# Bivalence: meaning theory vs. metaphysics

Peter Pagin

This paper is an attack on the Dummett-Prawitz view that the principle of bivalence has a crucial double significance, metaphysical and meaning theoretical. On the one hand it is said that holding bivalence valid is what characterizes a realistic view, i.e. a view in metaphysics, and on the other hand it is said that there are meaning theoretical arguments against its acceptability. I argue that these two aspects are incompatible. If the failure of validity of bivalence depends on properties of linguistic meaning, then there are no metaphysical consequences to be drawn. The case for this view is straightforward as long as we are discussing a language different from our own. But it seems that the distinction between failure because of meaning and failure because of reality cannot be applied to our own language, simply because our own language is just what we use to represent reality. I argue that this impression is illusory. In order to draw a conclusion about reality, meaning must be connected with truth in a non-trivial way, and precisely this cannot be done in the language for which the meaning theory itself is correct.

## 1. Bivalence as a metaphysical principle

According to Quine, acceptance of certain sentences as true, in a language that is equipped with features of cross reference, brings with it a metaphysical *commitment*. In the language of first order logic, this comes out in the acceptance of sentences involving objectual quantification. If I hold that the individual variables of my language take values in the domain of points of time, this commits me to the metaphysical acceptance of points of time as real entities. If time points exist, then e.g. the sentence “ $x(x = x)$” is true, with time points included in the domain of quantification, and if it is true, with such a domain, then time points exist.[1]

This general idea, that acceptance of certain forms of sentence, or certain principles, brings with it a metaphysical commitment, has been taken further by Michael Dummett. Dummett has been concerned with the question of *realism* regarding various subject matters. According to Dummett, belief in the existence of entities is less central to realism than belief in the *objectivity*, or *determinacy of facts*.[2] To believe in the determinacy of facts is to believe that each possible fact is either realized or not realized, or again that every state of affairs determinately either obtains or does not obtain. If you doubt the determinacy of facts, then you may also doubt that every fact stating proposition is either made true or made false by reality. To put it in a slightly misleading way, you don’t rule out the possibility that reality might have, as it were, gaps.

If we accept the simple equivalences

the state of affairs that *A* obtains iff *A*

the state of affairs that *A* does not obtain iff not *A*

then we might rephrase the statement

(1) each state of affairs obtains or does not obtain

1. See e.g. Quine (1960), p 242.
2. See e.g. Dummett (1980) pp xxvii-xxix.

as

(2) for each *p*, *p* or not *p*

If we read (2) as a version of (1), it should be read as quantifying over states of affairs. On this reading, (2) is concerned with states of affairs whether or not they correspond to any statement or proposition expressed by any of our sentences.

However, if we read the quantification substitutionally, then (2) becomes equivalent with the statement

(3) each substitution instance of the schema *A or not A* is true

and (3) is a statement of *the law of excluded middle*, a part of classical logic. This means that if you do take every state of affairs to determinately either obtain or not obtain, you should also accept the law of excluded middle as valid. And if you accept the law of excluded middle, then you should also take at least all states of affairs expressible by our sentences as determinately either obtaining or not obtaining.

To some people it will surely seem odd to say the validity of a logical law can depend on properties of reality. This might seem to make the law contingent, while it should be said that if something really is a logical law, then it holds by necessity. Fortunately, this issue can be bypassed. Dummett has not proposed to equate acceptance of the law of excluded middle with acceptance of realism in the sense of belief in the determinacy of facts. The main reason, or one main reason, is that questions of realism or anti-realism are concerned with distinct areas of reality, and each separately. On Dummett's view you can be a realist in one area, like mathematics, and an anti-realist in another, like the past. The law of excluded middle, however, is a logical law. If it is valid, then it is valid irrespective of the area of discourse. Dummett thinks that the law of excluded middle isn't valid, and that is precisely why he can think that it can vary from one area to the other whether all substitution instances, in that area, of the schema *A or not A* are true.

Dummett has, however, taken a further step. Instead of speaking of accepting the validity of a sentence schema, he has chosen to speak of acceptance or rejection, for each particular area, of the principle of *bivalence*, i.e. the principle that sentences (about states of affairs in some particular area), are either true or false. In fact, Dummett regards acceptance of bivalence not just as characteristic of the realist, but as actually defining what it is to be a realist. Similarly, rejection of bivalence by definition makes you an anti-realist, for some area, now delimited by the class of sentences that are concerned with that area. Although Dummett recognizes further aspects of realism and anti-realism, bivalence is taken as the main dividing line.[3] In this way realism becomes in a certain respect a *semantic* issue.

There are three interconnected reasons for this change of perspective. First, Dummett holds that traditional metaphysical disputes are framed in pictorial language, with pictures like that of an ethereal realm of abstract entities. These pictures should be replaced by more precisely stated claims. Secondly, the conflict about e.g. the reality of the past is a conflict about what makes our sentences about the past true, and therefore the question of the reality of the past will involve the question of the proper semantics for sentences about the past. The realist thinks that a sentence

3. In (1982) Dummett gives an overview of issues relevant for classifying views into realist or anti-realist.

can be made true by past facts even if it is impossible for us to know, and the anti-realist rejects this as incoherent. Similarly for other disputes over realism. Third, if the conflict is taken as a conflict about semantics, then it can actually be resolved. It will be resolved once it has been shown what a correct meaning theory for the language in question must be like.[4]

So the idea behind proposing bivalence as the criterion of realism is not to just ignore traditional metaphysical issues, but to give those issues a new form. The ultimate interest is still the nature of reality. And the interest in bivalence must depend on its connection with traditional metaphysics.

However, the change to a concern about bivalence induces a risk of distorting the original issues. What connection there is between bivalence, for some range of sentences, on the one hand, and the determinacy of facts in the corresponding area, on the other, depends on the *semantics* for those sentences, i.e. on how those sentences are interpreted. But it also depends on how well the meanings of the sentences are determined.

Suppose that you interpret sentences about the past according to a correspondence model, so that they are made true or false by past facts or events. The sentence ‘George Washington crossed the Potomac’ is then true, if it is true, in virtue of the event of Washington’s crossing the Potomac. Suppose further that all sentences concerned have a well determined meaning. If a (non-analytic) sentence has a well determined meaning, then it depends only on reality whether or not the sentence is true.

Under these assumptions realism and bivalence are equivalent. If past reality is fully determinate, so that all past states of affairs either obtained or didn’t obtain, and the meaning of the sentences are such that this and nothing else is needed for making them true or false, respectively, then each sentence is true or false. And similarly, if bivalence holds, so does realism.

However, if one of these two assumptions is dropped, then the equivalence is lost. Suppose that you give up the second assumption, about the determinacy of meaning. You may think e.g. that the application of a predicate like ‘… crossed …’ to a pair of objects, like Washington and the Potomac, is determinately either correct or incorrect only if it is within our present capacity to decide whether it is the one or the other. If we don’t have that capacity, then, on this view, the meaning of the predicate is underspecified for these arguments. This view leads to the rejection of bivalence. We cannot assert that each sentence is either true or false, since if we cannot verify, nor falsify, a particular sentence, then, on this view, its truth value is left indeterminate. In this case, however, there are no consequences for reality. Reality may be fully determinate. The failure of bivalence depends on meaning only, not on the world.

We get the same result if the first assumption is dropped. Suppose that sentences about the past are interpreted so that they are made true or false *not* by past facts or events, but by present evidence, i.e. by such present facts which we take as grounds for asserting those sentences, like facts about written documents or memories. Then again we have a reason to reject bivalence, for there is no guarantee that we will, for each sentence about the past, be in possession either of true-making evidence or of false-making evidence. And again nothing seems to follow about realism. If sentences about the past are not made true or false by past facts, then failure of bivalence implies nothing about the determinacy of past reality.

These objections seem straightforward, and I think they are straightforward, as long as we are allowed to speak independently of language and of reality. What is needed is to extend them from atomic sentences to logically complex sentences (see section 5). The philosophical diffi-

4. See Dummett (1991), chpts 1 and 16.

culty arises when we assume the language in question is our own language. We cannot pretend to treat the nature of reality as a completely separate issue from the semantics of our own language, for that is the language we use to talk about reality, and our ability to state facts about reality depends on the semantic properties of our language.

This circumstance can be used on Dummett's behalf to counter the two objections. Against the objection concerning truth and evidence it can be replied that we have no conception whatsoever about the reality of the past save as that which is represented by our sentences about the past. If these sentences cannot reasonably be given a semantics according to which they are true or false in virtue of past facts, then the notion of an independent past reality must be rejected, for it is only as the truth maker of our sentences that we have a reason to be realists about it. Moreover, we cannot then use the very sentence 'past facts are fully determinate' for stating an independent metaphysical truth, for this sentence must itself be evaluated according to the proposed alternative semantics.

Similarly, it might be replied to the first objection, regarding the underspecification of meaning, that there is no effective distinction to make here between failure in determinacy of meaning and failure in determinacy of reality. We don't have, again, any independent conception of the past, beyond what can be stated by our sentences about the past. So it can't be claimed that, while the meanings of our sentences about the past are not all fully determinate, past reality itself still is so.

These two replies should be treated rather differently. I think the first reply is not even fully coherent. If past reality is precisely what makes our sentences about the past true, then what makes those sentences true just cannot be present evidence, for our sentences about the past are not *about* present evidence (and in that case, what would the evidence be evidence for?). Rather, if we wish to maintain that sentences about the past are made true by present evidence, then they are made true by something else than the past state of affairs which they are about, and then it still appears that the failure of bivalence, if it depends on the lack of guarantee of the existence of present evidence, does not have any consequences for realism, in the sense of the determinacy of past facts.

The second reply, on the other hand, is to some extent justified. If we have agreed that the truth value of 'George Washington crossed the Potomac' is left indeterminate, for whatever reason, then I cannot just go on to claim that, nevertheless, either Washington *did* cross the Potomac or else he didn't.[5] For then I am just using the terms whose meanings were agreed to be problematic.

And it might also seem that I cannot speak in general terms about the reality of the past, for my general conception of the past seems to be determined by, or embodied in, all the particular things I can say about the past. If this is right, then I cannot claim that past reality is fully determinate, for I have no conception of past reality save as the totality of past states of affairs like the one concerning Washington and the Potomac.

If this is right, then we cannot make the general language/reality distinction in the case of our own language. If bivalence fails, we cannot, then, just blame language and save reality.

However, I don't think it is right, and the purpose of this paper is to explain why. The main idea can be illustrated by an example from vision. Suppose that wherever I look, I seem to observe an empty grey area in the middle. Naïvely, it appears to me as if there is some area of

5. That is, under the assumption that a disjunction is true if at least one of its disjuncts is.

greyness that travels around in the world. And if I have no conception of the world as distinct from what is represented by my eyes, then I also believe that there is such a grey travelling area. However, the situation is different if I have a *theory* of the representational function of the eye. Suppose that my theory tells me that because of inherent limitations of my eyes, the world must appear as having a grey area in the centre of what is seen. Suppose, further, that I *assert* that the world must be exactly what I can see, and that, therefore, there *is* an empty grey area travelling around. The problem with this claim is that, because I cannot *see* that the world is exactly as I see it (if *p*, then I can see that *p*), I cannot make it unless on the basis of cognitive resources that go *beyond* what is delivered by vision. Therefore, either I have some other means of knowing about the world, or else I have no basis for the claim.

I am going to argue that we have a very similar situation with the meaning theoretical argument, developed notably by Dummett and Dag Prawitz, against the validity of bivalence. There it is first argued that linguistic meaning must be of a verificationist kind, and then it is noted that bivalence cannot be justified from meanings (of that kind). However, if bivalence is also to have *metaphysical* relevance, then this argument needs to add some principle for inferring provability from truth (if *p*, then it is provable that *p*). But, or so I shall argue, such a principle is *not* assertible *within* a language of which Dummett-Prawitz semantic principles hold. It might be assertible in a *meta*-language as holding for the object language, but then the metaphysical issue is moved to the the meta-language, which need not itself be verificationist. The fact that bivalence cannot be *justified* from verificationist meanings does not by itself imply that it isn't valid, and the extra principle needed for this conclusion can be asserted only at the meta-level. But then it is without metaphysical significance, for then the failure of bivalence need not correspond to any failure of determinacy of reality.

In what follows I shall present Dummett's meaning theoretical argument (section 2), discuss the bearing of this argument on the concept of truth (section 3), and present the intuitionistic semantics for logical constants (section 4). Then, in section 5, I discuss what can be inferred from the fact that bivalence fails for an intuitionistic object language. In section 6 I move on to discuss what can be asserted about the relation about provability and truth, and therefore about bivalence, *within* an intuitionistic language. In section 7 I state my conclusions about meaning and realism.

## 2. Dummett´s argument

*The meaning theoretical argument against bivalence*, as I shall call it, is an argument with the following conclusion: the principle of bivalence cannot be justified in the canonical way, i.e. from the meanings of the logical constants, and this is a reason for regarding it as not universally, or logically, valid.

The first part of the argument is what is known as Dummett's argument. Dummett's argument, which I shall set out below, is an argument to the conclusion that the meaning of a declarative sentence cannot be its truth conditions. The second part consists in presenting an acceptable alternative to truth conditional semantics, viz. a verificationist, or intuitionist semantics. This semantics provides interpretations of the logical constants. The third part of the argument consists in showing that with this semantics, bivalence cannot be justified directly from those interpretations.

In this and the next section I shall present and discuss the first part, and then move on to the second part in section 4. The third part is then considered in sections 5-7.

Dummett's argument[6], in extremely condensed form, can be set out as follows:

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.[7]

For the time being, I shall assume that the argument is correct.[8] Hence I shall assume that *meaning*, whatever that precisely is, cannot be identified with truth conditions. Moreover, I shall assume that the reason why this is so, is that this would render impossible our public manifestation of knowledge of meaning, for at least some sentences.

As far as I have interpreted Dummett, what is required for manifestability, is that the so called central concept (of the right semantic theory) be *decidable*. The idea is this. For any given sentence, my interlocutor must be able to put me to the test: do I understand this sentence or not. What is this test like? I can, of course, sometimes give verbal explanations and definitions of terms contained in the sentence, but this, ultimately, just moves the question from the original sentence to the new sentences I use in the explanations and definitions. So this step can be dispensensed with, philosophically. Rather, since understanding is knowing meaning, and since meaning consists in the conditions under which the central concept applies to the sentence, I shall be tested for my command of these conditions. In case the central concept is truth, these are truth conditions. Since conditions aren't the kind of thing I can display physically, like on a platter, what I will be asked to do, is to decide whether the conditions are *satisified* or not. That is, I must give a verdict on the issue, and that verdict must be based on knowledge that is available to me. In case truth conditions are at issue, my verdict must be either that the sentence is true, or that it is false. In case the sentence is true, and I say it is true, I have manifested understanding, and if it is true, and I say it isn't, then I have manifested lack of understanding, or at least failed to manifest understanding.

Thus, when I am challenged for understanding, what I must do is either to make public the

---

6. Extracted primarily from Dummett (1976).

7. Some would want to qualify this, by adding that it only concerns truth conditions under the assumption that truth is bivalent. For a discussion, see section 3.

8. I don't think it is. As I see it, the great virtue of Dummett´s argument, and the philosophy it builds on, is that it presses the question what knowledge of truth conditions, and more generally knowledge of meaning, consists in. What I have come to see as its greatest weakness, on the other hand, is that, in the demand of manifestability, it builds on a preestablished and oversimplified conception of what is required for successful communication.

It should be noted that Prawitz is critical of Dummett's requirement of (full) manifestability, and thinks this should be replaced with a weaker requirement that a theory of meaning should have empirical import by way of consequences concerning linguistic behavior. See Prawitz (1994) pp 85-87. I agree with Prawitz's critical points, but Dummett's argument is not the issue in this paper.

knowledge I already have about whether the conditions are satisfied, or else acquire that knowledge. This is always possible, as long as the central concept is a *decidable* one, i.e. a property for which there is a general method for deciding whether or not the concept applies to any given object. For then I can acquire the knowledge needed by applying this method. It will be guaranteed that, within a finite amount of time, I will have reached the goal. But if the concept isn't decidable, then there are sentences for which I just have no idea how to acquire the knowledge needed. I know of no method of doing it. When that is the case, my understanding cannot be manifested.

Truth isn't a decidable concept. There is no general method for deciding whether a sentence is true.[9] Dummett has often given examples, such as some counterfactual conditionals, or sentences involving reference to inaccessible space-time regions, or quantification over infinite domains.[10] Dummett calls such sentences undecidable, and the point seems to be that we just don't know how to decide them. And then we cannot manifest our understanding of them.[11]

## 3. Dummett's argument and the concept of truth

If Dummett's argument is correct, as I presently assume, then *meaning* cannot be truth conditions. However, the argument is often taken to show just that meaning cannot be *bivalent* truth conditions, i.e. that meaning cannot be truth conditions, given that the principle of bivalence holds.

And this conclusion is further connected with a rejection of the corresponding concept of truth itself. If bivalence holds, then, as far as we know, a sentence may be true, even though it is impossible to know that it is true. Such a conception of truth has sometimes, e.g. in Crispin Wright's writings[12], been called a conception of truth as *verification trancendent*. That is, truth is verification transcendent if a sentence can be true even though it is impossible to verify. And it seems often taken for granted that this very concept of truth is just what Dummett's meaning theoretical argument rules out.

Both claims are mistaken, however, unless the present interpretation of Dummett's argument is fundamentally wrong. What that argument shows, if cogent, is that meaning cannot be taken to consist in truth conditions, on *any* (reasonable) conception of truth. For suppose we have the following definition of truth:

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

Truth is clearly not verification transcendent on the (TDEF) conception, since proofs are taken to be knowable (proofs are finite objects which can be recognized as proofs, and, moreover, recognized as proofs of what they prove). But *meaning* could not consist in (TDEF) truth conditions either. For even though the property of *being* a proof is a decidable one, the property expressed by "there is a proof of …" *isn't* decidable. There is no decision procedure for deter-

9. Strictly speaking, this can be categorically asserted only on the basis of Church's thesis, if decidability is an epistemological rather than a meta-mathematical concept. Given Church's thesis the undecidability of validity for predicate logic implies the undecidability of truth, simply because of the construction "'*x* is valid in predicate logic' is true".

10. E.g. Dummett (1976), p 81.

11. This is compatible, of course, with the possibility that we may later find such a method.

12. See e.g. Wright (1976), and the introduction to Wright (1987).

mining in general whether a sentence is provable. And decidability of the central semantic notion was exactly what the requirement of manifestability of knowledge of meaning required. Since there is no guarantee that for every sentence there is either a proof or a disproof, there would be no guarantee, on the (TDEF) conception, that each sentence would be such that knowledge of its meaning is publicly manifestable. And Dummett does require such a guarantee.

You might think that Dummett's argument strikes at bivalent truth conditions only because Dummett is considering undecidable sentences, where "undecidable" is understood as *absolutely* undecidable rather than *not effectively* undecidable. That would be an even worse mistake. Dummett presents examples of undecidable sentences, such as Goldbach's conjecture, and that would be in need of mathematical proof if absolute undecidablity was intended. Moreover, given the equivalence principle, or disquotation principle (*p* is true iff *p*), and given the rejection of verification transcendent truth, we get a contradiction from the claim that absolutely undecidable sentences exist. If truth isn't verification transcendent, then an unprovable sentence isn't true. An absolutely undecidable sentence has no proof and no disproof. So if there is such a sentence, neither that sentence nor its negation is true. Given the equivalence principle for truth, this would imply *not p and not not p*.

So, first, it does not seem that Dummett's argument provides any direct conclusion at all about the proper notion of truth, and second, if there were such a conclusion, it would cut across the bivalent and the (TDEF) conceptions of truth alike.[13]

It is true that Dummett also attacks the supposition that we can make sense of the idea that a sentence might be true although unverifiable:[14] we cannot grasp verification transcendent conditions for a sentence to be true. But this claim does go beyond the meaning theory. From the meaning theoretical argument we can conclude that no sentence apt for use in communication can have truth conditions as its meaning, and therefore cannot be said to *express* truth conditions. A further step is required for concluding that it isn't possible for speakers to *grasp* such truth conditions anyway, and associate them with sentences. And even if that step is taken, by means of the idea that we cannot think what we cannot express, the same conclusion could be applied to the (TDEF) conception of truth conditions as to the verification transcendent ones.

## 4. Intuitionistic semantics

Given that we must give up truth conditional semantics, what is the alternative? Well, here I shall just make the further assumption that the verificationist conception of meaning that has been developed within the intuitionistic tradition is the right, or at least best alternative. According to how it is often put by Dummett and Prawitz, knowing the meaning of a sentence is knowing what is, or counts as, a proof, or verification, of that sentence.[15] This can be put into the standard form in the following way. We have the general relation *x is a proof of y*. For each sentence we have conditions for objects to stand in the proof relation to this sentence. For each object *a*, there is the property *a is a proof of y*, which may and may not apply to the sentence.

---

13. It is not, in fact, clear that (TDEF) truth conditions are acceptable to Dummett, although they are to Prawitz. In the ongoing debate between the two, Dummett has objected to the idea of an objective realm of proofs, provided it is assumed to contain proofs of sentences of which we do not already know that they are decidable. That debate continues in this issue of this journal.

14. E.g. in Dummett (1991), pp 343-51.

15. See e.g. Prawitz (1987), pp 133-47.

The model then says that I know what the sentence means provided I can tell, for any given object *a*, whether the property *a is a proof of y* applies or doesn't apply to the sentence.

This conception of meaning satisfies the requirement of manifestability, provided the proof relation itself is decidable. And so it is in general conceived to be. Sometimes it may require some effort, but if I have the understanding required, I can always tell, after a finite amount of reflection or investigation, whether or not some particular object is a proof of some particular sentence.

This general idea connects well with what has become known as the Brouwer-Heyting interpretation of the logical constants. The modern version of it employs the distinction, due to Dummett and Prawitz[16], between canonical and non-canonical proofs. For each logical constant it is explained what counts as a canonical proof of a sentence with that constant as main operator. A proof in general of a sentence is conceived of as an effective method for arriving at a canonical proof of it.

A canonical proof of a conjunction *A*&*B* is a pair *a*,*b* , where *a* is a proof of *A*, and *b* is a proof of *B*, and a proof of a conjunction is then an effective method for arriving at such a pair. A canonical proof of a disjunction *A* *B* is either a pair 1,*a* , where *a* is a proof of *A*, or a pair 2,*b* where *b* is a proof of *B*. A canonical proof of a conditional $A \rightarrow B$ is a function which for any proof of *A* as argument gives a proof of *B* as value. A canonical proof of a negation $\neg A$ is a function which takes any proof of *A* into a proof of an absurd sentence (often represented by the absurdity constant ). A canonical proof of an existential sentence *xA* is a pair *a*,*b* such that *a* is a term and *b* a proof of $A(^a/x)$ (alternatively: such that *a* is an object and *b* a proof of *Ax* given the assignment of *a* to *x*). Finally, a canonical proof of a universal sentence *xA* is a function which takes any term *a* into a proof of $A(^a/x)$ (similar alternative).

These are then the *meanings* of the (first order) logical constants. For any logically complex sentence we can say something about its meaning on this basis, viz. about the general form of a canonical proof of that sentence.

The next question is how this relates to logical validity, or, more generally, logical consequence. There is no general agreement among intuitionists about how to define this notion. In Prawitz (1985) it is explained by way of the notions of valid inference rules and valid arguments. Logically valid arguments are built up in a systematic way from applications of valid inference rules, and are independent of proof systems (and thus interpretations) for atomic formulas. Valid inference rules are of two kinds. On the one hand there are the introduction rules, which are valid or justified by directly corresponding to canonical proofs (valid closed canonical arguments). On the other hand there are rules which are justified with respect to the introduction rules. The elimination rules are of this kind. Finally, it is said that *A* is a logical consequence of Γ if, and only if, there is a logically valid argument with *A* as conclusion and sentences in Γ as premisses.

A closely related idea, which dispenses with the reference to inference rules and arguments, is this: A consequence schema Γ, *A* with *A* and the sentences in Γ schematic, is logically valid iff there is a general proof that for each substitution instance there is a function which takes proofs of members Γ as arguments and gives a proof of *A* as value. Then a sentence *A* is a logical consequence of sentences Γ iff they instantiate a logically valid consequence schema. In particular, then, a sentence is logically valid iff it instantiates a schema for which there is a proof that each substitution instance has a proof.[17]

16. Dummett (1980a) pp 240-47, Prawitz (1974) pp 63-77.

As an example, take the schema $(A\&B) \rightarrow A$. Let’s say that RUN is a function or operation which, when applied to a method of obtaining a canonical proof of a particular sentence $A$, yields a canonical proof of $A$ which the methods provides. And let L be the function which takes every pair into the left element of that pair. Then the composition of RUN and L is a function which takes any proof of a conjunction as argument and gives as value the left element of a canonical proof of the conjunction, and that qualifies as a proof of any instance of the schema $(A\&B) \rightarrow A$. So every instance is logically valid.

And now it is straightforward to see that there is no corresponding entity available for $A \quad \neg A$, for there is no guarantee that (for every instance of this schema) there is either a proof of $A$ or a proof of $\neg A$. Hence, in this framework the law of excluded middle isn’t valid.

## 5. Truth and bivalence for an intuitionistic object language

These characterizations of logical validity are made without employing the concept of truth, and therefore it is not immediately clear how that concept should be connected with the intuitionistic meaning explanations, and the corresponding notions of validity. A natural idea, however, is to connect the concept of truth with provability.

In the present section I shall consider a truth definition, based on this idea, for an intuitionistic object language, i.e. for a language $L$ for which the intuitionistic meaning explanations hold. We will see that a truth definition be given such that bivalence fails for $L$, but such that the law of excluded middle remains valid in the meta-language. Because of this, no metaphysical consequences may be inferred from the failure of bivalence for $L$, since in virtue of the validity of excluded middle in the meta-language, we would be justified in claiming that facts are fully determinate.

Such a truth definition can be given, as can be seen from results about the relation between intuitionistic and classical logic. That intuitionistic propositional logic in a sense is *interpretable* in classical logic was first shown by Gödel in 1933.[18] Interpretability here means the existence of a mapping $^+$ from sentences of an intuitionistic system to sentences of a classical system such that the derivability relation is preserved:

$$\text{If} \quad \Gamma \vdash_I A, \quad \text{then} \quad \Gamma^+ \vdash_{C*} A^+$$

where C* is the classical system. The classical system is classical propositional logic extended to the modal system S4, where the necessity operator is intuitively interpreted as “it is provable that”. Gödel also conjectured the converse, which was later proved by McKinsey and Tarski[19]. With both the *if* and the *only if* part, intuitionistic logic is said to be *faithfully interpretable* in classical logic. This result was later extended to predicate logic.[20] It is proved by Prawitz and

17. If we dispense with the talk of schemata in favour of universal generalization, then we arrive at Per Martin-Löf’s conception (known by personal communication). According to his idea, a sentence is logically valid iff there is a proof of its universal generalization, with respect to all non-logical constants (of any type). Logical consequence reduces to logical validity of conditionals.

18. Gödel (1933). An introductory note by A.S. Troelstra, pp 296-300, to the reprint of Gödel (1933) provides a survey of later results. For further historical remarks I am indebted to this note and to material in Prawitz and Malmnäs (1968).

19. McKinsey and Tarski (1948).

20. Rasiowa and Sikorski (1953), Maehara (1954).

Malmnäs by considering natural deduction systems, using Prawitz normalization techniques.[21]

With ‘N’ as the modal operator, Prawitz and Malmnäs define the mapping ° as follows:

$(Pt_1t_2\ldots t_n)^\circ$ = N$Pt_1\ldots t_n$

(A → B)° = N(A° → B°)

(*A*  B)° = *A*°  *B*°

(*A*&*B*)° = *A*°&*B*°

(  *xA* )° = N(  *xA*°)

(  *xA* )° =  *xA*°

  ° =

Then the following theorem holds:

(D1) Γ Ê$_I$ *A* iff Γ° Ê$_{CS4}$ *A*°

This result can be used for providing a recursive truth definition for *L*. We replace ‘°’ by the truth predicate and the identity sign by the material biconditional, and treat the expressions on the right hand sides of the equations as *used* rather than as mentioned, thus belonging to a meta-language. We must assume that sentences which are equivalent according to the truth definition are intersubstitutable in the modal contexts, i.e. under the provability operator. This assumption is justified, since we are allowed to treat the clauses of the definition, or rather their instances, as provable themselves (in virtue of the definition).

Given this, the truth definition, via the definition of satisfaction, for *L* looks like this:

| | | | |
|---|---|---|---|
| T(*L*) | sat (s, $P_it_1,\ldots,t_n$) | iff | N($\boldsymbol{P_i}$((*s/x*)$\boldsymbol{t}_1$,…,(*s/x*)$\boldsymbol{t}_n$)) |
| | sat (s, *A* → *B*) | iff | N(sat (s, *A*) → sat (s, *B*)) |
| | sat (s, *A*  *B*) | iff | (sat (s, *A*)  sat (s, *B*)) |
| | sat (s, *A*&*B*) | iff | (sat (s, *A*) & sat (s, *B*)) |
| | sat (s,  ) | iff | |
| | sat (s,  *xA*) | iff | N(  $s' \underset{x}{\approx} s$ (sat (s', *A*)) |
| | sat (s,  *xA*) | iff |  $s' \underset{x}{\approx} s$ (sat (s', *A*)) |

Here sat is the satisfaction relation, ‘s’ and ‘s'’are variables ranging over sequences of objects, and ‘*s/x* α’ denotes the result of replacing, for any i, every free variable ‘$x_i$’ in α, by ‘$s_i$’, denoting the object assigned to ‘$x_i$’ by s. ‘$s' \underset{x}{\approx} s$’ means that s and s' differ at most in what they assign to ‘*x*’.

Negation is defined by way of implication and absurdity, in both object language and meta-language (i.e. ¬*p* =$_{def}$ *p* →  ), and falsity is equated with truth of negation. And truth is defined for closed sentences as satisfaction by all sequences.[22]

21. Prawitz and Malmnäs (1968). The main result is due to Malmnäs.

Then it can easily be shown that if '°' is accepted as a *translation* function, then Tarski's material adequacy condition is met by T(*L*). This results from showing, by induction over complexity, that

$$\text{sat}(s, A) \quad \text{iff} \quad s/x\,(A°)$$[23].

We can also see that bivalence does not hold logically for *L*. Because of the definitions of falsity and negation, and the implication clause, it holds that

$$(\text{T}(s) \text{ or F}(s)) \quad \text{iff} \quad (\text{T}(s) \text{ or N(not T}(s))$$

and for atomic *s* we further get

$$(\text{T}(s) \text{ or F}(s)) \quad \text{iff} \quad \text{N}p \text{ or N(not N}p)$$

and the right hand side does not hold for arbitrary atomic *p* in classical S4. So bivalence fails.[24]

However, on the admissible assumption that bivalence holds for the meta-language, we also have realism. So the failure of bivalence for *some* language does not have any direct metaphysical consequences, since the possibility remains open that bivalence will still hold for our *home* language, which is what counts.

There is, however, a question about how one should view the present conclusion. After all, why does bivalence fail? One reason might be metaphysical to begin with. You may think that for metaphysical rather than logical reasons, it fails to hold for every state of affairs that it determinately either obtains or doesn't obtain. I am not going to consider that possibility here, since that view isn't based on meaning theoretical considerations.

Another reason has to do with the concept of truth. It may be a general property of any adequate concept of truth that bivalence fails. I am not going to consider that either. I argued in section 5 that this doesn't follow from Dummett's meaning theoretical argument. Since this paper is concerned with the conclusions that can be drawn from meaning theoretical considerations, I will disregard other reasons why adequate concepts of truth may have that property.[25]

What is of interest in the present context is the failure of bivalence because of meaning. And on the assumption that bivalence does hold for the meta-language but fails for the object language, the most natural conclusion is that this is so because of the meanings of the sentences of the object language. But how should we view those meanings from the perspective of the meta-language? And how should we regard the intuitionistic meaning explanations of the logical constants. After all, they are rather abstract, and doesn't tell us exactly how the intuitionistic constants should be interpreted in a meta-language for which we don't already assume that the same explanations apply.

There were once attempts at placing meta-linguistic, or meta-mathematical, interpretations on intuitionistic sentences. Arend Heyting once suggested that a mathematical sentence *expresses* an expectation of finding a proof of it.[26] The meaning of  is then recursively

---

22. T(*L*) corresponds to the model theoretic thruth definition of Kripke semantics for intuitionistic logic. Oversimplifying, it can be said that Kripke semantics itself results by applying possible worlds semantics for S4 via the mapping °. Cf Tennant (1978) sections 5.8, 6,6.

23. Here "$A°$" is to be read as an abbreviation of the image of *A* under °, not as a term denoting that sentence.

explained by saying that a sentence $A \quad B$ expresses an expectation which is fulfilled just in case at least one of the expectations expressed by *A* and by *B* is fulfilled. And Prawitz and Malmnäs write:

> From a classical point of view, the difference between intuitionistic and classical logic may be said to be that intuitionistic logic fails to make any distinction between asserting that a formula holds and asserting the provability of the formula.[27]

It is, of course, quite implausible to suppose that an intuitionist could express the proposition that it is provable that *p*, by means of some sentence *s* of his language, without being capable of expressing the proposition that *p*. It is reasonable, again from a classical point of view, and in accordance with the ° interpretation above, to say that *s* is *true* just in case it is provable that *p*, but *s* does not express precisely that thought. The natural conclusion is that, although, from the clasical point of view, *s* does have truth conditions, its meaning, or the thought it expresses, cannot be identified with those truth conditions.

Rather, it is more reasonable to follow the later intuitionistic tradition, and speak, as did Heyting himself later[28], about the conditions under which sentences are *assertible*. In general, a sentence *s* is assertible by us just in case we possess a proof of it. And in the writings of Dummett, Prawitz and Martin-Löf, although in different ways, the idea that meaning is given by assertibility conditions has played a prominent, often central role.

My way of understanding this is that, again from a classical or realist point of view, we can say that intuitionistic propositions are *restricted versions* of corresponding classical ones. Take an ordinary predicate, like "… is a horse", and suppose we have a realistic grasp of this.The intuitionistic object language counterpart of it, a predicate "*H*…", would be true of things that are *provably* horses. The predicate would not express the property of being provably a horse, since it is implausible to treat it as having a conceptually complex meaning. Rather, it can more plausibly be seen as expressing a restricted version of the property of *being* a horse, restricted to entities that provably has the property. That is, the predicate expresses a restriced version of a concept, but it doesn't express the restriction itself. That is done by the complex expression ".. is

24. This truth definition allows that $(s \quad \neg s)$ can fail to be true, but doesn't allow it to be false, since that would amount to the truth of $\neg(s \quad \neg s)$.

Tim Williamson has objected to the definition of falsity, on the ground that you can equate falsity with truth of negation only when it is "real" negation, i.e. what counts as negation from the point of view of the meta-language. When the object language negation is intuitionistic and the meta-language negation classical, the object language negation isn't "real" in this sense. On Williamson's view, falsity of a sentence should rather be identified with non-truth, given that a proposition is expressed by the sentence, in which case bivalence does not fail for the object language *L*.

However, there are several principles intuitively connected with the concept of falsity. One that I find important is that speakers of a language should be able to express the conditions under which their sentences are false, and that principle is violated by Williamson's own suggestion, for there is not always any sentence of *L* that can express the conditions under which a sentence *s* (of *L*) isn't true ("isn't" is here understood classically).

Moreover, I don't think that the objection really touches the substance of the issue. We are concerned with semantic properties of a language that are connected with the more intuitive idea of a fully determinate reality, as represented in that language. Bivalence is connected with that idea, on condition that we *can* identify falsity with truth of negation. If another notion of falsity is selected, then bivalence isn't relevant for realism, in the present sense, but then we can simply replace bivalence, as a criterion of realism, with the one that *is* relevant, viz. the principle that either *s* is true or "$\neg s$" is true. That this principle fails for *L* is clear from the T(*L*) truth definition alone, irrespective of the definition of falsity.

provably a horse" of the meta-language.

And of a singular term *t* we would say that it refers to an entity *a* just in case *a* is, in a sense, *identifiable* as the referent of *t.* It is not easy to explain what identifiability in this sense would amount to, except that it would fulfil the requirement that an intuitionistic predicate combined with terms whose referents are identifiable is a sentence that is true provided it is provable that the property expressed by the predicate belongs to the referents of the terms. As a consequence, if two terms fulfilling this condition are coreferring, then the referents are provably identical.

This understanding of atomic sentences explains, again from a realist point of view, how bivalence can fail to hold for an atomic sentence like "Mildred is a horse", as understood intuitionistically. From the realist point of view, everything either is a horse or isn't a horse, but the predicate "… is a horse", on an intuitionistic interpretation, doesn't express the (realist's) concept of a horse. And the fact that it doesn't explains why the failure of bivalence has no consequences for realism.

## 6. Truth and bivalence for an intuitionistic home language

We cannot, of course, separate bivalence and realism in this way when we are speaking about our own language. Our predicate "… is a horse" expresses our concept of a horse, and that is not a restricted version of the concept of a horse. I shall now proceed on the assumption that the meaning theoretical argument applies to a language called "English*" that is pretty much like English, and that the logical particles of this language have intuitionistic explanations. In fact, no difference can be noticed between ordinary use of English and ordinary use of English*.[29] I prefer to use this device of English* because I find the complications arising from the assumption that English itself is an intuitionistic language difficult to manage (e.g., the entire reasoning in this paper would really have to be reviewed in the light of such an assumption). I am not making any assumptions about the validity of classical logic in English (and shall, accordingly, try to avoid intuitionistically unacceptable modes of reasoning). Our question is what the world looks like, so to speak, from within English*.

We are still concerned with the third step of the argument against bivalence: the step of connecting the concept of truth with the intuitionistic definitions. Without such a step there is really no conclusion to be made about realism. We know that there is no guarantee that, for arbitrary *p*, there is either a proof of *p* or a proof of *not p*, but without any further principle it doesn't follow that there is no guarantee that either *p* or *not p* is *true*.

We can add the principle that a sentence is *assertible* just in case we are in possession of a proof of what it says. But again, without any further principle we cannot conclude that a statement isn't *true*, just because it isn't assertible. In particular, the fact that we would not be justified, on the basis of the meanings of the logical particles, and a definition of validity in terms of proofs, to *assert* the law of excluded middle, does not by itself imply that there is any problem

---

25. It may be noted that both Prawitz and Martin-Löf, although in somewhat different ways, adhere to the (TDEF) conception of truth presented in section 3. See e.g. Martin-Löf (1984), p 11, Prawitz (1980), p 8, (1994) p 85.

26. Heyting, A, (1931), pp 113-115.

27. Prawitz and Malmnäs (1968), p 221.

28. Heying, A, (1971), chpt VII.

29. We may suppose that some speakers of *L* occasionally lapse into illegitimate employment of classical modes of reasoning.

about the *truth* of each instance of *A or not A*. In the previous section we gave a truth definition which did connect truth and provability, to the effect that bivalence had to be rejected. We need something analogous now.

What we need is a principle which provides a reason for doubting bivalence, given the non-assertibility of excluded middle. The obvious candidate for such a principle is

(PRO) if *A* is true, then there is a proof of *A*

Clearly, given that (PRO) holds, and if bivalence holds (and falsity equals truth of negation), then for each sentence *A* there is either a proof of *A* or a proof of *not A*. Since there is no general guarantee that this holds for every sentence *A*, we have a reason to doubt the validity of bivalence. So (PRO) is a principle of the kind we need.[30] Of course, (PRO) must be considered either as a sentence schema, or as implicitly expressing universal generality with respect to sentences or propositions.

Clearly, (PRO) is valid if truth is defined as

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

But we don't wan't the matter decided in advance by some biased notion of truth (and in fact it cannot be, for we cannot simply *decree* that the defined notion shall have metaphysical relevance). However, I think we can circumvent any problem arising from the possibility of competing concepts of truth, by simply assuming that the disquotational schema

(DQ) "*A*" is true iff *A*

holds for English* (and English). The disquotational schema should be valid for non-problematic sentences (e.g. not containing non-referring singular terms) on any reasonable conception of truth.[31] Formally, we can introduce the truth predicate in English* by means of specifying what counts as a canonical proof of a sentence of the form "*A* is true", so that the disquotational schema comes out valid. This is easy. Take some designated object or constant T, chosen for the purpose, and say that a canonical proof of *A is true* is a pair T,*a* where *a* is a proof of *A*. Then there is a function taking any proof of *A* into a proof of *A is true* (the function the forms the T-pair from the proof) and a function doing the converse (selecting the right element). The pair of these functions is the proof we need.

Given the disquotational schema we can reformulate the question as the question of the validity of

(PRO*) if *A* , then there is a proof of *A*

In fact, I am going to treat these two principles on a par, since with (DQ) the one is true if the

---

30. There are other candidates. For instance, there is the double negation of (PRO), which is classically equivalent, but intuitionistically weaker. As far as I can see, however, the arguments presented below against the assertibility of (PRO) apply equally to other candidates.

31. In Martin-Löf's type theory, (DQ) is ruled out, since there is an absolute difference between "propositions", to which the concept of truth is applied, and "judgements", which are about propositions. As far as I can see, this does not affect the present discussion.

other is.

Does (PRO), then hold in English*? This does seem *prima facie* plausible. For under the intuitionistic interpretation of “if, then”, each instance of (PRO) is provable, and hence true, provided there is a function *f* such that given any proof of the antecedent, a proof of *p*, *f* delivers a proof of the consequent, i.e. a proof that there is a proof of *p*. It seems that all that is needed for the proof of the consequent is existential generalization, to the conclusion that there is such an entity. But this is mistaken. The argument for the function *f* is just the proof of *p*, say *a*, *not* a proof that *a is* a proof of *p*. The assumption we make on this meta-level discussion, that *a* is a proof of *p*, is not itself an argument for the function *f*.

What the function *f* should do, for a particular *p*, is to take as argument an object *a*, and produce as value an ordered pair *b*,*c* such that *b* is a proof of *p*, and *c* is a proof that *b* is a proof of *p*. Is there such a function?

For any particular *p*, we have good reason to think that the answer is yes. Proofs (or verifications) are entities which, at least on the intuitionistic conception, are recognizable as such. The idea of a proof which cannot be effectively recognized as a proof is not acceptable in the intuitionistic tradition. Therefore, if there is a proof *b* of *p*, it is also possible to devise a proof *c* that *b* is a proof of *p*. And then there is a function which, for any proof *a* of *p* as argument, gives as value the pair *b*,*c*. This is what is needed.

So, based on the argument in the preceding paragraph, we can state that for each instance *p* of (PRO), there is a proof of *p*. Let’s call this argument *R*. Now, since we have realized that for each instance of (PRO), there is a proof of that instance, we should be able to conclude that we have a proof of (PRO) itself, viz. *R*. However, *R* is not a constructive argument. A constructive argument which is a proof of (PRO) is an argument which, for each instance of (PRO), actually *provides* a proof of that instance. *R*, on the other hand, does not, and only provides a *non*-constructive reason for believing that *there is*, for each instance, a proof of that instance. Because of that, “there is”, in the context of *R*, cannot be understood intuitionistically. And therefore, *R* is not really available as an acceptable argument in English*, where the logical particles have intuitionistic meanings.

A proof of (PRO) which is acceptable in English* is to be a function *F*, which, for each instance, does yield, as value, a proof of that instance. And this is impossible.

To see this, note first that this function must be constructive. It must be given by a finite set of rules or clauses that suffice for proving all instances. This requires that the two-place predicate “*x* is a proof of *y*” be formally well defined for all proofs and propositions. Since it is also required to be decidable, it must, by Church’s thesis, have a recursive definition.[32] But, by Gödel’s first incompleteness theorem, this is impossible. Any recursive definition of the proof relation can be represented in a formal system, and since the language in this case must be rich enough for arithmetic, it follows by Gödel’s theorem that there will be a sentence *q* (a Gödel sentence, or rather a Rosser sentence) in this language such that neither *q* nor the negation of *q* is provable by the axioms given, provided the system is consistent. Moreover, *q* is of the form *xAx*, where *A* is decidable. Since its negation isn’t provable, no instance $\neg At$ will be provable, and since *A* is decidable, this means that every instance *At is* provable. But then *xAx*, i.e. *q*, is true, and by means of this reflection we do have a proof of *q*. And this contradicts the assump-

32. That reliance on Church’s thesis may not be unproblematic has been stressed to me by Neil Tennant and Tim Williamson. As far as I understand the issue is complex and difficult, and presently I have no definitive answer to give.

tion that "*x* is a proof of *y*" was defined for all proofs and propositions.[33]

That the domain of proofs cannot be recursively defined is well known. Dummett himself has remarked that there cannot be such a thing.[34] The domain of proofs must be essentially open, or as Dummett expresses it, indefinitely extensible. New concepts, and new acceptable methods of proof can always be introduced. And for this reason, no constructive function can ever be sufficient for every possible proof, or for some proof of every possible proposition.

This means that (PRO) isn't provable. Moreover, given the intended *generality* of (PRO), and the requirement of a recursive definition of "*x* is a proof of *y*", there can in fact be no sentence in any language that expresses what was meant to be expressed by (PRO). (PRO) is meaningful only if the generality is restricted, or if the requirement of a recursively well defined predicate "*x* is a proof of *y*" is given up and the predicate understood as expressing our general, intuitive notion of a proof.

Since neither (PRO), nor any similar principle which relates truth to provability, can be asserted in English*, no reason can be provided in English* for doubting bivalence, or for doubting that all instances of *A or not A* are true. For then the fact that something isn't assertible in English* does not provide a reason for believing that it isn't true. And because of that, there would be no reason, based on the meanings of the logical constants and assertible in English*, for rejecting realism. To be sure, there might be no reason to affirm it either, but the kind of reason we use to have for doubting or rejecting realism would not be available.

## 7. Conclusion

In the case of the reality of the past, I complained that when a simple truth conditional idea of the meaning of sentences about the past is given up, then a failure of bivalence need not have any metaphysical consequences. The failure may depend on properties of the language only. The response, on Dummett's behalf, was that the past isn't anything beyond what we can speak of by means of our sentences about the past, and so if bivalence fails for these sentences, there is no further possibility that the past itself might yet remain fully determinate.

My reply, then, is that in order to show that bivalence *does* fail, for sentences about the past, because of meaning theoretical considerations, some version of (PRO), like

(PAST) If *p*, then there is a present verification of *p*.

where *p* can be replaced by sentences about the past, must be established. But if the instances of (PAST) have sentences about the past as constituent parts, then those instances must be gov-

33. Actaully, all that is required in this case is a predicate "*x* is a proof of *y*" defined for at least one proof of every provable proposition, not necessarily for every proof of every provable proposition. But this makes no difference for the incompleteness.

Some, although probably not any intuitionists, might object to the final part of the argument, i.e. that by means of reflecting over the incompleteness we can conclude that *q* is true. For there are non-standard models of arithmetic where the Gödel sentence is false.

However, even if this is accepted, a contradiction is easily derived by using the *reductio* assumption of the provability of (PRO) itself. For it follows from (PRO) that if *q* is true, then it is provable. By the incompleteness (and the existence of non-standard models) it isn't provable. So it isn't true. But this reflection gives us a proof of the negation of *q*, again contradicting the incompleteness, given the assumptions about the proof predicate.

34. Dummett( 1980b), pp 200-201.

erned by the same verificationist semantic principles as those sentences are governed by. To establish (PAST), then, a very general verification, of the kind discussed in section 6, is needed. But given that our language for speaking about the past is a rich as our language for speaking about the present, we have the same result as in section 6. There cannot be any definition of "*x* is a verification of *y*" that is defined well enough to make a verification of (PAST) possible.

Alternatively, we might go to a meta-level, and rather consider

(META-PAST) If *s* is true, then there is a present verification of *s*.

But then the problem is that in order to establish (META-PAST) we must provide a truth definition for sentences about the past. There is no general argument from meaning to truth unless the meaning of a sentence determines its truth conditions (and the concept of truth isn't just arbitrarily chosen). But if (META-PAST) is established by means of a truth definition, then there will have to be a meta-language in which we can also speak of the past, thus speak indepentently of past reality and of (object language) sentences about the past. And since that meta-language may have a bivalent semantics, we can retain realism about the past, and blame just the object language for the failure of bivalence. And so I conclude that the response, on Dummett's behalf, is flawed.

The Dummett-Prawitz claim was that the principle of bivalence has a double significance. On the one hand it needs a justification from the meanings of the logical vocabulary, which is its meaning theoretical significance, and on the other hand acceptance or rejection of it (for a range of statements) is acceptance or rejection of realism, respectively (for that range of statements, or for what they are about). But in order for such a double significance to be possible, meaning must be connected with truth in a relevant and non-trivial way. When the meanings are intuitionistic, this comes down to asserting (PRO). But, as argued in section 6, doing it *within* the language of which it shall hold isn't possible, and, as argued in section 5, doing it from *outside*, in a meta-language, isn't good enough.

This is not a challenge to intuitionism as such. You may think that intuitionism is the right framework for mathematics just because you think that there isn't any readymade, external and fully determinate mathematical reality. That is a reason for preferring the intuitionistic meaning explanations of mathematical vocabulary, and of logical vocabulary in mathematics. From the point of view of the present discussion, there is nothing problematic about that.

The problems arise only when metaphysical conclusions are to be derived from the semantics of an intuitionistic mathematical language. For the fact that there is *one* mathematical language for which bivalence fails does not imply that there is no other mathematical language for which it holds. It must be shown that it does fail for our mathematical home-language, and just that cannot be done by the meaning theoretical arguments.[35]

Department of Philosophy
Stockholm University

## References

Dummett, M, (1976), 'What is a theory of meaning (II)' in Evans G and McDowell J (eds), *Truth and Meaning*, Oxford University Press,Oxford.

— (1980), *Truth and Other Enigmas*, 2nd ed, Duckworth, London.

— (1980a), 'The philosophical basis of intuitionistic logic', in Dummett (1980).

— (1980b), 'The philosophical significance of Gödel's theorem', in Dummett (1980).

— (1982) 'Realism', *Synthese* 52, pp 55-112.

— (1991), *The Logical Basis of Metaphysics*, Duckworth, London.

Gödel, K, (1933), 'Eine interpretation des intuitionistischen Aussagenkalküls', *Ergebnisse eines matematischen Kolloquiums* 4, pp 39-40. Reprinted and translated in *Kurt Gödel: Collected Works*, Feferman & alt. (eds), Oxford University Press, Oxford 1986, pp 300-303.

Heyting, A, (1931), 'Die intuitionistische Grundlegung der Mathematik', *Erkenntnis* 2, 1931, pp 106-115, translated in Benacerraf, P. and Putnam, H. (eds), *Philosophy of Mathematics. Selected Readings*, Cambridge, Mass: Cambridge University Press, 1964.

— (1971), . *Intuitionism. An introduction*, 3rd edition, North-Holland, Amsterdam 1971. First edition from 1956.

Maehara, S, (1954), 'Eine darstellung der intuitionistischen Logik in der Klassischen', *Nagoya mathematical journal* 7, 45-64.

McKinsey, J.C.C, and Tarski, A, (1948), 'Some theorems about the sentential calculi of Lewis and Heyting', *Journal of Symbolic Logic* 4, pp 1-15.

Martin-Löf P, 1984, *Intuitionistic Type Theory*, Bibliopolis, Napoli.

Prawitz, D, (1974), 'On the idea of a general proof theory', *Synthese*, vol 27, pp 63-77.

---

35. This paper evolved from an earlier, rather different paper, written in the fall of 1996 (while I was fellow at SCASSS – The Swedish Collegium for Advanced Study in the Social Sciences) and called "The metaphysical axiom of intuitionism". The title referred to what is here called (PRO), and the paper consisted of an examination of two possible ways of arguing for (PRO). It concluded that both failed, and that therefore (PRO) should have to be regarded as an axiom. Two events forced me to reconsider the issue substantially.

First, by way of email comments Tim Williamson suggested that (PRO) could actually be justified by a way I had overlooked, namely from the intuitionistic interpretations of the logical constants. It was a shock to me to realize that I had overlooked what should have been first in line to consider. I soon set out to argue against the suggestion, along the present lines, but only later did I become clear about the central role of that argument, and its precise form.

Second, the first paper was presented at a Stockholm seminar (within the framework of an interdisciplinary Swedish project, *Meaning and Interpretation*, financed by a grant from the Anniversary Foundation of the Swedish Central Bank). Both Dag Prawitz and Per Martin-Löf were present. To my surprise, they accepted my arguments, and even in a sense my conclusion. (PRO), or rather what is here called (TDEF), was proposed by them as a definition of truth, without any explicit arguments behind it. So I had clearly missed my target.

Both events eventually led to the present paper. For very helpful comments on earlier drafts of it I am very much indebted to Matti Eklund, Per Martin-Löf, Göran Sundholm and Dag Westerståhl. Without their help (especially Dag's) this paper would have been much worse.

A penultimate draft was presented at the *Logic and Language* conference in London, in April 1998. For comments at the conference I am much indebted to Barry Smith, Neil Tennant and Tim Williamson.

The paper is intended as a tribute and challenge to my colleague and friend, and former supervisor, professor Dag Prawitz, and is a testimony to the power of the Dummett-Prawitz stand in philosophy, for only what is powerful deserves a careful attack. I hope I haven't missed my target this time.

— (1980), ‘Intuitionistic logic: a philosophical challenge’, in von Wright (ed), *Logic and Philosophy*, Martinus Nijhoff, The Hague.

— (1985), ‘Remarks on some approaches to the concept of logical consequence’, *Synthese* 62, pp 153-171.

— (1987), ‘Dummett on a theory of meaning and its impact on logic’, in Taylor, B, (ed), *Micahel Dummett: Contributions to Philosophy*, *Martinus Nijhoff*, Dordrecht.

— (1994), ‘Meaning theory and anti-realism’, in McGuinness, B, and Oliveri, G, (eds), *The Philosophy of Michael Dummett*, Kluwer, Dordrecht.

Prawitz, D and Malmnäs, P-E, (1968), ‘A survey of some connections between classical, intuitionistic and minimal logic’, in Schütte, K (ed), *Contributions to Mathematical Logic*, North-Holland, Amsterdam

Quine, W.V, (1960), *Word and Object*, The MIT Press, Cambridge, Mass.

Rasiowa, H, and Sikorski, R, (1953), ‘Algebraic treatment of the notion of satisfiability’, *Fundamenta Matematicae* 37, 193-200.

Tennant, N, (1978), *Natural Logic*, Edinburgh University Press, Edinburgh.

Wright, C, (1976), ‘Truth conditions and criteria’, in *Proceedings of the Aristotelian Society*, Supplementary Volume L. Reprinted in Wright (1987).

— (1987), *Realism, Meaning and Truth*, Basil Blackwell, Oxford.