Proof-Theoretic Validity

Stephen Read
University of St Andrews
email: slr@st-and.ac.uk

April 16, 2012

Abstract

The idea of proof-theoretic validity originated in the work of Gerhard Gentzen, when he suggested that the meaning of each logical expression was encapsulated in its introduction-rules, and that the elimination-rules were justified by the meaning so defined. It was developed by Dag Prawitz in a series of articles in the early 1970s, and by Michael Dummett in his William James lectures of 1976, later published as *The Logical Basis of Metaphysics*. The idea had been attacked in 1960 by Arthur Prior under the sobriquet ‘analytic validity’. Logical truths and logical consequences are deemed analytically valid by virtue of following, in a way which the present paper clarifies, from the meaning of the logical constants. But different logics are based on different rules, confer different meanings and so validate different theorems and consequences, some of which are arguably not true or valid at all. It seems to follow that some analytic statements are in fact false. The moral is that we must be careful what rules we adopt and what meanings we use our rules to determine.

Keywords: harmony, inferentialism, analytic, valid, introduction-rules, elimination-rules, negation, tonk; Gentzen, Prior, Prawitz, Dummett.

1 Analyticity

Could analytic statements be false? An analytic statement, or analytic truth, is often characterized as one which is true in virtue of the meaning of the words. So it seems to follow that analytic statements are *ipso facto* true. But must they be true? That question is many-ways ambiguous. To begin

---

*This work is supported by Research Grant AH/F018398/1 (Foundations of Logical Consequence) from the Arts and Humanities Research Council, UK.*
with, there are, of course, analytic falsehoods, statements which are false simply in virtue of their meaning. But that aside, analytic truths are often thought to be necessarily true. Yet even necessary truths could have been false, on some accounts of necessity. In S4, what is necessary, is necessarily so; but in other modal systems, like T, that does not follow, and some necessary truths might not have been necessary. In this sense, the question is whether analytic truths, though in fact true, might have been false.

But these are not the issue I want to address. For the question whether analytic statements must be true has a further sense. Consider Tarski’s discussion of semantically closed languages, for example. Tarski (1956a, p. 164-5) claimed that natural languages are universal and semantically closed, and that semantic closure leads to inconsistency, that is, that natural language is inconsistent. This seems at first rather puzzling. How can a language, a set of sentences or statements, be inconsistent? Is it not theories, based on languages, that are consistent or inconsistent? The fact which Tarski was invoking is that natural languages are interpreted languages—they consist not just in a set of sentences, but of interpreted, meaningful sentences.1 In virtue of that meaning, some of those sentences may force themselves on us, for example, the Liar sentence, ‘This sentence is not true’ (call it \( L \)). Why is \( L \) true in virtue of its meaning, that is, of what it says? Because we can prove it true. Suppose \( L \) were not true. Then things would be as it says they are. So it would be true. Hence, by reductio (if it weren’t true it would be true), \( L \) is true. But if \( L \) is true then \( L \) is also not true (since that is what it says), and so the meanings of certain terms in natural language, in particular, the word ‘true’, commit their users to assent to contradictory sentences. Yet the meaning of ‘not’ also commits its users to refusing to assent to contradictory sentences. Hence, simply in virtue of the meanings of words, users of such a language are subject to contradictory and inconsistent demands. The language itself is inconsistent. In this sense, certain statements may, although analytic, be false. ‘\( L \) is true’ and ‘\( L \) is not true’ are both analytic, but they cannot both be true, so one of them is false. Analytic statements can be false.

Boghossian (1997, p. 334) makes a useful distinction between metaphysical and epistemic analyticity:

“On [one] understanding, ‘analyticity’ is an overtly epistemological notion: a statement is ‘true by virtue of its meaning’ provided that grasp of its meaning alone suffices for justified belief in its truth. Another, far more metaphysical, reading of the phrase ‘true by virtue of meaning’ is also available, however, according to which a statement is analytic provided that, in some approxi-

---

1Tarski (1956a, p. 166). Note that even his formalized languages are interpreted (loc.cit.): “we are not interested here in ‘formal’ languages and sciences in one special sense of the word ‘formal’, namely sciences to the signs and expressions of which no material sense is attached.”
priate sense, it *owes its truth-value completely to its meaning*, and not at all to ‘the facts’.”

Boghossian believes that the metaphysical conception was rightly undermined by Quine’s arguments, but that the epistemic conception, grounding a notion of the *a priori*, survives them. Nonetheless, meaning does not confer knowledge, since some statements to which the speaker is committed *a priori* are false. Indeed, Boghossian says only that knowledge of meaning leads to a justified belief in the resulting statements. For without the metaphysical underpinning, that belief and its justification cannot make the statement true. Some analytic statements are false.

But if analytic commitments are not guaranteed to be true, wherein lies their truth when they are true? In particular, what makes logical truths true, when they are true, if not the meaning of the logical constants? We must return to this question in § 4.

Azzouni (2007) and Patterson (2007) have inferred from Tarski’s observation that natural language is inconsistent that it is in fact meaningless, if not trivial. Eklund (2002) and Scharp (2007), however, have drawn the more reasonable conclusion that inconsistent languages can be meaningful and useable, but stand in need of revision. The constitutive principles on which they are based, in particular, those for truth, pick out, for Scharp, inconsistent concepts which need further elucidation, like the pre-relativistic concept of mass. For Eklund, semantic competence “exerts pull” (as he puts it): speakers’ semantic competence disposes them to accept the untrue premises or invalid steps in arguments concerning inconsistent concepts. To reject those premises or inferences would indicate lack of competence with the concepts concerned.

Scharp cites another example, Dummett’s discussion of the term ‘Boche’. Dummett (1973, p. 454) claims that simply using this term in a language commits its practitioners to inferring that Germans are barbaric and prone to cruelty. Simply as a consequence of using a word with a certain meaning, one is committed, analytically, to assert falsehoods. That someone is German warrants the claim that they are a Boche, and that they are Boche implies that they are barbaric. It follows that Germans are barbaric, by the very meaning rules of the term ‘Boche’. But that is false. So analytic statements can be false.

This doctrine may sound reminiscent of dialetheism, the doctrine that there are true contradictions. But it adds a further twist to dialetheism. From the fact that the meanings of certain expressions have the consequence that they commit their users to assert to contradictory statements, dialetheists infer that those contradictions are true. In contrast, the position I am exploring reverses this inference. Anyone who uses such expressions would indeed be committed to asserting contradictory statements. But since con-
Tradictory statements cannot both be true, it is clearly a mistake to use such expressions. Indeed, this was surely Tarski’s response. What he effectively claimed was that both ‘\(L\) is true’ and ‘\(L\) is not true’, as formulated in a natural language, are analytic; that is, users of such languages are committed to assent to them both. That is precisely the respect in which natural languages are, in Tarski’s eyes, inconsistent. Consequently, Tarski’s response was to “abandon . . . the language of everyday life and restrict [himself] entirely to formalized languages.”\(^2\)

There are three positions one might adopt about a concept such as ‘true’:

1. that of Eklund and others (perhaps including Tarski): that the semantic paradoxes show that the concept encapsulated in natural languages is inconsistent and needs to be revised;

2. that of the dialetheists: that the concept we have is indeed inconsistent, but we must accept this fact and its consequences; and

3. my own position: that the paradoxes reveal that our theory of truth is mistaken, and more work is needed to find out how our concept of truth really works.

So too, in general, adopting certain logical conceptions would commit us to analytic consequences some of which are mistaken and unacceptable. Only the right conceptions of conjunction, implication and so on yields the right account of analyticity. If we are not careful, our analysis may over-generate,\(^3\) declaring invalid arguments to be valid.

2 Analytic Validity

We see, then that adopting a certain language may not only commit a speaker to believing certain statements, it may also commit a speaker to making certain inferences. Prior in his famous ‘Runabout Inference Ticket’ (Prior, 1960) described such a commitment as “analytic validity”. He didn’t like it, and argued that validity must instead be based on truth-preservation, not on meaning. Others claimed instead that the flaw lay in supposing that arbitrary rules can confer meaning on the relevant expressions. Arguably, ‘tonk’ had not been given a coherent meaning by Prior’s rules. Rather, whatever meaning tonk-introduction had conferred on the neologism ‘tonk’ was then contradicted by Prior’s tonk-elimination rule. But if rules were set down for a term which did properly confer meaning on it, then certain inferences would be “analytic” in virtue of that meaning. The question is:

\(^2\)Tarski (1956a, p. 165).

\(^3\)In Etchemendy’s phrase: Etchemendy (1990, p. 8).
what constraints must rules satisfy in order to confer a coherent meaning on the terms involved?

Dummett introduced the term ‘harmony’ for this constraint: in order for the rules to confer meaning on a term, two aspects of its use must be in harmony. Those two aspects are the grounds for an assertion as opposed to the consequences we are entitled to draw from such an assertion. Those whom Prior was criticising, Dummett claimed, committed the “error” of failing to appreciate “the interplay between the different aspects of ‘use’, and the requirement of harmony between them.”⁴ All aspects of an expression’s use reflect its meaning. But when they are in harmony, each of the aspects contributes in the same way.

Dummett is here following out an idea of Gentzen’s, in a famous and much-quoted passage where he says that “the E-inferences are, through certain conditions, unique consequences of the respective I-inferences.”⁵ The idea was that the meaning encapsulated in the introduction- or I-rule serves to justify the conclusions drawn by the elimination- or E-rule. But in fact, Dummett has loftier ambitions than this. He introduces the idea of the proof-theoretic justification of logical laws.⁶ He distinguishes three grades of proof-theoretic justification. The first is familiar, namely, the notion of a derived rule, that one can reduce a complex inference, and so justify it, by articulating it into a succession of simpler, more immediate, inference-steps. Then the simpler steps of immediate inference justify and validate the indirect consequence.

The second grade of proof-theoretic validity rests on another familiar notion, namely, that of an inference being admissible. We may not be able to derive the conclusion from the premise or premises, but we show that if the premises are derivable, so is the conclusion. The classic example of such a justification is cut-elimination, established by Gentzen for his sequent systems LJ and LK in his famous Hauptsatz. As a theorem about the elimination of Cut, this result takes a calculus which includes the Cut principle:

\[
\Gamma \Rightarrow \Theta, D, D, \Delta \Rightarrow \Lambda \\
\Gamma, \Delta \Rightarrow \Theta, \Lambda
\]

Cut

and shows that any derivation can be transformed into a derivation of the same end-sequent in which no application of the Cut rule is made. As a result about the admissibility of Cut, in contrast, one takes a calculus which does not include the Cut rule and shows that if the premises of the Cut rule are derivable in the system, then so is the conclusion.

⁴Dummett (1973, p. 396). Cf. Dummett (1978, p. 221): “If the linguistic system as a whole is to be coherent, there must be a harmony between these two aspects.” See also Dummett (1991, p. 215). He also put constraints on the introduction-rules. See §4 below.
⁵Gentzen (1969, p. 81). He also put constraints on the introduction-rules. See §4 below.
like Cut, which reflects the difference between axiomatic and sequent calculi on the one hand and natural deduction systems on the other. What is purportedly “natural” about natural deduction systems is that they work from assumptions, as Gentzen claimed was the natural method of reasoning in mathematics.\footnote{Gentzen (1969, p. 74 ff.). Cf. Gentzen (1932, p. 2 ff.).} Sequent calculi, by contrast, consist of apodeictic assertions, starting with axioms (aka basic sequents) and such that the end-sequent of a derivation is asserted categorically. Thus, whereas natural deduction derivations may be open, that is, contain open assumptions, sequent calculus derivations are always closed. The premises of a Cut inference can be interpreted as saying that their succedent-formulae follow from their antecedent-formulae as assumptions (given the tight connection between sequent calculi and the corresponding natural deduction system), but they say this categorically. In general, an admissible inference is one whose conclusion is provable (from closed assumptions) when its premises are provable. This will be generalized to the case of open assumptions in Dummett’s third grade of proof-theoretic justification, but first let us concentrate on the second grade.

When transferring the concept of an admissible inference from sequent calculus (or axiomatic systems) to the natural deduction case, caution is necessary; and at its heart is the notorious Fundamental Assumption. The Fundamental Assumption makes its first appearance in Dummett (1991, p. 252) in this form:

“Whenever we are justified in asserting [a] statement, we could have arrived at our entitlement to do so by [direct or canonical] means.”

Dummett attributes the Fundamental Assumption to Prawitz in this form:

“Prawitz expressly assumes that, if a statement whose principal operator is one of the logical constants in question can be established at all, it can be established by an argument ending with one of the stipulated introduction rules.” (\textit{loc.cit.})

The reference is to a series of articles on the “foundations of a general proof theory” which Prawitz published in the early 1970s. (Recall that Dummett’s book, though not published until 1991, had been originally given as the William James lectures at Harvard in 1976.) Prawitz’ idea was to find a characterization of validity of argument independent of model theory, as typified by Tarski’s account of logical consequence. Whereas Tarski’s analysis deems an inference valid when it preserves truth under all re-interpretations of the non-logical vocabulary, Prawitz, following Gentzen’s lead in the passage cited earlier, accounts an argument or derivation valid by virtue of the meaning or definition of the logical constants encapsulated...
in the introduction-rules. The idea is this: we take the introduction-rules as given. Then any argument (or in the general case, below, any argument-schema) is valid if there is a “justifying operation” (Prawitz, 1973, p. 233) ultimately reducing the argument to the application of introduction-rules to atomic sentences:

“The main idea is this: while the introduction inferences represent the form of proofs of compound formulas by the very meaning of the logical constants when constructively understood and hence preserve validity, other inferences have to be justified by the evidence of operations of a certain kind.”\(^8\)

The operations are essentially the reduction steps in the proof of normalization, and depend on the Fundamental Assumption. An argument is valid if either it reduces to a non-logical justification of an atomic sentence, or it reduces to an argument whose last inference is an introduction inference and whose immediate subarguments are valid.\(^9\)

The connection between the second grade of justification and the reduction to an atomic base is usefully spelled out by Humberstone (2010, §4.13). Take a natural deduction system in which no rule discharges assumptions (or closes open variables), e.g., with the sole connectives \(\land\) and \(\lor\). Say that an argument from \(\Gamma\) to \(A\) is proof-theoretically valid if for any atomic base \(\Pi\) (i.e., set of propositional variables, justified by what Dummett (1991, p. 204) calls “boundary rules”) for which, for each \(B \in \Gamma\), \(B\) can be derived from \(\Pi\) by I-rules alone, so can \(A\). Then (Humberstone, 2010, Theorem 4.13.3) “all and only the provable sequents of [the] natural deduction system for \(\land\) and \(\lor\) are proof-theoretically valid.” That is, the E-rules extend what can be proved by the I-rules alone only to infer what can be “introductively derived” from an atomic base of closed assumptions if their premises can be so derived.

Following Prawitz, Dummett’s third grade of proof-theoretic justification generalizes this account from closed arguments to open arguments, that is, argument-schemata.\(^10\) Argument-schemata either have open assumptions which will later be closed by rules which discharge those assumptions, or free variables which will later be the eigenvariables of a quantifier rule. In Prawitz’ discussion, these two cases are exemplified just by \(\rightarrow I\) and \(\forall I\). Such an argument-schema is valid if, besides the earlier cases, the result of closing the assumption (by a closed proof) or the eigenvariable (by substitution of

---

\(^8\)Prawitz (1973, p. 234). Prawitz (nd, p. 21) justifies the I-rules as follows: “To make an inference is to apply a certain operation to the given grounds for the premisses, and that the inference is valid is now defined just to mean that the result of applying this operation to grounds for the premisses is a ground for the conclusion, and hence it justifies the person in question in holding the conclusion true.”


\(^10\)Prawitz (2006, p. 511) calls them “argument skeletons”.

7
Let us consider a couple of examples, in this section, of implication, and in § 3, of negation. First, →I:

\[
\begin{array}{c}
[A] \\
\vdots \\
B \\
A \rightarrow B \rightarrow I
\end{array}
\]

inferring (an assertion of the form) \( A \rightarrow B \) from (a derivation of) \( B \), permitting the optional discharge of (several occurrences of) \( A \). But \( A \) and \( B \) may themselves contain occurrences of ‘→’, so we need not only to infer (assert) ‘→’-wffs, but also, as Dummett noted, to draw consequences from assertions of the form \( A \rightarrow B \). Consequently, we need a further rule of →E, which will tell us when we may do so. Accordingly, all further inferences should reduce, by the first grade of proof-theoretic justification, to application of →I and →E. The crucial task now is to identify the form of →E which is justified by →I by the second and third grades of proof-theoretic justification, so that all inference involving ‘→’ reduces to and is justified by the meaning conferred on ‘→’ by →I.

Whatever form →E has, there must be the appropriate justificatory operation of which Prawitz spoke. That is, we should be able to infer from an assertion of \( A \rightarrow B \) no more (and no less) than we could infer from whatever warranted assertion of \( A \rightarrow B \). We can represent this as follows:\(^{11}\)

\[
\begin{array}{c}
[A] \\
\vdots \\
B \\
A \rightarrow B \\
\vdots \\
C \\
C \rightarrow E
\end{array}
\]

That is, if we can infer \( C \) from assuming the existence of a derivation of \( B \) from \( A \), we can infer \( C \) from \( A \rightarrow B \). Then by the Fundamental Assumption, assuring us that there is a derivation \( \alpha \) of \( A \rightarrow B \) which terminates in an application of →I, we can apply the reduction operation:

\[
\begin{array}{c}
[A] \\
\vdots \\
B \\
A \rightarrow B \rightarrow I \\
\vdots \\
C \\
\alpha \rightarrow E
\end{array}
\]

reduces to

\[
\begin{array}{c}
A \\
\vdots \\
B \\
A \rightarrow B \\
\vdots \\
C \\
\alpha \rightarrow E
\end{array}
\]

inferring \( C \) directly from the grounds (the derivation \( \alpha \) of \( B \) from \( A \)) for asserting \( A \rightarrow B \). Here we have replaced the assumption of a derivation of

$B$ from $A$, marked by $\left[ \begin{array}{c} A \\ B \end{array} \right]$, with the actual derivation $\alpha$ of $B$ from $A$, written $\left[ \begin{array}{c} A \\ \alpha \\ B \end{array} \right]$.\(^{12}\)

So much for the general theory. What, however, is the rule $\to E$ which has here been justified? In particular, what does the minor premise actually say? It speaks of whatever ($C$) can be inferred from the assumption that $B$ can be inferred from $A$. In other words, if we had a proof of $A$, we could infer $B$ (and continue to infer whatever, $C$, from $B$):\(^{13}\)

\[
\begin{array}{c}
[B] \\
\vdots \\
A \to B \quad A \\
C
\end{array}
\]

Other things being equal,\(^{14}\) we can permute the derivation of $C$ from $B$ with the application of the elimination-rule, to obtain the familiar rule of Modus Ponendo Ponens (MPP):

\[
\begin{array}{c}
A \to B \\
B \\
\vdots \\
A_{\text{MPP}} \\
C
\end{array}
\]

Following Francez and Dyckhoff (2011), I dub the general procedure by which we obtain the $\to E$-rule from the I-rule, “general-elimination harmony”. It is the inverse of Prawitz’ reduction operation. The crucial idea is distilled by Negri and von Plato (2001, p. 6):

“**Inversion Principle**: Whatever follows from the direct grounds for deriving a proposition must follow from that proposition.”

Given a set of introduction-rules for a connective (in general, there may be several, as in the familiar case of ‘$\lor$’), the elimination-rules (again, there may be several, as in the case of ‘$\land$’) which are justified by the meaning so conferred are those which will permit a reduction operation of Prawitz’ kind. Each E-rule is harmoniously justified by satisfying the constraint that whenever its premises are provable (by application of one of the I-rules, by the Fundamental Assumption), the conclusion is derivable (by use of the assertion-conditions framed in the I-rule).

\(^{12}\)This notation is inspired by, but extends further, that introduced in Gentzen (1932), in passages which did not appear in the published version Gentzen (1935) (translated in Gentzen (1969)). See also von Plato (2008).

\(^{13}\)See Dyckhoff (1988) and Negri and von Plato (2001, p. 8).

\(^{14}\)In a particular case, we need to check that the premutation preserves the conditions on being a proof. For an example where that condition may fail, see § 4.
Elsewhere (Read, 2010), I have discussed the application of the general-elimination procedure to ‘∧’ and ‘∨’. These differ from the rules for ‘→’ in three ways: no assumptions are discharged, ∧I has two premises, and ∨I has two cases. The general conclusion reached was that, where there are m I-rules each with n_i premises (0 ≤ i ≤ m), there will be \( \prod_{i=1}^{m} n_i \) E-rules.

Take a formula $\vec{A}$ with main connective ‘$’ and immediate subformulae $\vec{A}$. Then the E-rules have the form:

$$
\begin{array}{c}
[\pi_{1j_1}] \\
\vdots \\
[\pi_{mj_m}]
\end{array}
\quad
\begin{array}{c}
\vec{A} \\
\vdots \\
\vec{C}
\end{array}
\xrightarrow{\text{SE}}
\begin{array}{c}
C \\
\vdots \\
C
\end{array}
$$

where each minor premise derives \(C\) from one of the grounds, \(\pi_{ij}\), for asserting $\vec{A}$. \(\pi_{ij}\) may be a wff (as in ∧I), or a derivation of a wff from certain assumptions (as in →I). Then the reduction operation permits one to infer \(C\) from one of the grounds for assertion of $\vec{A}$ whenever one can infer \(C\) from $\vec{A}$ itself:

$$
\begin{array}{c}
\alpha_1 \\
\vdots \\
\alpha_n
\end{array}
\quad
\begin{array}{c}
[\pi_{1j_1}] \\
\vdots \\
[\pi_{mj_m}]
\end{array}
\quad
\begin{array}{c}
\vec{A} \\
\vdots \\
\vec{C}
\end{array}
\xrightarrow{\text{SI}}
\begin{array}{c}
\beta_1 \\
\vdots \\
\beta_m
\end{array}
\quad
\begin{array}{c}
C \\
\vdots \\
C
\end{array}
\xrightarrow{\text{SE}}
\begin{array}{c}
\alpha_{ij} \\
\vdots \\
\alpha_{ij}
\end{array}
\quad
\begin{array}{c}
\pi_{ij_i} \\
\vdots \\
\pi_{ij_i}
\end{array}
\quad
\begin{array}{c}
\beta_i \\
\vdots \\
\beta_i
\end{array}
\quad
\begin{array}{c}
C \\
\vdots \\
C
\end{array}
$$

Having one minor premise in each E-rule drawn from among the premises for each I-rule ensures that, whichever I-rule justified assertion of $\vec{A}$ (here it was the \(i\)-th), one of its premises can be paired with one of the minor premises to carry out the reduction.\(^{15}\)

### 3 The Fundamental Assumption

Before we turn to the second example, we need to give further consideration to the Fundamental Assumption. This is the claim that if there is a proof of $\vec{A}$, there is a proof of $\vec{A}$ in which the final step is an application of SI. In the case of open arguments, this is false. Given a derivation of $\vec{A}$ by means of an open argument (i.e., an argument-schema), with open assumptions or free variables, there is no such guarantee. However, proof-theoretic justification of the third grade reduces to the second grade by closing those open assumptions and free variables. This is reflected in Prawitz’ theorem on the form of normal deductions.\(^{16}\) Define a thread as a sequence of successive formulae in a derivation running from an assumption to the end-formula,
and let a *branch* be an initial segment of a thread which ends in the minor premise of an application of an E-rule (or in the end-formula if there is no such minor premise). Prawitz’ result shows that, given a proof, there is a proof in which each branch is divided into two parts, an E-part and an I-part, separated by a minimum formula. Each formula in the E-part is major premise of an E-rule, and each formula in the I-part (except the last) is premise of an I-rule. Reduction of proof-theoretic justification of the third to the second grade depends on this result. Moreover, this result itself depends on the existence of what Dummett (1977, p. 112) calls “permutative reductions”.

The proof of normalization consists in the removal of maximum formulae from derivations. A maximum formula is a formula which is both the conclusion of an I-rule and major premise of an E-rule, as was $\vec{A}$ above. Dummett, in a graphic expression, calls such a part of a derivation a “local peak” (Dummett, 1991, p. 248), so that normalization is the “levelling of local peaks”. However, before Prawitz’ “justifying operations” can be applied to local peaks, a preliminary permutative reduction is needed:

“The other reduction steps are auxiliary, being principally concerned to rearrange the order in which the rules are applied, so that a proof in which a sentence is introduced by an introduction rule, and only later removed by means of an elimination rule in which it is the major premiss, can be transformed into one in which the elimination rule is applied immediately after the introduction rule to form a local peak.” (Dummett, 1991, p. 250)

What we have in general is called by Prawitz a “maximum segment” (Prawitz, 1965, p. 49), what we might graphically describe as a “local plateau”\(^ {17} \)
\( ^{17}\) a succession of occurrences of the same formula as minor premises of applications of E-rules separating its introduction by an I-rule and elimination by the corresponding E-rule. The procedures by which we reduce plateaux to peaks are Dummett’s permutative reductions. They correspond to the reductions of rank in the proof of Cut-elimination. By successively permuting the application of an I-rule for a connective with the E-rule for which its conclusion is a minor premise, we can eventually bring the formula in question face to face with the corresponding application of the E-rule for the same connective. Then what Dummett calls the “fundamental reduction step” which we described at the end of § 2 can be applied to eliminate the maximum formula entirely.

What this shows, however, is that the Fundamental Assumption is more than just an assumption. It requires proof that the rules as a whole are such that each branch has this hour-glass shape, with a minimum formula

\( ^{17}\)Gentzen calls it a “hillock” (Hügel): see von Plato (2008, p. 247).
separating the branch into an E-part and an I-part. That is a holistic matter, depending on the interaction between the various rules and connectives. It cannot be simply assumed, but must be proved.

The other example which deserves exploration is negation. Prawitz makes little mention of negation in his papers on general proof theory. Where it does appear, \( \neg A \) is treated by definition as \( A \rightarrow \bot \), where the “absurdity constant” \( \bot \) is governed solely by an elimination-rule, from \( \bot \) infer anything:

\[
\frac{1}{A} \ \bot \text{E}
\]

As far as I know, every other author follows Prawitz’ lead except Dummett (1977, ch. 4); thereby, the harmony of ‘\( \neg \)’ devolves on the harmony of ‘\( \rightarrow \)’ (and ‘\( \bot \)’).

Gentzen (1969, II 2.21, 5.2) treats negation in two ways in his published paper, as primitive and as defined, but using the absurdity constant \( \bot \) in both cases. In the MS previously mentioned, he treated ‘\( \neg \)’ as primitive without appeal to \( \bot \). As introduction-rule, he took \textit{reductio ad absurdum} (dubbed by him ‘R’) in this form:

\[
\begin{array}{c}
[A] \\
\vdots \\
B \\
\hline
\neg A \\
\end{array}
\]

\[
\frac{[A]}{\neg B} \ \bot
\]

What elimination-rule does this justify? We can infer from \( \neg A \) all and only what we can infer from its grounds. There is one I-rule with two premises \( m = 1, n_1 = 2 \), so there will be two E-rules, one for each premise of the I-rule:

\[
\begin{array}{c}
[A] \\
\vdots \\
B \\
\hline
\neg A \\
\end{array}
\]

\[
\begin{array}{c}
[A] \\
\vdots \\
\neg B \\
\hline
\bot
\end{array}
\]

\[
\neg A \\
\hline
C
\]

\[
\neg E_1 \quad \text{and} \quad \neg A \\
\hline
C \\
\neg E_2
\]

Applying the simplification as before, where we infer \( C \) from assuming the existence of derivations, respectively, of \( B \) and of \( \neg B \) from \( A \), we obtain:

\[
\begin{array}{c}
[B] \\
\vdots \\
\neg A \\
\hline
A \\
\hline
C
\end{array}
\]

and so

\[
\begin{array}{c}
[B] \\
\vdots \\
\neg A \\
\hline
A \\
\hline
C
\end{array}
\]

and

\[
\begin{array}{c}
[\neg B] \\
\vdots \\
\neg A \\
\hline
A \\
\hline
C
\end{array}
\]

and so

\[
\begin{array}{c}
[\neg B] \\
\vdots \\
\neg A \\
\hline
A \\
\hline
C
\end{array}
\]
The second of these is simply a special case of the first, and so we have justified, by considerations of ge-harmony, a form of \textit{ex falso quodlibet} (EFQ) as the matching elimination-rule for ‘\neg’ (dubbed ‘V’ by Gentzen):

\[
\frac{\neg A}{B} \frac{A}{\neg E}
\]

We need to check, however, that this rule does accord harmoniously with \neg I and permit a reduction of Prawitz’ kind. So suppose we have an assertion of \neg A justified by \neg I, immediately followed by an application of \neg E:

\[
\begin{array}{c}
\alpha_1 \\
\neg B \\
\alpha_2
\end{array}
\begin{array}{c}
\beta \\
\neg A \\
\beta
\end{array}
\begin{array}{c}
\neg E
\end{array}
\]

If we now close the open assumptions of the form A in \alpha_1 and \alpha_2 with the derivation \beta, we obtain:

\[
\begin{array}{c}
\beta \\
A \\
\alpha_1
\end{array}
\begin{array}{c}
\beta \\
A \\
\alpha_2
\end{array}
\begin{array}{c}
\neg B \\
C \\
\beta \\
A
\end{array}
\begin{array}{c}
\neg E
\end{array}
\]

Our worry, and it was Gentzen’s worry too, is that we still have an occurrence of the wff \neg B, major premise of an application of \neg E and possibly inferred by \neg I. Indeed, since A and B are independent, the degree of \neg B may be greater than that of \neg A. So has a suitable reduction been carried out?\footnote{Dummett (1977, p. 154) seems not to address this worry.}

Gentzen’s solution, described in the MS but not appearing in the published version (where the contradictories B and \neg B were replaced by ‘\bot’), is first to perform a new kind of permutative reduction on the original derivation of \neg A, so that it concludes in a single application of \neg I. So suppose that the derivation of \neg A concludes in successive applications of \neg I:

\[
\begin{array}{c}
[A,B] \\
\beta_1 \\
\beta_2
\end{array}
\begin{array}{c}
[A] \\
\alpha_2 \\
\alpha_1
\end{array}
\begin{array}{c}
\neg B \\
C \\
\neg I
\end{array}
\begin{array}{c}
\neg A
\end{array}
\]

The detour through \neg B is unnecessary. The derivation can be simplified as follows:

\[
\begin{array}{c}
[A] \\
\alpha_2 \\
B
\end{array}
\begin{array}{c}
A, [A] \\
\beta_1
\end{array}
\begin{array}{c}
[A] \\
\alpha_2 \\
B
\end{array}
\begin{array}{c}
A, [A] \\
\beta_2
\end{array}
\begin{array}{c}
\neg C \\
\neg A
\end{array}
\begin{array}{c}
\neg I
\end{array}
\]

\[
\frac{\neg A}{A} \frac{\neg E}{B}
\]
By successive simplifications of this kind, we can ensure that \( \beta_1 \) does not conclude in an application of \( \neg I \) and so \( \neg B \) in (†) is not a maximum formula.

### 4 Fregean Absolutism

We have now shown the analytic validity of \( \rightarrow I \), MPP, RAA (or \( \neg I \)) and EFQ (or \( \neg E \)). \( \rightarrow I \) and \( \neg I \) are justified directly, in defining the meaning of ‘\( \rightarrow \)’ and ‘\( \neg \)’, while MPP and EFQ are justified indirectly, as admissible rules, by showing that whenever their premises are derivable, so too is their conclusion. Consequently, whatever we can prove with these rules is analytic, by proof-theoretic justification of the first grade, in particular:

\[
\frac{\neg A}{A \rightarrow \neg A} \quad (1)
\]

\[
\frac{B \rightarrow A \rightarrow I}{A \rightarrow (B \rightarrow A) \rightarrow I(1)}
\]

\[
\frac{\neg A \rightarrow B}{A \rightarrow B \rightarrow I(2) \neg E}
\]

\[
\frac{\neg A \rightarrow (A \rightarrow B) \rightarrow I(1)}{-E}
\]

But by my lights, (*) \( A \rightarrow (B \rightarrow A) \) and (**) \( \neg A \rightarrow (A \rightarrow B) \), as general forms of inference, are false. As Ackermann put it:

“One would reject the validity of the formula \( A \rightarrow (B \rightarrow A) \), since it permits the inference from \( A \) of \( B \rightarrow A \), and since the truth of \( A \) has nothing to do with whether a logical connection holds between \( B \) and \( A \).”\(^{19}\)

Indeed, EFQ is also invalid, allowing one apparently to infer anything whatever from a pair of contradictories, for in conjunction with other plausible theses, it can be shown that EFQ leads to the validation of invalid arguments.\(^{20}\)

Where does the mistake lie? Arguably, it lies in the introduction-rules postulated for ‘\( \rightarrow \)’ and ‘\( \neg \)’ and the meanings thereby conferred on these connectives. That’s what results in the theses (*) and (**) which follow analytically from those rules, as we saw in §§ 2-3.

It is tempting to think that there’s nothing in themselves wrong with (*) and (**); what is wrong is to suppose those theses hold of ‘if’ and ‘not’. That is the usual moral drawn from the Paradoxes of Material Implication.\(^{21}\)

---

\(^{19}\) Ackermann (1956, p. 113): “So würde man die Allgemeingültigkeit einer Formel \( A \rightarrow (B \rightarrow A) \) ablehnen, da sie den Schluß von \( A \) auf \( B \rightarrow A \) einschließt und da die Richtigkeit von \( A \) nichts damit zu tun hat, ob zwischen \( B \) und \( A \) ein logischer Zusammenhang besteht.”

\(^{20}\) See, e.g., Read (1988, ch. 2).

\(^{21}\) See, e.g., Haack (1978, p. 200).
'¬', but they are paradoxical (counter-intuitive, that is, false) as applied to our pre-theoretic conceptions of implication and negation, since there are false implications with false antecedent and others with true consequent.

This is a mistake. Take ¬I first. As we saw, ¬I justifies EFQ, that is, the claim that each statement is paired with another, its contradictory, which together entail everything. To be sure, A and 'If A then everything is true' together entail that everything is true, and so entail everything. But we can deny A without claiming that if A then everything is true. Not every falsehood is that false! A and ¬A cannot both be true, but their conjunction does not necessarily bring the heavens crashing down, as Tennyson’s Lady of Shalott feared:

```
Out flew the web and floated wide;
The mirror cracked from side to side;
'The curse is come upon me', cried
The Lady of Shalott.
```

We can justify denying A simply by showing that A implies a falsehood, without necessarily showing it implies everything. The contradictory of A is the weakest proposition inconsistent with A, and 'If A then everything is true' is too strong.

The concept ‘If A then everything is true’ is often referred to as Boolean negation. It is a dangerous and unhelpful concept, which threatens to trivialize any theory. One does not need to be a dialetheist to want to reject it. Whether a theory contains both A and ¬A (where ‘¬’ is Boolean negation), for some A, is in general undecidable. But closing under consequence (as one does to form a theory) results in the trivial theory if both such statements are in it. This is unhelpful. A better account of negation is given by De Morgan negation.

Jean van Heijenoort contrasted absolutism with relativism in logic:

“Absolutism, . . . is the doctrine that there is one logic, that this logic is what has become known as classical logic, and, moreover, that such a logic is all-embracing and universal. Relativism is the opposite doctrine.” (van Heijenoort, 1985, p. 75)

He attributed absolutism to Kant, Frege and Russell. Kant seems in strange company as an adherent of classical logic. But in light of this, and the fact that van Heijenoort recognises the two doctrines as “tendencies”, one might adapt the title of absolutism to the belief in one universal logic, whatever it is, which I have elsewhere termed “logical monism” (Read, 2006).

---

We may well tolerate other logics, e.g., intuitionistic or dialetheic logic, or even classical logic with its commitment to EFQ, but that does not mean we accept them, or believe that they are suitable for the task of determining logical validity. They give the wrong answer about certain arguments, either validating invalid arguments (over-generation), or invalidating valid arguments (under-generation). Nonetheless, this raises the question why these arguments are valid or invalid, even though certain logics give a contrary verdict. They are invalid when the premises can be true while the conclusion is false; they are valid when this is impossible. Classical and intuitionistic logic fail as logics of the conditional since the truth of the consequent does not suffice to make a conditional true, nor does the falsity of the antecedent. There are false conditionals with false antecedent, e.g., ‘If I didn’t post the letter I burned it’, assuming I posted it; and with true consequent, e.g., ‘If second order logic is undecidable it is incomplete’, since other logics, e.g., first-order logic, are undecidable but still complete. So the compound conditionals, ‘If I posted the letter, then if I didn’t post it I burned it’ and ‘If second-order logic is incomplete, then if it’s undecidable then it’s incomplete’ (of the form (**) and (*) respectively) are false, since they have (or could have) true antecedent and false consequent.

A first thought concerning what is wrong with \(\rightarrow I\) (as given above) is that it allows vacuous discharge of the assumption, as in the above proof of (*). Prawitz (1965, p. 84) showed that we can circumvent this restriction and still prove (*), by use of the rules for ‘\(\wedge\)’ and ‘\(\vee\)’:

\[
\begin{align*}
A & \quad \bar{B} \\
A \wedge B & \quad \wedge I \\
A & \quad \wedge E \\
B \rightarrow A & \quad \rightarrow I(2) \\
A \rightarrow (B \rightarrow A) & \quad \rightarrow I(1)
\end{align*}
\]

\[
\begin{align*}
\bar{A} & \quad (1) \quad \bar{B} \\
A \vee (B \rightarrow A) & \quad \vee I \\
A & \quad (3) \quad \bar{B} \\
B \rightarrow A & \quad \rightarrow E \quad (4)
\end{align*}
\]

\[
\begin{align*}
A & \quad \bar{A} \\
B \rightarrow A & \quad \rightarrow I(2) \\
A & \quad \rightarrow I(1)
\end{align*}
\]

\[
\begin{align*}
A & \quad B \\
A \wedge B & \quad \wedge I \\
A & \quad \wedge E \\
B \rightarrow A & \quad \rightarrow I(2) \\
A \rightarrow (B \rightarrow A) & \quad \rightarrow I(1)
\end{align*}
\]

Of course, these proofs are not in normal form, but, more importantly, cannot be normalized. \(\rightarrow I\), with this restriction, does not interact holistically with the rules for ‘\(\wedge\)’ and ‘\(\vee\)’ to allow normalization. When the maximum formula is removed by the reduction step, the later application of \(\rightarrow I\) to derive \(B \rightarrow A\) is no longer legitimate if vacuous discharge is disallowed. Prawitz (1965, p. 84 n.2) floats the idea that one could simply restrict proofs to proofs in normal form, but realises that this would prevent us chaining proofs together in ways we find convenient and natural. For chaining two proofs in normal form together might result in a non-normal derivation, for example:

\[
\begin{align*}
A & \quad \bar{B} \\
A \wedge B & \quad \wedge I \\
B \rightarrow (A \wedge B) & \quad \rightarrow I(1)
\end{align*}
\]

Chaining these proofs together and eliminating the maximum formula \(B \rightarrow \)
\((A \land B)\) results in the non-normalizable derivation of \(B \rightarrow A\) from \(A\) given above.

The moral is that harmony is not enough to guarantee validity. Harmony ensures that the consequences of an assertion are no more and no less than the meaning encapsulated in the introduction-rule warrants. But that meaning may itself be corrupt. An example is the I-rule for • in Read (2000, §2.8), a formal Liar whose assertion warrants its own denial. One might try to bar such monsters by a restriction on the form of I-rules. The •I-rule breaches Dummett’s proposed constraint on I-rules, his “complexity condition”:

“The minimal demand we should make on an introduction rule intended to be self-justifying is that its form be such as to guarantee that, in any application of it, the conclusion will be of higher logical complexity [i.e., degree] than any of its premises and than any discharged hypothesis.” (Dummett, 1991, p. 258)

As with many proposed solutions to the Liar, out goes the baby with the bath-water: e.g., Gentzen’s rule R (i.e., \(\neg I\)) does not satisfy Dummett’s condition, yet seems otherwise harmless. Moreover, the complexity condition does not help with our present difficulty, since \(\rightarrow I\) satisfies the condition, yet leads directly to (*).

Prawitz (1985, 2006), and following him, Schroeder-Heister (2006), try to restrict further the proof-theoretic validity which results from the I-rules by relativizing it to the existence of a justification procedure.²⁴

“My approach is now to let the arguments for which validity is defined consist of argument skeletons together with proposed justifications of all the inferences that are non-canonical.” (Prawitz, 2006, p. 514)

As we have seen, whether there are such justifications, or reductions, will depend on the form of the E-rules and on the interaction of the various rules for the different connectives (in the end, on the Fundamental Assumption). But if the E-rules are in harmony with the I-rules (and the Fundamental Assumption can be accepted for argument-schemata as in §3) and so admissible, then they are justified and the appropriate reductions do exist. In that case, the justification cannot be denied, and the consequent derivations are analytically valid, even if they fail to preserve truth. So such a relativization is ineffective.

What has been realised over the past forty years or so (in the theory of relevance logic, in linear logic, and in the theory of sub-structural logics generally) is that one needs to distinguish two different ways of combining premises. They may be used side by side to entail a conclusion, or one may be applied to the other. For example, \( A \land B \) follows from \( A \) and \( B \) in tandem, whereas \( B \) follows from applying \( A \to B \) to \( A \). In this way, we can develop a theory of \( \to \) and \( \neg \) which does not have such problematic consequences as (*) and (**). What is valid is then analytically valid in virtue of the meanings of \( \to \), \( \neg \) and so on. This is not the place to spell out the details, which require making the assumptions in the rules explicit and the way they are combined to yield the conclusion. There is nothing new in this: the task is to find the right (formulation of the) rules, and hence the real meaning of the logical terms, by checking their consequences and revising accordingly.25 The moral we can draw, however, is that what logic one is committed to depends on the meaning one gives to the logical particles, encapsulated in the rules for their assertion, and so adopting the wrong logic may result in asserting falsehoods. The latter is something to be avoided.

5 Conclusion

The prevailing account of validity was given shape by Tarski, saying that an inference is valid if it preserves truth through all substitutions for the non-logical vocabulary. This is a reductionist enterprise, attempting to reduce the real modality in the criterion, ‘the premises cannot be true without the conclusion’, to the possibility of re-interpretation. It is a hostage to fortune whether the reduction is available (as Etchemendy (1990) observed), and at best turns the obscure question of validity into the more obscure one of the possibility of true premise and false conclusion. In contrast, the idea of proof-theoretic validity is that validity of inference is based on rules of proof and the meanings of the logical constants encapsulated in those rules. As in Tarski’s account, a division is required between the logical and descriptive vocabulary. In Tarski-validity, the essential idea is truth-preservation regardless of the meaning of the descriptive terms, while in proof-theoretic validity it is proof in accordance with the meanings of the logical terms given by the proof-rules.

Thus a proof-theoretically valid inference is analytically valid in virtue of the meanings of the logical constants specified by the rules for their application. Prawitz wrote:

“Once we know the condition for correctly asserting a sentence, we also know when to accept an inference and when to accept that a sentence follows logically from a set of premisses.”26

---

25See, e.g., Read (1988, ch. 1).
However, despite Prawitz’ use of ‘know’ here, analytic truth, and analytical validity, does not guarantee truth or validity. Use of an expression with a certain meaning can commit its user to the a priori assertion of falsehoods or to the endorsement of invalid inferences. The classic examples are Tarski’s claim that natural languages are inconsistent, by virtue of the meaning they give to the term ‘true’, and Prior’s observation that validity can be trivialized by adoption of his rules for the term ‘tonk’. An understanding of ‘true’ (if Tarski is right) and of ‘tonk’ commits their users to a justified belief in the correctness of certain statements and of certain inferences. But justified belief is famously not enough to guarantee truth. Hence Tarski refrains from using ‘true’ in its natural language guise and opts for an account of truth in stratified formal languages; Prior and others refrain from using ‘tonk’ with Prior’s rules.

Prawitz, and following him, Dummett, set out to articulate the mechanisms by which the rules for a logical term result in an analytical commitment to the validity of the resulting inference. The core idea is that all aspects of the term’s meaning should be in harmony. Meaning can be conferred by I-rules, E-rules, and by the proofs composed of such rules. But when the rules are in harmony, they all determine the same meaning. In particular, the whole meaning is then contained in the introduction-rule or rules. The elimination-rule is in harmony and is justified by the meaning so conferred if it is admissible, that is, if its conclusion is provable (without the rule, from some non-logical base) whenever its premises are. This is Dummett’s second grade of proof-theoretic justification. Even here, there is a crucial assumption, what Dummett terms the “Fundamental Assumption”, namely, that the elimination-rule can be permuted with the other E-rules so that its major premise is an application of the corresponding I-rule. In the familiar systems, this is true, but nonetheless it is a fact which needs to be proved. In general, it is an holistic assumption about the interaction between the rules for the different logical terms, so that local plateaux (or “hillocks”) can be reduced to mere peaks.

Moreover, natural deduction proofs do not consist only of closed proofs, but also of open proofs, that is, of derivations of conclusions from assumptions, and derivations generalizing on free variables. The Fundamental Assumption is simply false of open proofs. To apply the Fundamental Assumption in order to show the E-rules admissible, it needs to be shown that the proof can be articulated into a succession of branches, each ending in the application of an E-rule, and then closing off the open assumptions and free variables in each branch to obtain a closed proof for which the Fundamental Assumption is available.

What is good about the notion of proof-theoretic validity is that it recognizes that what rules one adopts determines the meaning of the logical terms of other arguments or argument schemata is thus valid by the very meaning of the logical constants . . . ; it is valid by definition so to speak.”
involved and commits one to accepting certain inferences as valid. What is bad is to infer from this that those inferences really are valid. Proof-theoretic validity serves an epistemological function to reveal how those inferences result from the meaning-determining rules alone. But it cannot serve the metaphysical function of actually making those inferences valid. Validity is truth-preservation, and proof must respect that fact.

Nonetheless, this is not to equate validity with preservation of truth through arbitrary replacement of the non-logical vocabulary. That is hostage to fortune through the richness or poverty of the vocabulary and availability of suitable models, and so can also result in the validation of invalid inferences (and failure to validate valid ones). Moreover, it is impredicative and unworkable since it equates validity with membership of a class of arguments all of which are valid, as Etchemendy (2008) showed. Validity is necessary truth-preservation, in itself dependent on the meanings of the constituent propositions. To that extent, Eklund and Scharp are right: be careful what you wish for—or choose to mean—for you may receive it, even though it may not be what you want, or what is true.

References


Prawitz, D. (n.d.). Validity of inferences. Revised version of a paper presented to the 2nd Launer Symposium on the Occasion of the Presentation of the Launer Prize 2006 to Dagfinn Follesdal.


