

Semantics and the Justification of Deductive Inference*

Ebba Gullberg
ebba.gullberg@philos.umu.se

Sten Lindström
sten.lindstrom@philos.umu.se

Umeå University

Abstract

Is it possible to give a justification of our own practice of deductive inference? The purpose of this paper is to explain what such a justification might consist in and what its purpose could be. On the conception that we are going to pursue, to give a justification for a deductive practice means to explain in terms of an intuitively satisfactory notion of *validity* why the inferences that conform to the practice coincide with the valid ones. That is, a justification should provide an analysis of the notion of validity and show that the inferences that conform to the practice are just the ones that are valid. Moreover, a complete justification should also explain the purpose, or point, of our inferential practice. We are first going to discuss the objection that any justification of our deductive practice must use deduction and therefore be circular. Then we will consider a particular model of justificatory explanation, building on Kreisel's concept of informal rigour. Finally, in the main part of the paper, we will discuss three ideas for defining the notion of validity: (i) the classical conception according to which the notion of (bivalent) *truth* is taken as basic and validity is defined in terms of the preservation of truth; (ii) the constructivist idea of starting instead with the notion of (a canonical) *proof* (or *verification*) and define validity in terms of this notion; (iii) the idea of taking the notions of *rational acceptance* and *rejection* as given and define an argument to be valid just in case it is irrational to simultaneously accept its premises and reject its conclusion (or conclusions, if we allow for multiple conclusions). Building on work by Dana Scott, we show that the last conception may be viewed as being, in a certain sense, equivalent to the first one. Finally, we discuss the so-called paradox of inference and the informativeness of deductive arguments.

*[Sten] It is a great pleasure to congratulate Wlodek on the occasion of his 60th birthday! He is a very special friend who has taught me more about philosophy—and how to do philosophy—than anyone else.

1 Introduction

In this paper we are going to consider the question whether it is possible for the members of a community of reasoners to justify their own deductive practice. In particular, is it possible for us to give a justification of our own practice for constructing deductive arguments? We assume that this practice is rule-governed and that it conforms to classical logic. The question is now whether we can justify the basic rules of inference that govern our practice of deductive reasoning.

In view of the conceptual analysis of Gentzen [13], it is plausible that a central aspect of our deductive practice can be represented by a system of natural deduction for first-order predicate logic with introduction and elimination rules for the logical constants \wedge , \vee , \rightarrow , \forall , \exists , and \neg .¹ According to Gentzen's analysis, the intuitively correct inferences involving the logical constants of predicate logic are broken down into atomic steps in such a way that each atomic step involves only one logical constant.² There is for each logical constant an *introduction rule* that allows the introduction of the constant in a proof and an *elimination rule* that allows its elimination:

$$\begin{array}{ll}
 (\wedge\text{-I}) & \frac{A \quad B}{A \wedge B} & (\wedge\text{-E}) & \frac{A \wedge B}{A} \quad \frac{A \wedge B}{B} \\
 (\vee\text{-I}) & \frac{A}{A \vee B} \quad \frac{B}{A \vee B} & (\vee\text{-E}) & \frac{A \vee B \quad \begin{array}{c} [A] \\ \vdots \\ C \end{array} \quad \begin{array}{c} [B] \\ \vdots \\ C \end{array}}{C} \\
 (\rightarrow\text{-I}) & \frac{\begin{array}{c} [A] \\ \vdots \\ B \end{array}}{A \rightarrow B} & (\rightarrow\text{-E}) & \frac{A \quad A \rightarrow B}{B} \\
 (\forall\text{-I}) & \frac{A(x)}{\forall x A(x)} & (\forall\text{-E}) & \frac{\forall x A(x)}{A(t)} \\
 (\exists\text{-I}) & \frac{A(t)}{\exists x A(x)} & (\exists\text{-E}) & \frac{\exists x A(x) \quad \begin{array}{c} [A(x)] \\ \vdots \\ C \end{array}}{C} \\
 (\neg\text{-I}) & \frac{\begin{array}{c} [A] \\ \vdots \\ \perp \end{array}}{\neg A} & (\neg\text{-E}) & \frac{A \quad \neg A}{\perp}
 \end{array}$$

¹Here, \wedge , \vee , \rightarrow , \forall , \exists , and \neg are formal counterparts of the expressions 'and', 'or', 'if ... , then ...', 'every', 'some', and 'not' as they are used in ordinary mathematical practice.

²For a detailed analysis and development of Gentzen's ideas, see Prawitz [22, 23]. The description given here is based on Prawitz's presentation.

In addition, the system contains the elimination rule for \perp :

$$(\perp\text{-E}) \quad \frac{\perp}{A}$$

together with all instances of the schema:

$$(\text{LEM}) \quad A \vee \neg A \quad (\text{Law of Excluded Middle})$$

as axioms.³

Informal mathematical proofs start out from assumptions, which are later eliminated or discharged. This argumentative structure of our informal deductive practice is not captured by the classical logical systems of Frege-Hilbert type. The ambition of Frege and Hilbert was to characterise the set of logically *provable* formulas rather than to analyse the concept of a *proof*. Gentzen's [13, p. 68] ambition, however, was another:

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

In contrast, I intended first to set up a formal system which comes as close as possible to actual reasoning. The result was a '*calculus of natural deduction*' [...]

There is an analogy between Gentzen's analysis of the notion of intuitively valid *proof* and Turing's analysis of the notion of *computation*. According to Turing's analysis every effective computation can, when sufficiently analysed, be broken down into operations performable by a Turing machine (*Turing's thesis*). Similarly, according to Gentzen's analysis any intuitively valid proof (involving the standard logical constants only) can, when sufficiently analysed, be broken down into the atomic steps of Gentzen's system of natural deduction. In analogy with Turing's thesis, we might call this latter thesis *Gentzen's thesis*. Turing's and Gentzen's theses are the results of conceptual analyses and cannot be proved in the strict mathematical sense.⁴ Here, we shall assume that Gentzen's thesis is correct i.e., that there is a distinctive part of our ordinary deductive practice that can be adequately represented by a system of natural deduction for classical predicate logic of the kind presented by Gentzen.⁵

What we are looking for, when we ask for a justification of the above rules, is an argument showing that they are correct in the sense that all inferences that can be

³The system without $(\perp\text{-E})$ and (LEM) yields so-called *minimal logic*, while the system with $(\perp\text{-E})$ but without (LEM) yields *intuitionistic logic*. The full system, of course, yields *classical predicate logic*.

⁴Turing's and Gentzen's analyses are examples of what Kreisel [17] calls *informal rigour*. See section 3.

⁵Usually we will be concerned with propositional logic only, making occasional reference to predicate logic. Sometimes we also consider non-classical ("intensional") connectives, in addition to the classical ones.

constructed by means of these rules are valid. An inference is *valid* if its conclusion is a *logical consequence* of its premises. So one can say that the central notion of logic is the notion of logical consequence. However, it is far from obvious how this concept should be defined. Classical logicians usually define logical consequence in terms of the preservation of truth from premises to the conclusion. Another idea is to define it in terms of preservation of assertability: if one is justified in asserting the premises of a valid argument, then one is also justified in asserting the conclusion. A third intuition is that it is irrational to accept the premises of a valid argument and, at the same time, reject its conclusion. In this paper we are going to compare these three ideas for defining logical consequence.

2 Justification and rule-circularity

Some philosophers, e.g. Susan Haack [14, 15], have argued that fundamental rules of deduction cannot be justified in a way that is not question-begging. Often when we want to justify something, we find it natural to do this by means of a deductive argument. However, deductive arguments are ultimately built up from basic rules of inference. So it seems impossible to argue deductively for these rules without making use of the very same rules. Consider, for example, the following proof that the rule modus ponens, i.e. (\rightarrow -E), is truth-preserving:

(1)	$\text{True}(A) \wedge \text{True}(A \rightarrow B)$	Premise
(2)	$\text{True}(A)$	\wedge -E
(3)	$\text{True}(A \rightarrow B)$	\wedge -E
(4)	$\text{True}(A \rightarrow B) \rightarrow (\text{True}(A) \rightarrow \text{True}(B))$	truth-table for \rightarrow
(5)	$\text{True}(A) \rightarrow \text{True}(B)$	(3), (4) modus ponens
(6)	$\text{True}(B)$	(2), (5) modus ponens
(7)	$(\text{True}(A) \wedge \text{True}(A \rightarrow B)) \rightarrow \text{True}(B)$	(1)–(6) \rightarrow -I

This argument is *rule-circular*: we have used modus ponens in order to prove that modus ponens is truth-preserving. Similarly, any deductive argument with the conclusion that a basic rule R is correct must either use R or some derived rule whose proof will presuppose the use of R .⁶ Moreover, non-deductive methods of justification, like appeal to intuition or induction, appear to be non-starters. It seems, therefore, that any justification of our basic deductive rules has to be rule-circular.

A circular argument of the ordinary kind, what Boghossian [2] calls a *grossly circular* argument, has the conclusion as one of its premises. The term rule-circularity is only applicable to arguments that purport to establish the validity of a rule of inference. Such an argument is rule-circular if it uses that very rule in one of its argumentative steps. A grossly circular argument is obviously question-begging and gives no support for its conclusion. Couldn't the same charge be levelled at

⁶Here, we assume that the basic rules of a deductive practice are *independent* of each other, i.e., none of the basic rules can be derived from the others.

rule-circular arguments? If the correct use of a rule of inference always required a justification of the rule itself, any rule-circular argument would presuppose its own conclusion and would therefore be grossly circular. It seems reasonable, however, that we can be entitled to use some rules, the basic ones, without prior justification.⁷ It would be too strong to demand that the entitlement to use a rule of inference always required the possession of a justification of that rule, since such a demand—it appears—leads to an infinite regress.⁸

It seems that we can still level two objections against rule-circular justifications. Firstly, they are powerless of persuading a sceptic of the correctness of a rule. Secondly, it appears that we can give rule-circular justifications for patently unsound rules of inference. In connection with the first objection, Michael Dummett [9] has suggested that we make a distinction between what he calls *suasive* and *explanatory* arguments. A *suasive* argument is one in which we believe or accept the premises and are convinced or persuaded by the argument to accept the conclusion. This is the kind of argument that a justification of a basic rule of inference cannot be because of circularity. If we genuinely doubt the correctness of our own deductive practice we cannot be convinced by a rule-circular justification. In an *explanatory* argument, on the other hand, we do not doubt the validity of our basic rules of inference. What the argument does is to *explain* why it is reasonable for us to believe that they are valid. If we accept this distinction, and also accept that explanatory arguments, at least in some sense, can provide us with justifications of things we already believe, then it seems clear that our rules of deductive inference can be justified by giving an explanation of why they work the way they are supposed to. For example, one might argue that a deductive practice achieves its purpose if acting in accordance with the practice can never lead from true premises to false conclusions. Suppose that we can show that this is indeed the case for a given practice and that we can also explain why this is so. Then it seems that we can justify the given practice in the sense that we can explain what its purpose is and why this purpose will be achieved as long as we act in accordance with the practice.

The second objection, the so-called *bad company objection*, says that it is possible to give rule-circular justifications also for rules that are unsound. An example that is often cited concerns Prior's [26] connective 'tonk' which is supposed to be governed by the following pair of inference rules:

$$\text{(tonk-I)} \quad \frac{A}{A \text{ tonk } B} \qquad \text{(tonk-E)} \quad \frac{A \text{ tonk } B}{B}$$

Imagine a community of reasoners whose deductive practice includes the connective 'tonk' governed by these rules. It seems that they should be able to produce a rule-circular argument for the validity of the rule tonk-I along the following lines:

⁷Cf. Boghossian [2, 3].

⁸Cf. Lewis Carroll [5].

- | | | |
|-----|---|--------------------------|
| (1) | $\text{True}(A)$ | Premise |
| (2) | $\text{True}(A) \text{ tonk } \text{True}(A \text{ tonk } B)$ | (1) tonk-I |
| (3) | $\text{True}(A \text{ tonk } B)$ | (2) tonk-E |
| (4) | $\text{True}(A) \rightarrow \text{True}(A \text{ tonk } B)$ | (1)–(3) \rightarrow -I |

A rule-circular argument for the validity of tonk-E could be given along similar lines.

By means of the ‘tonk’ rules, we can infer any sentence from any other sentence. Given that there is at least one logical truth, one can even show that A is logically true for any A . In other words, any logical system that contains the connective ‘tonk’ is “exploding”, i.e., it proves any sentence whatsoever. It can therefore seriously be doubted whether any deductive practice could contain a connective like ‘tonk’. Or rather a practice involving ‘tonk’ would not deserve the label ‘deductive’. A rule-circular justification for a rule of inference is supposed to be given from the standpoint of a community of reasoners. In the case of ‘tonk’, there appears to be no place for such a community. Hence, there is no rule-circular justification for the validity of the rules for ‘tonk’.

Still there is another worry, namely that it would be impossible to rationally criticise our own deductive practice. But, as we will see below (section 4.2), this is not necessarily the case. It is possible that philosophical considerations would lead us to accept a different system of logic than the one we were initially trying to justify. If this happens and we take our arguments seriously, we could find ourselves forced to revise our logical practice in order to make it consistent with our most basic assumptions.

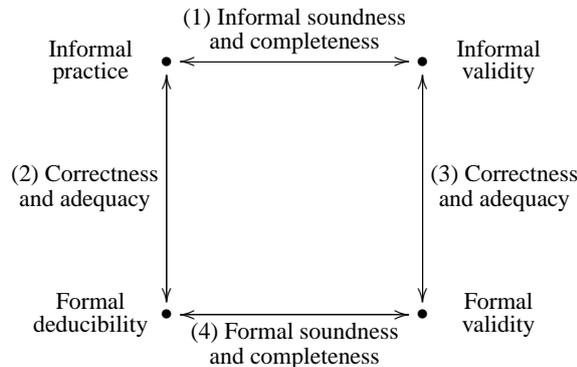
3 Justification and informal rigour

According to the standard, or received, view, a *formal logic* has both a *proof-theoretic* and a *semantic* part. The proof-theoretic part consists of a formal language together with a deductive system (for example, a logical calculus of propositional or first-order predicate logic) formulated therein. The semantic part is usually taken to consist of a Tarski-style model-theoretic semantics, but there are also other alternatives, for example a proof-theoretic semantics in the style of Prawitz [25]. Providing an informal deductive practice with a justification might be taken to involve the following steps:

- *Informal semantic analysis*: This means providing an informal notion of *valid inference* and arguing that the aim of the practice is to construct inferences that are valid in this sense. A deductive practice is *informally sound* if every chain of reasoning in accordance with the practice corresponds to an argument that is valid in the informal sense. If every argument that is valid in the informal sense can be reproduced as a piece of reasoning in accordance with the practice, we say that the practice is *informally complete*.

- *Formalisation*: This step involves constructing a *formal deductive system* and arguing that the informal practice can be represented (correctly and adequately) within this system. A deductive system \mathcal{D} is *correct* relative to an informal deductive practice if each of its primitive rules of inference corresponds to a gap-free inferential step in accordance with the practice. \mathcal{D} is *adequate* with respect to a certain deductive practice if, for each chain of reasoning in accordance with the practice, there is a deduction in \mathcal{D} that captures its form.⁹
- *Formal semantic analysis*: Here we define an exact notion of *formal validity* and argue that this notion can be taken to represent (correctly and adequately) the notion of informal validity. The formal notion of validity is *correct* if every formally valid inference corresponds to arguments in natural language that are informally valid. It is *adequate* if for every natural language argument that is valid in the informal sense, there is a formal language argument that is formally valid.
- *Soundness and completeness proofs*: Finally, we provide, if possible, a mathematical proof that the deductive system \mathcal{D} is sound and complete with respect to the notion of formal validity. \mathcal{D} is *sound*, if whenever a formula A is provable in \mathcal{D} from a set of premises Γ , A is a logical consequence of Γ according to the formal semantics. \mathcal{D} is (strongly) *complete* with respect to the formal semantics if the converse implication holds.

These steps can be represented by the following picture:



Given a deductive practice, a notion of informal validity, and a logic consisting of a formal deductive system with a formal semantics, we may formulate the following hypotheses (compare the numbered arrows in the picture):

⁹The terminology of “correctness” and “adequacy” is due to Shapiro [34]. Shapiro stresses that correctness “is a vague matter, not an all-or-nothing affair. If a deductive system \mathcal{D} is (more or less) correct, then each deduction in \mathcal{D} (more or less) corresponds to a legitimate, or valid, derivation in ordinary reasoning”. He adds: “Like correctness, adequacy is a vague matter, especially if we are limiting its scope to certain kinds of arguments.” [34, p. 661].

- (1) *Informal soundness and completeness hypothesis*: The notions of *informal logical proof* and *informal logical derivation* (from assumptions) are sound and complete with respect to informal notions of logical validity and logical consequence, respectively.
- (2) *Proof-theoretic representation hypothesis*: The notions of *intuitive logical proof* and *intuitive logical derivation* (from assumptions) are (correctly and adequately) represented by the notions of *formal proof* and *formal derivation*, respectively.
- (3) *Semantic representation hypothesis*: The notions of *informal logical truth* and *informal logical consequence* are (correctly and adequately) represented by the corresponding formal notions of logical validity and logical consequence, respectively.
- (4) *Formal soundness and completeness hypothesis*: The notions of *formal logical proof* and *formal logical derivation* (from assumptions) are sound and complete with respect to the formal notions of logical validity and logical consequence.

Arguing for the claims (1)–(3) is basically a matter of philosophical analysis. Formal soundness and completeness, on the other hand, are mathematical claims that demand mathematical proofs. In this connection, Kreisel [17] has made a distinction between *informal* and *formal rigour*. Formal rigour consists in constructing formal deductive proofs and derivations in accordance with fixed (formal) rules. Informal rigour, on the other hand, extends formal rigour by appealing to uncontroversial properties of our intuitive notions. Informal rigour is applied, for instance, when we by means of careful conceptual analysis lay down axioms for some informal notion. Dedekind’s analysis of our concept of the system of natural numbers, leading to the formulation of the Dedekind-Peano axioms, or Zermelo’s analysis of the iterative notion of set, leading eventually to the formulation of the axioms of ZF set theory, are examples of informal rigour. The claim that a given intuitive notion can be represented by a corresponding precise mathematical one, we may call a *representation hypothesis*. Kreisel argued that we sometimes can *prove* a representation hypothesis, by means of informal rigorous argumentation.

A famous application of Kreisel’s notion of informal rigour is his proof that for first-order languages, the informal notion of *logical validity* is extensionally equivalent to the exact notion of *model-theoretic validity*. Informal logical validity, Kreisel analyses as truth in *all* interpretations, where the domain of an interpretation may be either a set or a proper class. For example, the so-called *standard model* for set-theory $\langle V, E \rangle$, where V is the collection of all sets (“the cumulative hierarchy”) and E is the relation of membership between sets, is an interpretation in this informal sense. It cannot be thought of as a set-theoretic entity, since its domain is not a set. Intuitively, a set-theoretic sentence α is *true* just in case it is true in the standard model. Set-theoretic truth is an informal, or intuitive, notion

that cannot be studied by set-theoretic means. The same holds for the notion of informal logical validity. The following relationship is an obvious consequence of the definitions of these notions:

- (1) If α is a set-theoretic sentence, then α is logically valid (in the informal sense) only if α is true.

However, the following statement is far from obvious:

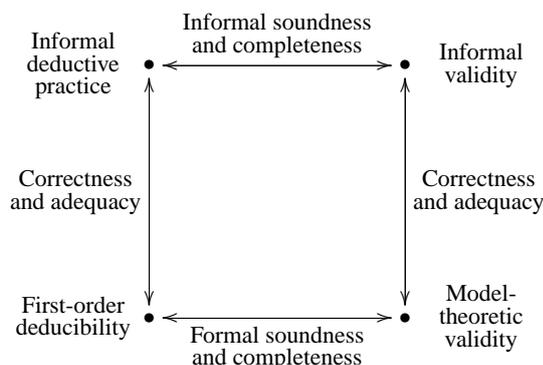
- (2) If α is a set-theoretic sentence, then α is model-theoretically valid only if α is true.

Consider a given first-order language. Let D be the set of all sentences in this language that are theorems (provable) in (a given system of) first-order predicate logic (with identity). Let Val be the set of informally valid sentences and let val be the set of all sentences that are model-theoretically valid. Now, Kreisel shows that:

$$Val = val = D$$

Proof. Clearly, $Val \subseteq val$. By Gödel's completeness theorem for first-order logic, $val \subseteq D$. By the intuitive soundness of first-order logic, $D \subseteq Val$. Hence, the three notions are extensionally equivalent. \square

Suppose that we are considering a fragment of natural language for which deductive reasoning can be (correctly and adequately) represented by first-order logic. In view of Kreisel's result, the following diagram commutes for such a fragment:



The scope of this result is not clear, since it is not obvious how much of informal deductive practice can be represented within first-order logic. In particular, it is not clear whether first-order set theory (ZFC) is adequate for the representation of ordinary mathematical reasoning. Some philosophers—Quine [27] is one example—claim that it is. But others, like Shapiro [32], have argued that only full second-order logic has the resources for adequately representing informal mathematical reasoning. However, completeness does not hold for second-order logic,

i.e. there is no formal deductive system which is complete for full second-order logic. Moreover, there is a simple example, due to Vann McGee [19], showing that the equivalence between informal logical validity and model-theoretic validity need not hold when we increase the expressive resources of the object language:

Consider the language which is obtained from the language of first-order set theory by adding a generalised quantifier $\exists_{\text{abs inf}}$ as a new logical constant, where $(\exists_{\text{abs inf}} x)\alpha(x)$ means:

for absolutely infinitely many $x, \alpha(x)$.

The concept of the *absolutely infinite* is due to Cantor. Intuitively a class is absolutely infinite if it is bigger than any set, i.e., if it is a proper class. Hence, $(\exists_{\text{abs inf}} x)\alpha(x)$ is true in an interpretation \mathcal{I} if and only if the collection of all members of the domain of \mathcal{I} satisfying α is a proper class.

Consider now the sentence

$$(\exists_{\text{abs inf}} x)(x = x).$$

This sentence is true in the universe V of sets: there are absolutely infinitely many things in the universe of sets. However, there is no (set) model in which it is true, since any model has a set as its domain. Hence,

$$\neg(\exists_{\text{abs inf}} x)(x = x)$$

is an example of a *false sentence that is true in all (set) models*. That is, it is false but valid in the model-theoretic sense. It is clear from this that the implication

Truth in all (set) models \rightarrow logical validity,

fails for the language in question. This is so, since logically valid sentences must be true.

For a certain fragment of natural language (“the first-order fragment”), however, we seem to have a justification of sorts for first-order logic in view of the equivalence of the notions of first-order provability, intuitive logical validity, and model-theoretic validity. For this fragment, it appears that the proof-theoretic and semantic concepts are all interwoven into a coherent, mutually supporting structure.

4 Three types of justification of deductive logic

In the following, we are going to pursue three ideas for defining logical consequence and justifying deductive inference:

- (i) in terms of truth-preservation,
- (ii) in terms of assertability- or verifiability-conditions, and
- (iii) in terms of rational acceptance and rejection.

4.1 Truth-theoretic justification

In this section, we take the classical notions of truth and falsity as given and use them to explain the meaning of the logical constants and the validity of the logical rules of inference. In accordance with classical truth-conditional semantics—going back to Frege and the early Wittgenstein—the meanings of the logical constants are specified in terms of *truth-conditions*, or in the case of the quantifiers, in terms of *satisfaction-conditions*. The logical validity of inferences is then explicated in terms of the *preservation* of truth (or satisfaction) from premises to conclusion. Given the truth-conditions (or, in the more general case, satisfaction-conditions) for the logical constants, one can then show that the rules of inference of classical logic are *sound*, i.e. truth-preserving. It is assumed that this kind of argument is available to a sufficiently sophisticated reasoner who can thereby give a theoretical explanation (an *explanatory justification* in Dummett’s sense) of her own deductive practice. Such an explanation will be rule-circular, but will lead to a theoretical understanding of the given practice. By such an explanation the reasoner can, it is hoped, achieve a coherence between her deductive practice and her logical theory.

Let us now outline how the above line of reasoning can be made explicit.¹⁰ We assume that every sentence in the language under study has exactly one of the truth-values *True* and *False*. According to the classical truth-conditional view, the meaning of a declarative sentence is its truth-conditions, and the meaning of a meaningful part of a sentence is the contribution it makes to the truth-conditions of sentences of which it is a part. To understand the meaning of a meaningful expression is to know its meaning, i.e., in the case of a sentence to know its truth-conditions, and in the case of a meaningful subsentential expression to know its contribution to the truth-conditions of sentences in which it is a part.

The question now arises how we should understand the notion of truth-conditions. One way, going back at least to Carnap [4], is to explicate the truth-conditions of a sentence as an abstract entity, namely as a function from possible worlds (possible circumstances, situations) to truth-values. The idea is that the meaning (or to use Carnap’s terminology, the *intension*) of a declarative sentence that does not involve context-sensitive elements like tense, indexicals or demonstratives, is the function which specifies under which circumstances, in which “possible worlds”, the sentence is true. We may speak of such a function as the (Carnapian) *proposition* expressed by the sentence. Equivalently we can identify the Carnapian proposition expressed by a sentence with the set of possible worlds (or situations) where the sentence is true.

According to classical possible worlds semantics, every meaningful expression E of a language (without context-dependent elements) has an intension which specifies for every possible world u , an appropriate *extension* for the expression E in the world u . If E is a singular term, then the extension of E is the object that E

¹⁰There are various ways of developing truth-conditional semantic: the one sketched here being only one of the alternatives, another one being to follow Davidson’s [6] truth-theoretic approach inspired by the early work of Tarski on semantic truth definitions.

refers to (relative to u), if E is an n -ary predicate expression, then the extension of E is an n -ary (set-theoretic) relation. Finally, if E is a sentence, the extension of E is its truth-value (relative to u). We write $\llbracket E \rrbracket$ for the intension of E and $\llbracket E \rrbracket_u$ for the extension of E in the possible world u . It is assumed that the semantics is *compositional* in the sense that the intension of a complex expression is always a function of the intensions of its parts (the *principle of intensionality*).

In terms of this semantical framework we can specify the meaning (truth-conditions) of the classical logical connectives as follows:

For any possible world u ,

- (1) $A \wedge B$ is true in u iff A is true in u and B is true in u .
- (2) $A \vee B$ is true in u iff A is true in u or B is true in u .
- (3) $A \rightarrow B$ is true in u iff either A is false in u or B is true in u (or both).
- (4) $\neg A$ is true in u iff A is false in u .

(Or, if we define $\neg A$ as $A \rightarrow \perp$, we stipulate that \perp is false in u , for any u .)

A connective is *extensional* (or *truth-functional*) if for any world u , the truth-value in u of a sentence built up by means of the connective from other sentences is a function of the truth-values of these sentences. For example, a binary connective $*$ is extensional (or truth-functional), if there exists a truth-function t_* such that for any possible world u ,

(Ext) $\llbracket A * B \rrbracket_u = t_*(\llbracket A \rrbracket_u, \llbracket B \rrbracket_u)$.

The classical connectives \wedge , \vee , \rightarrow , \perp defined above are obviously extensional.

In addition to extensional connectives, the language may contain various intensional connectives satisfying the principle of intensionality. For instance, if $*$ is an intensional binary connective, then there is a function F_* such that for all sentences A, B ,

(Comp) $\llbracket A * B \rrbracket = F_*(\llbracket A \rrbracket, \llbracket B \rrbracket)$.

Within this framework, it seems reasonable to say that a binary connective $*$ is meaningful only if there exists a function F_* (specifying the meaning, or intension, of $*$) satisfying the condition (Comp). At the end of this section, we show that ‘tonk’ is not meaningful in this sense. The assumption that ‘tonk’ has a meaning leads to a contradiction.

Next, we turn to the notion of *logical consequence*. The following conditions are usually considered necessary for a sentence A to be a logical consequence of a set Γ of premises:

- (1) it is impossible for all the premises in Γ to be true but the conclusion A to be false; in other words, the inference from the premises to the conclusion is necessarily truth-preserving.
- (2) it is in virtue of the meanings of the logical constants that the inference from premises to conclusion is necessarily truth-preserving.

We assume that these requirements are also jointly sufficient for logical consequence. Accordingly, we say that A is a *logical consequence* of Γ in the *intuitive sense* if, and only if, the conditions (1) and (2) are both satisfied. We say that an argument (or an inference) is *intuitively valid* if its conclusion is a logical consequence (in the intuitive sense) of its premises.

Once we have agreed on an intuitive analysis of the notion of logical consequence, we can turn to the task of giving an exact mathematical characterisation. Suppose that our language \mathcal{L} contains the standard Boolean connectives $\wedge, \vee, \rightarrow, \neg$ and in addition some intensional connectives, for instance a unary connective \Box and a binary connective $*$. A *model* for \mathcal{L} is a structure $\mathfrak{M} = \langle U, \llbracket \cdot \rrbracket \rangle$, where U is a non-empty set, $\llbracket \cdot \rrbracket$ is a function from the sentences of \mathcal{L} into 2^U (i.e., the set of all functions from U to the set $\{0, 1\}$, where 0 and 1 represent the truth-values False and True, respectively), and the following conditions are satisfied:

- (1) $\llbracket A \wedge B \rrbracket_u = 1$ iff $\llbracket A \rrbracket_u = \llbracket B \rrbracket_u = 1$
- (2) $\llbracket A \vee B \rrbracket_u = 1$ iff $\max(\llbracket A \rrbracket_u, \llbracket B \rrbracket_u) = 1$
- (3) $\llbracket A \rightarrow B \rrbracket_u = 1$ iff not: $\llbracket A \rrbracket_u = 1$ and $\llbracket B \rrbracket_u = 0$.
- (4) $\llbracket \neg A \rrbracket_u = 1$ iff $\llbracket A \rrbracket_u = 0$.

U is called the *domain* or *universe* of the model and the elements of U are called *possible worlds*. For each sentence A , $\llbracket A \rrbracket$ is the intension of A in \mathfrak{M} and, for $u \in U$, $\llbracket A \rrbracket_u$ is the extension (or truth-value) of A at u (relative to \mathfrak{M}). We say that A is true at the world u in the model \mathfrak{M} (in symbols, $\mathfrak{M}, u \models A$) if $\llbracket A \rrbracket_u = 1$ holds in \mathfrak{M} . The model \mathfrak{M} is said to be classical with respect to the connectives \Box and $*$ if there are functions $F_\Box : 2^U \rightarrow 2^U$ and $F_* : 2^U \times 2^U \rightarrow 2^U$ such that:

- (5) $\llbracket \Box A \rrbracket = F_\Box(\llbracket A \rrbracket)$.
- (6) $\llbracket A * B \rrbracket = F_*(\llbracket A \rrbracket, \llbracket B \rrbracket)$.

Intuitively, these conditions mean that the connectives \Box and $*$ are intensional in \mathfrak{M} .

Let \mathbb{K} be a non-empty class of models. We say that A is a \mathbb{K} -*consequence* of Γ (in symbols, $\Gamma \models_{\mathbb{K}} A$) if for every model \mathfrak{M} in \mathbb{K} and every world u in \mathfrak{M} , if $\mathfrak{M}, u \models B$ for all sentences B in Γ , then $\mathfrak{M}, u \models A$. When \mathbb{K} is the class of all models, we say that A is a *tautological consequence* of Γ .

It is easy to see that the rules of classical natural deduction for propositional logic are sound relative to $\models_{\mathbb{K}}$ for any class \mathbb{K} of models. Moreover, by the strong completeness theorem for propositional logic, A is a tautological consequence of Γ if and only if A is formally derivable from Γ in the system of classical natural deduction.

We end this section with a discussion of the connective ‘tonk’. We assume that the object language contains a connective ‘tonk’ and that there is a non-empty class

of models \mathbb{K} validating the tonk-rules. Then we have for any model \mathfrak{M} in \mathbb{K} and any world u in \mathfrak{M} :

- (a) if $\llbracket A \rrbracket_u = 1$, then $\llbracket A \text{ tonk } B \rrbracket_u = 1$,
- (b) if $\llbracket A \text{ tonk } B \rrbracket_u = 1$, then $\llbracket B \rrbracket_u = 1$.

Hence, for all A, B

- (c) if $\llbracket A \rrbracket_u = 1$, then $\llbracket B \rrbracket_u = 1$.

But $\llbracket \top \rrbracket_u = 1$ ($\top =_{\text{df}} \neg \perp$). Thus $\llbracket B \rrbracket_u = 1$ for all B . In particular, $\llbracket \perp \rrbracket_u = 1$, which contradicts the definition of truth in a model. Hence, the ‘tonk’-connective is not meaningful within the truth-conditional semantics presented here.

4.2 Verificationist justification

In order to accept the kind of justification of deductive inference that was given in the previous section, one must accept the assumptions on which the argument depends. This seems to involve the adoption of *semantic realism*, characterised by Miller [20] as the view that the understanding of a sentence consists in grasping its truth-conditions and that these conditions are potentially *recognition-transcendent*, i.e. there can be cases where the truth-conditions of a sentence obtain without us having any means of recognising that they do. Opponents of semantic realism, e.g. Dummett [8], have argued that this view is incompatible with the idea, going back to Wittgenstein, that our understanding of the meaning of a sentence has to be something that we are capable of *manifesting* through the way we use the sentence (the *manifestability requirement*). Their argument runs along the following lines:

Suppose that we can understand the sentences of a language by grasping their truth-conditions, i.e. by knowing what it is for the sentences to be true. If communication and learning is to be possible, there must be some way for us to *manifest* that we have understood a sentence and to *observe* that someone else has understood. Sometimes this can be accomplished by the use of a synonymous sentence. But sooner or later we reach a point where we can no longer state our understanding explicitly in words. The idea then is that we must manifest the knowledge that constitutes our understanding of a certain sentence through some practical ability connected to the way we use this sentence. When the sentence is decidable, in the sense that we have a procedure for determining whether it is true or false, we can say that our understanding can be manifested through our ability to apply this decision procedure and place ourselves in a position where we can recognise the truth-value of the sentence. However, there are sentences—for instance many sentences of mathematical theories, or sentences about the past—that we think we understand even though there is no method for determining their truth-value. For such sentences it seems that the knowledge that, according to semantic realism, constitutes our understanding of them goes beyond what can be manifested through any practical ability. We simply cannot find any linguistic behaviour that

corresponds to the ability of grasping what it is for these undecidable sentences to be true in cases when we have no means of recognising that they are. So we must either reject the idea that the understanding of a sentence consists in grasping its truth-conditions, or redefine truth in a way that does not make it recognition-transcendent.

Miller [20] distinguishes between a strong and a weak version of this so-called manifestation argument. According to the strong version, advocated by Dummett, semantic realism is false and must be rejected.¹¹ The weak version, on the other hand, only claims that we have no reason to prefer semantic realism to semantic anti-realism. Perhaps this is the best way to understand the argument: not as a conclusive argument forcing us to abandon semantic realism, but rather as presenting a *challenge* to the semantic realist, namely, to explain how our knowledge of truth-conditions is manifested in our use of language.

If this challenge is to be taken seriously, we need an alternative to truth-conditional semantics, an explanation of meaning and logical consequence that avoids recognition-transcendence. In the special case of mathematics, Dummett [8] and Prawitz [24] suggest that the most natural starting point for such an alternative is to use the notion of *proof* instead of truth as fundamental. On this conception, it is not the preservation of (classical) truth from premises to conclusion that makes a mathematical argument valid, but the preservation of provability: the validity of an argument is a guarantee that if we can prove the premises, then we can also prove the conclusion.

The practical ability through which knowledge of the meaning of a mathematical statement can be manifested is the ability to recognise a proof of the statement when one is presented to us. We should note that this is not the same as to say that understanding lies in the ability to actually construct a proof or to decide whether a statement is provable or not. Dummett explicitly points out that “our understanding of a statement consists in a capacity, not necessarily to find a proof, but only to recognise one when found”.¹² It seems to be a part of the intuitive notion of a proof that if something is a proof of a statement A , then it is possible to know that it is.¹³ Hence, if knowledge of the meaning of A manifests itself in the ability to recog-

¹¹Cf. Prawitz [24] and Pagin [21] for analyses of Dummett’s version of the manifestability argument.

¹²[10, p. 70].

¹³Cf. Kreisel [16, pp. 201-202], “The *sense* of a mathematical assertion denoted by a linguistic object A is intuitionistically determined (or understood) if we have laid down what constructions constitute a *proof* of A , i.e., if we have a construction r_A such that, for any construction c , $r_A(c) = 0$ if c is a proof of A and $r_A(c) = 1$ if c is not a proof of A : the logical particles in this explanation are interpreted truth functionally, since we are adopting the basic intuitionistic idealization that we can recognize a proof when we see one, and so r_A is decidable. (Note that this applies to *proof*, not *provability*.)” Compare also Dummett [10, p. 110], “The fundamental idea is that the grasp of the meaning of a mathematical statement consists not in a knowledge of what has to be the case, independently of our means of knowing whether it is so, for the statement to be true, but in an ability to recognize, for any mathematical construction, whether or not it constitutes a proof of the statement; an assertion of such a statement is to be construed, not as a claim that it is true, but as a claim that a proof of it exists or can be constructed.”

nise a proof (or a refutation) of A when one is presented to us, then it seems that our knowledge of the meaning of A can be manifested in our use of A . It appears, therefore, that an approach to meaning based on the notion of proof of the kind proposed by Dummett and Prawitz will satisfy the manifestability requirement.

Let us now explain, in rough outline, what a theory of meaning of the proposed kind might look like. For simplicity we focus on the language of arithmetic. The so-called *BHK-interpretation* (Brouwer-Heyting-Kolmogorov) of intuitionistic mathematics consists in explaining the meaning of a sentence A by defining, in a recursive manner, what a direct, or *canonical*, proof of A would consist in. Given that we know what constitutes a canonical proof of an atomic sentence, we can define the notion of canonical proof of complex sentences in the following way (which we can also see as a specification of the meaning of the logical constants):¹⁴

To form a canonical proof of A it is necessary and sufficient to have

$A \wedge B$	canonical proofs of A and of B ;
$A \vee B$	a canonical proof of A or of B ;
$A \rightarrow B$	a procedure which yields a canonical proof of B when applied to a canonical proof of A ;
$\neg A$	a procedure which transforms any canonical proof of A into a canonical proof of a contradiction (\perp);
$\forall xA(x)$	a procedure which yields a canonical proof of $A(\mathbf{n})$ when applied to any numeral \mathbf{n} ;
$\exists xA(x)$	a canonical proof of $A(\mathbf{n})$ for some numeral \mathbf{n} .

It is assumed that nothing is a proof of \perp .

In this definition, the notion of a *procedure* is taken as primitive. Someone who understands this notion together with the above definition, can also be credited with an understanding of the notion of a canonical proof. If we identify the *meaning* of a statement A with the property P_A of being a canonical proof of A , then a person who knows the meaning of A , should in principle be able to decide whether any construction has the property P_A of being a proof of A , or not. It is through this ability that the person's understanding of A manifests itself.

The BHK-interpretation immediately provides us with justifications of the introduction rules for the logical constants \wedge , \vee , \rightarrow , \forall and \exists . These rules can be viewed as prescriptions for constructing canonical proofs. For example, the rule (\rightarrow -I), can be read as saying that any procedure which yields a canonical proof of B when applied to a canonical proof of A , yields a canonical proof of $A \rightarrow B$. We can also show, in a more indirect way, that the elimination rules for \wedge , \vee , \rightarrow , \forall and \exists are valid under the BHK-interpretation. Consider, for instance, (\rightarrow -I), i.e., *modus ponens*. Given canonical proofs of the premises A and $A \rightarrow B$, one can obtain a canonical proof of B by concatenating the canonical proof of A with the

¹⁴See Prawitz [24, p. 26].

the canonical proof of $A \rightarrow B$. Moreover, the the BHK-interpretation validates the rules (\neg -I) and (\neg -E) for negation. Hence, all of Gentzen’s natural deduction rules for intuitionistic predicate logic are validated by the BHK-interpretation.

However, we cannot show that the law of excluded middle, $A \vee \neg A$, is valid on the BHK-interpretation. One has the right to assert a particular instance $A \vee \neg A$ of the law of excluded middle, only if one can construct either a canonical proof of A or a canonical proof of $\neg A$. But we have no guarantee that we always will be able to accomplish this. Let A be an undecided sentence, i.e., a sentence that we have not been able yet either to prove or disprove. For instance, A could be the sentence “There exists an odd perfect number”. Then there is no guarantee that we will ever find a proof of A . Nor is there a guarantee that we will find a proof of $\neg A$. Hence, we have no right to assert $A \vee \neg A$. Thus, the law of excluded middle cannot be justified solely on the basis of the BHK-interpretation. It should be pointed out that $\neg\neg(A \vee \neg A)$ is valid on the BHK-interpretation, so we can never hope to find a statement A for which $\neg(A \vee \neg A)$ holds. In other words, we can never expect to find, on the basis of the BHK-interpretation, an actual counterexample to the law of excluded middle.

An important part of Gentzen’s proof-theoretic analysis of our deductive practice is his discovery of a certain symmetry between the introduction rules and the elimination rules for the logical constants. As Prawitz [23] puts it:

... the corresponding introductions and eliminations are *inverses* of each other. The sense in which an elimination, say, is the inverse of the corresponding introduction is roughly this: the conclusion obtained by an elimination does not say anything more than what must already have been obtained if the major premiss of the elimination was inferred by an introduction. [...] In other words, a proof of the conclusion of an elimination is already “contained” in the proofs of the premises when the major premiss is inferred by introduction. We shall refer to this by saying that the pairs of corresponding introductions and eliminations satisfy the *inversion principle*.

The inversion principle is closely connected with Gentzen’s [13, p. 80] seminal idea that the meaning of a logical constant is determined by its introduction rule in a system of natural deduction:

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.’

The connection with the BHK-interpretation is immediate as soon as we view the introduction rules as prescriptions for constructing canonical proofs. All the other rules of inference in a system of natural deduction must, on this interpretation, be justifiable on the basis of the meaning provided for the logical constants by the introduction rules. However, it is easy to see that the law of excluded middle

is not so justifiable. Thus, the Gentzen-Prawitz view that all rules of inference be justifiable on the basis of the meaning given to the logical constants by the introduction rules leads to a *revision* of our logical practice (that we have assumed to be in conformity with classical logic).

Gentzen’s proof-theoretic analysis can be applied in a straightforward way to the connective ‘tonk’.¹⁵ It is easily seen that the rules of ‘tonk’ do not satisfy the inversion principle: the elimination rule for ‘tonk’ cannot be justified on the basis of its introduction rule. Hence, ‘tonk’ does not satisfy Gentzen’s and Prawitz’s requirements on a meaningful logical connective. As Tennant [37, pp. 637–638] puts the point:

Prior’s introduction rule [tonk-I] confers on A [tonk] B the meaning ‘ A regardless whether B ’. By contrast, Prior’s elimination rule [tonk-E] would confer on A [tonk] B the meaning ‘Regardless whether A , B ’ [...] There is no logical operator that these two rules, taken together, succeed in characterizing [...]

The proof-theoretic theory of meaning advocated by Dummett and Prawitz has many attractive features. However, it is revisionist in nature. Adherents of this approach advocate abandoning classical logic in favour of some weaker logic satisfying the inversion principle. However, we were looking for a justification of existing practice, which we assumed to be in accordance with classical logic. Hence, we cannot be satisfied with the kind of approach considered in this section.

4.3 Epistemic justification

In this section, we will try to give a justification of classical propositional logic without assuming, at the outset, a notion of truth or proof. Instead, we will start out from the idea of a sentence A being a logical consequence of a set Γ if and only if the following two conditions are satisfied:

- (1) it is incoherent to accept all the sentences of Γ and simultaneously reject A ;
- (2) it is in virtue of the meaning of the logical constants that (1) holds.¹⁶

That is, we will start out from the *epistemic attitudes of acceptance and rejection* that an epistemic agent might have towards a sentence A . At a given time a rational agent might either:

¹⁵Cf. Belnap [1] and Tennant [37].

¹⁶In his book *Rational Belief Systems* [11], Brian Ellis develops an epistemic approach to logical validity along these lines: “[...] as validity is here understood, an argument is valid iff there is no rational belief system in which its premises are accepted and its conclusion rejected. [...] it follows from my requirements on validity that to understand a statement sufficiently for all purposes of logic we need to know its acceptability conditions, i.e., the conditions under which it may be accepted or rejected by an ideally rational man.” Ellis’s belief systems are essentially what we here call *belief states*. Restall [28] contains an epistemic approach to (multiple-conclusion) logical consequence which is similar to the one described here. He does not mention the equivalence between the epistemic semantics and Scott’s two-valued truth-value semantics (see Theorems 1 and 2, below), which is the main point of this section.

- (1) *accept* A ,
- (2) *reject* A , or
- (3) hold A in *suspence*, i.e., neither accept nor reject A .

It is not possible for a rational agent, simultaneously, both to accept and reject one and the same sentence.

We assume that a certain language \mathcal{L} be given which is closed under the sentential connectives $\wedge, \vee, \rightarrow, \neg$. The *belief state* s of an agent (at a certain time) is represented by a system $[\Gamma; \Delta]$ consisting of the set Γ of all the sentences that the agent accepts in s (the *acceptance set* of s) together with the set Δ of all the sentences that she rejects in s (the *rejection set* of s). We shall assume that both sets Γ and Δ are finite. The intuitive motivation for this requirement is that we are only concerned with finite agents.

Moreover, we assume that there is a set \mathbb{K} of all possible belief states. We may think of the elements of \mathbb{K} as the states that are epistemically possible for a given agent. That $[\Gamma; \Delta] \in \mathbb{K}$ means intuitively that it is possible for the agent to accept all the sentences in Γ and, at the same time, reject all the sentences in Δ . Such a non-empty set \mathbb{K} of belief states, we call a *belief system*.

Next we might ask what requirements it is reasonable to impose on the belief system \mathbb{K} . In other words we are going to formulate requirements that an agent's belief system should satisfy in order to be rational. We have already mentioned two such requirements:

- (R1) The acceptance set Γ and the rejection set Δ of a belief state $[\Gamma; \Delta] \in \mathbb{K}$ should both be finite.
- (R2) If $[\Gamma; \Delta]$ is a belief state in \mathbb{K} , then $\Gamma \cap \Delta = \emptyset$.

We find it natural to impose two additional requirements:

- (R3) If $s = [\Gamma; \Delta] \in \mathbb{K}$ and $\Gamma' \subseteq \Gamma$ and $\Delta' \subseteq \Delta$, then $s' = [\Gamma'; \Delta'] \in \mathbb{K}$. In this case, we may say that s' is a *substate* of s (and s is an *extension* of s'). Thus, any substate of a belief state is a belief state.
- (R4) If $s = [\Gamma; \Delta] \in \mathbb{K}$ and A is any sentence in \mathcal{L} , then either

$$s' = [\Gamma \cup \{A\}; \Delta] \in \mathbb{K}$$
 or

$$s'' = [\Gamma; \Delta \cup \{A\}] \in \mathbb{K}.$$

That is, it should be possible for an agent being in the state $[\Gamma; \Delta]$ to add a sentence either to his acceptance set or to his rejection set.

Finally, we assume that the connectives $\wedge, \vee, \rightarrow, \neg$ satisfy:

- (\wedge) $[A \wedge B, \Gamma; \Delta] \in \mathbb{K}$ iff $[A, B, \Gamma; \Delta] \in \mathbb{K}$
- (\vee) $[\Gamma; A \vee B, \Delta] \in \mathbb{K}$ iff $[\Gamma; A, B, \Delta] \in \mathbb{K}$
- (\rightarrow) $[\Gamma; A \rightarrow B, \Delta] \in \mathbb{K}$ iff $[\Gamma; A, B, \Delta] \in \mathbb{K}$
- (\neg) $[\Gamma, \neg A; \Delta] \in \mathbb{K}$ iff $[\Gamma; A, \Delta] \in \mathbb{K}$

We say that a belief system \mathbb{K} is *coherent* if it satisfies the conditions (R1)–(R4) and (\wedge) , (\vee) , (\rightarrow) , (\neg) above.

Suppose \mathbb{K} is a coherent belief system. We say that Γ \mathbb{K} -entails Δ (in symbols, $\Gamma \vdash_{\mathbb{K}} \Delta$) if and only if $[\Gamma; \Delta] \notin \mathbb{K}$. Intuitively, Γ \mathbb{K} -entails Δ iff there is no belief state s in \mathbb{K} such that Γ is the acceptance set of s and Δ is the rejection set of s . It is easily seen that \mathbb{K} -entailment is a Scott-type (multiple-conclusion) *consequence relation*, i.e., it satisfies the following conditions:¹⁷

- (R) If $\Gamma \cap \Delta \neq \emptyset$, then $\Gamma \vdash_{\mathbb{K}} \Delta$.
- (M) If $\Gamma \vdash_{\mathbb{K}} \Delta$, $\Gamma \subseteq \Gamma'$, $\Delta \subseteq \Delta'$, then $\Gamma' \vdash_{\mathbb{K}} \Delta'$.
- (C) If $\Gamma, A \vdash_{\mathbb{K}} \Delta$ and $\Gamma \vdash_{\mathbb{K}} \Delta, A$, then $\Gamma \vdash_{\mathbb{K}} \Delta$.

Moreover, we have for the connectives:

$$\begin{aligned}
 (\wedge) \quad & \frac{\Gamma, A, B \vdash_{\mathbb{K}} \Delta}{\Gamma, A \wedge B \vdash_{\mathbb{K}} \Delta} \\
 (\vee) \quad & \frac{\Gamma \vdash_{\mathbb{K}} A, B, \Delta}{\Gamma \vdash_{\mathbb{K}} A \vee B, \Delta} \\
 (\rightarrow) \quad & \frac{\Gamma, A \vdash_{\mathbb{K}} B, \Delta}{\Gamma \vdash_{\mathbb{K}} A \rightarrow B, \Delta} \\
 (\neg) \quad & \frac{\Gamma \vdash_{\mathbb{K}} A, \Delta}{\Gamma, \neg A \vdash_{\mathbb{K}} \Delta}
 \end{aligned}$$

Each of these inference rules works both ways: the assertions above and below the double line are said to be equivalent.

We have seen that every coherent belief system \mathbb{K} gives rise to a consequence relation satisfying the usual laws of classical propositional logic. It is a well-known result of Scott [30, 31] that every consequence relation has a two-valued truth-value semantics. Let us outline this result.

We first introduce the notions of a valuation and a semantics. A *valuation* v is a function which assigns to any sentence A one of the truth values 0 (false) or 1 (true). A valuation v is *Boolean* if it satisfies the requirements

$$\begin{aligned}
 v(A \wedge B) = 1 & \quad \text{iff} \quad v(A) = v(B) = 1, \\
 v(A \vee B) = 0 & \quad \text{iff} \quad v(A) = v(B) = 0, \\
 v(A \rightarrow B) = 0 & \quad \text{iff} \quad v(A) = 1 \text{ and } v(B) = 0, \\
 v(\neg A) = 1 & \quad \text{iff} \quad v(A) = 0.
 \end{aligned}$$

A (two-valued) *semantics* is a non-empty set of valuations. A semantics is Boolean if and only if each of its members is Boolean.

¹⁷This notion was introduced by Dana Scott [30, 31] as a generalisation of Tarski's notion of a *consequence operation*.

Theorem 1 (Scott). Let \vdash be a consequence relation. Then there exists a semantics V such that for all finite Γ, Δ : $\Gamma \vdash \Delta$ if and only if for all $v \in V$, if $v(A) = 1$ for all $A \in \Gamma$, then $v(B) = 1$ for some $B \in \Delta$. Moreover, V is Boolean if \vdash is Boolean, i.e., satisfies the conditions $(\wedge), (\vee), (\rightarrow), (\neg)$.

Hence any coherent belief system \mathbb{K} determines a Boolean consequence relation $\vdash_{\mathbb{K}}$. Moreover, there exists a Boolean semantics V such that $\vdash_{\mathbb{K}}$ is the consequence relation determined by V .

Now we might define: $\Gamma \vdash \Delta$ (Δ is a *classical consequence* of Γ) if and only if for every Boolean valuation v , whenever v assigns the value 1 (true) to all sentences in Γ , then v also assigns 1 to some sentence in Δ . Then we have:

Theorem 2. $\Gamma \vdash \Delta$ if and only if $\Gamma \vdash_{\mathbb{K}} \Delta$ for every coherent belief system \mathbb{K} .

In other words, the epistemic semantics presented here gives the same propositional logic as classical two-valued semantics.¹⁸

The above justification of classical logic in terms of the epistemic notions of acceptance and rejection can be generalised to the case of predicate logic.¹⁹

Tor Sandqvist has pointed out that the approach presented here only requires of a rational agent that she be coherent, namely that she not accept a set Γ of sentences while at the same time rejecting any set Δ which is a logical consequence of Γ . However, it is not required that the agent accept sentences that are logical consequences of sentences that she accept. Hence, the agent need not be *logically omniscient*. This is as it should be, we think. Sandqvist, however, also points out that an agent that is rational in our sense is not even required to accept known consequences of premises that she accept (*closure under known consequences*). This Sandqvist regards as unintuitive. Arguably, an agent with a coherent belief system who has a bivalent notion of truth, ought to satisfy closure under known consequences. Assume namely that the agent knows that A logically implies B . Then she can reason as follows: I know that A logically implies B . Hence, I know that A cannot be true unless B also be true. Thus, accepting the truth of A commits me to accepting the truth of B .

However, it is not part of the approach described here that the agent be in possession of a notion of truth—let alone a bivalent one. That she does is consistent with the approach, but not entailed by it. In view of Theorem 2, however, any coherent agent may be interpreted as having a bivalent notion of truth. Moreover, for an agent having such a notion, it seems reasonable to require that she satisfies closure under known logical consequences.

¹⁸Notice that the justification we have given here is not acceptable to an intuitionist, since we have freely made use of classical reasoning in the metalanguage. In his dissertation [29], Tor Sandqvist has given an account, based on inferentialist ideas, of the meanings of the standard logical operators of predicate logic that validates classical logic and does not presuppose a recognition-transcendent notion of truth. Sandqvist's approach only makes use of constructively valid reasoning in the metalanguage.

¹⁹Cf. Gabbay [12, Chapter 0].

5 The paradox of inference

Our main concern so far has been to justify deduction by showing that it is a valid form of reasoning. Given that such a justification is successful, it could be interpreted as saying that the conclusion of a deductive argument in some sense is already contained in the premises. However, then it seems hard to explain how deductive inference can yield “new” knowledge. If the conclusion is already contained in the premises, one is tempted to infer that someone who knows the premises of a valid argument also knows the conclusion. But do we really know all the logical consequences of what we know?²⁰

The answer to this question lies in the distinction between *logical consequence* and *logical inference*. Logical consequence is an objective relation between statements. A may be a logical consequence of B without us knowing that this is the case. Logical inference, on the other hand, is a process by which we gain knowledge that some statement is in fact a logical consequence of some premises.

Still there may be a problem whether someone, after having gone through a complicated mathematical proof, really can be said to know the conclusion of the proof. Perhaps one could point to inference being a *reliable process* of gaining new truths. Hence, on a reliabilist conception of knowledge, someone who has actually proved a theorem from known premises has arrived at the conclusion by a reliable process, and can therefore be said to have knowledge of it.

6 Conclusion

If the explanation of the validity of the laws of classical logic given in section 4.3 in terms of the epistemic attitudes of acceptance and rejection is correct, then we seem to have been able to justify our actual deductive practice in a way that does not presuppose a primitive notion of truth or questionable metaphysical assumptions. We could still have access to a notion of truth. This notion could then be explained in a deflationary manner by the T-schema:

(T) ‘ A ’ is true iff A ,

suitably restricted to avoid semantic paradoxes; or defined à la Tarski [36] or Kripke [18].

Acknowledgements

We are grateful to Tor Sandqvist for excellent advice and helpful discussions. To one of us, Joseph Almog has been a source of inspiration and knowledge for years. His influence is especially strong in section 3. An earlier version of this paper was discussed at the “Friday Seminar” at Umeå University and the Logic and Language

²⁰Cf. Sundholm [35] for a discussion of this so-called paradox of inference.

Seminar at Stockholm University. We thank the participants of these seminars for criticisms and valuable suggestions. We are especially grateful to Kathrin Glüer-Pagin, Per Martin-Löf, Peter Pagin, Dag Prawitz, Levi Spectre, Fredrik Stjernberg, Bertil Strömberg, and Pär Sundström. Lindström’s work was supported by The Bank of Sweden Tercentenary Foundation via the research project “The Ontology and Epistemology of Mathematics”.

References

- [1] Belnap, N. D., ‘Tonk, plonk and plink’, *Analysis* **22** (1962): 130–134. Reprinted in Strawson, P. F. (ed.), *Philosophical Logic*, Oxford University Press, 1967: 132–137.
- [2] Boghossian, P., ‘Knowledge of logic’, in Boghossian, P. and Peacocke, C. (eds.), *New Essays on the A Priori*, Clarendon Press, Oxford, 2000: 229–254.
- [3] Boghossian, P., ‘How are objective epistemic reasons possible?’, *Philosophical Studies* **106** (2001): 1–40.
- [4] Carnap, R., *Meaning and Necessity. A Study in Semantics and Modal Logic*, Chicago: University of Chicago Press, 1947. Second edition with supplements, 1956.
- [5] Carroll, L., ‘What the tortoise said to Achilles’, *Mind* **4** (1895): 278–280.
- [6] Davidson, D., *Inquiries into Truth and Interpretation*, Oxford University Press, 1984.
- [7] Dummett, M., *Truth and Other Enigmas*, Harvard University Press, 1978.
- [8] Dummett, M., ‘The philosophical basis of intuitionistic logic’, in [7]: 215–247.
- [9] Dummett, M., ‘The justification of deduction’, in [7]: 290–318.
- [10] Dummett, M., ‘What is a theory of meaning? (II)’, in *The Seas of Language*, Clarendon Press, 1993: 34–93.
- [11] Ellis, B., *Rational Belief Systems*, Blackwell, 1979.
- [12] Gabbay, D., *Investigations in Modal and Tense Logics with Applications to Problems in Philosophy and Linguistics*, Reidel, 1976.
- [13] Gentzen, G. ‘Untersuchungen über das logische Schliessen’, *Mathematische Zeitschrift* **39** (1934): 176–210, 405–431. Translated in Szabo, M. E. (ed.) *The Collected Papers of Gerhard Gentzen*, North-Holland, 1969.
- [14] Haack, S., ‘The justification of deduction’, *Mind* **85** (1976): 112–119.
- [15] Haack, S., ‘Dummett’s justification of deduction’, *Mind* **91** (1982): 216–239.
- [16] Kreisel, G., ‘Foundations of intuitionistic logic’, in Nagel, E., Suppes, P., and Tarski, A. (eds.) *Logic, Methodology and Philosophy of Science, Proceedings of the 1960 International Congress*, Stanford University Press, (1962): 198–210.

- [17] Kreisel, G., ‘Informal rigour and completeness proofs’, in Lakatos, I. (ed.) *Problems in the philosophy of mathematics*, North-Holland, 1967: 138–186.
- [18] Kripke, S., ‘Outline of a theory of truth’, *Journal of Philosophy* **72** (1975): 690–716.
- [19] McGee, V., ‘Two problems with Tarski’s theory of consequence’, *Proceedings of the Aristotelian Society* **92** (1992): 273–292.
- [20] Miller, A., ‘What is the manifestation argument?’, *Pacific Philosophical Quarterly* **83** (2002): 352–383.
- [21] Pagin, P., ‘Truth, proofs and compositionality’, under review.
- [22] Prawitz, D., *Natural Deduction: A Proof-Theoretical Study*, Stockholm Studies in Philosophy 3, Almqvist & Wiksell, 1965.
- [23] Prawitz, D., ‘Ideas and results in proof theory’, in J. E. Fenstad (ed.) *Proceedings of the Second Scandinavian Logic Symposium*, North-Holland, 1971: 235–308.
- [24] Prawitz, D., ‘Meaning and proofs: on the conflict between classical and intuitionistic logic’, *Theoria* **43** (1977): 2–40.
- [25] Prawitz, D. ‘Logical consequence from a constructivist point of view’, in [33]: 671–695.
- [26] Prior, A. N., ‘The runabout inference ticket’, *Analysis*, **21** (1960): 38–39. Reprinted in Strawson, P. F. (ed.), *Philosophical Logic*, Oxford University Press, 1967: 129–131.
- [27] Quine, W. V., *Philosophy of Logic* (Second edition), Harvard University Press, 1986.
- [28] Restall, G., ‘Multiple Conclusions’, in Hajek, P., Valdes-Villanueva, L. and Westerståhl, D. (eds.), *Logic, Methodology and Philosophy of Science: Proceedings of the Twelfth International Congress*, Kings College Publications, 2005: 189–205.
- [29] Sandqvist, T., *An Inferentialist Interpretation of Classical Logic*, Doctoral dissertation, Uppsala Prints and Preprints in Philosophy, 2005:1.
- [30] Scott, D., ‘On engendering an illusion of understanding’, *Journal of Philosophy* **68** (1971): 787–807.
- [31] Scott, D., ‘Completeness and axiomatizability in many-valued logic’, in L. Henkin (ed.), *Proceedings of the Tarski Symposium, Proceedings of Symposia in Pure Mathematics* **25**, American Mathematical Society, Providence, 1974: 412–435.
- [32] Shapiro, S., *Foundations without Foundationalism: A Case for Second-Order Logic*, Oxford University Press, 1991.
- [33] Shapiro, S. (ed.), *The Oxford Handbook of Philosophy of Mathematics and Logic*, Oxford University Press, 2005.
- [34] Shapiro, S., ‘Logical consequence, proof theory, and model theory’, in [33]: 651–670.

- [35] Sundholm, 'Proof theory and meaning', in D. Gabbay and F. Guentner (eds.) *Handbook of Philosophical Logic*, **III**, D. Reidel, 1986: 471–506.
- [36] Tarski, A., *The Concept of Truth in Formalized Languages*, in Corcoran, J., (ed.) *Logic, Semantics, Metamathematics* (second edition), Hackett Publishing Company, 1983: 152–278.
- [37] Tennant, N., 'Rule-circularity and the justification of deduction', *The Philosophical Quarterly* **55** (2005): 625–648.