DAG PRAWITZ

The Fundamental Problem of General Proof Theory

Abstract. I see the question what it is that makes an inference valid and thereby gives a proof its epistemic power as the most fundamental problem of general proof theory. It has been surprisingly neglected in logic and philosophy of mathematics with two exceptions: Gentzen's remarks about what justifies the rules of his system of natural deduction and proposals in the intuitionistic tradition about what a proof is. They are reviewed in the paper and I discuss to what extent they succeed in answering what a proof is. Gentzen's ideas are shown to give rise to a new notion of valid argument. At the end of the paper I summarize and briefly discuss an approach to the problem that I have proposed earlier.

Keywords: Proof theory, Proof, Valid inference, Valid argument, Gentzen's natural deduction, Intuitionism.

1. Introduction

One may expect proof theory to have something to say about what a proof is. When Hilbert coined the term *Beweistheorie*, later on translated as *proof theory*, this expectation was grandly formulated as a request: “we must make the concept of the specific mathematical proof itself the object of investigation just as also the astronomer pays attention to his place of observation, the physicist must care about the theory of his instrument, and the philosopher criticizes reason itself”.¹ Already a few years before, he had called attention to “how necessary it is to study the nature of the mathematical proof in itself”, which was to be the task of “an important, new field of research”.²

However, as this new field developed, it was not associated with any such conceptual analysis. For long time proof theory came in fact to be identified

This paper is quite different from the one presented at the conference on General Proof Theory (to be found, somewhat modified, in the proceedings published on line; [23]), but there are some overlaps; in particular, both contain a presentation of a new notion of valid inference.

¹ Hilbert [6].

² Hilbert [5], where the passage quoted above also occurs.

Special Issue: General Proof Theory
Edited by **Thomas Piecha** and **Peter Schroeder-Heister**

with the more limited *Hilbert program* to prove the consistency of mathematics. In reaction to that, the term *general proof theory* was proposed as the name of a field where proofs were studied in their own right, and was contrasted to *reductive proof theory* where the study of proofs was a tool used for reductive aims; the latter was proposed as a more general characterization of Hilbert's program and its modification after Gödel's incompleteness results.³

My proposal was inspired by Georg Kreisel, who had criticised the main trend of proof theory of that time for its limited program and had suggested different and broader aims, pointing out that with some care many results in the field could actually be stated in such a way that it got a wider bearing.⁴ To me, Gentzen's Hauptsatz, in particular when formulated as a normalization theorem for natural deduction, had seemed to constitute a prime example of a contribution to a proof theory with a genuinely broader aim than Hilbert's program; this had been a main point of my monograph "Natural Deduction".⁵ At least for a period, Kreisel came to share and develop this view of mine. In a review of the English translation of Gentzen's collected work, he saw Gentzen's dissertation as the "*germs for a theory of proofs*" where "*proofs, expressed by formal derivations, are principal objects of study*".⁶ In the paper "A survey of proof theory II", he contrasted such a proof theory with the focus of his earlier survey paper [11], which had been on "*proof theory as a tool for studying logical consequence*".⁷

In my plea for general proof theory, I suggested a number of obvious topics: the question of defining the concept of proof, investigations of the structure of proofs, the representation of proofs by formal derivations, and the finding of identity criteria of proofs that answered the question when two derivations represent the same proof.⁸ Shortly afterwards, I presented a paper with the ambitious title "Towards a foundation of a general proof theory".⁹ It approached the concept of proof by considering arguments composed of arbitrary inferences and proposed a definition of what it is for such

³Prawitz [18].

⁴His paper "A survey of proof theory" [11] is one among his many publications on this theme.

⁵Prawitz [17]. The point was further developed in [18].

⁶Kreisel ([12], p. 242), where he also says that as he sees it now ("guided by D. Prawitz's reading of Gentzen"), "the single most striking element of Gentzen's work occurs already in his doctoral dissertation".

⁷Kreisel ([13], p. 109).

⁸Prawitz ([18], p. 237).

⁹Prawitz [19].

an argument to be valid. The tentative character of the approach was emphasized, and today I think, for reasons soon to be explained, that that approach as developed there is not viable.

But I still think that the problem of defining the concept of proof or the validity of inference is the most fundamental problem of general proof theory, a problem that the foundation of that discipline should try to tackle. It has been surprisingly neglected in logic and philosophy of mathematics. Proofs have certainly since Greek antiquity been seen as establishing mathematical theorems, required when making mathematical assertions, and as being built up by inferences. But inferences can be valid or invalid, and the question what it is that makes an inference valid and thereby gives a proof its epistemic power is seldom considered.

The classical notion of valid inference, defined in terms of truth preservation, is obviously not the relevant notion here. The *fact* that the conclusion of an inference follows from the premisses does not in itself make the use of the inference in a proof legitimate. For the use to be legitimate the inference must yield a ground for the conclusion and thereby make the conclusion warranted or known to hold. The crucial question is what it is that makes an inference have this feature.

There are two notable exceptions to what I said about the neglect of this fundamental question about proofs and valid inferences: Gentzen's remarks about what justifies the rules of his system of natural deduction, including works inspired by them, and proposals in the intuitionistic tradition about what a proof is. In the rest of this paper, I shall review them and discuss to what extent they succeed in answering the question that I am raising here. In the part that deals with ideas of Gentzen, I shall present a new notion of valid argument, and at the end I shall summarize and briefly discuss an approach that I have proposed earlier. Although I cannot offer a solution of the problem raised, I hope that the paper will succeed in drawing attention to an important problem.

2. Two Ideas of Gentzen About the Justification of Inferences

In his doctoral dissertation,¹⁰ Gentzen presented two ideas about how the rules of his system of natural deduction were justified. They were accentuated and somewhat developed in my own doctoral dissertation,¹¹ and are

¹⁰Gentzen [4].

¹¹Prawitz [17].

quoted quite often nowadays. Gentzen suggested, firstly, that the introduction rule for a logical constant was justified by being seen as determining the meaning to be attached to the constant and, secondly, that the corresponding elimination rule was justified by being in accord with this meaning. The second idea was exemplified by observing in effect that an application of the rule of implication elimination did not yield anything above what could be obtained from proofs of the premisses without applying this rule if the major premiss had been inferred by an introduction. The point is illustrated by the two deductions below:

$$\begin{array}{c}
 [A] \\
 \Pi \\
 B \\
 \hline
 A \supset B \quad \Delta \\
 A \\
 \hline
 B
 \end{array}
 \qquad
 \begin{array}{c}
 \Delta \\
 [A] \\
 \Pi \\
 B
 \end{array}$$

The right deduction is here obtained from the left one by substituting the deduction Δ of the minor premiss A for the free assumptions of A in the deduction Π of B from A that become bound (discharged) in the left deduction when $A \supset B$ is inferred by introduction. This operation constitutes an example of the reductions that I introduced in order to show how the natural deductions can be transformed to a certain normal form.

In reference to Gentzen's first idea, one may ask: How does the introduction rule for a logical constant determine the meaning of the constant? It specifies one way, but of course not the only way, in which a sentence with the logical constant as its outer sign can be rightly inferred. So what is special with this particular way of inferring the sentence? The idea can be indicated by saying that the introduction rules specify the *canonical form* of proofs in the same way as we can say that the numerals constitute the canonical representation of the natural numbers—we present exhaustively the natural numbers by saying that they can be written in the form of numerals. Similarly the meanings of the logical constants are given by saying that when a sentence is provable, its proof can in principle be put in the form of a proof whose last inference is an introduction, which is what is called a proof in canonical form.

This way of describing how the introduction rules determine the meanings of the logical constants also gives a criterion for what it is for the elimination rules to be in accord with these meanings: a proof ending with an elimination rule should be possible to transform to canonical form. In my dissertation, I first characterized the relation between an introduction rule and the corresponding elimination rule by saying that the latter is in

a sense the *inverse* of the former: if the major premiss of an elimination is inferred by introduction, the deductions of the premisses of the elimination taken together already “contain” a deduction of the conclusion without the use of this elimination. In this *inversion principle*, as I called it, borrowing a term from Lorenzen, the term “contain” was used metaphorically. The idea was spelled out more precisely by the reductions that I defined for the elimination rules, as exemplified above. By them a deduction of the conclusion of an application of an elimination was obtained from the deductions of the premisses without the use of the elimination provided that the major premiss had been inferred by introduction; the use of such reductions was shown to terminate in a normal deduction, which in the case of proofs, that is, deductions not depending on assumptions, would be in canonical form.

3. The Idea of Valid Argument

This was how I understood and developed the remarks that Gentzen made after having defined his system of natural deduction. For this to be a general way of justifying deductive reasoning, it should be possible to generalize these ideas so as to apply not only to natural deductions. How to do this was the question that I wanted to answer in the paper mentioned above [19].

To this end, I considered arbitrary chains of inferences put in the general form of natural deduction, that is, in tree-form where inferences could bind assumptions and variables (that is, discharge assumptions and have so-called proper variable restrictions, respectively). They were called *arguments*. An argument was said to be *closed* when all assumptions and variables occurring in the argument were bound and *open* otherwise.

The generalization took the form of a definition of what it was for such an argument to be valid. It was assumed that there was a given set of introduction rules, not necessarily those of Gentzen. In the way described above, they were viewed as giving meanings to the logical constants by stipulating what was to be counted as canonical arguments for compound sentences. The meanings of the atomic sentences were supposed to be given by a base $\mathcal{B}$ that singled out the predicates and the individual terms of the language that the validity notion was defined for and stipulated what was to count as canonical arguments for the atomic sentences.

The validity of an argument was made relative to such a base $\mathcal{B}$ and a set $\mathcal{R}$ that assigned to the inference rules applied in the argument, except the introduction rules, reductions of the same general kind as the ones assigned to Gentzen’s elimination rules, that is, transformations of arguments to other

arguments for the same sentences depending on no additional assumptions. It could then be defined what it was for an argument to reduce to another argument relative to such a set $\mathcal{R}$ of reductions in the same way as one defined what it was for a natural deduction to reduce to another one.

The notion of validity of an argument was based on three principles concerning the validity of different kinds of arguments. Two principles were more or less already formulated for natural deductions in the previous section when it was said that the meanings of the logical constants, and hence the meanings of compound sentences, were given by stipulating the introduction rules as (a) *one way* of proving compound sentences, and furthermore as (b) the *canonical way* of proving them—(a) motivates principle (1) and (b) motivates principle (2):

- (1) A closed argument in canonical form for a compound sentence is valid relative to $\mathcal{B}$ and $\mathcal{R}$, if and only if, its immediate sub-arguments are.
- (2) A closed argument in non-canonical form is valid relative to $\mathcal{B}$ and $\mathcal{R}$, if and only if, it reduces relative to $\mathcal{R}$ to a closed argument in canonical form that is valid relative to $\mathcal{B}$ and $\mathcal{R}$.

The third principle was implicit in Gentzen's example of how implication elimination was justified when it was taken for granted that a deduction from an assumption A remains valid when the assumption is replaced by a deduction of A . If we similarly regard an argument with free variables not bound by any inference as an argument schema that remains valid when the free variables are replaced by closed terms, we get the principle:

- (3) An open argument $\mathcal{A}$ is valid relative to $\mathcal{B}$ and $\mathcal{R}$, if and only if, all substitution instances of $\mathcal{A}$ are valid relative to $\mathcal{B}$ and any (consistent) extension $\mathcal{R}^*$ of $\mathcal{R}$; a substitution instance being obtained by substituting first closed terms for the free variables in $\mathcal{A}$ and then closed arguments valid relative to $\mathcal{B}$ and $\mathcal{R}^*$ for the free assumptions in $\mathcal{A}$.

Extensions of the given set $\mathcal{R}$ of reductions are considered in (3) because we need to assign reductions to inference rules that are applied in the arguments substituted for the free assumptions when not already given by $\mathcal{R}$.

Provided that the inference rules taken as introductions satisfy, as Gentzen's introduction rules do, the condition that the premisses and the assumptions bound by the inference are of less complexity than the conclusion, these three principles together with the stipulation that the closed canonical arguments for atomic sentence given by the base $\mathcal{B}$ are valid outright can

be taken as an inductive definition of what it is for an argument to be valid relative to a base $\mathcal{B}$ and a set $\mathcal{R}$ of reductions.

This was in effect the notion of valid argument defined in my paper. The notion was taken up, slightly modified, discussed in detail, and essentially agreed to by Michael Dummett in his book *The Logical Basis of Metaphysics*.¹² It was also the object of another substantial discussion by Peter Schroeder-Heister, who modified the notion in some essential respects.¹³

As I see it now, the notion of valid argument in its different versions has two main shortcomings. Firstly, the validity of an argument relative to a base $\mathcal{B}$ and a set of reductions $\mathcal{R}$ may depend essentially on the reductions in $\mathcal{R}$ and not at all on the inferences that make up the argument.¹⁴ The justification of Gentzen's elimination rules discussed in the preceding section makes use of reductions of a particularly restrictive kind, and the deductions to which they are applied are crucial. When applying such a reduction to a deduction $\mathcal{D}$, the result is actually a deduction of the conclusion of $\mathcal{D}$ of the kind guaranteed by the inversion principle (Section 2), that is, a deduction "contained" in the deductions of the premisses of the last inference of $\mathcal{D}$. By generalizing this feature, after having made the notion of containment precise, it is possible to get a quite different notion of validity that is not open to the objection now considered. I shall turn to this in the next section.

There is a second reason why the notion of closed valid argument does not offer a plausible analysis of the intuitive concept of deductive proof. A deductive proof of a sentence A establishes conclusively the truth of A ; it gives its possessor a conclusive ground for asserting A , thereby making the assertion of A warranted. A proof does so as it stands, without any further additions; otherwise it is a "proof with gaps" that does not become a real proof until the gaps are filled in. But to be in possession of a closed valid argument for a sentence A does not in general make the assertion of A warranted, if it is not known that the argument is valid, and to establish this validity may require a proof of great complexity. Admittedly to

¹²Dummett [3]. For a discussion of how Dummett's notion of valid argument is related to mine, see Prawitz [20].

¹³Schroeder-Heister [24] in a volume that grew out of a conference on *Proof-Theoretic Semantics* held in Tübingen in 1999. For a discussion of the relations between different versions of the notion of valid argument and the notion of BHK-proof, see Prawitz [22].

¹⁴Dummett does not relativize validity to a set of given reductions but allows in what corresponds to the principle (2) above any kind of effective transformation. The validity may then depend on the possibility of such a transformation and not at all on the inferences of the argument.

have an argument $\mathcal{A}$ and a set $\mathcal{R}$ of reductions such that $\mathcal{A}$ is valid relative to $\mathcal{B}$ and $\mathcal{R}$ means that one is in possession of a method by which one is able to transform $\mathcal{A}$ to canonical form, and according to how the meanings of sentences are now explained, canonical arguments constitute indeed the canonical way of proving assertions. However, it remains that until one has applied this method, one does not know by just being in possession of the method that it will finally yield a canonical argument as a result.

Furthermore, it should be noted that the definition of the notion of valid argument is by recursion over the complexity of sentences, while one expects the intuitive notion of proof to be explained inductively, presupposing a notion of valid inference. The arguments are certainly built up inductively by applying inferences, but the notion of valid argument does not presuppose a notion of valid inference. Such a notion can instead be given in terms of valid argument, saying that an inference is valid if it preserves the validity of arguments, but this seems to turn the natural order of explanation upside down. I shall return to these questions in the last part of the paper.

4. A New Notion of Valid Argument: Analytical Validity

The containment spoken of in the inversion principle for Gentzen's elimination rules is of an implicit kind because, except for the case of conjunction, a deduction of the conclusion of an elimination whose major premiss is inferred by introduction is not literally a part of the deductions of the premisses.

But a deduction of the conclusion is literally a part of what can be obtained from the deductions of the premisses by compositions of them and substitutions of terms for free variables. By noting this, the notion of containment is easily made precise.

The notion can be defined not only for natural deductions but also for arguments in general. This is conveniently done in two steps. Let us first say that the argument $\mathcal{A}$ is *immediately extracted* from the set Σ of arguments if and only if either

- (a) $\mathcal{A}$ is an argument in Σ or a sub-argument of some argument in Σ , or
- (b) $\mathcal{A}$ is the result of substituting a term for the occurrences of a free variable in an argument in Σ , or
- (c) $\mathcal{A}$ is the result of composing two arguments $\mathcal{B}$ and $\mathcal{C}$ in Σ , that is, $\mathcal{A}$ is the result of replacing some free assumptions B in $\mathcal{C}$ by $\mathcal{B}$, which can be written:

$$\mathcal{A} = \begin{array}{c} \mathcal{B} \\ [B] \\ \mathcal{C} \end{array}$$

We can then define an argument $\mathcal{A}$ to be *contained* in a set Σ of arguments, if there is a sequence of arguments $\mathcal{A}_1, \mathcal{A}_2, \dots, \mathcal{A}_n$ where $\mathcal{A}_n = \mathcal{A}$ and for each $i \leq n$, $\mathcal{A}_i$ is immediately extracted from $\Sigma \cup \{\mathcal{A}_1, \mathcal{A}_2, \dots, \mathcal{A}_{i-1}\}$. I shall say that the argument $\mathcal{A}$ is contained in the argument $\mathcal{B}$ or that $\mathcal{B}$ contains $\mathcal{A}$, if $\mathcal{A}$ is contained in $\{\mathcal{B}\}$.

The reductions that occur in the definition of valid argument in the previous section can now be restricted to those where the result of the reduction is, in the now defined sense, contained in the argument that it is applied to. Still better, we can dispense altogether with reductions in the definition of validity, and require in clause (2) of the definition of valid argument that a closed, non-canonical argument of A contain a valid, closed argument for A in canonical form. The other clauses can be left as they are. To distinguish this new validity notion¹⁵ from the previously defined ones, we may call it *analytical validity*.¹⁶ The recursion clauses in the definition of this notion then become:

1. A closed argument in canonical form for a compound sentence is *analytically valid*, if and only if, its immediate sub-arguments are.
2. A closed argument in non-canonical form for a sentence A is *analytically valid*, if and only if, it contains an *analytically valid*, closed, and canonical argument for A .
3. An open argument is *analytically valid*, if and only if, all substitution instances of $\mathcal{A}$ are analytically valid; a substitution instance being obtained by substituting first closed terms for the free variables in $\mathcal{A}$ and then closed arguments analytically valid for the free assumptions in $\mathcal{A}$.

Here the relativization to a base $\mathcal{B}$ is left implicit. An argument that is analytically valid relative to all bases $\mathcal{B}$ may be called logically valid.

A condition clearly equivalent to the one given in clause 2 is that the set of immediate sub-arguments contains an analytically valid, closed, and canonical argument for A . A slightly weaker condition that I shall return

¹⁵The notion has grown out in conversations with Peter Schroeder-Heister from remarks made in [22].

¹⁶The idea here is of course quite different from the one involved in Kant's notion of analytical truth—his notion of containment is one between predicates, while the present one is between arguments.

to below is that it contains the immediate sub-arguments of an analytically valid, closed, and canonical argument for A .

The definition of this new notion of analytical validity is still using recursion over the complexity of sentences and is not presupposing a notion of validity for inferences. Such a notion, *analytical validity of inferences*, may be defined again as preservation of the property in question. In other words, an inference $\mathcal{I}$ (specified by its premisses, conclusion, variables bound by the inference, and assumptions allowed to be bound by the inference) is analytically valid (relative to a base $\mathcal{B}$), if and only if, all arguments whose last inference is an instance of $\mathcal{I}$ are analytically valid (relative to $\mathcal{B}$), when the immediate sub-arguments are analytically valid (relative to $\mathcal{B}$).

All applications of the intuitionistic inference rules of natural deduction are easily seen to be analytically valid (relative to any base $\mathcal{B}$). For introductions this follows immediately from clauses 1 and 3. It may be instructive to verify an example of the validity of an elimination. Let us consider implication elimination (modus ponens), and assume that $\mathcal{A}$ is a valid argument for A and $\mathcal{B}$ a valid argument for $A \rightarrow B$; I drop the prefix “analytical” before valid in contexts where I only speak of this new notion of validity. We have to show that the argument

$$\mathcal{C} = \frac{\frac{\mathcal{A}}{A} \quad \frac{\mathcal{B}}{A \rightarrow B}}{B}$$

is valid. In case $\mathcal{C}$ is an open argument, this amounts to showing that any closed substitution instance $\mathcal{C}^*$ of $\mathcal{C}$ (in the sense of clause 3) is valid. $\mathcal{C}^*$ may be written

$$\frac{\frac{\mathcal{A}^*}{A^*} \quad \frac{\mathcal{B}^*}{A^* \rightarrow B^*}}{B^*}$$

where the asterisk stands for the substitution in question (if $\mathcal{C}$ is already closed, drop the asterisk).

Since $\mathcal{A}$ and $\mathcal{B}$ are valid, so are the closed arguments $\mathcal{A}^*$ and $\mathcal{B}^*$ (by clause 3). Hence, $\mathcal{B}^*$ either is or contains a closed, canonical, and valid argument $\mathcal{B}^{*\circ}$ for $A^* \rightarrow B^*$ (by clause 2). This means that $\mathcal{B}^*$ or $\mathcal{B}^{*\circ}$ must have as immediate sub-argument a valid argument $\mathcal{B}_1$ for B^* from A^* (by clause 1). Let $\mathcal{B}_2$ be

$$\frac{\mathcal{A}^*}{[\mathcal{A}^*]} \mathcal{B}_1$$

that is, the closed argument for B^* obtained by substituting $\mathcal{A}^*$ for all free assumptions $\mathcal{A}^*$ in $\mathcal{B}_1$. By clause 3, $\mathcal{B}_2$ is valid, and by clause c (in the definition of immediate extraction), $\mathcal{B}_2$ is contained in $\{\mathcal{A}^*, \mathcal{B}_1\}$ and hence also in $\mathcal{C}^*$ (by the transitivity of containment). We have shown that $\mathcal{C}^*$ contains a closed, valid argument for B^* , and that it is therefore valid according to the following useful lemma.

LEMMA. *If a closed, non-canonical argument $\mathcal{A}$ for a sentence A contains a closed analytically valid argument $\mathcal{A}^\circ$ for A , then $\mathcal{A}$ contains a canonical argument for A that is closed and analytically valid, and therefore $\mathcal{A}$ itself is analytically valid.*

The correctness of the lemma is seen by noting that $\mathcal{A}^\circ$ must either be canonical or contain in its turn a valid, closed, canonical argument for A (by clause 2), which is also contained in $\mathcal{A}$ (because of the transitivity of the containment relation). Therefore, $\mathcal{A}$ is valid by clause 2.

The analytical validity of Gentzen's elimination rules reflects much better than their validity in the previously defined sense Gentzen's idea about how the rules are justified, in particular his idea that we are using the major premiss only "in the sense afforded it by the [corresponding] introduction". For validity in the previous sense, what mattered was that the non-canonical argument obtained by applying an elimination rule could be transformed to a valid argument in canonical form by applying a reduction; how the reduction operation looked was left essentially open, which meant that the meaning of the major premiss of the last inference of the argument did not necessarily matter. When we now for analytical validity demand that the non-canonical closed argument contain a closed, canonical, and analytically valid argument for the sentence in question, we get a condition whose satisfaction depends on what the major premiss means, that is, on the nature of a canonical argument for it, as is illustrated in the verification above of the analytical validity of implication elimination.

It is also easy to see that applications of the rule of mathematical induction in the form

$$\frac{\begin{array}{c} [A(x)] \\ A(0) \quad A(sx) \end{array}}{A(a)}$$

which binds the variable x and assumptions $A(x)$, are analytically valid relative to an arithmetical base whose numerical terms are the numerals (formed from 0 by applying the successor operation s) and which takes as canonical all arguments for atomic sentences that can be obtained by applying

the usual arithmetical inference rules for atomic sentences. Considering a closed instance of an application of this rule where the conclusion is $A(n)$, and assuming that we have valid arguments $\mathcal{A}_1$ and $\mathcal{A}_2(x)$ for the induction base and the induction step, respectively, we can compose $\mathcal{A}_1$ and a suitable number of instances of $\mathcal{A}_2(x)$ until we get an argument for the conclusion $A(n)$. The obtained argument is valid and is contained in the original one, which hence is valid by the lemma above.

Although all the intuitionistic inference rules of natural deduction are analytically valid, not all intuitionistically derivable inferences are. For the previously defined notion of validity, it is easy to see that an inference is valid if its conclusion is derivable from its premisses by valid inference rules. The same holds of course for the traditional definition of valid inference, which identifies validity with logical consequence, since that relation is transitive. Analytical validity is quite different and is a much narrower notion. It can be illustrated with inferences of the form

$$\frac{A \rightarrow (B \rightarrow C)}{B \rightarrow (A \rightarrow C)}$$

which can be seen to be valid relative to a suitable set of reductions.¹⁷ The conclusion is of course derivable from the premiss by analytically valid inferences, but the inference itself is not analytically valid for arbitrary A , B , and C , which can be seen by letting A be the same as $B \rightarrow C$. The one-step argument that assumes $B \rightarrow C$ and then applies $\rightarrow$ I is a closed valid argument for the premiss, but it does not contain an argument for the conclusion. This one-step argument does not either contain an argument for $(B \rightarrow C) \rightarrow C$ from B . Thus, when clause 2 is weakened as described above, the inference remains invalid.

Not even the trivial inference

$$\frac{A \vee B}{B \vee A}$$

is analytically valid; an analytically valid, closed argument for $A \vee B$ need not contain a canonical argument for $B \vee A$. However, it does contain an analytically valid, closed argument for A or for B , which is sufficient to make the inference valid if we use the weaker variant of clause 2 in the definition of validity.

Analytically valid inference is a very narrow concept in a seemingly interesting way, if one is interested in what it is for an inference to be legitimate in a proof. A validity notion of inference which makes any inference valid

¹⁷See Schroeder-Heister [24], who considers this example in detail.

whose conclusion is derivable from its premisses by valid inferences is not a plausible candidate for what it is to be an acceptable inference in a proof. It would make it legitimate to infer any theorem directly from axioms, thus making any one-step argument for a theorem from axioms a proof, regardless of how complicated a real proof would be. In proofs actually occurring in mathematical practice, we may make inference steps as big as competent persons are able to follow, which of course varies with the context. An objective definition of what it is for an inference to be legitimate in a proof can of course not reflect this dependency on the context, but must make legitimate inference steps reasonably small. Analytical validity of inference seems to be a concept that fits this requirement, but this remains to be further investigated. It can be noted that besides Gentzen's introductions and eliminations, inferences that put together a number of eliminations into one inference, like the inference

$$\frac{A \quad B \quad A \rightarrow (B \rightarrow C)}{C}$$

are analytically valid.

5. Proofs in the Intuitionistic Tradition

The term *proof* occurs frequently in the intuitionistic tradition, but what has been taken to be a proof in this tradition has varied considerably. Heyting's main use of the term *proof* was epistemic, but he also said, "A proof is a mathematical construction which can itself be treated mathematically".¹⁸ This ambiguity came to live on in the intuitionistic tradition.

Heyting's explanation of his epistemic notion of proof was closely linked to his notion of proposition: A proposition expresses an intention of (finding) a construction that satisfies certain conditions, while "a proof of a proposition consists in the realization of the construction required by the proposition".¹⁹ A proof in the epistemic sense is accordingly an act, and it clearly makes an assertion warranted; "the assertion of a proposition signifies the realization of the intention [expressed by the proposition]", according to Heyting.²⁰

The so-called BHK-interpretation or BHK-explanation by Troelstra [27] or Troelstra and van Dalen [28], which tells us "what forms proofs of logically

¹⁸Heyting [8]. Heyting's different uses of the term *proof* are investigated in detail in [25].

¹⁹Heyting ([8], p. 14).

²⁰Heyting ([7], p. 247).

compound statements take in terms of the proofs of the constituents”, was certainly meant to supplement Heyting’s explanation of the epistemic notion of proof. Here we may seem to be offered recursive clauses in a definition of what a proof is. But what are presented are rather recursive clauses in a definition of Heyting’s notion of the construction intended by a compound first order proposition. That this was not the intention of [27] is seen for instance from his demanding of a proof of an implication $A \rightarrow B$ that it consists not only of a construction which transforms any proof of A into a proof of B but also of the insight that the construction has this property; the insight is obviously thought to be a requirement needed in the case of epistemic proofs. [28] dropped the insight component, and so a proof of an implication came for them to stand for just a construction, a mathematical object, not for the finding of the object; nevertheless other things seem to indicate that it is anyway the epistemic notion they had in mind.

Howard’s paper “The Formulae-as-Types Notion of Construction”²¹ wants explicitly to develop the intuitionistic notion of a construction and describes in effect the build-up of terms in an extended lambda calculus, typed by formulas, which denote constructions of propositions expressible in first order predicate logic or Heyting arithmetic. Furthermore, an isomorphism between these terms and Gentzen’s natural deductions is indicated.

Martin-Löf’s intuitionistic type theory²² continues Howard’s approach and exhibits explicitly the process of building up a construction of a proposition. He calls these constructions “proofs”, but in addition, his theory introduces another kind of proofs, proofs of judgements, as he calls them, expressing that an object a is a proof (construction) of a proposition A , written $a : A$.²³ As can be expected because of the isomorphism indicated by Howard, the rules for building up proofs of propositions are like Gentzen’s introduction and elimination rules in natural deduction. In later works, he emphasizes that proof of proposition, in contrast to proof of judgement, is not an epistemic notion.²⁴

²¹Howard [10], privately circulated already in 1969.

²²Martin-Löf [14], but also several earlier papers.

²³When confusions may occur because of this double use of the term *proof*, he reserves the word *construction* for proofs of propositions. Diller and Troelstra [1] suggested that Martin-Löf’s proofs of propositions are called proof-objects; presumably to indicate that such a proof is not what they took a proof to be.

²⁴Martin-Löf [16], in particular. Nowadays he usually refers to proofs of propositions as proof-objects, following the suggestion of Diller and Troelstra; see the previous footnote.

No doubt, Martin-Löf's proofs of propositions correspond to Heyting's constructions intended by propositions, while the processes of constructing them, exhibited in Martin-Löf's type theory, correspond to what Heyting called proofs of propositions, equated by Heyting with the realizations of the constructions intended by propositions. In contrast to Martin-Löf, Heyting did not see the need to prove that an object is the construction intended by a proposition. In a later work, he said that what establishes a mathematical theorem is a successful construction, and continued: "The proof of the theorem consists in this construction itself, and the steps of the proof are the same as the steps of the mathematical construction".²⁵ In other words, he compared the steps of a realization of a construction to the inference steps of a proof, and as one does not usually prove that an inference step establishes its conclusion, he did not count with proofs of the fact that a step in the construction of an object results in the object intended.

Summing up, a proof in the intuitionistic tradition is either a construction intended by a proposition, or a realization of such a construction, or a demonstration of the judgement that a certain constructed object is the construction intended by a proposition. Proofs in the first sense, constructions intended by propositions, are the terms in which the meanings of propositions are determined. They may be viewed as truth-makers, as Göran Sundholm has suggested²⁶—the existence of such an object is what makes the proposition in question true. The condition for something to be the construction intended by a proposition may then be seen as the truth-condition of the proposition, not radically different from classical truth-conditions, as Martin-Löf [15] has emphasized. Indeed, it can be seen as the construction of the corresponding classical truth-condition—while a classical truth-condition is from a realistic point of view a condition that may be inaccessible to us, concerning a world independent of us, the intuitionistic truth-condition concerns something that we may construct and that then comes in our possession. Thus, the first notion of proof is not primarily an epistemic one, but is first of all concerned with the meaning of propositions.

The epistemic power of proofs in the second sense, realizations of intended constructions, is of course plain, if with Heyting we take the assertion of a proposition to say that one has found the intended construction. Furthermore, to take inferences to consist of construction steps in such a realization is to define them as valid, so there is no more any question of what makes them have epistemic power. However, this is a non-standard conception of

²⁵Heyting [9].

²⁶Sundholm [26].

proof and inference. Even in intuitionistic mathematics one makes arguments and infers judgements about the outcome of a construction step, as comes out in Martin-Löf's type theory.

Proofs in the third sense, demonstrations of judgements about constructions being of a certain type, are of course epistemic, but what gives these proofs their epistemic power and what it is for their inferences to be valid are questions that remain unanswered in the intuitionistic tradition.

6. Constructions as Grounds and Inferences as Operation on them

Since the point of making an inference in a proof is to get a ground for the conclusion, the inference cannot be legitimate unless it provides such a ground, given that the premisses have been satisfactorily proved and have thus been provided with grounds. But what is a ground for an assertion?

In mathematics the ground for an assertion is often equated with a proof of the assertion. But if we see proofs as chains of valid inferences and want to explain valid inference in terms of its yielding a ground for the conclusion, we must of course seek an explanation of the notion of ground that is independent of the notion of proof. I shall end by summarizing and briefly commenting on an approach to the problem of defining the notions of ground and valid inference that I have proposed in previous publications.²⁷

The explanation of propositions in terms of constructions intended by them seems to offer an account of the notion of ground independent of the notion of proof, because it seems right to say that when one has realized or found the intended construction, one is in possession of a ground for asserting the proposition. I have therefore proposed that the ground for the assertion of a proposition is to be identified with the construction intended by the proposition.²⁸ It is not to be ruled out that this way of explaining propositions is open also when the logical constants have their classical usage, but what construction is intended by a proposition must of course sometimes differ depending on whether the proposition is understood classically or constructively.

To satisfy the requirement that it is by the very inference act, the act of inferring a conclusion from a number of premisses without further additions, that we get a ground for the conclusion, it seems necessary to understand

²⁷The proposal is most developed in Prawitz [21].

²⁸The suggestion is in accordance with Dummett's idea, as found e.g. in Dummett [2] and [3], to identify the meaning of a sentence with the condition for using it warrantably with assertive force.

such an act as something more than just a speech act. Just to *assert* “ B because of $A_1, A_2, \dots, A_n$ ” cannot possibly in itself give a ground for the assertion of B , it seems, even if the premisses $A_1, A_2, \dots, A_n$ have been satisfactorily proved. My proposal is that an inference act should be taken as also involving an operation on the grounds one takes oneself to have for the premisses. An inference is then naturally defined as valid, when this operation applied to any grounds for the premisses always yields a ground for the conclusion. It thereby becomes literally true that when one makes a valid inference, thus carrying out the operation involved, one gets in possession of a ground for the conclusion.

It may be argued that for an inference act to be valid it is not enough that it gives the agent a ground for the conclusion, the agent should also know that the ground is a ground for the conclusion and that the inference has the property of always yielding a ground for the conclusion when applied to grounds for the premisses. However, I think that when we make an assertion, we do not usually also assert that the ground that we take ourselves to have for the assertion is such a ground. To require that we also have a ground for such an assertion is to start on a regress. Similarly, when making an inference, we do not usually also assert that the inference is valid; again to avoid a regress, it must be possible to make a valid inference without first making sure that the inference is valid, since it seems that we cannot hope to achieve the latter in general without making some kind of inferences.

Nevertheless, it remains that when we make a conscious inference, we take ourselves to become justified in asserting the conclusion by having got a ground for the assertion. Therefore, if a valid inference is to be understood in the way suggested as involving an operation that yields a ground for the conclusion when applied to grounds for the premisses, we must be able to recognize in some way that the operation has this property. Clearly, for this to be possible, the operation must have a limited complexity. In the same direction speaks the fact that without any such limits, the composition of valid inferences is again a valid inference, which goes against what was said in Section 4 about plausible notions of validity. The approach outlined above must consequently be supplemented with a restriction on the operations involved in valid inferences. A possibility that would need further exploration is that the restriction could be made in line with the ideas behind the notion of analytical validity.

Open Access. This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give

appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- [1] DILLER, J., and A. S. TROELSTRA, Realizability and intuitionistic logic, *Synthese* 60:253–282, 1984.
- [2] DUMMETT, M., The philosophical basis of intuitionistic logic, in H. E. Rose et al. (eds.), *Logic Colloquium '73*, North-Holland Publishing Company, Amsterdam, 1975, pp. 5–40.
- [3] DUMMETT, M., *The Logical Basis of Metaphysics*, Duckworth, London, 1991.
- [4] GENTZEN, G., Untersuchungen über das logische Schließen, *Mathematische Zeitschrift* 39:176–210, 1935.
- [5] HILBERT, D., Axiomatisches Denken, *Mathematischen Annalen* 78:405–415, 1918.
- [6] HILBERT, D., Neubegründung der Mathematik, Erste Mitteilung, *Abhandlungen aus dem mathematischen Seminar der Hamburgischen Universität* 1:157–177, 1922.
- [7] HEYTING, A., Die intuitionistische Grundlegung der Mathematik, *Erkenntnis* 2:106–115, 1931.
- [8] HEYTING, A., *Mathematische Grundlagenforschung, Intuitionismus, Beweistheorie*, Springer, Berlin, 1934.
- [9] HEYTING, A., Intuitionism in mathematics, in R. Klibansky (ed), *Philosophy in the Mid-Century*, La Nuova Italia, Florence, 1958, pp. 101–115.
- [10] HOWARD, W., The formulae-as-types notion of construction, in J. Seldin et al. (eds.), *To H. B. Curry: Essays on Combinatory Logic, Lambda Calculus and Formalism*, Academic Press, London, 1980, pp. 479–490.
- [11] KREISEL, G., A survey of proof theory, *Journal of Symbolic Logic* 33:321–388, 1968.
- [12] KREISEL, G., Book reviews, The Collected Papers of Gerhard Gentzen, *The Journal of Philosophy* 68:238–265, 1971.
- [13] KREISEL, G., A survey of proof theory II, in J. Fenstad (ed.), *Proceedings of the Second Scandinavian Logic Symposium*, North-Holland, Amsterdam, 1971, pp. 109–170.
- [14] MARTIN-LÖF, P., *Intuitionistic Type Theory*, Bibliopolis, Napoli, 1984.
- [15] MARTIN-LÖF, P., Truth of a proposition, evidence of a judgement, validity of a proof, *Synthese* 73:407–420, 1987.
- [16] MARTIN-LÖF, P., Truth and knowability: on the principles C and K of Michael Dummett, in H. G. Dales and G. Oliveri (eds.), *Truth in Mathematics*, Clarendon Press, Oxford, 1998, pp. 105–114.
- [17] PRAWITZ, D., *Natural Deduction. A Proof-Theoretical Study*, Almqvist & Wiksell, Stockholm, 1965. (Republished by Dover Publications, New York, 2006.)
- [18] PRAWITZ, D., Ideas and results in proof theory, in J. Fenstad (ed.), *Proceedings of the Second Scandinavian Logic Symposium*, North-Holland, Amsterdam, 1971, pp. 237–309.
- [19] PRAWITZ, D., Towards a foundation of general proof theory, in P. Suppes et al. (eds.), *Logic, Methodology and Philosophy of Science IV*, North-Holland, Amsterdam, 1973, pp. 225–250.
- [20] PRAWITZ, D., Meaning approached via proofs, *Synthese* 148:507–524, 2006.

- [21] PRAWITZ, D., Explaining deductive inference, in H. Wansing (ed.), *Dag Prawitz on Proofs and Meaning* (Outstanding Contributions to Logic vol. 7), Springer, Cham, 2015, pp. 65–100.
- [22] PRAWITZ, D., On the relation between Heyting’s and Gentzen’s approaches to meaning, in T. Piecha and P. Schroeder-Heister (eds.), *Advances in Proof-Theoretic Semantics* (Trends in Logic vol. 43), Springer, Cham, 2016, pp. 5–25.
- [23] PRAWITZ, D., Gentzen’s justification of inferences, in T. Piecha and P. Schroeder-Heister (eds.), *General Proof Theory. Celebrating 50 Years of Dag Prawitz’s “Natural Deduction”*. *Proceedings of the Conference held in Tübingen, 27–29 November 2015*, URI: <http://dx.doi.org/10.15496/publikation-10394>, University of Tübingen, 2016, pp. 263–276.
- [24] SCHROEDER-HEISTER, P., Validity concepts in proof-theoretic semantics, *Synthese* 148:525–571, 2006.
- [25] SUNDHOLM, G., Questions of Proof, *Manuscrito (Campinas)* 16:47–70, 1993.
- [26] SUNDHOLM, G., Existence, proof and truth-making: A perspective on the intuitionistic conception of truth, *Topoi* 13:117–126, 1994
- [27] TROELSTRA, A. S., Aspects of constructive mathematics, in J. Barwise (ed.), *Handbook of Mathematical Logic*, North-Holland Publishing Company, Amsterdam, 1977, pp. 973–1052.
- [28] TROELSTRA, A. S., and D. VAN DALEN, *Constructivism in Mathematics, vol. 1*, North Holland Publishing Company, Amsterdam, 1988.

D. PRAWITZ
Stockholm University
Stockholm
Sweden
`dag.prawitz@philosophy.su.se`