step into a normal proof. The reason for the approach chosen in this paper is that it involves little extra effort and yet provides us with important additional results, viz., the correctness and completeness of our forms of inference. (p. 38)

By following this same line of thought in the *Untersuchungen*, Gentzen could have established the correctness and completeness of the forms of inference of the calculus LK simultaneously with the eliminability of its cut rule. It is evident in *Gentzen 1932* that he did not consider nonconstructive proofs of normalization results to be in any way illegitimate and that he even recognized the additional value that they usher in. So it would have been perfectly natural for Gentzen to provide the 'slick' proof of the *Hauptsatz*, if not in place of his constructive one, at least in addition to it, for the sake of the 'important additional results' obtainable along the way.

The relevant observations were *technically* and *methodologically* within Gentzen's reach in 1935: The fact that LK and its cut-free fragment are sound and complete is guaranteed by the proofs in section V of their deductive equivalence with the Hilbert calculus (for which Gödel had proved semantic completeness five years earlier) and with NK (which Gentzen had carefully designed to be complete in his more traditional sense), and the template of inferring normalization from such results appeared already in Gentzen's own earlier work. But on Gentzen's *conception* of logic, the parallel between the sequence ⟨Theorem I, Theorem II, Theorem III⟩ of the *1932* paper and the sequence ⟨LK soundness, cut-free completeness, cut-elimination⟩ breaks down. Theorems I and II of *Gentzen 1932* were Gentzen's verification that the notion of logical consequence is fully analyzed by the formal rule 'cut'. Because the notion of logical consequence appears again in this exact form in the immanent features of the calculus LK, the question of the completeness of that logical system was not for Gentzen a question about how the system corresponds with something beyond itself, but a question about the ability of its analytic fragment to keep pace with its internal consequence relation. If one thinks of cut-elimination as the completeness of the analytic methods with respect to the synthetic notion of logical consequence, then the idea of inferring the *Hauptsatz* from the 'semantic completeness' of those methods does not arise.

5. Concluding remarks

In a recent series of monographs and papers,[29] Jean-Yves Girard has stressed a technical coincidence between semantic completeness and the subformula property.

The semantic completeness of a logical system is the property that all closed formulas of the system that are true under every interpretation are provable. Traditionally, one summarizes Gödel's theorem by saying that first-order quantification theory is complete in this sense. Trivially, the same result can be expressed in the setting of higher-order quantification theory by saying that all closed, true Π^1 formulas are provable. (First-order quantificational expressions are generally not closed in this setting, because their predicate symbols are unbounded—so in higher order logic, where one quantifies over sets, sets of sets, etc., one constructs the closure of first-order expressions simply, by prefixing such expressions with the universal quantifiers needed to bind each predicate symbol.) Now consider an instance of first-order *numerical* quantification (a *numerical* Σ^0_1 formula) $\exists n A(n)$ and note that, by Dedekind's definition of natural number $x \in \mathbb{N} \Leftrightarrow \forall X((X(0) \,\&\, \forall y(X(y) \supset X(Sy))) \supset X(x))$, this formula can be translated into pure quantification theory as the Π^1 formula $\exists x(x \in \mathbb{N} \,\&\, A(x))$. In the same way, *numerical* Π^0_1 formulas can be rewritten as Σ^1 formulas of pure quantification theory.[30] Gödel's *1931* incompleteness theorem exhibits a numerical Π^0_1 formula that is true but unprovable, which implies that the completeness phenomenon does not extend beyond the logical Π^1 class—i.e., there is a true but unprovable formula already in the class

[29] *1999*, *2001*, *2003*, *The Blind Spot*

[30] One obtains $\forall x(x \in \mathbb{N} \supset A(x))$, and the suppressed second-order universal quantifier in $x \in \mathbb{N}$ becomes a second-order existential quantifier after prenex operations.

Σ^1.[31]

'Now', Girard notes, 'if we formulate logic in sequent calculus, we discover that the subformula property holds for the same class Π^1, and fails outside' (*2003*, p. 139). In other words, although the *Hauptsatz* extends to higher-order logic, the cut-free fragment of the logical system fails to be analytic beyond the class Π^1. He claims that 'such a property cannot be an accident' and concludes that 'the subformula property is the actual completeness' (Ibid).

Girard proceeds to explain how it is possible to think of completeness as an internal property of logical systems. Much of the conceptual exposition of his *ludics* program[32] is devoted to making vivid such a construal, and the accompanying announcements are conscientiously iconoclastic: He calls the syntax/semantics distinction 'schizophrenic', refers to research in the orthodox paradigm as 'scholasticism', and suggests that the entire field of foundational research that fits that paradigm can 'be compared to the attraction of political and religious systems which retain their believers long after having failed in practice' (*1987*, p. 36)

This is not the time for me to weigh in on the relative benefits and drawbacks of the various conceptions of logic that have been proposed. But it does seem to dampen the iconoclastic overtones of Girard's work to consider that his 'internal' conception of completeness already flourished in Gentzen's thought. Of course, Gentzen was not driven to the view by reflecting on the coincidental scope of the subformula property and the semantic completeness phenomenon. The construals of 'cut' as consequence and of its eliminability as completeness were native to Gentzen's approach to thinking about logical systems. Another point can be added in favor of the naturalness of Gentzen's view: Because cut-elimination concerns the coordination of various parts of the immanent features of a logical system, its proof is of substantially lower computational complexity than proofs of semantic completeness, which concern the correspondence of entire logical systems with a set-theoretical realm that transcends them. The entire proof of the *Hauptsatz* can in fact be formalized and carried out in PRA.

The details of Herbrand's thought appear to be much more difficult to reconstruct than those of Gentzen's. But it is fitting to conclude with one striking similarity between the two. In section 6 of the fifth chapter of *Recherches sur la théorie de la démonstration*, Herbrand pointed out as a consequence of his fundamental theorem that 'the only rules of reasoning that are needed are the rules of passage, the two rules of generalization and the generalized rule of simplification. The rule of implication [= *modus ponens*] drops out'. Considerable attention has been drawn to the first of two comments he then made: 'Because of the difficulties that the rule of implication might create in certain demonstrations that proceed by recursion on proofs, we consider this result most important'.[33] But Herbrand's second comment addresses the conceptual significance of

[31]Here I have followed Girard's exposition in section 2.A.2 of *The Blind Spot*.

[32]For example, the absurdist appendix of *Locus Solum*.

[33]See, for example, p. 45 of *Mostowski 1965*.

the result more than its importance in applications: 'It shows, moreover, that the rule of implication, whose origin, after all, is in the classical syllogism, is not necessary in building logic' (p. 175). This second point is repeated in an abstract that Herbrand prepared to accompany his thesis:

> [T]he theorem in question permits us to show that the system of rules of reasoning can be changed profoundly while still remaining equivalent to the original ones, so that the rule of the syllogism, the basis of Aristotelian logic, is of no use in any mathematical argument' (*1931a*, p. 276).

Gentzen went beyond Herbrand in pointing out that the rule variously known as 'syllogism' (here and in Hertz's writing), '*modus ponens*', 'implication', 'detachment', and 'cut' not only was the basis of Aristotelian logic, but is a self-sufficient formalization of the intuitive concept of logical consequence. One can only conjecture that the 'profundity' that Herbrand saw in its eliminability derives from the same coordination of analysis and synthesis that motivated Gentzen.

Acknowledgments

The participants in my spring 2009 proof theory seminar at the University of Notre Dame, with their comments and questions, helped me clarify my reading of Gentzen. Graham Leach-Krouse, Neil Tennant, and three anonymous referees for *History and Philosophy of Logic* offered helpful suggestions on earlier drafts of the paper.

References

Belnap, N. 1962. 'Tonk, plonk and plink', *Analysis*, **22**(6), 130-4.

Bernays, P. 1965. 'Betrachtungen zum Sequenzen-kalkul', in *Contributions to logic and methodology in honor of J. M. Bochenski*, ed. by A. T. Tymieniecka, Amsterdam: North Holland.

Buss, S. R. (ed.) 1998. *Handbook of Proof Theory*, New York: North Holland.

Buss, S. R. 1998a. 'Introduction to proof theory', *Buss 1998*, 1-78.

Bynum, T. W. (ed.) 1972. *Conceptual Notation and Related Articles*, New York: Oxford University Press.

Dreben, B. and J. van Heijenoort. 'Introductory note to *Gödel 1929*, *Gödel 1930*, and *Gödel 1930a*', *Feferman et al. 1986*, 44-59.

Etchemendy, J. 1988. 'Tarski on Truth and Logical Consequence', *The Journal of Symbolic Logic* **53**(1), 51-79.

Ewald, W. and W. Sieg (eds.) 2010. *David Hilbert's Lectures on the Foundations of Arithmetic and Logic, 1917-1933*, vol. 3. Springer.

Feferman, S., J. W. Dawson, Jr., S.C. Kleene, G. H. Moore, R. M. Solovay, and J. van Heijenoort (eds.) 1986. *Kurt Gödel, Collected Works, Vol I: Publications 1929-1936*, New York: Oxford University Press.

Fenstad, J. E. (ed.) 1970. *Thoralf Skolem: Selected Works in Logic*, Oslo: Universitetsforlaget.

Franks, C. 2009. *The Autonomy of Mathematical Knowledge: Hilbert's Program Revisited*, Cambridge: Cambridge University Press.

Frege, G. 1879. *Begriffsschrift, eine der arithmetischen nachgebildete Formelsprache des reinen Denkens* Halle: L. Nebert. Translated by T. W. Bynum as *Conceptual Notation: a formula language of pure thought modeled upon the formula language of arithmetic* in *Bynum 1972*, 101-208.

Frege, G. 1882. 'Über die wissenschaftliche Berechtigung einer Begriffsschrift', *ZPPK* **81**, 48-56. Translated by T. W. Bynum as 'On the scientific justification of a conceptual notation' in *Bynum 1972*, 83-9.

Frege, G. 1883. 'Über den Zweck der Begriffsschrift', *Sitzungsberichte der Jenaischen Gesellschaft für Medicin und Naturwissenschaft, JZN* **16**, 1-10. Translated by T. W. Bynum as 'On the aim of the "Conceptual Notation"' in *Bynum 1972*, 90-100.

Frege, G. 1884. *Die Grundlagen der Arithmetik*. Translated by J. L. Austin as *The Foundations of Arithmetic*. Evanston: Northwestern University Press. Cited according to sectioning.

Gentzen, G. 1932. 'Über die Existenz unabhängiger Axiomensysteme zu unendlichen Satzsystemen', *Mathematische Annalen* **107**, 329-50. Translated as 'On the existence of independent axiomsystems for infinite sentence systems' in *Szabo 1969*, 29-52.

Gentzen, G. 1934–35. 'Untersuchungen über das logische Schliessen'. Gentzen's doctoral thesis at the University of Göttingen, translated as 'Investigations into logical deduction' in *Szabo 1969*, 68-131.

Gentzen, G. 1936. 'Die Widerspruchsfreiheit der reinen Zahlentheorie', *Mathematische Annalen* **112**, 493-565. Translated as 'The consistency of elementary number theory' in *Szabo 1969*, 132-213.

Gentzen, G. 1938. 'Die gegenwärtige Lage in der mathematischen Grundlagenforschung', *Forschungen zur Logik und zur Grundlegung der exakten Wissenschaften*, New Series **4**, 5-18. Translated as 'The present state of research into the foundations of mathematics' in *Szabo 1969*, 234-51.

Gentzen, G. 1938a. 'Neue Fassung des Widerspruchsfreiheitsbeweises für die reine Zahlentheorie', *Forschungen zur Logik und zur Grundlegung der exakten Wissenschaften*, New Series **4**, 19-44. Translated as 'New version of the consistency proof for elementary number theory', in *Szabo 1969*, 252-86.

Girard, J. Y. 1987. *Proof Theory and Logical Complexity.* Vol. 1, Napoli: Bibliopolis.

Girard, J. Y. 1999. 'On the meaning of logical rules I: syntax vs. semantics', in Berger and Schwichtenberg, eds., *Computational Logic*, Heidelberg: Springer-Verlag, 1999, 215-72.

Girard, J. Y. 2001. *Locus Solum: from the rules of logic to the logic of rules*, *Mathematical Structures in Computer Science* **11**(3), 301-506.

Girard, J. Y. 2003. 'From foundations to ludics', *The Bulletin of Symbolic Logic*, **9**(2), 131-68.

Girard, J. Y. *The Blind Spot*, English translation of lectures in Rome, 2004.

Gödel, K. 1929. 'Über die Vollständigkeit des Logikkalküls', Gödel's doctoral thesis at the University of Vienna, translation by S. Bauer-Mengelberg and Jean van Heijenoort as 'On the completeness of the calculus of logic' reprinted in *Feferman et al. 1986*, 60-101.

Gödel, K. 1930. 'Die Vollständigkeit der Axiome des logischen Functionenkalküls', *Monatshefte für Mathematik und Physik* **37**, 349-60, translation by S. Bauer-Mengelberg as 'The completeness of the axioms of the functional calculus of logic' reprinted in *Feferman et al. 1986*, 102-23.

Gödel, K. 1930a. 'Über die Vollständigkeit des Logikkalküls', *Die Naturwissenschaften* **18**, 1068. Translated by J. W. Dawson Jr. as 'On the completeness of the calculus of logic' in *Feferman et al. 1986*, 124-5.

Gödel, K. 1931. 'Über formal unentscheidbare Sätze der *Principia Mathematica* und verwandter Systeme I', *Monatshefte für Mathematik und Physik* **38**, 173-98, translation by J. van Heijenoort as 'On formally undecidable propositions of *Principia Mathematica* and related systems I' reprinted in *Feferman et al. 1986*, 144-95.

Goldfarb, W. (ed.) 1971. *Jacques Herbrand: Logical Writings.* Cambridge: Harvard University Press.

Herbrand, J. 1930. *Recherches sur la théorie de la démonstration.* Herbrand's doctoral thesis at the University of Paris. Translated by W. Goldfarb, except pgs. 133-88 trans. by B. Dreben and J. van Heijenoort, as 'Investigations in proof theory' in *Goldfarb 1971*, 44-202.

Herbrand, J. 1931a. 'Unsigned note on *Herbrand 1930*', *Annales de l'Université de Paris* **6**, 186-9. Translated by W. Goldfarb in *Goldfarb 1971*, 272-6.

Hertz, P. 1929. 'Über Axiomensysteme für beliebige Satzsysteme', *Mathematische Annalen*, **101**, 457-514.

Hilbert D. 1929. 'Probleme der Grundlegung der Mathematik', *Mathematische Annalen* **102**, 1-9.

Hilbert, D. and W. Ackermann. 1928. *Grundzüge der theoretischen Logik*, Berlin: Springer.

Johansson, I. 1937. 'Der Minimalkalkül, ein reduzierter intuitionistic Formalismus'. *Compositio Mathematica* **4**, 119-136.

Kanger, S. 1957. 'Provability in logic', Almqvist and Wiksell: Stockholm. Reprinted in *Collected Papers of Stig Kanger with Essays on his Life and Work*, vol 1, G. Holmström-Hintikka, S. Lindtröm, and R. Sliwinski, eds., Dordrecht: Kluwer, 2001, 8-41.

Mostowski, A. 1965. 'Theorems of Herbrand and of Gentzen', in Mostowski *Thirty Years of Foundational Studies*, Helsinki: Acta Philosophica Fennica, 1965, 43-50.

Schroeder-Heister, P. 2002. 'Resolution and the origins of structural reasoning: Early proof-theoretic ideas of Hertz and Gentzen', *The Bulletin of Symbolic Logic* **8**, 246-265.

Skolem, T. 1923. 'Einige Bemerkungen zur axiomatischen Begründung der Mengenlehre' *Matematikerkongressen i Helsingfors 4-7 Juli 1922, Den femte skandinaviska matematikerkongressen, Redogörelse.* Helsinki: Akademiska Bokhandeln, 217-301. Translated by S. Bauer-Mengelberg as 'Some remarks on axiomatized set theory' in *van Heijenoort 1967*, 290-301.

Szabo, M. E. 1969. *The Collected Papers of Gerhard Gentzen*, London: North Holland.

Tarski, A. 1936. 'O pojciu wynikania logicznego', *Przeglad Filozoficzny* **39**, 58-68. Translated by J. H. Woodger as 'On the concept of logical consequence' in J. H. Woodger, ed., *Logic, Semantics, Metamathematics: Papers from 1923 to 1938 by Alfred Tarski*, Oxford: Oxford University Press. 1956.

Van Heijenoort, J. (ed.). 1967. *From Frege to Gödel: A sourcebook in mathematical logic 1879-1931*, Cambridge: Harvard University Press.

Van Heijenoort, J. 1976. 'Set-theoretic semantics', reprinted in *van Heijenoort 1985a*, 43-53.

Van Heijenoort, J. 1985. 'Jacques Herbrand's work in logic in its historical context', in *van Heijenoort 1985*, 99-121.

Van Heijenoort, J. 1985a. *Selected Essays*, Naples: Bibliopolis.

Von Plato, J. 2008. 'Gentzen's proof of normalization for natural deduction', *The Bulletin of Symbolic Logic*, **14**(2), 240-57.

Von Plato, J. 2009. 'Gentzen's logic', in *Handbook of the History of Logic*, vol. 5, ed. by D. M. Gabbay and J. Woods, Amsterdam: North Holland.

Wang, H. 1974. *From Mathematics to Philosophy*, London: Routledge & Kegan Paul.