DiVA

Institutional Repository of Stockholm University

<http://su.diva-portal.org/smash/>

This is an author produced version of a **contribution to a collection**:

[Logicism, Intuitionism, and Formalism — What has Become of Them? Springer Netherlands. Dordrecht 2008.)

This contribution has been peer-reviewed but may not include the final publisher proof-corrections or pagination.

Citation for the published contribution:

[Peter Pagin]

[Intuitionism and the anti-justification of bivalence]

[Logicism, Intuitionism, and Formalism — What has Become of Them? Springer Netherlands. Dordrecht 2008.) Pp]

[ISBN: 978-1-4020-8925-1]

[URL: <https://www04.sub.su.se:2119/content/k267t02475936531/resource-secured/?target=fulltext.pdf&sid=lilqa1fhhub5dnrautkucoe&sh=www.springerlink.com>]

Published with permission from: **[Springer Netherlands]**.

Intuitionism and the anti-justification of bivalence*

Peter Pagin

June 12, 2006

Abstract

Dag Prawitz has argued (Prawitz 1998) that it is possible intuitionistically to prove the validity of ‘ $A \rightarrow$ there is a proof of $\ulcorner A \urcorner$ ’ by induction over formula complexity, provided we observe an object language/meta-language distinction. In the present paper I mainly argue that if the object language with its axioms and rules can be represented as a formal system, then the proof fails. I also argue that if this restriction is lifted, at each level of the language hierarchy, then the proof can go through, but at the expense of virtually reducing the concept of a proof to that of truth in a non-constructive sense.

1. Background

A couple of years ago (Pagin 1998) I argued against Dag Prawitz and Michael Dummett that the principle of bivalence, i.e. the principle that every sentence is either true or false, cannot have both a metaphysical and a meaning-theoretical significance. This *double significance view* is a conjunction of two claims. The first claim is that acceptance of bivalence for a particular area of discourse is the principal mark of a *realist* view of that area. Correspondingly, rejection of bivalence is the principal mark of an *anti-realist* view. Even though there are other ingredients in realism and anti-realism, and even though there are shades of realism and anti-realism, the stand on bivalence is the main dividing line.¹

The bivalence criterion was not meant to introduce a new idea distinct from traditional metaphysics, but as a way of articulating traditional metaphysical ideas in a clearer and more precise way.² The realist idea to be captured by the bivalence criterion is that of a determinate and mind-independent reality. Mind-independence is here equal to independence of being *knowable* (by humans or at least by finite minds). A certain domain D of reality is *mind-independent*

*A predecessor of the paper was presented at the Uppsala conference in August 2004. I benefited there especially from comments by John Burgess, Dag Prawitz, Stewart Shapiro and Göran Sundholm. In addition, the paper improved significantly thanks to comments from an anonymous referee. My overall philosophical debt to Dag Prawitz should be evident from the paper.

¹For an overview and classification of positions, see Dummett 1982.

²The general idea is spelled out in the last chapter of Dummett 1991.

if it is possible that there be unknowable facts in D . D is also *determinate* if all states of affairs of D either obtain or don't obtain.

To a reality domain D , which may be a domain such as the past, mathematical reality, or counterfactual reality, there corresponds an area of discourse, A_D , of sentences describing states of affairs in D . Assuming that everything in D is expressible in A_D , we can say that a state of affairs in D obtains just in case a sentence s in A_D that describes it is *true*, and it fails to obtain if s is false. It will also fail to obtain if s is neither true nor false, but I shall not here take account of that option. Therefore, every state of affairs in D either obtains or does not obtain just in case every sentence in the corresponding discourse area A_D is either true or false. Hence, the following general equivalence of bivalent reality:

(BR) A domain D is determinate iff A_D is bivalent

The realism issue for a domain D will not arise if we know that for any sentence s in A_D it is either the case that we can know that s is true or the case that we can know that s is false. For, if we know that, then we do know that D is *not* mind-independent. An example might be, for each person, the domain of present sensations, reflected in sentences like 'I am in pain'. Disputed domains are those where we lack any such guarantee. For such disputed domains we have two alternatives. Either we can hold on to bivalence, in which case we will also hold that domain to be mind-independent, for then every sentence of that area of discourse is true or is false regardless of whether we can know it or not. This is realism. Or we can reject mind-independence for the domain, and as a consequence we will then have to reject bivalence, for if a sentence is true just in case knowably true, and false just in case it is knowably false, and there is no guarantee that it be either knowably true or knowably false, then there is no guarantee that it is either true or false. So the issue of determinacy is closely connected with that of mind dependence. Let's say that

(KA) An area of discourse A_D is *knowable* iff it holds of every sentence s in A_D that if s is true, then it can be known that s is true.

I shall here take falsity to be equivalent with truth of negation, so we don't need a separate clause about knowability of falsity. We can now define mind-independence:

(DK) A domain D is *mind-independent* iff A_D is not knowable

For disputed domains, determinacy implies mind-independence, and acceptance of bivalence for the corresponding area of discourse amounts to affirming both. We can sum up the first conjunct of the double significance view as

(B1) A disputed domain D is *metaphysically real* iff A_D is bivalent.

The second conjunct in the double significance view is the claim that there are *meaning-theoretical reasons* for rejecting bivalence. One can put it by saying

that there is a meaning-theoretical *anti-justification* of bivalence. Rejecting bivalence here does not consist in claiming that there are sentences that are neither true nor false, but rather in claiming that there are reasons for believing that not every sentence is either true or false. By classical logic, if not every sentence is either true or false, then some sentence is neither. However, this does not hold by intuitionistic logic, and that is what counts in the present context, since only intuitionistic logic is justifiable according to the meaning-theoretical requirements.

We can then state the second conjunct of the double significance view as

- (B2) For any disputed domain D , there are meaning-theoretical reasons for believing that A_D is not bivalent.

I turn now to these meaning-theoretical reasons. What I shall call *Dummett's argument* is an argument to the conclusion that meaning cannot consist in truth conditions, at least as truth conditions are normally understood. In a very condensed form, it can be set out as follows³:

1. Knowledge of the meaning of a sentence is publicly manifestable.
2. Public manifestation of knowledge of the meaning of a sentence consists in exercising the ability to tell whether the central semantic concept applies to that sentence or not.
3. If the central semantic concept is the concept of truth (and therefore knowledge of meaning is knowledge of truth conditions) and the sentence is not effectively decidable, then there is no ability to tell whether that concept applies to the sentence.
4. Hence knowledge of truth conditions is not (always) publicly manifestable.
5. Hence knowledge of meaning is not knowledge of truth conditions.
6. Hence meaning is not truth conditions.

This is the way I have interpreted Dummett. Granting the argument for the sake of discussion, we have the requirement that the central semantic concept must be *decidable*. I granted this much in Pagin 1998, and I also granted that best alternative semantic concept would be the relation *x is a proof of y* .⁴ I assumed that this relation is decidable. If it is decidable, a semantic theory that has it as the central semantic concept would meet Dummett's manifestability requirement.

Further, I assumed that the meanings of the logical constants are given in terms of what counts as a proof, or canonical proof, of sentences with those constants as main operator, as is done in the intuitionistic tradition. Then, if a sentence is logically valid just in case all sentences of the same logical form have a proof, instances of

³Extracted mainly from Dummett 1976.

⁴In this case the concept is dyadic. What most directly compares with the monadic concept of *truth* is the relativized monadic concept *x is a proof of $\dots$* , one for each purported proof x .

(EM) A or not A

are not logically valid.

Given the equivalence principle, or disquotation principle, i.e. the principle that every instance of

(EP) ‘ A ’ is true iff A

is true, or correct in some appropriate sense, and again equating falsity with truth of negation, it also holds that the principle of bivalence cannot be logically justified: it is not derivable from the meaning explanations of the logical constants together with the equivalence principle, because (EM) is not a logically valid schema.

As a conclusion, bivalence is not logically justifiable. This is the basis for the claim that bivalence is not a *valid* principle, or that there is no guarantee that it is valid, or again that there is no reason for believing in its validity. This amounts to a *weak* rejection of bivalence: we lack positive reasons for believing in its validity. The *strong* rejection is the claim that there are reasons for believing that it is *not* valid. Part of the reason for thinking that bivalence is not valid is precisely that there is no general justification of its validity. This is then a meaning-theoretical rejection, or meaning-theoretically motivated rejection, and this is the second conjunct of the double significance thesis.

Then putting (B1) and (B2) together, we have as a consequence of the double significance thesis that

(B3) For any disputed domain D , there are meaning-theoretical reasons to think that D is not metaphysically real.

(B3) is a remarkable claim. How can reflection over the requirements on language for communication give such a metaphysical result?

In Pagin 1998 I argued that it cannot. There is a gap in the reasoning. The mere fact that bivalence cannot be *logically* justified, i.e. justified from the meanings of the logical constants, does not by itself imply that it isn’t valid in the sense that every instance is *true*. Reality might be such as to make either A or *not* A true, for some or for every disputed domain, even if we don’t have any general guarantee that this is the case, i.e. even if we don’t have any proof that this holds. It might be that for some particular sentence s of an area A_D s is true but not provable/verifiable. Or it may be that for each s of A_D either s or its negation is in fact provable, even though this general fact itself is *not* provable.⁵

In order to close this gap, truth and provability need to be connected in a non-trivial way. The most natural candidate is something like

(P) If A is true, then there is a proof of A

⁵Neil Tennant’s view, in Tennant 1996, is in agreement with this claim. According to Tennant, principles that are valid according to intuitionistic relevant logic are *analytic*, while others, like The Law of Excluded Middle, if valid at all, are not justifiable from meaning explanations and therefore *synthetic*.

(where I write ‘ A ’ instead of ‘the proposition that A ’). In the present terminology, if (P) holds for a particular domain D , then D is mind-dependent, assuming that proofs are knowable. Thus, arguing for the general validity of (P), for any domain D , is in effect arguing that reality in general is mind-dependent.

On some definitions of truth, e.g.

(TP) A is true $=_{def}$ there is a proof of A

proposed by Prawitz as well as by Per Martin-Löf, (P) holds by definition.⁶ But since we cannot simply *define* reality to be mind-dependent, we would need to start from a more neutral notion or property of truth. The equivalence principle itself is sufficient, if we decide to consider truth of sentences rather than propositions. Let’s do that. Then, given some restrictions, e.g. because of reference failure or for the sake of avoiding semantic paradoxes, the equivalence principle should hold for any acceptable conception of truth. Accepting this, we would have, for any sentence $\ulcorner A \urcorner$,

(P*) If A , then there is a proof of $\ulcorner A \urcorner$

I argued in Pagin 1998 that (P*) isn’t intuitionistically acceptable.⁷ First, it is not enough to give a semantics for some particular object language and show (P*) to hold for that object language (or a fragment of it), for what counts for the *realism* issue is whether bivalence and thus excluded middle holds for the *home* language, the language that is *used*. The home language may be bivalent and yet be the meta-language for a non-bivalent object language.⁸ So (P*) has to be proved for the home language itself, or for some fragment of it that comprises the area of discourse under discussion.

Secondly, I argued that there cannot be an intuitionistically acceptable proof of the validity of (P*) for the home language. A proof of (P*) would, according to the intuitionistic meaning explanations of the logical constants, be a function F that, for any sentence s , as the antecedent of the instance of (P*), for a proof of s as argument, gives a proof of the consequent ‘there is a proof of s ’ as value. Such a proof of the consequent would be a pair $\langle a, b \rangle$ such that a is an object and b a proof of ‘ x is a proof of s ’ given the assignment of a to x (alternatively, that a is a term and b a proof of ‘ a is a proof of s ’).⁹

I argued that there cannot be such a general function, on two assumptions. The first assumption was Church’s Thesis, according to which any decidable property is recursive, and the second assumption that the relation *x is a proof of y* is decidable. On these assumptions, since the language of arithmetic, together with the axioms of arithmetic, is part of the home language, Gödel’s

⁶See Martin-Löf 1984, p 11, Prawitz 1980, p 8, Prawitz 1994, p 85.

⁷In Pagin 1998 (P) is called ‘(PRO)’ and (P*) ‘(PRO*)’. I was also not making the distinction between proof of a sentence and proof of a proposition explicit.

⁸This can be achieved e.g. by the method of interpreting the intuitionistic language into a classical modal language, and is carried out in Pagin 1998, section 5.

⁹Hence, accepting (P*) simply as an *axiom*, in which case the one-step inference from zero premises is the (unreducible) proof, is not in accordance with the intuitionistic meaning explanations.

incompleteness results apply. The first incompleteness theorem implies that there is no exhaustive recursive relation *x is a proof of y*. Any recursive relation is representable in a formal system for arithmetic. Gödel showed that for any such formal system G we can construct a universally quantified sentence s of which there is no formal derivation in G , but which is informally proved by the observation that each instance of s is provable in G . Hence, no formula of a formal system can represent the relation *x is a proof of y*.

Hence, again, if the relation *x is a proof of y* is to be decidable, there is no well-defined collection of *all proofs*, and therefore no function F that could be a proof of (P^*) . Finally, applying (P) to (the universal closure of) (P^*) itself, we get

- (1) If (P^*) is true, then there is a proof of (P^*)

By contraposition, from the result that there is no proof of (P^*) , we get the conclusion that (P^*) is not true. Thus, from an intuitionistic perspective, (P^*) should in fact be rejected.¹⁰

2. Professor Prawitz's reply I: the proof relation

Professor Prawitz accepted the challenge. His reply, Prawitz 1998, has two main claims. First, the proof relation, *x is a proof of y*, cannot reasonably be regarded as decidable, and isn't required to be. Second, (P^*) can in fact be proved. I shall look at the first claim in this section.

According to Prawitz (Prawitz 1998, 306), the proof relation is not recursive. If the reason for this were only the incompleteness results, then there is an alternative view. You can think that there isn't one but many proof relations. A proof relation R is a subset of $P \times S$, the product of a domain P of proofs and a domain S of sentences, or of a product $P \times Q$ of a domain of proofs and a domain of propositions, depending on whether we think of proofs as related to sentences or to propositions. Let's assume it is sentences. Then we could say that for any proof relation R there is another proof relation $R' \subset P' \times S$, where $P \subset P'$. The restriction of R' to $P \times S$ is identical with R , but the expanded domain P' of proofs contains elements that by R' are proofs of sentences in the original domain S . We would say, as did Dummett in his reflections on Gödel (Dummett 1980), that the domain of proofs is *indefinitely extensible*. It would be open now to claim that although each of the members in the hierarchy of proof relations *is* recursive, there is no largest relation, no comprehensive proof relation.

Prawitz seems to take the opposite view, that there *is* a comprehensive proof relation, which, consequently, is not recursive. His reasons are not only related to the incompleteness results, but directly to the requirements on proofs by the

¹⁰It has been objected, and was so objected at the Uppsala conference, that Church's Thesis is not intuitionistically valid. However, the reason for rejecting Church's Thesis comes from Brouwer's theory of real numbers as choice sequences (the theory of the creative subject). The problems for (P^*) I point to arise already with arithmetic, where as far as I know intuitionists accept Church's Thesis.

intuitionistic meaning explanations. A proof of an implication $A \rightarrow B$ is to be a function that transforms any proof of A into a proof of B , and it may clearly be difficult to determine of a particular construction that it has this property.¹¹

However, if we give up the idea that the proof relation is decidable, what happens with the manifestation requirements that plays a crucial role in Dummett's meaning-theoretical argument? Prawitz says

Similarly, knowing a method for transforming proofs to other proofs and knowing the meaning of an implication $A \rightarrow B$, I should be able to decide whether the method *as I know it* amounts to a direct verification of the implication. The method may in fact transform any proof of A into a proof of B , but one cannot in general require that I should know a procedure that allows me to establish this fact. However, knowing the meaning of an implication, I should be able to decide for a given method known to me whether *in virtue of its properties known to me* the method amounts to a direct verification of $A \rightarrow B$. The required capacity is something else than a capacity to decide for an arbitrary object whether it is a direct verification of a given sentence (Prawitz 1998, 306-37, italics in the original).

Thus, according to Prawitz, since the proof relation isn't recursive, and hence not decidable in any clear sense, one cannot require that a speaker be able to manifest his understanding of a sentence by displaying an ability to decide for any given object, whether or not that object is a proof, or is a canonical proof, of the sentence. The requirement should be weaker.

What is the weaker requirement? Prawitz suggests that the speaker should be able to decide, for an object a and a sentence s whether a , *in virtue of the properties known to the speaker*, is a canonical verification of s . But what does 'in virtue of the properties known to the speaker' mean here? On one interpretation the requirement is as strong as the original one. For I may be aware of the properties of the object that are in fact such that in virtue of having them, i.e. as a consequence of having those properties, the object is a canonical proof of a $A \rightarrow B$, even though I am not able to verify that this *is* a consequence of having them. I can see how the object is constructed out of certain operators, and in virtue of being so constructed it is a proof of $A \rightarrow B$, but this I am not at the time able to tell. Hence, on this interpretation it is not clear that Dummett's requirement is weakened at all, and if it is, certainly not sufficiently.

On another interpretation, the requirement is clearly too weak. On this interpretation the phrase 'in virtue of the properties known to the speaker' refers to *proof properties the speaker is aware of*. That is, if the speaker is aware that a has the property of transforming any proof of A into a proof of B , then the speaker should also know that a is a proof of $A \rightarrow B$. However, this is an

¹¹Although I don't object to Prawitz's reason, it may be noted that it goes against an old intuitionistic view, expressed by Georg Kreisel by the aphorism 'We can recognize a proof when we see one' (Kreisel 1962, 202).

almost empty requirement. If I have no idea of what it requires of an object to be a function that transforms any proof of A into a proof of B , then I will never be able to tell when something is a proof of $A \rightarrow B$, but if I am required only to say that something is a proof of $A \rightarrow B$ in case I am aware that it transforms any proof of A into a proof of B , then I do meet the manifestation requirements nonetheless.

So this interpretation is too weak. What is needed is some interpretation intermediate in strength between the first and the second, weak enough not to require impossible capacities, but strong enough to rule out truth conditional semantics. It is not clear that there is such an intermediate manifestability requirement to be had.

The rejection of the demand that the proof relation be decidable does, however, take away the basis of my 1998 argument against the provability of (P^*) . Is it enough to make it provable?

3. Professor Prawitz's reply II: the provability of (P^*)

Professor Prawitz accepted my demand that (PRO^*) be intuitionistically provable as holding of our home language. He says

The upshot of this seems to be that the metaphysical relevance of the meaning-theoretical argument against bivalence for a certain language hinges upon two things: firstly, that the language is *our* language for speaking about the part of reality in question, and, secondly, that the meaning theory for this language adequately describes our use of this language, in particular what our sentences mean and when they are true (1998, 309-10; italics in original).

Prawitz does hold on to the (TP) truth definition, by which truth is the same as the existence of proof. (P) is an immediate consequence of (TP), but Prawitz thinks that Tarski's condition of material adequacy, of which the equivalence principle (EP) is a special case, should be satisfied. Because of this, both (P) and (P^*) should be correct for the home language.

On Prawitz's view, this demand can in fact be met, and he goes into some detail in indicating how this is to be done. First, Prawitz requires a strict separation of linguistic levels, in order to avoid the liar paradox, in the format

(S) There is no proof of (S).

Liar reasoning gives the result that there is and there is not a proof of (S). Therefore, we need an object language / meta-language separation, according to Prawitz. There are of course other ways to avoid the liar paradox, but in the present context the Tarskian is as good as any other, and I shall here follow Prawitz. He adopts the picture of the home language as made up of a hierarchy of object language and metalanguages. Three are needed for Prawitz's construction, an object language, a meta-language and a meta-meta-language, and I shall here simply call them L , ML and MML . As usual, ML is an extension of L , and MML an extension of ML .

I shall use a notation that is slightly different from Prawitz's. I shall use ' A ', ' B ' etc. as substitutional variables for sentences of L . I shall use ' $P(\dots, \dots)$ ' as relational expression in ML denoting the proof relation between objects and sentences of L , and boldface ' $\mathbf{P}$ ' as the corresponding expression of MML (denoting a relation between objects and sentences of ML). ' a ', ' b ' etc. are variables of ML ranging over a domain of proofs of sentences of L , and correspondingly ' $\mathbf{a}$ ', ' $\mathbf{b}$ ' etc. as variables in MML ranging over a domain of proofs of sentences of ML .

I shall use ' $\circ$ ' for forming structural-descriptive names in ML of expressions in L , and ' $\bullet$ ' for forming structural-descriptive names in MML of expressions in ML . Note that since MML is an extension of ML , ML names of expressions in L are also MML names of expressions in L . I shall use ' $\forall$ ' and ' $\exists$ ' as objectual quantifier symbols for ML and MML alike.

With this much machinery, we can state the formal counterpart to the (P*) schema as

$$(2) \quad A \rightarrow \exists a(P(a, A^\circ))$$

which is a sentence schema of ML . The claim that each instance of (2) is provable belongs, as Prawitz points out, to MML . With the addition of ' $[\dots]$ ' as a substitutional universal quantifier of MML , we can state that claim as

$$(3) \quad [A](\exists \mathbf{b})(\mathbf{P}(\mathbf{b}, (A \rightarrow \exists a(P(a, A^\circ)))^\bullet))$$

which in the present notation corresponds to Prawitz's (PRO+):

For any sentence A in English^*_L , there is an α such that
 Proof($\alpha, A \rightarrow (\exists p)\Pi(p, \mu(A))$).¹²

I suppose we could simplify things by having the substitutional quantifier already in ML . Then it would be sufficient to prove

$$(4) \quad [A](A \rightarrow \exists a(P(a, A^\circ)))$$

Be that as it may. According to Prawitz, proving (PRO+) is sufficient for answering the challenge. He gives a sketch of how (PRO+) can be proved by induction over complexity of L sentences, and carries out the induction step for conjunction. As he points out, for this to work we must add semantic clauses for the expression of the proof relation, i.e. ' $P(\dots, \dots)$ '. This must be done in accordance with the meaning explanations for logical constants of L . Thus, translated into the present notation, Prawitz says in the case of conjunction that a canonical proof of a sentence

$$(5) \quad P(c, (A \& B)^\circ)$$

¹²Here ' $(\exists p)\Pi(p, \mu(A))$ ' is a name in the meta-meta-language English^*_{MML} of a sentence in the meta-language English^*_{ML} which says that there is a proof of the sentence A of the object language English^*_L . ' $A \rightarrow (\exists p)\Pi(p, \mu(A))$ ' is the structural-descriptive MML name of the corresponding ML conditional.

of ML is to be the same as a canonical proof of

$$(6) \quad P(\mathbb{L}(\text{RUN}(c)), A^\circ) \ \& \ P(\mathbb{R}(\text{RUN}(c)), B^\circ)$$

The reason is this. A canonical proof of a conjunction $(A \& B)^\circ$ in L (named in ML), is a pair $\langle a, b \rangle$, where a is a proof of A° and b is a proof of B° . A proof in general is a *method* for arriving at a canonical proof. To accommodate this idea we need a function which takes a proof, i.e. a method, as argument, and delivers the corresponding canonical proof as value. We also need to enrich ML with a function symbol denoting this function, and ‘ $\text{RUN}(\dots)$ ’ is such a function symbol. We also need symbols for the *projection functions* that extract the left and right members of a pair, respectively, and ‘ $\mathbb{L}$ ’ and ‘ $\mathbb{R}$ ’ are such function symbols (thus, ‘ $\mathbb{L}(\langle a, b \rangle) = a$ ’ is true in ML).

Then, equating (5) and (6) amounts to saying that a proof that something c is a proof of a conjunction is the *same* as a proof that the elements of the canonical counterpart of c are proofs of the conjuncts. The alternative would be to include as a separate inference rule for MML that we can infer (5) from (6), (and another that the converse holds as well). In that case, a canonical proof of a sentence of the form of (5) would proceed by way of one application of this inference rule from a canonical proof of a sentence of the form of (6). In this case, the difference is of small significance, but as we shall, see, it will be more important in the case of the quantifiers.

Given the equation of (5) and (6), Prawitz’s idea of the induction step for conjunction is pretty straightforward. I shall sketch it here. For the details, the reader is referred to Prawitz 1998. First, we assume we have a proof $\mathbf{h}$ in MML of a sentence $(A \& B)^\bullet$ in ML . Since $\mathbf{h}$ is a method for arriving at a canonical proof of $(A \& B)^\bullet$ and such a canonical proof is a pair, we can from $\mathbf{h}$ extract a proof $\mathbf{c}$ of $A^\bullet$ and a proof $\mathbf{d}$ of $B^\bullet$. Now we can make use of the induction hypothesis. We infer by the induction hypothesis that there is a proof $\mathbf{e}$ of $(A \rightarrow \exists a(P(a, A^\circ)))^\bullet$ and a proof $\mathbf{g}$ of $(B \rightarrow \exists b(P(b, B^\circ)))^\bullet$.

Since by definition a canonical proof of an implication is a function that takes any proof of the antecedent into a proof of the consequent, we can apply $\mathbf{e}$ to $\mathbf{c}$, and $\mathbf{e}(\mathbf{c})$ is then a proof of $(\exists a(P(a, A^\circ)))^\bullet$. Similarly, $\mathbf{g}(\mathbf{d})$ is a proof of $(\exists b(P(b, B^\circ)))^\bullet$.

By definition, a canonical proof of an existential sentence is a pair where the first element is a term that replaces the variable, and the second element is a proof of the embedded sentence after the substitution. Therefore, from $\mathbf{e}(\mathbf{c})$ we can extract a pair $\langle u, \mathbf{m} \rangle$, where $\mathbf{m}$ is a proof of $(P(u, A^\circ))^\bullet$. Similarly, from $\mathbf{g}(\mathbf{d})$ we can extract a pair $\langle v, \mathbf{n} \rangle$, where $\mathbf{n}$ is a proof of $(P(v, B^\circ))^\bullet$. By forming the pair $\langle \mathbf{m}, \mathbf{n} \rangle$ we get a proof of the conjunction $(P(u, A^\circ) \ \& \ P(v, B^\circ))^\bullet$.

Now we apply the equivalence of (5) and (6). In virtue of this equivalence, the proof $\langle \mathbf{m}, \mathbf{n} \rangle$ of $(P(u, A^\circ) \ \& \ P(v, B^\circ))^\bullet$ is also a proof of $(P(\langle u, v \rangle, (A \& B)^\circ))^\bullet$. Now we can form the pair of pairs $\langle \langle u, v \rangle, \langle \mathbf{m}, \mathbf{n} \rangle \rangle$, which is a proof of the existential sentence $(\exists a(P(a, (A \& B)^\circ)))^\bullet$.

So, given the induction hypothesis, from the assumption of having a proof $\mathbf{h}$ of $(A \& B)^\bullet$ in ML we have derived a proof of $(\exists a(P(a, (A \& B)^\circ)))^\bullet$. Discharging

the assumption with an implication introduction, we have a proof (depending on the induction hypothesis) of $(A \& B \rightarrow \exists a(P(a, (A \& B)^\circ)))^\bullet$. The derivation we have carried out corresponds to a series of operations in *MML* on the assumed proof $\mathbf{h}$, building up a complex construction $\mathbf{q}(\mathbf{h})$, which is a term in *MML* denoting the proof of $(\exists a(P(a, (A \& B)^\circ)))^\bullet$. Discharging the assumption corresponds to abstracting on the argument, and we therefore have $\lambda x \mathbf{q}(x)$ as a proof term of *MML* denoting the proof of $(A \& B \rightarrow \exists a(P(a, (A \& B)^\circ)))^\bullet$. With the reasoning carried out in *MML* there is a proof expressible in *MML* of the *MML* sentence

$$(7) \quad P(\lambda x \mathbf{q}(x), (A \& B \rightarrow \exists a(P(a, (A \& B)^\circ)))^\bullet)$$

Existential generalization gives

$$(8) \quad \exists \mathbf{b}(P(\mathbf{b}, (A \& B \rightarrow \exists a(P(a, (A \& B)^\circ)))^\bullet))$$

which concludes the induction step, as carried out by Prawitz.

4. Problems with Prawitz's construction

Is Prawitz's induction step for conjunction correct as it stands, and can it be extended to a complete induction proof of (P^*) ? I am going to answer both questions in the negative, on the condition of a certain assumption. The assumption is that each level in the hierarchy is taken to come with a decidable set of axioms and rules, so that it can be represented as a formal system. In order to see the problem, consider gödel sentences in L . It is assumed that L contains the language and standard axioms of arithmetic. Under these assumptions, which I think Prawitz does make, there are gödel sentences in L . Suppose that $S = (\forall x Gx)^\circ$ is a gödel sentence in L , given some gödel numbering. Then a proof of S will as usual involve the observation that every instance is provable by the proof rules in L , and this reasoning can be carried out in ML . That is, we will have as a truth in ML

$$(9) \quad \forall t^\circ \exists a(P(a, (Gt)^\circ))$$

We may suppose that the proof rules of ML allow the derivation of (9). We may also assume that the proof rules of ML allow the (correct) inference from (9) to

$$(10) \quad \forall x Gx$$

i.e. to S itself. This can be stated in ML , since ML is an extension of L .

It might now seem, since we have a proof of S , that we can go on to claim that there *is* a proof of S , i.e. to claim

$$(11) \quad \exists a(P(a, (\forall x Gx)^\circ))$$

This would be a mistake, however. The proof variable ' a ' in (11) ranges over proofs in the ML domain of proofs, and these are proofs constructible *in* L . The proof of S carried out in ML is not, however, constructible in L , precisely since

S is a gödel sentence for L . Therefore, (11) is false: there is (on the assumption that the set of axioms and proof rules for L is consistent) no element in the domain of discourse of ML that is a proof of S .

By contrast, we *can* correctly say in MML that there is a proof of S , since the proof variables of MML range over proofs constructible in ML (as well as in L). That is, we can correctly say in MML

$$(12) \quad \exists \mathbf{a}(P(\mathbf{a}, (\forall xGx)^\bullet))$$

But now, the fact that (12) is assertible but (11) is not, has negative consequences for (P^*) . For consider the instance of (P^*)

$$(13) \quad \exists \mathbf{b}(P(\mathbf{b}, (\forall xGx \rightarrow \exists a(P(a, (\forall xGx)^\circ))^\bullet)))$$

(13) is true if there is an element in the MML domain of discourse that is a proof of the ML sentence

$$(14) \quad \forall xGx \rightarrow \exists a(P(a, (\forall xGx)^\circ))$$

Such a proof $\mathbf{b}$ is to be a function that maps any proof $\mathbf{a}$ of the antecedent, i.e. S , in the MML domain, on a proof $\mathbf{b}(\mathbf{a})$ of the consequent, i.e. $\exists a(P(a, (\forall xGx)^\circ))$, again in the MML domain. But there is no such proof $\mathbf{b}$, because there is a proof of the antecedent in the domain but no proof (at all) of the consequent. Hence, (13) is false. It immediately follows that (P^*) itself is false, since (13) is an instance of (P^*) .

The argument above, if correct, constitutes a refutation of (P^*) . On the other hand, Prawitz has provided a sketch of a proof of (P^*) . So what has gone wrong? The crucial part of Prawitz's induction steps are those corresponding to the equating of (5) and (6). In the case of the universal quantifier, the corresponding part would be a principle, derived from the intuitionistic meaning explanation of the universal quantifier, saying that a canonical proof of a sentence

$$(15) \quad P(c, (\forall xAx)^\circ)$$

of ML is the same as a canonical proof of

$$(16) \quad \forall t^\circ (P((\text{RUN}(c))(t^\circ), (At)^\circ))$$

where ' t° ' ranges over closed singular terms of L . This is to say that a canonical proof that c is a proof of $(\forall xAx)^\circ$ is the same as a canonical proof that the canonical proof extractable from c , when applied to a closed term t° of L , yields a proof of $(At)^\circ$.

The problem with equating (15) and (16) is that this constitutes a *reflection principle* for L itself. For suppose we have a canonical proof of (16). Then, again, we have shown that there is a certain construction c which when applied to a term t° of L yields a proof of $(At)^\circ$. But now, if the predicate A happens to be the gödel predicate G , then there is indeed, for every term t° a proof $c(t^\circ)$ of $(Gt)^\circ$, but still no proof in the domain of ML , i.e. constructible in L , of $(\forall xGx)^\circ$. Yet, if a canonical proof of the gödel instance of (16) also is a

canonical proof of the corresponding gödel instance of (15), there is, contrary to assumption, a proof in the domain of ML of the gödel sentence.¹³

The situation is obscured by the natural assumption that if ‘ $c(t^\circ)$ ’ denotes a proof constructible in L , for any L numeral t° , then so does ‘ c ’ itself (or ‘ $\lambda x(c(x))$ ’). But this cannot be right, for it seems that ‘ $c(t^\circ)$ ’ does denote a proof, although not a canonical proof, of $(Gt)^\circ$ in L , for any numeral t° . For the gödel predicate ‘ G ’ is a decidable predicate, which means that it can be proved in ML that

$$(17) \quad \exists a(P(a, (\forall x(Gx \vee \neg Gx))^\circ))$$

The proof a is a method for getting a canonical proof, which in turn is a function which, when applied to a term t° yields a proof of the corresponding instance. That is

$$(18) \quad \exists a(\forall t^\circ(P((\text{RUN}(a))(t^\circ), (Gt \vee \neg Gt)^\circ))$$

Suppose that e is such a proof, and u° a term. Then we have

$$(19) \quad P((\text{RUN}(e))(u^\circ), (Gu \vee \neg Gu)^\circ)$$

but $\text{RUN}(e)(u^\circ)$ is not a canonical proof. The canonical proof of $(Gu \vee \neg Gu)^\circ$ introduces the disjunction from either a proof of $(Gu)^\circ$ or a proof of $(\neg Gu)^\circ$. As usual, the canonical proof can be extracted from the non-canonical, or indirect, proof (in this case by iterating instances of the proof of the induction step in u). That is, it is provable in ML that

$$(20) \quad P(\text{RUN}((\text{RUN}(e))(u^\circ)), (Gu)^\circ) \vee P(\text{RUN}((\text{RUN}(e))(u^\circ)), (\neg Gu)^\circ)$$

This generalizes into

$$(21) \quad \forall t^\circ(P(\text{RUN}((\text{RUN}(e))(t^\circ)), (Gt)^\circ) \vee P(\text{RUN}((\text{RUN}(e))(t^\circ)), (\neg Gt)^\circ))$$

Now, the proof that there is a proof of $(Gt)^\circ$ for any numerical term t° makes use of the construction of G° and theorems such as the Diagonal Lemma, the proof of which is not available in L itself. By these means it can be proved, in ML , that the assumption

$$(22) \quad \exists a(P(a, (\exists x(\neg Gx))^\circ))$$

leads to a contradiction. From this it is quickly inferred in ML that

$$(23) \quad \forall t^\circ(\neg P(\text{RUN}((\text{RUN}(e))(t^\circ)), (\neg Gt)^\circ))$$

And from (23) and (21) it can be proved in ML that

¹³Prawitz (1998, 318) draws attention to the realizability formula ‘ $A \leftrightarrow \textcircled{R}A$ ’ of Nelson 1947. As far as I understand, beside the difference Prawitz himself points to, there is a further difference that is crucial. The counterparts to (15) and (16) in Nelson 1947 are equivalent, simply because the counterpart to (15) is defined to be an abbreviation of the counterpart to (16). This means that, in primitive notation, ‘ $\forall xAx$ ’ does not even occur as a proper sub-expression. Hence, there is no substantial issue of equating or not.

$$(24) \quad \forall t^\circ (P(\text{RUN}((\text{RUN}(e))(t^\circ)), (Gt)^\circ))$$

By the extrapolated Prawitz reflection principle, i.e. the equating of (15) and (16), a canonical proof in ML of (24) is also a canonical proof of

$$(25) \quad P(\lambda t^\circ (\text{RUN}((\text{RUN}(e))(t^\circ))), (\forall x Gx)^\circ)$$

(25), however, is false. The lambda term ‘ $\lambda t^\circ (\text{RUN}((\text{RUN}(e))(t^\circ)))$ ’ does not denote any proof constructible in L . This would be a free variable construction in L , giving a proof of $(Gt)^\circ$ for arbitrary term t° . There is no such thing. Instead, there is a different construction for each term, even though each of these constructions can be extracted from the induction proof of $(\forall x (Gx \vee \neg Gx))^\circ$, given a term t° . That this is the case, again, is proved by a free variable construction in ML , in the manner hinted above. This free variable construction is, in turn, denoted by a lambda term in MML .

This refutation of the reflection principle, and thus of Prawitz’s induction proof of (P^*) has proceeded on the assumption that the set of rules and axioms of L is decidable, and therefore can be regarded as a formal system. The same was assumed of each level in the hierarchy, although the total hierarchy itself can not be seen as a formal system. But suppose we give up that assumption about L . Will Prawitz’s construction then go through?

5. Consequences of Prawitz’s construction

Giving up the assumption that L together with its axioms and rules of inference is a formal system allows us to accept the reflection principle involved in the universal quantifier step of Prawitz’s induction proof of (P^*) , and then I can see no reason to think that the induction proof will not go through at this level, i.e. for sentences of L . I believe it will.

If we don’t treat L as a formal system we can regard the lambda term ‘ $\lambda t^\circ (\text{RUN}((\text{RUN}(e))(Gt)^\circ))$ ’ as denoting a proof in L , e.g. a separate axiom. Similarly, for each sentence S of L of which it can be proved in ML that it is a gödel sentence for the original set of rules, S is an axiom in the new, expanded set of axioms. The expanded set of axioms is not recursive, but it is still recursively enumerable, since each axiom corresponds to a theorem provable in ML , assuming that ML , with its rules and axioms, still is a formal system.

If the set of rules and axioms of ML is decidable, ML can still be treated as a formal system, and then has its own incompleteness and its own gödel sentences. Because of this, however, we will again have a counter-example to the Prawitz induction proof when we want to move it one level up, and state in $MMML$ that for each sentence of ML there is a proof constructible in MML . In order to eliminate the counter-example at this level, we need to accept a reflection principle for ML . The rules for ML will therefore not make a formal system, either. And, as far as I can see, to make the induction proof go through in full generality, i.e. for all levels, the same must be done at each level. Therefore *no* level in the linguistic hierarchy comes with a decidable set of rules.

But this has consequences already at L . For among the new axioms of ML

that are added because of the reflection principle for ML , are sentences about proofs in L , and this in turn induces even more new axioms for L , because of the reflection principle for L . There will then also be even more new axioms for ML and for L because of the reflection principle for MML , and so on. So, a sentence is an axiom in L if it is a consequence of what is provable at *some level or other* in the hierarchy. But since no level will have a decidable set of rules and axioms, this means that the set of axioms of L will not even be recursively enumerable.

As a consequence, it is not clear what remains of the concept of a proof over and above the concept of being true in a non-constructive sense. If in effect we say that any sentence that is *true* is to be added as an axiom of L , then provability is simply reduced to truth. This seems to be the ultimate consequence: by availing oneself of the means to prove (P^*) in its full generality, if only level by level, provability reduces to truth in a non-constructive sense. Or at least, it is not clear to me how this consequence can be avoided. Of course, because the equivalence principle (EP) holds for ‘true’, (P^*) is clearly correct in case of such a reduction.

6. Conclusion

Since there is reason to think that we cannot, for *every* sentence A , come to know either that A is true or that A is false, we also have a reason to doubt the principle of bivalence, *provided* there is reason to think that

If a sentence A is true, it is also provable/knowable.

This is intuitively what (P^*) says. In the absence of a reason for holding on to (P^*) , or at least an interesting restricted version of it, there is no reason to doubt bivalence either. That is, we have been given no justification of the claim that it isn’t valid.

What does hold, if the reasons for intuitionistic semantics are correct, is that bivalence does *not* have the status of a quasi-logical principle. In a classical semantics it does, since it is justifiable from the law of excluded middle together with the equivalence principle (EP), and the equation of falsity with truth of negation. But that it isn’t a principle that enjoys the status of quasi-logical validity is not a reason for thinking that it isn’t valid. In fact, it is precisely because of this that Dummett’s connection between bivalence and realism is to the point. Belief in bivalence is an expression of realism because bivalence is *not* anti-justifiable by meaning theory and logic alone.

Department of Philosophy
Stockholm University

References

- Dummett, M., 1976, ‘What is a theory of meaning? (II)’, in G. Evans and J. McDowell (eds.), *Truth and Meaning*, Oxford University Press, Oxford. Reprinted in Dummett 1998, 34–93. Page references to the reprint..
- Dummett, M., 1980, ‘The philosophical significance of Gödel’s theorem’, in *Truth and Other Enigmas*, 186–201. Duckworth, London, 2nd edn., Originally published in *Ratio* V, 1963:140–55.
- Dummett, M., 1982, ‘Realism’, *Synthese* 52:55–112.
- Dummett, M., 1991, *The Logical Basis of Metaphysics*, Harvard University Press, Cambridge, Mass.
- Dummett, M., 1998, *The Seas of Language*, Clarendon Press, Oxford.
- Kreisel, G., 1962, ‘Foundations of intuitionistic logic’, in E. Nagel (ed.), *Logic, Methodology and Philosophy of Science I*, North-Holland, Amsterdam.
- Martin-Löf, P., 1984, *Intuitionistic Type Theory*, Bibliopolis, Napoli.
- Nelson, D., 1947, ‘Recursive functions and intuitionistic number theory’, *Transactions of the American Mathematical Society* 61:307–68.
- Pagin, P., 1998, ‘Bivalence: meaning theory vs metaphysics’, *Theoria* LXIV:157–86.
- Prawitz, D., 1980, ‘Intuitionistic logic: a philosophical challenge’, in G. H. von Wright (ed.), *Logic and Philosophy*, Martinus Nijhoff Publishers, The Hague.
- Prawitz, D., 1994, ‘Meaning theory and anti-realism’, in B. McGuinness and G. Oliveri (eds.), *The Philosophy of Michael Dummett*, 79–89, Kluwer.
- Prawitz, D., 1998, ‘Comments on Peter Pagin’s paper’, *Theoria* LXIV.
- Tennant, N., 1996, ‘The law of excluded middle is synthetic a priori, if valid’, *Philosophical Topics* 24:205–29.