Bivalence and Determinacy

Rumfitt, Ian

BIVALENCE AND DETERMINACY*

Ian Rumfitt

Abstract

The principle that every statement is bivalent (i.e. either true or false) has been a bone of philosophical contention for centuries, for an apparently powerful argument for it (due to Aristotle) sits alongside apparently convincing counterexamples to it. I analyse Aristotle's argument (§§1-2), showing that it relies crucially on the logical laws of Excluded Middle and Proof by Cases. Even given these logical laws, however, the argument only shows that every *determinate* statement is true or false, where a determinate statement 'says one thing', i.e. has univocal truth-conditions. In the light of this analysis, I examine three sorts of problem case for bivalence. Future contingents, I contend, are bivalent (§3). Certain statements of higher set theory, by contrast, are not. *Pace* the intutionists, though, this is not because Excluded Middle does not apply to such statements, but because they are not determinate (§§4-6). Vague statements too are not bivalent, in this case because the law of Proof by Cases does not apply (§§7-8). I show how this opens the way to a solution to the ancient Paradox of the Heap or *Sorites* (§9) that draws on quantum logic.

Keywords

Bivalence, future contingents, Continuum Hypothesis, Georg Kreisel, vagueness, quantum logic, Paradox of the Heap, Sorites Paradox

---

* I am grateful to the editor of this volume for his comments on a draft.

The Principle of Bivalence is the thesis that every statement is either true or false. By a statement, I mean a declarative utterance or inscription that expresses a complete thought. Statements, in this sense, need not be assertions: the antecedent of the conditional 'If Tom is in Germany, he is in Berlin' expresses the thought that Tom is in Germany and hence qualifies as a statement, even though a speaker who affirms the whole conditional does not assert that Tom is in Germany. The notion of expressing a complete thought is not completely clear and, as the sequel will show, the truth of the Principle depends, among other factors, on how it is clarified. But it is clear enough to identify our topic.

The Principle of Bivalence has been a bone of philosophical contention for centuries. An apparently powerful argument for it sits alongside apparently convincing counterexamples to it. I start with the classic argument for the Principle.

1. The Simple Argument for Bivalence

The argument I have in mind is all but explicit in Aristotle. Everybody remembers Aristotle's explanations of truth and falsity in *Metaphysics* Γ.7, but it is sometimes forgotten that they form part of an argument for Bivalence:

> Of one subject we must either affirm or deny any one predicate. This is clear, in the first place, if we define what the true and the false are. To say of what is that it is not, or of what is not that it is, is false, while to say of what is that it is, and of what is not that it is not, is true; so that he who says of anything that it is, or that it is not, will say either what is true or what is false (*Metaphysics* 1011b25; cf. Plato, *Cratylus* 385b2).

The idea behind these definitions is this: an utterance is true if it says that things are somehow, and they are thus; it is false if it says that things are somehow, and they are not thus. The non-nominal quantifier 'somehow' is one that English speakers understand, but in tracing out the implications of the proposed definitions, we shall need a rigorous statement of the rules governing its use. What are those rules?

In his book *Objects of Thought* (1971), Arthur Prior gave an answer. Let us suppose that a standard first-order formalized language is enriched with universal and existential quantifiers whose attendant variables replace complete well-formed formulae. Let us also suppose that the introduction and elimination rules for these two new 'propositional' quantifiers are analogues of the rules that govern the corresponding quantifiers into name position, aside from differences consequential upon the different syntactic categories of the associated variables. Then, for example, assuming that the underlying propositional logic is classical, the formulae '$\forall P\,(P \vee \neg P)$' and '$\neg\exists P(P \wedge \neg P)$' will be logical theorems. What Prior noticed is that if we read '$\forall P$' as 'However things may be said or thought to be', read '$\exists P$' as 'There is a way things may be said or thought to be', and read the associated variables as 'they are thus', then the theorems of this system emerge as logical truths under the proposed interpretation. Thus the theorem '$\forall P\,(P \vee \neg P)$' says 'However things may be said or thought to be, either they are thus or they are not thus', which a classical logician will take to be a logical truth. Reversing the translation, we can formalize Aristotle's definitions of truth and falsity as follows, where '$\forall u$' is an ordinary objectual quantifier ranging over utterances (which I shall henceforth take to include inscriptions), and '$\exists P$' is the existential propositional quantifier lately explained:

$$(T) \qquad \forall u(\text{True}\ (u) \leftrightarrow \exists P(\text{Say}\ (u, P) \wedge P))$$

$$(F) \qquad \forall u(\text{False}\ (u) \leftrightarrow \exists P(\text{Say}\ (u, P) \wedge \neg P)).$$

(*T*) and (*F*), then, capture Aristotle's definitions of truth and falsity in a way that enables us rigorously to trace out their implications.\[1]/

Aristotle presents his formulae as definitions, and (*T*) and (*F*), if they are correct at all, can serve as explicit definitions of 'true' and 'false' as these notions apply to utterances.\[2]/ Just for this

[1] Read strictly, Aristotle affirms only the right-to-left halves of (*T*) and (*F*), but the converse conditionals are implied by his advertising his formulae as *definitions* of truth and falsity. If I say 'This result is clear if we define what an Abelian group is. If a group is commutative it is Abelian', my second sentence has the force of a biconditional.

[2] Corresponding formulae could serve as definitions of truth and falsity as these notions apply to states of belief and to other bearers of truth and falsehood. For some discussion of the range of truth-bearers, see Rumfitt 2011 §2.

reason, though, it might seem as if (*T*) and (*F*) cannot be correct: they appear to conflict with Alfred Tarski's theorem that truth is indefinable (Tarski 1935). Tarski's theorem applies to a formalized language, *L*, containing a negation operator $\neg$ and a device $\ulcorner\ \urcorner$ which, when applied to any well-formed expression of *L*, yields a singular term designating that expression. It is assumed that the syntax of the language is strong enough to prove the diagonal lemma: for any formula $A(x)$ in *L*, with *x* free, there is a formula *B* in *L* such that *B* is equivalent to $A(\ulcorner B \urcorner)$. According to Tarski, a *truth-predicate* for *L* is a one-place predicate $Tr(\xi)$ such that, for any closed formula *A* of *L*, $Tr(\ulcorner A \urcorner)$ is equivalent to *A*: Tarski assumes that *L* contains no context-sensitive expressions, which is why he can take a truth-predicate for *L* to be a unary predicate of *L*'s type sentences. What Tarski then proves is that no truth-predicate *for L* can be a predicate *in L*. For suppose $Tr(\xi)$ were a truth-predicate for *L* in *L*. Then $\neg Tr(x)$ would be a well-formed formula of *L* with *x* the only free variable. By the diagonal lemma, there would exist a closed formula *D* in *L* such that *D* is equivalent to $\neg Tr(\ulcorner D \urcorner)$. Since $Tr(\xi)$ is a truth-predicate for *L*, we would also have that $Tr(\ulcorner D \urcorner)$ is equivalent to *D*, so that $Tr(\ulcorner D \urcorner)$ and $\neg Tr(\ulcorner D \urcorner)$ would be equivalent. This contradiction reduces to absurdity the supposition that $Tr(\xi)$ is a truth-predicate in *L*. So no truth-predicate for *L* can be in *L* (Tarski 1935, 249-51). Although Aristotle does not of course attain modern standards of explicitness in such matters, he does seem to advance, in Greek, an account of a truth-predicate which applies (among other things) to Greek utterances and inscriptions, so Tarski's theorem appears to cast doubt on his account. It also casts doubt on (*T*) and (*F*), insofar as the predicates they purport to define are predicates of a semi-formalized version of English which apply, *inter alia*, to utterances and inscription in that semi-formal language.

Given Tarski's assumptions, his proof of the indefinability theorem is unassailable, but one of those assumptions is highly contestable. Tarski assumes that if $Tr(\xi)$ is a truth-predicate for *L*, then any closed formula *A* of *L* is equivalent to $Tr(\ulcorner A \urcorner)$. He thereby presupposes that any closed formula of *L* has truth-conditions. A formula will have truth-conditions, though, only if it expresses a complete thought, so Tarski's argument presupposes that any closed formula of the relevant language expresses such a thought. The very formula that Tarski uses in proving his theorem, however, casts doubt on that presupposition. That formula 'says of itself' that it is not true. To see why Tarski's presupposition is doubtful, then, let us consider a corresponding English inscription, λ:

The inscription on the top line of page 5 of this essay is not true.

When applied to λ, (*T*) yields

1. $\text{True}(\lambda) \leftrightarrow \exists P(\text{Say}(\lambda, P) \wedge P)$.

Now *if* λ says anything, what it says is that the inscription on the top line of page 5 of this essay is not true, so on the assumption that λ does succeed in saying something, (1) yields

2. True (λ) ↔ the inscription on the top line of page 5 of this essay is not true.

By inspection we also have

3. λ = the inscription on the top line of page 5 of this essay.

By Leibniz's Law, (2) and (3) together entail

4. $\text{True}(\lambda) \leftrightarrow \neg\text{True}(\lambda)$.

As in Tarski's proof of his indefinability theorem, (4) is a contradiction, but we need not take this as reducing to absurdity the assumption that semi-formal English contains a predicate with the intended sense of 'True'. In order to move from (1) to (2), we need the assumption that λ expresses a thought and we may, instead, take the contradiction to refute that assumption. Tarski entirely overlooks this possibility. He writes that we 'wish to use the term "true" in such a way that *all* [T-equivalences in the form "*S* is true if and only if *P*"] can be asserted, and we shall call a definition of truth "adequate" if [and only if] *all* these equivalences follow from it' (Tarski 1944, 344, emphasis added). But what this sort of case reveals is that our commitment to assert such T-equivalences is provisional. We shall assert '*S* is true if and only if *P*' only when we believe that *S* expresses a thought; in particular, we shall assert (2) only when we believe this of λ. For all that Tarski says, the above deduction is simply a proof by *Reductio* that λ fails to express a thought, i.e. that $\neg\exists P\ \text{Say}(\lambda, P)$. On this view, (*T*) and (*F*) transform ‘the semantic paradoxes…into sound arguments for constraints on what can say what in what contexts’ (Williamson 1998, 19).

Whether this approach gives the best solution to the Paradox of the Liar is a large issue that I cannot address here.\[3]/ It is clear, though, that the theory of truth and falsity that comprises (*T*), (*F*) and their logical consequences is formally consistent. As Timothy Williamson has observed, we can show this ‘by constructing an unintended model…in which formulas are treated as referring to truth-values, the propositional quantifiers range over truth-values, and all formulas of the forms “Say (A, c, P)”, “True (A, c)”, and “False (A, c)” are treated as false’ (Williamson *op. cit.*, 14).

Assuming that the logic of this theory is classical, the Principle of Bivalence seems, at first blush, to be a theorem of it. For, given an arbitrary statement, u, we can reason as follows:

(1)	$\exists P$ Say (u, P)	Definition of statement
(2)	Say (u, P)	(1), existential instantiation
(3)	$P \vee \neg P$	Excluded middle
(4)	P	Assumption
(5)	Say (u, P) $\wedge$ P	(2), (4), $\wedge$-introduction
(6)	$\exists P$ (Say (u, P) $\wedge$ P)	(5), existential generalization
(7)	True (u)	(6), (*T*)
(8)	True (u) $\vee$ False (u)	(7), $\vee$-introduction
(9)	$\neg P$	Assumption
(10)	Say (u, P) $\wedge$ $\neg P$	(2), (9), $\wedge$-introduction
(11)	$\exists P$ (Say (u, P) $\wedge$ $\neg P$)	(10), existential generalization
(12)	False (u)	(11), (*F*)
(13)	True (u) $\vee$ False (u)	(12), $\vee$-introduction
(14)	True (u) $\vee$ False (u)	(3), (8), (13) $\vee$-elimination, with discharge of assumptions (4) and (9)

[3] For approaches to the Liar along these general lines, see Prior 1971 chap. 6, Kneale 1972, Mackie 1973 chap. 6, Parsons 1974, Smiley 1993, Glanzberg 2001, and Rumfitt 2014.

Since u was an arbitrarily chosen statement, we may generalize to reach the conclusion ‘Every statement is either true or false’. I shall call this deduction the *Simple Argument* for Bivalence. It spells out the argument that is implicit in the passage I quoted from Aristotle.

2. Problems with the Simple Argument: determinacy of sense

The Simple Argument may seem to be conclusive, at least if the underlying logic is classical. In fact, though, there is a serious problem with it.

The Principle of Bivalence, I said, is the thesis that every statement is either true or false, but the ‘or’ here is exclusive, not the logician’s *vel*. That is, the Principle might more explicitly be formulated as ‘Every statement is either true or false but not both’. However, even in classical logic, which has ‘$\neg\exists P(P \wedge \neg P)$’ as a theorem, ($T$) and ($F$) leave open the possibility that a statement is both true and false. A statement might be true by virtue of saying that P in a circumstance where P, and at the same time false by virtue of saying that Q in a circumstance where not Q. Where ‘P’ and ‘Q’ are distinct formulae, the truth of ‘$P \wedge \neg Q$’ may well be a logical possibility. Even given classical logic, then, nothing in the Aristotelian definitions (T) and (F) excludes the possibility that a statement is both true and false.

What would exclude that possibility is the further assumption that the statement has a *determinate sense*. Any statement, we said, expresses *a* complete thought. But if it has a determinate sense, it will express just one such thought: there will be such a thing as *the* thought it expresses. While there may be alternative ways of expressing that thought, those ways cannot diverge in truth value. If u has a determinate sense, then, it will meet the following condition:

$$(D) \qquad \forall P \forall Q(\mathrm{Say}\ (u, P) \wedge \mathrm{Say}\ (u, Q) \rightarrow (P \leftrightarrow Q)).$$

If u has a determinate sense it will meet stronger conditions too, such as the following necessitated form of (D):

$$\Box\ (\forall P \forall Q(\mathrm{Say}\ (u, P) \wedge \mathrm{Say}\ (u, Q) \rightarrow \Box\ (P \leftrightarrow Q))).$$

In the present essay, though, I am not concerned with the modal aspects of statements, so we may focus on (*D*), which I shall take to be the defining condition for *u* to be *determinate*. Given (*T*) and (*F*), and the premiss that *u* satisfies (*D*), it is straightforward to show that *u* cannot be both true and false.

How might a statement fail to be determinate? How could an utterance say something without there being one thing that it says? We should set aside as irrelevant a shallow sense in which this is possible. If I say 'It is wet and cold', I may be reported as having said that it is wet, and as having said that it is cold. In the intended sense, though, 'Say (u, P)' means 'The *whole* of what *u* says is that *P*'; partial reports of an utterance's content are beside the point.

Even in the intended sense of 'Say (u, P)', however, (*D*) may fail. As we shall see in more detail below, special problems attend the use of classical logic when this is applied to statements involving unrestricted quantification over sets. The usual semantic justification for the use of classical logic when reasoning with first-order statements presumes that the domain of quantification constitutes a set. In standard set theory, however, there is no set of all sets. An alternative justification for the use of classical logic in set theory rests on the view that apparently unrestricted quantification over sets is implicitly restricted to the members of a standard model for set theory. Since there are many such models, however, this view casts doubt upon (*D*). For what, on this view, does the statement 'There exists a strongly inaccessible cardinal' say? It 'says' that there is such a cardinal in the smallest standard model $V_{\kappa 1}$; but it equally 'says' that there is such a cardinal in the second smallest model $V_{\kappa 2}$, and so forth. The claims that our statement 'says' are not even materially equivalent. The first inaccessible, κ_1, is a member of $V_{\kappa 2}$ but not of $V_{\kappa 1}$ so, on the first reading, the statement is false whereas it is true on the second reading. I put the word 'say' in scare quotes because it is somewhat strained to use the term in this way: we want to ask, 'Well, *which* does it say?' On the view in question, though, the only available sense for '*u* says that *P*' is that of 'An interpretation of *u* under which it says that *P* is as legitimate as any other interpretation'. The existence of equally legitimate ways of interpreting *u* should not prevent us from classifying *u* as a statement: it is not that it fails to say anything. All the same, in such a case no *unique* content can be ascribed to it.

For statements that exhibit this kind of indeterminacy, (*T*) is inadequate as an account of truth. If the only available sense for '*u* says that *P*' is 'An interpretation of *u* under which it says that *P* is as legitimate as any other interpretation' then the only available sense of the formula '$\exists P(\text{Say}\ (u, P) \wedge P)$'

is ‘An interpretation which renders u true is as legitimate as any other interpretation’; but to say so much is to say something weaker than ‘u *is* true’. Similarly, (F) is inadequate as an account of falsity. When the only available sense for ‘u says that P’ is as above, ‘$\exists P(\text{Say}\ (u, P) \wedge \neg P)$’ means ‘An interpretation which renders u false is as legitimate as any other interpretation’. In such a case, then, the Simple Argument fails to establish that u is either true or false (even in the weak sense in which ‘or’ means *vel*). Its conclusion says only that u is either such that it could, with maximal legitimacy, be taken to be true, or could, with equal legitimacy, be taken to be false.

How should we define truth and falsity when there is a serious possibility of some statements being indeterminate? The natural answer is this. We count a statement as true if, *however* it may legitimately be taken to say that things are, they are thus; we count it as false if, *however* it may legitimately be taken to say that things are, they are not thus. In other words, when indeterminacy of sense is a serious possibility we need to replace (T) and (F) by the following:

$$(T^*) \qquad \forall u(\text{True}\ (u) \leftrightarrow (\exists P\ \text{Say}\ (u, P) \wedge \forall Q(\text{Say}\ (u, Q) \rightarrow Q)))$$

$$(F^*) \qquad \forall u(\text{False}\ (u) \leftrightarrow (\exists P\ \text{Say}\ (u, P) \wedge \forall Q(\text{Say}\ (u, Q) \rightarrow \neg Q))).$$

Like (T) and (F), (T*) and (F*) may be regarded as explicit definitions.

The theory whose axioms are (T*) and (F*) in a logical system permitting quantification into sentence position is again consistent. In fact, this new theory is consistent even if every utterance in the domain of quantification says something. This may be shown by constructing an unintended interpretation in which each utterance both ‘says’ that P and ‘says’ that not P, so that every formula ‘$\exists P\ \text{Say}\ (u, P)$’ is true, while every formula of the form ‘True (u)’ or ‘False (u)’ is false. In this new theory, the semantic paradoxes are transformed into sound arguments to show that a paradoxical utterance expresses no *unique* thought.

When we took truth and falsity to be defined by (T) and (F), we needed the assumption that u is determinate in order to show that u is not both true and false. Now that truth and falsity are defined by (T*) and (F*), we do not need the assumption of determinacy to show that truth and falsity are contraries. Under the new definitions, however, we need determinacy in order to show that u is true *vel* false. For now that ‘true’ is defined by (T*), we cannot move from line (6) of the Simple Argument, viz. '$\exists P$ (Say $(u, P) \wedge P$)', to line (7), viz. 'True (u)'. Similarly, now that 'false' is defined by (F*), we are blocked from moving from line (11) to line (12). Given the assumption that u satisfies (D), however, we can recast the deduction as follows:

(1)	$\exists P$ Say (u, P)	Definition of statement
(2)	Say (u, P)	(1), existential instantiation
(3)	$\forall Q(\text{Say }(u, P) \wedge \text{Say }(u, Q) \rightarrow (P \leftrightarrow Q))$	(D), universal instantiation
(4)	$\forall Q(\text{Say }(u, Q) \rightarrow (P \leftrightarrow Q))$	(2), (3)
(5)	$P \vee \neg P$	Excluded middle
(6)	P	Assumption
(7)	$\forall Q(\text{Say }(u, Q) \rightarrow Q)$	(4), (6)
(8)	$\exists P$ Say $(u, P) \wedge \forall Q(\text{Say }(u, Q) \rightarrow Q)$	(1), (7) $\wedge$-introduction
(9)	True (u)	(8), (T*)
(10)	True (u) $\vee$ False (u)	(9), $\vee$-introduction
(11)	$\neg P$	Assumption
(12)	$\forall Q(\text{Say }(u, Q) \rightarrow \neg Q)$	(4), (11)
(13)	$\exists P$ Say $(u, P) \wedge \forall Q(\text{Say }(u, Q) \rightarrow \neg Q)$	(1), (12) $\wedge$-introduction
(14)	False (u)	(13), (F*)
(15)	True (u) $\vee$ False (u)	(9), $\vee$-introduction
(16)	True (u) $\vee$ False (u)	(5), (10), (15) $\vee$-elimination, with the discharge of assumptions (6) and (11)

(I have elided some elementary logical steps at lines (4), (7) and (12).) I shall call this latest deduction the *Revised Argument*. Its conclusion is a restricted version of Bivalence: every *determinate* statement is either true or false.

Miroslava Andjelković and Timothy Williamson (2000) claim to be able to deduce the thesis that every statement satisfies (D) from (T*). If their argument worked, (T*) would entail (T) and (F*) would entail (F); moreover, the Simple Argument would apply to every statement and we would have no need for the Revised Argument. Andjelković and Williamson reason as follows. Where s and t are type sentences in a given language L, let sEt be the sentence formed by writing s, then a sign E in L that means ‘if and only if’, and finally t. (It is assumed that L contains at least one expression meaning ‘if and only if’.) sEt, they claim, will always express a biconditional proposition whose first component is what s says and whose last component is what t says. That is, where $<s, c>$ is the (possible) utterance of sentence s in the context c,

(E1) $\forall s \forall t \forall c \forall P \forall Q\ [\text{Say}\ (<s, c>, P) \wedge \text{Say}\ (<t, c>, Q) \rightarrow \text{Say}\ (<sEt, c>, P \leftrightarrow Q)]$.

Andjelković and Williamson further claim that if the utterance of a sentence s in a context says something, then an utterance in that same context of a complex sentence in which an occurrence of s is followed by E, which in turn is followed by a second occurrence of s, is true:

(E2) $\forall s \forall c \forall P\ [\text{Say}\ (<s, c>, P) \rightarrow \text{True}\ (<sEs, c>)]$.

Now a special case of (E1) is

(1) $\forall s \forall c \forall P \forall Q\ [\text{Say}\ (<s, c>, P) \wedge \text{Say}\ (<s, c>, Q) \rightarrow \text{Say}\ (<sEs, c>, P \leftrightarrow Q)]$.

Furthermore, (T*) yields

(2) $\forall s \forall c \forall P \forall Q\ [\text{Say}\ (<sEs, c>, P \leftrightarrow Q) \rightarrow (\text{True}\ (<sEs, c>) \rightarrow (P \leftrightarrow Q)]$.

From (1) and (2) we get

(3) $\forall s \forall c \forall P \forall Q\ [\text{Say}\ (<s, c>, P) \wedge \text{Say}\ (<s, c>, Q) \rightarrow (\text{True}\ (<sEs, c>) \rightarrow (P \leftrightarrow Q))]$.

But (E2) and (3) yield

(4) $\forall s \forall c \forall P \forall Q\ [\text{Say}\ (<s, c>, P) \wedge \text{Say}\ (<s, c>, Q) \rightarrow (P \leftrightarrow Q)]$.

Since any possible utterance is the utterance of a sentence in a context, (4) amounts to the claim that any possible statement satisfies (*D*).

How does this argument fare when the only available interpretation of '*u* says that *P*' is 'An interpretation of *u* under which it says that *P* is as legitimate as any other interpretation'? When 'says' has that meaning, (E2) and (1) cannot both be correct. (1) implies that, in a given context, a speaker-hearer may give non-equivalent, but equally legitimate, interpretations to the two occurrences of *s* in *sEs*. (E2), by contrast, presumes that those two occurrences will be interpreted in equivalent ways: if <*sEs*, *c*> is always true, then an utterance of *sEs* must either be heard as saying that $P \leftrightarrow P$ or be heard as saying that $Q \leftrightarrow Q$ (where *P* and *Q* are the legitimate interpretations of *s*). The latter presumption may well be correct: there may be a convention in the relevant language that precludes switching from one legitimate interpretation of *s* to another when interpreting the two occurrences of *s* in the complex sentence *sEs*. But if such a convention is in place, (1) will be false.

Andjelković and Williamson are right to say that (E1) logically entails (1). That point, however, merely shows that one's first thought—that (E1) is a wholly exceptionless truth about biconditionals—may need to be revised. It is an interesting question how it is best emended. One suggestion is this. Let us call a context *regular* when it is understood that repeated occurrences of a single expression type in the same sentence must be interpreted in the same way.\[4]/ To ensure that (E2) is true, we must restrict the range of the variable '*c*' to regular contexts. When the range of '*c*' is so restricted, though, (E1) needs to be revised by adding the conjunct $s \neq t$ to its antecedent; this ensures that (1) is not among its consequences. In any case, whether or not this way of preserving what is right in (E1) is optimal, the Andjelković-Williamson argument fails. Indeterminate statements are a real possibility and (*T*), (*F*), and the Simple Argument all need to be revised to take account of them.

The Revised Argument shows, indeed, from where challenges to Bivalence must come. A non-bivalent statement must either be indeterminate—i.e. fail to satisfy (*D*)—or express thoughts to which classical logic does not apply. In the latter case, the inapplicable laws will almost certainly be Excluded Middle (invoked at line (5)) or the rule of $\vee$-elimination (alias 'Proof by Cases') that is needed at the last step. One reason for thinking that this latest reconstruction of the argument for

[4] I write of expressions in general, for puzzle cases of this kind are not confined to complete sentences. In any regular context, 'Princeton = Princeton' will be true, even though 'Princeton' could equally legitimately be taken to refer to Princeton Borough or to Princeton Township, and Princeton Borough ≠ Princeton Township (David Lewis's example; see Lewis 1988, 128).

Bivalence is on the right lines is that the most interesting challenges to the bivalence of a given statement proceed by questioning its determinacy, or the application to it of Excluded Middle, or of Proof by Cases. With that in mind, I turn to consider three putative counterexamples to Bivalence in the light of our analysis.

3. Future contingents

The first of these putative counterexamples is not, I think, especially persuasive, but it needs to be mentioned because it looms so large in the history of the topic.

On the traditional reading, Aristotle put forward a counterexample to Bivalence in the ninth chapter of his early treatise *De Interpretatione*.\[5]/ 'A sea-battle will take place tomorrow' expresses a complete thought, and so qualifies as a statement. But to ascribe truth to that statement, Aristotle seems to argue, would imply that it is settled that battle will be joined; and to ascribe falsehood to it would imply that it is settled that battle will be avoided. So, at a time when it is not settled whether battle will be joined, we cannot say that the statement is either true or false.

Aristotle's discussion of the case is convoluted and hard to interpret. Part of the difficulty stems from his switching between two formulations of Bivalence. We can see this switch in the passage already quoted from the *Metaphysics*. The passage concludes with a good formulation of the Principle:

(1) He who says…anything…will say either what is true or what is false.

But it starts with a potentially misleading formulation:

(2) Of one subject we must either affirm or deny any one predicate.

[5] The traditional reading has been challenged. The debate continues, but classic contributions include Łukasiewicz 1922, Anscombe 1956, Kneale and Kneale 1962 chap. II §4, Hintikka 1964, and von Wright 1984.

(2) may be read as equivalent to (1). However, mentioning the speech acts of affirmation and denial muddies the waters, for (2) is more naturally read as expressing a principle that is not equivalent to Bivalence:

(3) Given any statement, we must either be entitled to affirm it or be entitled to deny it.

Principle (3) is false, and future contingents are among many counterexamples to it. In order to be entitled to affirm 'There will be a sea-battle tomorrow', a speaker needs some ground for that assertion; he needs evidence that battle will be joined. Similarly, in order to be entitled to deny 'There will be a sea-battle tomorrow', he needs evidence that battle will not be joined. In a case where the admirals have not decided whether to fight, there may be no such evidence, so Aristotle's case is a plausible—if rather sketchy—counterexample to (3). But it is not thereby a counterexample to (1), which is the correct formulation of Bivalence. The correct formulation does not mention affirmation or denial; it mentions only truth and falsity.

Can the Revised Argument be applied to show that 'There will be a sea-battle tomorrow' is either true or false? I think it can. The relevant instance of Excluded Middle is in order: there is nothing amiss in the claim 'Either there will be a sea-battle tomorrow or there won't be'. Moreover, 'There will be a sea-battle tomorrow' seems to have a determinate sense. As uttered on 1 January 2015, what it says is that a sea-battle will take place on 2 January 2015. If a sea-battle does take place on the latter date, the utterance is true; if no sea-battle takes place then, it is false. In order for the utterance to be true, it is not necessary that battle was inevitable on 1 January; and in order for the utterance to be false, it is not necessary that battle was then precluded. So the claim that the statement is either true or false carries no deterministic implications.

At least, this is so if we understand 'true' and false' 'atemporally'—i.e. in a sense that does not permit significant tensing.\[6]/ It is this understanding of these words that (*T**) and (*F**) articulate—for the quantification in (*T**) and (*F**) is itself tenseless. In this atemporal sense, the statement's truth does not imply that a battle was determined by 1 January, and its falsity does not imply that a battle

[6] This paragraph and the next summarise the argument of von Wright 1984, to which I refer the reader for further elaboration.

was then precluded. When 'true' and 'false' are taken in this sense, Bivalence is consistent with the future's being open.

What further muddies the waters here, and gives the present attack on Bivalence some spurious plausibility, is that some English speakers use 'true' in a way that *does* permit significant tensing. They say things like 'It is already true (now, on 1 January 2015) that there will be a total solar eclipse, visible from the Faeroe Islands, on 15 March 2015'—meaning thereby that the occurrence of such an eclipse on that date is already settled, fixed, or determined. Now in *this* sense of 'true', the Revised Argument would be fallacious and its conclusion false. Whilst it is now determined that there either will or will not be a sea-battle tomorrow, it does not follow that it is either now determined that there will be a battle tomorrow or now determined that there will not be. To the contrary: if the admirals have yet to decide whether to fight, that conclusion will be false. The Revised Argument fails because (T*) does not capture this temporal sense of 'true' (and (F*) does not capture the corresponding sense of 'false'). But the failure of the Argument, and the falsity of its conclusion, when 'true' and 'false' are understood temporally are irrelevant to the Argument and to the Principle in the sense articulated here, in which 'true' and 'false' are understood atemporally. When the Principle is understood in that latter way, future contingents pose no threat to it whatever.

In his essay 'Truth', Michael Dummett considers a statement that may be regarded as an infinitary version of Aristotle's example of the sea-battle: 'A city will never be built on this spot' (Dummett 1959, 16). Dummett takes a statement's content to be given by the commitments that a speaker who affirms it thereby incurs (i.e. by the 'requirements' that must be satisfied if his assertion is to have been correct: *op. cit.*, 22). He also holds that a speaker incurs a commitment only if there is some finite bound on the time by which that commitment will either have been fulfilled or not. Let v be an utterance of 'A city will never be built on this spot', made on day d at place π. Then the commitments of v are these: that there be no city at π on day $(d + 1)$; that there be no city at π on day $(d + 2)$; … Or, using Prior's quantifiers into sentence position:

$$\forall P \forall n\ [\text{Commit}\ (v, P) \leftrightarrow (P \leftrightarrow \text{there is no city at } \pi \text{ on day } d + n)].$$

Now if we think of content in these terms, it is plausible to count a statement as true if all its commitments are fulfilled:

(T^D) $\quad \forall u(\text{True}\ (u) \leftrightarrow \forall P(\text{Commit}\ (u, P) \rightarrow P))$.

Similarly, it is it is plausible to take a statement to be false if one of its commitments is not fulfilled:

(F^D) $\quad \forall u(\text{False}\ (u) \leftrightarrow \exists P\ (\text{Commit}\ (u, P) \wedge \neg P))$.

So in the present case we shall have

True $(v) \leftrightarrow \forall n$ (there is no city at π on day $d + n$)

and

False $(v) \leftrightarrow \exists n$ (there is a city at π on day $d + n$).

These specifications of v's truth- and falsity-conditions seem to be correct: Dummett's requirement that commitments are met or unmet in a finite time does not stop (T^D) and (F^D) from delivering the desired results in this case. Indeed, given classical logic, we can now prove that v is bivalent. A theorem of first-order classical logic is $\ulcorner \forall x \neg Fx \vee \exists x Fx \urcorner$. In particular, then, where $\ulcorner Cn \urcorner$ symbolizes $\ulcorner$There is a city at π on day $(d + n)\urcorner$ we have $\ulcorner \forall n \neg Cn \vee \exists n Cn \urcorner$,\[7]/ which combines with our conditions for v's truth and falsity to yield the conclusion that v is either true or false.

Dummett, however, insists that we cannot assert the bivalence of v. In part this reflects his rejection of classical logic, and I shall return to the validity of $\ulcorner \forall x \neg Fx \vee \exists x Fx \urcorner$ in the next section. His claim that we cannot assert v's bivalence also rests, though, upon a thesis about truth. There is, he writes, an 'important feature of the concept of truth which is not expressed by the law "It is true that *p* if and only if *p*" and which we have so far left quite of account: that a statement is true only if there is something in the world in virtue of which it is true' (Dummett 1959, 14). The things in virtue of which a statement is true, he explains, are 'the sort of fact we have been taught to regard as justifying us in

[7] In fact, this particular instance of Excluded Middle may be justified otherwise than by appeal to classical logic; see *n*.8 below.

asserting it' (*op. cit.*, 16). Let us call such a fact a *ground* of the statement. In these terms, Dummett is making the following claim:

(T^G) $\quad \forall u(\text{True }(u) \rightarrow \text{a ground for } u \text{ obtains}).$

Dummett holds that there is a corresponding thesis about falsehood. Let us say that a fact is an *anti-ground* of a statement if its obtaining justifies us in denying the statement. Then he also maintains

(F^G) $\quad \forall u(\text{False }(u) \rightarrow \text{an anti-ground for } u \text{ obtains}).$

(T^G) and (F^G) together entail

(Biv^G) $\quad \forall u((\text{True }(u) \vee \text{False }(u)) \rightarrow (\text{a ground for } u \text{ obtains}) \vee (\text{an anti-ground for } u \text{ obtains}))$

and, in the case of *v*, we are not entitled to assert the consequent of the relevant instance of (Biv^G):

> We are entitled to say that a statement *P* must be either true or false, that there must be something in virtue of which either it is true or it is false, only when *P* is a statement of such a kind that we could in a finite time bring ourselves into a position in which we are justified either in asserting or in denying *P*; that is, when *P* is an effectively decidable statement. This limitation is not trivial: there is an immense range of statements which, like 'Jones was brave' [said of a man who died without facing danger], are concealed conditionals, or which, like 'A city will never be built here', contain—explicitly or implicitly—an unlimited generality, and which therefore fail the test (Dummett 1959, 16-17).

We should agree with Dummett that we may not be entitled to assert that either a ground or an anti-ground for *v* obtains. A ground for *v* is the sort of fact that justifies us in asserting 'A city will never be built here'. We know what sort of facts these are: the place in question is too cold to support a city, or it is too hot, or it lacks water, etc. Let us suppose, then, that none of the features afflicts π. The locus of *v*, in other words, is a place where a city *could* well be built. On that supposition, no ground for *v* obtains. An anti-ground for *v* is the sort of fact that justifies us in denying 'A city will never be built here'. We know what sort of facts these are too: population growth makes the urbanization of the place inevitable, or plans to build a city are already laid, etc. Let us suppose that none of these facts obtains either. In that case, we shall not be entitled to say that either a ground or an anti-ground for *v* obtains.

But does it follow that we cannot assert that *v* is bivalent? I think not. Dummett maintains that (T^G) is a general constraint on the application of the predicate 'true': truth is an 'epistemically constrained' notion. In 'Truth', however, he offers no defence of this claim and it is striking that (T^G) is not needed to derive the intuitively correct truth-conditions for *v*. It is, then, open to a philosopher to defend Bivalence by maintaining that *v* is a counterexample to (T^G) and hence to (Biv^G). Indeed, *v* is a particularly strong form of counterexample to (T^G), for *v* may be true even though a ground for *v never* obtains. It might be that π always remains a place where a city *could* be built—so that no one is ever justified in asserting 'A city will never be built here'—while, as a matter of fact, no one ever happens to build a city there—so that 'A city will never be built here' is true. This is why it is helpful to place Dummett's example alongside Aristotle's. Today, on 1 January, we can assert that 'There will be a sea-battle tomorrow' is either true or false. Because the admirals have not decided whether to fight, no one today can assert that the statement is true and no one today can assert that it is false. Tomorrow, though, some people will be able to assert either that the statement is true or that it is false; by 2 January, either a ground or an anti-ground will have obtained. From a perspective that rejects (T^G), the proper moral of Dummett's example is that even this eventual obtaining of either a ground or an anti-ground is inessential to a statement's bivalence. We may assert that *v* is either true or false even though we know that neither a ground nor an anti-ground for *v* may ever obtain. Dummett's case brings out something interesting, then, but it is not counterexample to Bivalence.

4. The continuum hypothesis: the intuitionist analysis

A second class of putative counterexamples is more problematic.

When Cantor hypothesized 'There is no set strictly intermediate in cardinality between the integers and the real numbers (the continuum)', it seems he succeeded in expressing a thought. Many philosophers, however, resist the claim that this statement—Cantor's celebrated Continuum Hypothesis, CH for short—is bivalent. Gödel (1940) showed that the truth of CH was consistent with the axioms of Zermelo-Fraenkel set theory including the Axiom of Choice (assuming that those axioms are themselves consistent). Cohen (1966) proved that its falsity (i.e. the truth of its negation) was also consistent with the same axioms (under the same assumption). These results do not refute the claim that CH is either true or false, but they do cast doubt upon it. *Contra* Dummett, a statement may be true even though there are no (and never will be) grounds for asserting it. But many philosophers believe that a true (false) statement must have a *basis*: if a statement is true (false), there must be something that *makes* it true (false), even if we can never find out whether that basis obtains. The results of Gödel and Cohen appear to show that, in the case of CH, there is no basis for its truth or falsity in the generally accepted axioms of set theory, and it is wholly unclear where else such a basis might be.

But, if CH is not bivalent, where does the Revised Argument go wrong when applied to it? The intuitionists hold that we are not entitled to assert that CH is bivalent; on their view, the fallacious step in the Revised Argument is the appeal to the classical logical law of Excluded Middle at step (5). They identify 'the principle of excluded middle with the principle of the solvability of every mathematical problem' (Brouwer 1927, 491). Since no one is entitled to assert that the continuum problem is solvable, no one is entitled to assert line (5) of the Argument in the relevant case. As I now argue, though, the intuitionists do not properly capture the nature of mathematical doubts about the bivalence of statements like CH.

The semantic theory that Arend Heyting (1934) developed for the language of intuitionist mathematics reflects Brouwer's identification of Excluded Middle with universal solvability. According to that theory,

> the meaning of each [logical] constant is to be given by specifying, for any sentence in which that constant is the main operator, what is to count as a proof of that sentence, it being assumed that we already know what is to count as a proof of any of the constituents. (Dummett 2000, 8).

Specifically, Heyting stipulated that a proof of $\ulcorner A \vee B \urcorner$ is anything that is a proof either of *A* or of *B*, and that a proof of $\ulcorner \neg A \urcorner$ is a construction of which we can recognize that, applied to any proof of *A*, it will yield a proof of a contradiction. A statement counts as intuitionistically valid if the semantic principles guarantee it to be provable no matter which atomic statements are provable. So a statement of the form $\ulcorner A \vee \neg A \urcorner$ will be valid only if either A or $\ulcorner \neg A \urcorner$ is provable, i.e. only if the problem of deciding A is solvable. On this semantics, to take $\ulcorner A \vee \neg A \urcorner$ to be valid, no matter what mathematical statement A might be, would precisely be to postulate that every mathematical problem is solvable.

At first blush, the Heyting semantics seems to give us what we want. Since we cannot assert that the continuum problem is solvable, we cannot assert the relevant instance of Excluded Middle in line (5). So, in intuitionistic logic, the Revised Argument for the bivalence of the Continuum Hypothesis barely gets started.

There is a snag, though. Consider Goldbach's Conjecture: 'Every even number greater than two is the sum of two primes'. At the time of writing, no one has a proof or a refutation of this conjecture, or a demonstration that a proof or refutation must exist. If the statement is false, a refutation will exist. For if the statement is false, there will be a counterexample to the Conjecture, and in that event it will be possible in principle to identify all the prime numbers less than the counterexample and then verify that no pair of them has the counterexample as its sum. However, there is at present no reason to assert that the Conjecture must have either a proof or a refutation. For all we know, there may be no counterexample to it, but at the same time no uniform reason why every even number greater than two is the sum of two primes, nor even a finite partitioning of those numbers with a uniform reason for each partition. Given the Heyting semantics, then, we are at present unable to assert $\ulcorner GC \vee \neg GC \urcorner$ where GC states Goldbach's Conjecture. This is a problem, however, for almost every mathematician believes that we *can* assert $\ulcorner GC \vee \neg GC \urcorner$, even in the face of the recognition that GC may not be decidable. For this reason, the intuitionist seems to give the wrong explanation of why CH is not bivalent.

What grounds the widespread belief that we can assert $\ulcorner GC \vee \neg GC \urcorner$? In articulating these grounds, it helps to spell out some notions. Let us say that a property φ is *definite* (with respect to a domain) if each member of that domain is either φ or not φ. And let us say that a domain D is *determinate* when, for any property φ that is definite with respect to D, the following thesis holds when the range of the quantifiers is restricted to D:

(*Det*) $\quad \forall x \varphi x \vee \exists x \neg \varphi x$.

The thought underlying (*Det*) is this. For a domain to be determinate is for it to be determinate which objects belong to it. Now where φ is definite with respect to D, each member of D is either φ or not φ. If it is also determinate which objects belong to D, then the question 'Is every member of D a φ?' must have one of its expected answers. If that answer is 'yes', then $\forall x \varphi x$. If the answer is 'no', then some member of D is not φ, so that $\exists x \neg \varphi x$. Either way, we have $\ulcorner \forall x \varphi x \vee \exists x \neg \varphi x \urcorner$.

These notions enable us to spell out the basis for the widespread conviction that we are entitled to assert $\ulcorner GC \vee \neg GC \urcorner$. Let the domain of quantification, D, be the even natural numbers greater than two; and let $\ulcorner Fn \urcorner$ be the statement $\ulcorner n$ is the sum of two primes$\urcorner$. Then Goldbach's Conjecture is expressed by the formula $\ulcorner \forall x\, Fx \urcorner$. Now it is agreed on all hands (even by intuitionists operating under the Heyting semantics) that each even number greater than two is either the sum of two primes or it is not. In our terms, F is a definite property of the members of the domain D. Moreover, it is natural to hold that the domain of even numbers is determinate: that is, it is determinate which objects are even numbers. For example, 2, 4, and 6 are even numbers, while 1, 3, and Julius Caesar are not. We can, then, assert $\ulcorner \forall x Fx \vee \exists x \neg Fx \urcorner$. Even in intuitionistic logic, though, $\ulcorner \exists x \neg Fx \urcorner$ entails $\ulcorner \neg \forall x Fx \urcorner$, so $\ulcorner \forall x Fx \vee \exists x \neg Fx \urcorner$ entails $\ulcorner \forall x Fx \vee \neg \forall x Fx \urcorner$, i.e., $\ulcorner GC \vee \neg GC \urcorner$.\[8]/

The contentious premiss of this argument is the claim that the domain of even numbers—and, with it, the domain of all natural numbers—is determinate. Whether these domains are determinate or not is a deep question in the philosophy of arithmetic that I cannot discuss here.\[9]/ Whatever the answer may be, though, our analysis shows where Brouwer's doubts about CH's bivalence diverge from those of mainstream mathematicians. The problem does not lie in the use of intuitionistic logic *per se*. In that logic, some instances of Excluded Middle may be asserted. Rather, the divergence stems from Brouwer's contention, which is embodied in the Heyting semantics, that $\ulcorner A \vee \neg A \urcorner$ may be asserted only when A is decidable. For Brouwer, we cannot assert that CH is bivalent *because* we cannot assert that CH is decidable. For most mathematicians who have doubts about the bivalence of CH, though, that cannot be the proper explanation: they take themselves to be entitled to assert that

---

[8] A parallel argument establishes the truth of $\ulcorner \forall n \neg Cn \vee \exists n Cn \urcorner$, where $\ulcorner Cn \urcorner$ means $\ulcorner$There is a city at place π on day $(d + n)\urcorner$. (An empirical property of the natural numbers may still be definite in the specified sense.) The truth of $\ulcorner \forall n \neg Cn \vee \exists n Cn \urcorner$ was the main premiss of the positive argument given in §3 for the bivalence of Dummett's example *v*. That example may be regarded as an ingenious hybrid of Goldbach's Conjecture and Aristotle's Sea Battle.

[9] For an interesting discussion, see Field 1998.

Goldbach's Conjecture is bivalent even though they cannot assert that the Conjecture is decidable. If CH is not bivalent, its being so does not lie simply in its being undecidable.\[10]/

5. Failures of bivalence in set theory: a better analysis

I think people are right to doubt the bivalence of many statements of set theory but I do not think that those doubts offer any support to Brouwer's intuitionism. Rather, I want to argue, the doubts are well-founded because many set-theoretic statements are indeterminate: they fail to be bivalent because they do not satisfy condition (*D*).

When a mathematician utters the sentence 'There is no set strictly intermediate in cardinality between the integers and the real numbers (the continuum)', what is he saying? He is clearly saying something about sets, but what are sets? The word 'set', as it comes from his mouth, is surely a theoretical term, and the theory that implicitly defines it—the theory that endows it with sense—can only be set theory. In uttering CH, then, the mathematician is saying something about the mathematical structure that is characterized by the axioms of set theory.

An account of the content of set-theoretic statements along these lines seems to me to be the only one that is remotely plausible.\[11]/ There is, however, a problem with the formulation I have given: there is no such thing as *the* mathematical structure (in the singular) that is characterized by the axioms of standard set theory. The point is that even second-order set theory, ZFC^2, is not fully categorical. Its axioms leave (at least) the 'height' of the set-theoretic universe undetermined, so two models of ZFC^2 need not be isomorphic. As Zermelo showed, the theory is only *quasi-categorical*: given any two models, one will be isomorphic to an initial sub-model of the other (Zermelo 1930). More exactly, each model of ZFC^2 has the form V_κ, where κ is a strongly inaccessible cardinal. (The hierarchy of sets V_α is defined as follows for each sort of ordinal α: $V_0 = \varnothing$; $V_{\alpha+1} = \wp(V_\alpha)$ for each successor ordinal

---

[10] Cfr. Benacerraf and Putnam: 'It is instructive to compare [and contrast] set theory with number theory. In number theory too there are statements that are neither provable nor refutable from the axioms of present-day mathematics. Intuitionists might <argue> that that this shows (not by itself, of course, but together with other considerations) that we do not have a clear notion of "truth" in number theory, and that our notion of a "totality of all integers" is not precise. Most mathematicians would reject this conclusion. Yet most mathematicians feel that the notion of an "arbitrary set" is somewhat unclear' (Benacerraf and Putnam 1983, 19).

[11] For a persuasive elaboration of the account, see now Issacson 2011.

α+1; $V_\lambda = \bigcup_{\beta<\lambda} V_\beta$ for each limit ordinal λ. For present purposes, the cardinal number of a set A may be taken to be the least ordinal equinumerous with A.)

What are the consequences of this? One is that some theorems of first-order classical logic are hard to justify when the domain of quantification is understood to range over all sets. An example of such a theorem is (*Det*) from §4:

$$\forall x\varphi x \vee \exists x\neg\varphi x.$$

I argued there that we are entitled to assert (*Det*) whenever the domain of quantification is determinate and φ is a definite property. Even when φ is definite, however, it is hard to see what entitles us to assert $\ulcorner\forall x\varphi x \vee \exists x\neg\varphi x\urcorner$ when the variable 'x' ranges over absolutely all sets. Certainly the justification given in §4 does not apply. That justification presumed that it was a determinate matter which objects are sets. But it is hard to see how this can be a determinate matter if (a) set theory is what determines which objects satisfy 'set' whereas (b) set theory is not categorical.

Is the use of classical logic illegitimate, then, when reasoning about the totality of all sets? Not necessarily. For one might maintain that statements apparently involving quantification over all sets need to be tamed before we can do any systematic reasoning with them and that, once we tame them, it is legitimate to apply classical logic to them. The sort of 'taming' I have in mind is already implicit in Zermelo's writings. The very idea of non-categoricity presumes a multiplicity of non-isomorphic models of ZFC^2; but the domains of those models are themselves sets and, as such, are determinate. Thus 'what appears as an "ultra-finite non- or super-set" in one model is, in the succeeding model, a perfectly good, valid set with both a cardinal number and an ordinal type, and is itself a foundation stone for the construction of a new domain' (Zermelo 1930, 1233). Now when the domain of quantification is restricted to the members of some set, the laws of classical logic will apply. What is peculiar about statements apparently about all sets, on this view, is not that their variables range over a collection that is not a set. It is, rather, that the semantics of such statements must be understood in relation to a whole sequence of domains of quantification each of which is a set: 'To the unbounded series of Cantor ordinals there corresponds a similarly unbounded double-series of essentially different set-theoretic models, in each of which the whole classical theory is expressed…This series reaches no true completion in its unrestricted advance, but possesses only relative stopping-points, just those "boundary numbers" [i.e. strongly inaccessible cardinals] which separate the higher model types from the lower' (*ibid.*).

This way of understanding statements involving apparent quantification over all sets may justify the use of classical logic in reasoning with them. It does so, however, at the price of rendering certain set-theoretic statements indeterminate. For we have here an example of the situation considered earlier, where the only available sense that can be attached to '*u* says that *P*' is 'An interpretation of *u* under which it says that *P* is as legitimate as any other interpretation'. *Ex hypothesi*, the only legitimate interpretations of *u* are those which vindicate the classical logic by restricting the domain of quantification to the members of a standard model of ZFC^2. Under this hypothesis, what is our mathematician saying when he utters CH? Where κ_1 is the first strongly inaccessible cardinal, one maximally legitimate interpretation of his statement is that there is no set in $V_{\kappa 1}$ that is strictly intermediate in cardinality between the integers and the reals. But, where κ_2 is the second strongly inaccessible cardinal, another maximally legitimate interpretation of his statement is that there is no set in $V_{\kappa 2}$ that is strictly intermediate in cardinality between the integers and the reals. And so forth. In making a set-theoretic statement, a speaker presents himself as saying something about the universe of sets, but there is no such thing—no unique such thing, anyway. The various determinations of that universe generate various, but equally legitimate, interpretations of set-theoretic statements.

For this reason, many set-theoretic statements fail to satisfy condition (*D*). Let *u* be the statement 'There is at least one strongly inaccessible cardinal', let '*P*' abbreviate 'There is at least one strongly inaccessible cardinal in $V_{\kappa 1}$' and let '*Q*' abbreviate 'There is at least one strongly inaccessible cardinal in $V_{\kappa 2}$'. An instance of (*D*) is

(**) $\mathrm{Say}\ (u, P) \wedge \mathrm{Say}\ (u, Q) \rightarrow (P \leftrightarrow Q).$

On the view we are considering, (**) is false. On that view, the only available interpretation that sustains a classical logic for '*P*', '*Q*' etc. is one whereby 'Say (*u*, *P*)' means 'An interpretation of *u* under which it says that *P* is as legitimate as any other interpretation'. Under that interpretation, the antecedent of (**) is true. Because our set-theoretic axioms fail to single out a unique intended 'universe' of sets, one maximally legitimate interpretation of *u* is that there is at least one strongly inaccessible cardinal in $V_{\kappa 1}$, while another is that there is at least one strongly inaccessible cardinal in $V_{\kappa 2}$. The consequent of (**), however, is false: P is false (κ_1 is not a member of $V_{\kappa 1}$) whereas Q is true.

The Revised Argument cannot be applied, then, to establish that an arbitrarily selected statement of set theory is bivalent. Since some set-theoretic statements do not satisfy (D), we cannot always affirm line (3) of the Argument. Indeed, if the theory that gives content to the notion of 'set' is simply ZFC^2, many set-theoretic statements are *not* bivalent, u of the previous paragraph being a case in point. u 'says' that P and 'says' that Q, where Q but not P. Since u says that P, and not P, (T*) entails that u is not true. Since u says that Q, and Q, (F*) entails that u is not false. So u is neither true nor false.

Unlike Brouwer's analysis, the present ground for saying that certain set-theoretic statements are not bivalent is entirely consistent with holding that undecidable statements of number theory such as Goldbach's Conjecture *are* bivalent. The crucial difference is that there are theories which provide a fully categorical axiomatization of the natural number structure: our comprehension of such a theory provides the basis for a conception of the domain of natural numbers as being determinate in the sense of §4. Second-order Peano arithmetic (PA^2) is the most famous example of a categorical axiomatization of the natural number structure. Some philosophers doubt if we really understand full second-order quantification over infinite domains, but even if we cannot grasp PA^2, there are weaker theories—indeed, theories whose logic is weaker than full second-order logic—that characterize the natural numbers up to isomorphism. By the Löwenheim-Skolem Theorem, the logic of a categorical axiomatization of number theory cannot be first-order. But consider 'ancestral logic', a system in which first-order logic is supplemented with introduction and elimination rules for an operator * which maps a relation to its ancestral (see Myhill 1952). (The *-operator is counted as a logical constant.) In this system, we may formulate a theory, T, in which the axioms of first-order Peano arithmetic, PA^1, are supplemented by the axiom $\ulcorner \forall x(x = 0 \vee S^*0x) \urcorner$, where S is the relation of immediate succession. Unlike PA^1, T characterizes the natural numbers up to isomorphism: since the new axiom requires each member of the domain to be only finitely many steps of immediate succession from zero, it excludes the non-standard numbers. However, the conceptual resources needed to make sense of T are much weaker than those needed to understand PA^2. In particular, it is not necessary to make sense of the idea of an arbitrary set of natural numbers.

Many contemporary set theorists regard ZFC^2 as too weak a theory to characterize the notion of 'set'. Most of the stronger theories that have been proposed do not threaten the above argument that not all set-theoretic statements are bivalent, for the theories in question remain only quasi-categorical. Vann McGee (1997), however, has proposed a new axiom that renders the resulting set theory categorical. McGee works in a system, ZFU^2, in which the first-order quantifiers are permitted to range over *Urelemente*—objects that are not sets—as well as over sets. He proposes as a further set-theoretic axiom the postulate that these *Urelemente* form a set (call this axiom 'McG'); he then proves that the theory ZFU^2 + McG gives a categorical characterization of the pure sets—i.e. the sets that may be formed from the empty set using the standard set-theoretic operations. More precisely, McGee shows that there is a theorem of ZFU^2 + McG which, under the intended interpretation of the language of ZFU^2, says the following: any structure that could, given the axioms of ZFU^2 + McG, serve as the interpretation of 'pure set' and 'is a member of', will be isomorphic to the intended structure of the pure sets.\[12]/

McGee's result is striking, and one commentator has welcomed it as opening 'the door to a structuralist account of set-theoretic truth on which every sentence of [the language of] pure set theory is assigned a determinate truth value' (Uzquiano 2002, 181). But have we reason to accept McGee's axiom? So far from its being an evident truth, I think it is highly doubtful.

Let us consider the notion of an ordinal, i.e. the order-type of a well-ordered set. Ordinals can be represented as sets. However, there are many equally good ways of so representing them. For the reasons expounded by Paul Benacerraf in 'What Numbers Could Not Be' (Benacerraf 1965), this suggests forcibly that they are not themselves sets. If an ordinal were a set, there would have to be a fact as to which set it is. Given the multiplicity of equally good set-theoretic representations of the ordinals, it is hard to see what that fact could be.\[13]/ We shall do better to follow Christopher Menzel and treat ordinals as *properties* of well-ordered sets: an ordinal is something that isomorphic well-ordered sets have in common (Menzel 1986, 43).

---

[12] That is, McGee shows that the theory of pure sets in ZFU^2 + McG is 'internally categorical' in the sense of Walmsley 2002, section IV.

[13] For this application of Benacerraf's argument see Menzel 1986, 41-3. Of course, if the natural numbers *are* ordinals, this is not a new application of Benacerraf's argument but a reiteration of it. In 'What Numbers Could Not Be', though, Benacerraf seemed to think of a natural number as something abstracted from its ordinal and cardinal applications.

On this view, an ordinal is an *Urelement*, so McG (together with Separation) implies that there is a set **On** of all ordinals. At this point, however, the Burali-Forti Paradox looms (see Menzel *op. cit.*, 38). Certainly, that Paradox does not need the assumption that ordinals are sets. All we need is that the putative set **On** is well-ordered by the relation < of being strictly less than: since an ordinal just *is* the order-type of a well-ordered set, it follows immediately from this theorem that **On** possesses an ordinal, β. Now any ordinal is the order-type of the set of ordinals strictly less than itself. So β is also the order-type of the set A of all ordinals that are strictly less than β. Since A and **On** share an order-type, they must be isomorphic. It is easy to show, though, that no well-ordered set is isomorphic to any of its proper initial segments, so that A and **On** must be identical. Since **On** is the set of all ordinals, $\beta \in$ **On**, whence $\beta \in A$. But then, by the definition of A, it follows that $\beta < \beta$. Menzel (1986 and 2014) has explored ways of avoiding outright contradiction here by restricting some of the other axiom of ZF. The fact remains, though, that McG's having this implication renders it highly doubtful. Nice as it might seem to have a categorical characterization of the universe of sets, we have reason not accept McGee's theory.

6. Kreisel's argument for the bivalence of CH

The argument of the previous section shows that some set-theoretic statements are not bivalent. Others are, though. Further confirmation of our analysis comes from the sense it makes of a famous argument purporting to show that CH, in particular, is bivalent.

The argument I have in mind was given by Georg Kreisel in section 1 of his paper 'Informal Rigour and Completeness Proofs' (1967). Kreisel's point of departure is Gödel's well-known essay 'What is Cantor's Continuum Problem?' (Gödel, 1964). In that essay, Gödel was concerned to distinguish between the sort of independence that the Parallel Postulate has from the other axioms of Euclidean geometry, and the sort of independence that CH has from the axioms of ZFC. Gödel's account of the difference is informal and somewhat vague; Kreisel aims to characterize it more precisely 'in terms of higher-order consequence' (Kreisel 1967, 79). He duly begins section 1 by remarking that second-order logic has a good claim to be the implicit logic of the working mathematician:

> The familiar classical structures (natural numbers with the successor relation, the continuum with a denumerable dense base, etc.) are definable [i.e. may be characterized up to isomorphism] by *second*-order axioms, as shown by Dedekind. Zermelo showed that his cumulative hierarchy up to ω or $\omega + \omega$, or $\omega + n$ (for fixed n) and other important ordinals is equally definable by second-order formulae. Whenever we have such a second-order definition there is associated a schema in first-order form (in the language considered): For instance, in Peano's [induction] axiom
>
> $$\forall P[\{P(0) \wedge \forall x\, [P(x) \rightarrow P(x + 1)]\} \rightarrow \forall x\, P(x)]$$
>
> one replaces the second-order quantifier $\forall P$ by a list of those P which are explicitly defined in ordinary first-order form (from + and $\times$, for instance). A moment's reflection shows that the evidence of the first-order axiom schema derives from the second-order schema: the difference is that when one puts down the first-order schema one is supposed to have convinced oneself that the specific formulae used (in particular, the logical operations) are well defined in any structure that one considers (*op. cit.*, 85-6).

This point enables Kreisel to put his finger on the relevant difference between the two cases of independence. Geometry, too, admits of a second-order axiomatization, and '*the parallel axiom is not even a second-order consequence*' of the other geometric axioms (*op. cit.*, 88, Kreisel's emphasis). So there really are different geometries, in some of which the Parallel Postulate holds, in others of which it does not. Consequently, it makes no sense to ask whether the Postulate is true or false *simpliciter*; all one can say is that it holds in some geometries and does not hold in others. For the question of its truth or falsity to be meaningful, we would need to give some extra-axiomatic specification of the meaning of the geometric theory's primitive terms—for example, by stipulating that a straight line is the path that a light-ray traverses *in vacuo*. Once that stipulation had been made, it would make sense to ask whether the Parallel Postulate is true or false, and the question would be settled by empirical investigation (our best theory of the path of light-rays says that the Postulate is false). *Per contra*, 'CH is only independent of the *first-order* schema (associated with the axioms) of Zermelo-Fraenkel' (*ibid.*)—i.e. of the schema that is associated with the second-order Zermelo-Fraenkel axioms in the way in which the formula with ‘*P*’ as a schematic letter is associated with the second-order induction axiom displayed above.

Kreisel surely identified an important point of difference between the independence of the Parallel Postulate from the other axioms of classical geometry and the independence of CH from the axioms of ZFC. But how, exactly, is that point supposed to establish that CH is bivalent? Here is the crucial passage:

> Let $\mathcal{Z}$ be Zermelo’s axiom with the axiom of infinity, and let CH be the (canonical) formulation of the continuum hypothesis in the following form: if C_ω is the collection of hereditarily finite sets without individuals, $C_{\omega+1} = C_\omega \cup \wp(C_\omega)$, $C_{\omega+2} = C_{\omega+1} \cup \wp(C_{\omega+1})$, CH states that
>
> $$X \subset C_{\omega+1} \rightarrow (\| X \| \leq \| C_\omega \| \vee \| X \| = \| C_{\omega+1} \|),$$
>
> which is expressed by means of quantifiers over $C_{\omega+2}$. As Zermelo pointed out, if we use the current set-theoretic definition $Z(x)$ of the cumulative hierarchy, in any model of $\mathcal{Z}$, this formula Z defines a C_σ for a limit ordinal $\sigma > \omega$. Consequently we have
>
> $(\mathcal{Z} \vdash_2 \text{CH}) \vee (\mathcal{Z} \vdash_2 \neg\text{CH})$ (*op. cit.*, 87-8).

Although the single turnstile ‘$\vdash_2$’ is standardly used to signify deducibility in axiomatic second-order logic, Kreisel does not (or should not) mean this: Weston (1976, 289-90) extends an argument of Tharp’s to show that neither CH nor ¬CH is a theorem of axiomatic ZFC^2. The proper conclusion of his argument is that either $\mathcal{Z} \models_2$ CH or $\mathcal{Z} \models_2 \neg$CH, where ‘$\models_2$’ signifies the model-theoretic consequence relation of full second-order logic.\[14]/

Kreisel’s argument is compressed, but it may be spelled out as follows. Let us begin by considering a theory, *T*, formulated in a second-order language *L*, and let us suppose that *T* is

[14] Kreisel gives his argument for Zermelo’s original (1908) axiom system Z, which lacks the Axiom of Replacement, rather than for the more familiar system ZF, which includes it. (Kreisel clearly uses ‘Z’ to stand for Zermelo’s original system rather than for ZF itself. He observes (1967, 88) that while CH is determined by the second-order axioms of Z, Replacement is not; this would make no sense if ‘Z’ meant ZF.) This renders his argument problematical, for Z^2 is not even quasi-categorical. Indeed, in the context of Z^2, standard formulations of the axiom of infinity turn out to be non-equivalent, and some interpretations of Z^2 + Infinity have non-well-founded models (see Uzquiano 1999). A charitable exposition of Kreisel’s argument cuts through these complexities by taking it to apply to ZFC, not Z.

categorical—i.e. that all its models are isomorphic. Consider now two such models, M and N. A general theorem tells us that the same closed sentences of L are true in the two models. That is, for any closed sentence σ of L, $M \models_2 \sigma$ if and only if $N \models_2 \sigma$. The proof of this theorem is in essence straightforward. Since T is categorical, there is an isomorphism f from M to N. One then shows by induction on complexity that, for any expression e of L, the semantic value of e with respect to N is the result of applying f to the value of e with respect to M. This shows that if $N \models_2 \sigma$ then $M \models_2 \sigma$, for any closed sentence σ of L. Since there is also an isomorphism from N to M, a parallel argument establishes the converse.

As we have seen, because ZFC^2 is only quasi-categorical, there are closed sentences in its language that are true in some of its models and false in others. Kreisel's insight, though, was to recognize that CH is not subject to this variation in truth-value between models. As he remarks in the last passage quoted from him, CH may be formulated using quantification over sets at levels $V_{\omega+2}$ and below. All the models of ZFC^2, *when restricted to those levels*, are isomorphic, so the argument of the previous paragraph shows that, for any models M and N of ZFC^2, $M \models_2$ CH if and only if $N \models_2$ CH. The argument rests on particular features of CH; it turns on the fact that CH quantifies only over sets that lie so low in the set-theoretic hierarchy that every model of ZFC^2 will include them.

The argument as presented so far, however, does not take us all the way to Kreisel's conclusion. The quasi-categoricity of ZFC^2 combines with the particular features of CH to yield

(1) For any models of ZF^2 M and N, $M \models_2$ CH if and only if $N \models_2$ CH.

Kreisel's eventual conclusion is the bivalence of CH:

(B) CH is either true or false.

How are we supposed to move from (1) to (B)?

The natural reconstruction runs as follows. A familiar principle of classical model theory says that, for any model, any closed sentence is either true in the model or false in it. So, in particular,

(2) For any model M of ZFC^2, either $M \models_2$ CH or $M \models_2 \neg$CH.

Together, (1) and (2) entail

(3) Either $M \models_2$ CH for any model M of ZFC^2, or $M \models_2 \neg$CH for any model M of ZFC^2.

An abbreviated way of writing (3) is

(4) Either ZFC$^2 \models_2$ CH or ZFC$^2 \models_2 \neg$CH

which is what Kreisel wrote—or meant to write—in the last line quoted from him. But in order to get from (4) to (*B*) we need

(5) If ZFC$^2 \models_2$ CH then CH is true

and

(6) If ZFC$^2 \models_2 \neg$CH then CH is false.

What justifies (5) and (6)?

I think (5) and (6) may be justified as specializations, to the present case, of our general principles (*T**) and (*F**). According to (*T**), a statement will be counted as true if and only if, however it may legitimately be taken to say that things are, they are thus. As we have seen, the various models of ZFC^2 correspond to the various things that set-theoretical statements could legitimately be taken to say. So (*T**) yields

(5+) CH is true if and only if ZFC$^2 \models_2$ CH

from which (5) follows. Similarly, according to (*F**), a statement will be counted as false if and only if, however it may legitimately be taken to say that things are, they are not thus. As before, the models of ZFC^2 correspond to the ways set-theoretical statements could legitimately be taken to say that things are. So (F*) yields

(6+) CH is false if and only if $ZFC^2 \models_2 \neg CH$

from which (6) follows.

In saying this, I am not endorsing Kreisel's argument. One doubtful step is line (2). Although (2) follows from a principle of classical model theory, the present application of that theory has been contested. Since the pure hereditarily finite sets are isomorphic to the set N of natural numbers, line (2) assumes that both $\wp(N)$ and $\wp(\wp(N))$ exist as well-defined sets, over which it is legitimate to quantify classically. The discussion of Kreisel's argument among philosophers of mathematics has focused on whether such quantification is legitimate (see especially Feferman 2009 and forthcoming, and Koellner 2010). All the same, our analysis confirms that his argument has the right 'shape' to justify the bivalence of CH. In effect, Kreisel argues that CH is determinate (despite the general indeterminacy in the sense of 'set') so that line (3) of the Revised Argument can be asserted.\[15]/

Even those set theorists who doubt the bivalence of some set-theoretic statements generally take classical logic for granted in doing set theory. Our analysis shows that their position is coherent. There can be failures of determinacy, and hence statements that are neither true nor false, without any deviations from classical logic. It is noteworthy, however, that without the Principle of Bivalence, one cannot justify the classical logical laws on the basis of the specifications of the meanings of the connectives in the familiar classical truth-tables, even given a classical metalogic. Thus the truth-tables for '$\vee$' and '$\neg$' entail that $\ulcorner A \vee \neg A \urcorner$ is true if A is either true or false, but we need Bivalence to infer

[15] Kreisel's claim that CH is bivalent is entirely consistent with the philosophical thesis (cf. the beginning of §4) that a statement can be true (false) only if something *makes* it true (makes it false)—i.e. that truth and falsehood require a *basis*. If Kreisel's argument works at all, it shows that either $ZFC^2 \models_2 CH$ or $ZFC^2 \models_2 \neg CH$. So the basis of the truth or falsehood of CH (as the case may be) lies in the axioms of ZFC^2. Since the consequence relation of full second-order logic cannot be completely axiomatized, it does not follow that either $ZFC^2 \vdash_2 CH$ or $ZFC^2 \vdash_2 \neg CH$, where '$\vdash_2$' signifies deducibility in axiomatic second-order logic. Indeed, the Tharp-Weston result shows that this claim is false. If our knowledge of set-theory were confined to what we could deduce from ZFC^2 in axiomatic second-order logic, it would follow that neither CH nor $\neg$CH has a *ground* in the sense of §3. That is to say: we shall never be entitled to assert CH or to deny it. However, that is no threat to CH's bivalence, for we found reason to reject Dummett's claim that a statement can be true only if it has a ground and false only if it has an anti-ground. The case of CH, then, is one where it is important to distinguish between *basis* and *ground*.

from that that any instance of $\ulcorner A \vee \neg A \urcorner$ is true. This need not disturb a set-theorist who is wedded to classical logic. Even if called upon to justify his adherence to that logic, there are alternative justifications—indeed, alternative semantic justifications—that he can pray in aid.\[16]/ Given a weak metalogic, Bivalence is sufficient to establish the validity of the classical logical laws, but it is far from being necessary.

7. The challenge of vague statements

Another class of putative counterexamples to Bivalence are statements in which a vague predicate attaches to a borderline case of the property that it signifies. Let us call such a *vague statement*. Imagine a hundred transparent tubes of paint, $a_1, \ldots, a_{100}$, that steadily and almost imperceptibly change from clearly red to clearly orange as the subscript increases. Then consider the corresponding sequence $A_1, \ldots, A_{100}$ of statements, in which the statement A_n says that the tube a_n is red. Bivalence tells us that each of these statements is either true or false. Now if a statement attributing redness to an object a is true, and if b is redder than a, then a statement attributing redness to b will also be true. So Bivalence implies that there is a number N such that all of the statements $A_1, \ldots, A_N$ in our sequence are true and all of the statements $A_{N+1}, \ldots, A_{100}$ are false. It implies, in other words, that there is a cut-off point in the sequence at which the statements switch from being true to being false. But that in turn implies that there is a cut-off point at which the tubes switch from being red to being not red—a grossly implausible conclusion. Yet the only premisses needed to reach that conclusion were Bivalence and a highly intuitive assumption about how the predicate 'red' is to be applied.

One way of resisting the claim that vague statements are bivalent runs parallel to the treatment of set-theoretic statements recommended in §§5 and 6. On this approach, only sentences with completely precise conditions of truth may replace the propositional variables 'P', 'Q' etc., so the applications of classical logic in the Revised Argument are unexceptionable. When the variables are interpreted in this way, vague statements will not be determinate in sense: a vague statement may be interpreted equally legitimately as saying that P and as saying that Q, where 'P' and 'Q' are inequivalent. Some of these precise specifications of its sense come out true and others come out false,

[16] See Rumfitt 2015, chapter 9 for one of them.

so (T^*) and (F^*) entail that the statement is neither true nor false. This conclusion is consistent with the soundness of the Revised Argument, whose conclusion says only that every determinate statement is either true or false.

This approach yields a supervaluational treatment of vagueness, and so exhibits what it at once the great strength and the great weakness of any such treatment, its concern to preserve classical logic.\[17]/ Just as set-theorists use classical logic to construct their proofs, even when they doubt or deny the bivalence of some of the statements in those proofs, so ordinary thinkers apply the classical laws to vague statements, even as they doubt or deny *their* bivalence. All the same, there are grounds for querying the application of classical logic to vague statements that have no counterpart in the case of set theory. For by applying classical logic to vague statements we can derive apparently absurd conclusions, without making any appeal to Bivalence.

This is shown, of course, by versions of the *Sorites Paradox*. Let us revert to our hundred tubes of paint and the statements that attribute redness to each of them. The ground for denying that all of those statements are bivalent was that Bivalence entails the existence of a sharp boundary to the red tubes, and that consequence is absurd. But if that consequence is absurd, we ought to be able to assert its negation. Now there will be a sharp boundary to the red tubes if some tube in the sequence is red and its successor is not, i.e. if $\ulcorner A_n \wedge \neg A_{n+1} \urcorner$ is true for some n. So the claim that there is a sharp boundary may be formulated as a long disjunction $\ulcorner[(A_1 \wedge \neg A_2) \vee \ldots \vee (A_{99} \wedge \neg A_{100})]\urcorner$, and the claim that there is none as its negation. That is, we ought to be able to assert the following:

(1) $\neg[(A_1 \wedge \neg A_2) \vee \ldots \vee (A_{99} \wedge \neg A_{100})]$.

In the situation described, though, we are also given that a_1 is red and that a_{100} is not red. So it seems that we ought also to be able to assert

(2) A_1

and

[17] See Fine 1975 for the classic exposition and Varzi 2007 for a recent survey of the logical issues.

(3) $\neg A_{100}$.

Now if we suppose that a_{99} is red, i.e. if we suppose

(4) A_{99}

then the rule of conjunction-introduction applied to (3) and (4) would yield

(5) $A_{99} \wedge \neg A_{100}$,

which, after 99 applications of $\vee$-introduction, yields

(6) $(A_1 \wedge \neg A_2) \vee \ldots \vee (A_{99} \wedge \neg A_{100})$

which directly contradicts (1). Given (1) and (3), then, supposition (4) stands refuted, so by *Reductio* we may assert

(7) $\neg A_{99}$.

By repeating this inferential sub-routine a further 98 times, we reach

(8) $\neg A_1$

which contradicts (2). This, then, is the initial paradox. We have some reason to accept the trio of postulates (1), (2), and (3), but we also have an apparently valid deduction showing that the trio is inconsistent. It may be noted that the form of *Reductio* that is applied in reaching line (7)—and that is re-applied at the corresponding later steps—is acceptable to an intuitionist. So the trio comprising (1), (2), and (3) is inconsistent in intuitionistic logic as well as in classical logic.

How should we react to this apparent demonstration of inconsistency? Since the case is one in which (2) and (3) are clearly true, it seems that we must take it as showing that (1) is false. In other words, we would appear to be entitled—indeed, compelled if the question of (1)'s truth arises—to make a further application of the relevant form of *Reductio* and infer the negation of (1), namely,

(9) $\neg\neg[(A_1 \wedge \neg A_2) \vee \ldots \vee (A_{99} \wedge \neg A_{100})]$.

In classical logic, however, (9) is equivalent to

(10) $(A_1 \wedge \neg A_2) \vee \ldots \vee (A_{99} \wedge \neg A_{100})$.

This, however, seems to land us in a yet more acute paradox, which Crispin Wright has called the *Paradox of Sharp Boundaries* (see Wright 2007). For formula (10) says that at some point in the sequence a red tube is immediately followed by a non-red tube, and this seems to ascribe a sharp boundary to the red tubes. Wright calls this conclusion 'unpalatable' and it does indeed seem to be something we are reluctant to accept. Given classical logic, however, it follows from the premisses (2) and (3). That is, given only the premisses 'Tube a_1 is red' and 'Tube a_{100} is not red', classical logic yields a conclusion which seems to say that there is a sharp boundary to the red tubes in the sequence $a_1, \ldots, a_{100}$. Little wonder that many philosophers have taken the *Sorites* to cast doubt on whether classical logic can be applied to deductions involving vague predicates.

Wright (*op. cit.*) takes this form of the paradox as a ground for switching from classical to intuitionist logic. That logic validates all the steps up to line (9), but it does not validate the elimination of double negation needed to reach (10). Wright's response, though, is not founded on any semantic analysis of vague expressions; as Dummett (2007) noted, the Heyting semantics certainly will not do.\[18]/ Because the choice of logic is in play, it would help to have a semantic model for vague statements against which intuitions about the validity of arguments involving them can be tested. But what might that model be?

[18] See, though, Rumfitt 2016 for a semantic theory which, if accepted, would vindicate Wright's choice of intuitionistic logic as providing the standards for assessing arguments involving vague predicates.

8. The semantics and logic of vague statements

What is characteristic of vague expressions? An attractive general account was put forward by C. S. Peirce. In the entry 'Vague (in logic)' that he contributed to Baldwin's *Dictionary of Philosophy and Psychology*, Peirce explained the word 'vague' as meaning:

> Indeterminate in intention.
>
> A proposition is vague when there are possible states of things concerning which it is intrinsically uncertain whether, had they been contemplated by the speaker, he would have regarded them as excluded or allowed by the proposition. By intrinsically uncertain we mean not uncertain in consequence of any ignorance of the interpreter, but because the speaker's habits of language were indeterminate; so that one day he would regard the proposition as excluding, another as admitting, those states of things. Yet this must be understood to have reference to what might be *deduced* from a perfect knowledge of his state of mind; for it is precisely because those questions never did, or did not frequently, present themselves that his habit remained indeterminate (Baldwin 1901-2, vol. 2, 748).

In remarking that a statement like 'This is red' might exclude different states of affairs from day to day, Peirce points to an adaptability that vague terms confer upon a language. Precisely because the accepted sense of a vague statement is indeterminate, a speaker is free to render it more determinate by laying it down, as it might be, that the statement is to be taken as excluding certain possibilities whose status as excluded or allowed had hitherto been left unsettled. These determinations resemble decisions more closely than they resemble discoveries, so that indeterminacy of sense goes hand in hand with a form of semantic 'freedom' whose importance for the analysis of vagueness Mark Sainsbury has stressed. The word 'red' is associated with no boundary between the red things and the rest. But

> it can be permissible to draw a line even where it is not mandatory to do so. No one can criticize an art materials shop for organizing its tubes of paint on various shelves, including one labelled 'red' and another 'orange', even though there is a barely detectable, or perhaps even in normal circumstances undetectable, difference between the reddest paint on the shelf

> marked 'orange' and the most orange paint on the shelf marked 'red'. Hence one can consistently combine the following: *red* draws no boundaries, that is, there is no adjacent pair in the series of tubes of paint such that the nature of the concept, together with the colour of the tube, requires one to apply *red* to one member of the pair but withhold it from the other; yet one can draw a boundary to the reds, that is, one may behave consistently with the nature of the concept in drawing a line between adjacent pairs.\[19]/

This seems right, as far as it goes, but I would add two points. First, many determinations of the sense of vague terms will be partial: even after the determination has been made, there will be some possible states of affairs of which it remains unsettled whether the statement is understood to exclude them. Sainsbury's paint shop owner may determine the senses of 'red' and 'orange' so that precisely one of these predicates applies to every tube of paint in his shop. But there may be other tubes, redder than the reddest tube on his 'orange' shelf but more orange than the most orange tube on his 'red' shelf, whose classification remains undetermined even after he has settled how all the tubes currently in his shop are to be classified.

Second, in further determining the sense of a vague term, we often settle the truth-value of a complex statement without settling the truth-values of the components. Determination need not proceed from atoms to molecules. In particular, this holds for disjunctive statements: we often settle that a disjunction is true without settling which disjunct is true. A simple example of this will be familiar to readers who have been examiners in British universities. Undergraduate degrees in Britain are classified, and vague principles relate the candidate's average numerical mark to his or her eventual class of degree. Thus the examiners' deliberations are directed towards making the vague senses of 'first-class', 'upper second-class' etc. sufficiently determinate (in the given context) that one (and only one) of these predicates applies to each of that year's finalists. Now suppose that there are two finals candidates, *A* and *B*, whose performance is borderline first-class. Suppose, in particular, that *A*'s average numerical mark is 69.25 while *B*'s is 69.2. Eventually, the examination board will have to determine the sense of the predicates of classification sufficiently precisely that each atom—in this case, each statement that ascribes a class to a candidate—is either true or false. But there may be

[19] Sainsbury 1990, 259-60. Here and in the next quotation, I have written 'orange' where he writes 'yellow'.

stages in the determination where it is settled that a disjunction is true, but where it is left open which disjunct is true. Thus, having re-read the scripts of *A* and *B*, the examining board may decide that their work is so similar in overall quality that the difference in their average mark reflects no real difference in the standard attained, so that the two candidates must be classified alike. Moreover, it may well reach this decision before settling which degrees *A* and *B* are to receive. This decision, then, amounts to determining the senses of 'first-class' and 'upper second-class' so as to verify the disjunction 'Either *A* and *B* both receive first-class degrees, or *A* and *B* both receive upper second-class degrees' in advancing of determining which disjunct is true. Examples to the same effect could easily be multiplied.

This suggests that we might model the meanings of vague statements by reference to a space of *partial determinations of sense*, or *pds*'s for short; a pds will partially determine the senses of all the vague expressions in the relevant language. When a given pds renders a statement true, I shall say that it *verifies* the statement. As in the example of the examiners' classification, we allow that a pds may verify a complex statement even though it does not verify any of its parts. With the notion of verification in play, we may lay down the condition for consequence for a language containing vague terms: B follows from $A_1,\ldots, A_n$ just in case any pds that verifies all of $A_1,\ldots, A_n$ also verifies B. This account takes as given the extra-linguistic circumstances, which of course help to determine whether a given pds verifies a given statement. For the present, though, I shall suppress the way those circumstances vary in order to focus on the particular bearing that vagueness has on consequence.

A pds verifies a statement when it renders it true. In a similar spirit, we say that a pds falsifies a statement when it renders it false. It is natural to use the notion of falsification alongside that of verification in analysing cases of vagueness. The customary sense of the word 'red' is partial: it is indeterminate or 'intrinsically uncertain' (in Peirce's phrase) whether the term applies to our tube a_{50}. All the same, a grasp of that customary sense enables a speaker *both* to assert 'Pillar boxes are red' and to deny 'Oranges are red'. That is to say: the customary sense of 'red' is marked out both by its verifying the first statement and by its falsifying the second. Because the customary sense is partial, we cannot infer that a statement is falsified from the premiss that it is not verified. In order to characterize a pds adequately, we need to say both which statements it verifies and which it falsifies.

How are we to do this? Let us first define a dyadic relation $\perp$ of *incompatibility* between pds's as follows:

$x \perp y$ if and only if there is some possible statement A such that x falsifies A and y verifies A.

I postulate that incompatibility is (1) irreflexive and (2) symmetric. Postulate (1) says that no pds both verifies and falsifies a single statement. It amounts, in other words, to a requirement that every pds avoid the incoherence of entitling a speaker who uses terms in accordance with it simultaneously to assert and deny a given statement. As for postulate (2), let us assume that a speaker who denies A is committed to asserting A's negation, and that a speaker who asserts A is committed to denying A's negation. Given these assumptions, if x falsifies A and y verifies A, then y falsifies $\ulcorner$not $A\urcorner$ and x verifies $\ulcorner$not $A\urcorner$. That is, given these assumptions, if $x \perp y$ then $y \perp x$. Postulate (2), then, is sustained by the assumption that any possible statement has a negation with the following properties: any speaker who denies the original statement is committed to asserting the negation; and any speaker who asserts the original statement is committed to denying the negation. These properties are widely accepted attributes of negation.

This suggests that the space of pds's is an *orthoframe*—that is, a structure $<X, \perp>$ comprising a non-empty set X of pds's, and a symmetric and irreflexive relation $\perp$ between members of X. I use the notation $|A|$ to signify the set of pds's which verify the statement A. It is then straightforward to define the notion of falsification. Let us write $x \perp Y$ to mean that $x \perp y$ for every y belonging to a set Y, and consider the relation $x \perp |A|$. This will obtain when x is incompatible with every pds that verifies A; that is, when for *every* verifier y of A, x falsifies a statement which y verifies. This condition is met only when x falsifies A. Conversely, if x falsifies A, x will certainly be incompatible with any verifier of A. Given the meaning we have attached to $\perp$, then, 'x falsifies A' is equivalent to '$x \perp |A|$'. Although this obviates the need for separate semantic axioms concerning the conditions in which complex statements are falsified, the definition of the $\perp$-relation still presupposes a grip on the notion of denying a statement that does not reduce to asserting its negation.

Indeed, with the notion of falsification in play, we can lay down the semantic axiom for negation. A partial determination of sense will verify A's negation precisely when it falsifies A itself:

(N) $\quad |\neg A| = |A|^{\perp}$.

(I write $U^{\perp}$ for $\{x: x \perp y$ for all $y \in U\}$.) The axiom for conjunction is also straightforward. A partial determination of sense will verify a conjunction just in case it verifies both conjuncts, so again we have

$$(C) \qquad |A \wedge B| = |A| \cap |B| .$$

However, the analogous axiom for disjunction would be wrong. That axiom would be

$$(D1) \qquad |A \vee B| = |A| \cup |B| ,$$

but the case of the examination candidates is a counter-example to (D1). We there had a partial determination of sense that verifies a disjunction without verifying either disjunct.

How can we find a correct axiom for disjunction? The key is to attend to the topological properties of those sets that comprise a statement's verifiers. When Y and Z are subsets of X in an orthoframe $<X, \perp>$, let us say that Y is *$\perp$-closed in Z* if the following condition is met:

For all $x \in Z$, $x \notin Y$ only if there exists $y \in Z$ such that $y \perp Y$ and not $x \perp y$.

I now argue that whenever sets Y and Z are semantic values of statements (i.e. whenever there are statements A and B such that Y is $|A|$ and Z is $|B|$), the following condition is met:

($\perp$-closure) If $Y \subseteq Z$, then Y is $\perp$-closed in Z.

I shall give the argument for particular choices of A and B; it will be clear how it generalizes.

Let us take A to be the statement 'Tube a_{n+1} is red' and B to be the statement 'Tube a_n is red', and let $Y = |$'Tube a_{n+1} is red'$|$ and $Z = |$'Tube a_n is red'$|$. Since tube a_{n+1} is less red than the tube a_n, any determination of the sense of 'red' that renders A true also renders B true, so $Y \subseteq Z$. In order to establish $\perp$-closure, then, we need to show that Y is $\perp$-closed in Z.

What is it for Y to be $\perp$-closed in Z? It comes to this. Consider an arbitrary pds, x, that verifies the statement 'Tube a_n is red'. We need to show that if x does not verify 'Tube a_{n+1} is red' then there is a pds, y, compatible with x, that verifies 'Tube a_n is red' but falsifies 'Tube a_{n+1} is red'. There are two cases to consider. If x falsifies 'Tube a_{n+1} is red', then x itself is such a y, for any pds is compatible with itself. If x does not falsify 'Tube a_{n+1} is red', then x leaves that statement undetermined—neither verified nor falsified. Now if a pds leaves a statement undetermined, there will be a further determination of it that renders it false. Such a further determination of x will of course verify 'Tube a_n is red' (as x does) and will be compatible with x, so this further determination serves as the desired y. Either way, then, such a pds y exists, so the condition for Y to be $\perp$-closed in Z is met.

Although I have given the argument only for a particular choice of statements, it generalizes. The crux of the argument is that, if a pds leaves a statement undetermined, there will be a further determination of that pds that renders the statement false. I claim that this is a generally acceptable principle. Indeed, it is really just a manifestation of the 'semantic freedom' that Sainsbury identified as being a mark of a vague term such as 'red' and that follows from Peirce's conception of vagueness as indeterminacy of sense. To quote Sainsbury's formula once more, while 'red' itself draws no boundary, 'yet one can draw a boundary to the reds, that is, one may behave consistently with the nature of the concept in drawing a line between adjacent pairs'. On the Peircean view, such freedom is a mark of all vague concepts. If that is right, the argument for $\perp$-closure will generalize to all vague statements.

We call a set $\perp$-closed (*simpliciter*) if it is $\perp$-closed in the whole space of pds's, X. Since the verifiers of any statement form a subset of X, the verifiers of each statement must be $\perp$-closed (*simpliciter*). If sets Y and Z are $\perp$-closed, so are $Y \cap Z$ and $Y^{\perp}$. So, for complex statements built up using $\wedge$ and $\neg$ alone, the verifiers of every statement will be $\perp$-closed so long as the verifiers of each atom form a $\perp$-closed set.

This suggests the following definitions. Let us call the triple $<X, \perp, \xi>$ a *vagueness frame* if $<X, \perp>$ is an orthoframe and ξ is a non-empty collection of $\perp$-closed subsets of X such that (1) ξ is closed under set intersection and the orthocomplementation operation $^{\perp}$, and (2) if $Y, Z \in \xi$, then $Y \subseteq Z$ only if Y is $\perp$-closed in Z. We then call the quadruple $<X, \perp, \xi, V>$ a *vagueness model* if V is a function which assigns a member of ξ to be the verifiers of each atomic statement of the relevant language and which respects axioms (C) and (N) above. Since $Y \cap Z$ and $Y^{\perp}$ are $\perp$-closed whenever Y and Z are, V will assign members of ξ to complex statements built up using conjunction and negation. We can then say that a conclusion follows logically from some premisses if every vagueness model that verifies all the premisses also verifies the conclusion.

Our account of the constraints on the assignment function V is incomplete, for we have as yet said nothing about disjunction. Assuming contraposition in the meta-logic, the condition for Y to be $\perp$-closed is that $x \in Y$ whenever $\forall y\,(y \perp Y \rightarrow y \perp x)$. That is to say, it is equivalent to the condition: $Y^{\perp\perp} \subseteq Y$. By the symmetry of $\perp$, any set is such that $Y \subseteq Y^{\perp\perp}$, so Y is $\perp$-closed if and only if $Y^{\perp\perp} = Y$.

This analogy points the way towards the correct semantic axiom for disjunction. We surely want the entailments from A to $\ulcorner A \vee B \urcorner$ and from B to $\ulcorner A \vee B \urcorner$ to come out valid, so $|A \vee B|$ must include $(|A| \cup |B|)$. Since the verifiers of any statement are $\perp$-closed, it is natural to take $|A \vee B|$ to be the smallest $\perp$-closed set that includes $(|A| \cup |B|)$. Thus we reach

(D) $\quad |A \vee B| = (|A| \cup |B|)^{\perp\perp}$.

Since $(Y \cup Z)^{\perp\perp} = (Y^{\perp} \cap Z^{\perp})^{\perp}$, (*D*) is equivalent to $|A \vee B| = (|A|^{\perp} \cap |B|^{\perp})^{\perp}$.

What logic does the proposed semantic theory validate? My definitions of a vagueness frame and of a vagueness model are exactly parallel to Goldblatt's definitions of a quantum frame and a quantum model, and it is quantum logic that turns out to be sound and complete with respect to the recommended notion of consequence (see Goldblatt 1974, 32-4). Quantum logic is usually formalized as a 'binary' system, in which consequence is taken to relate a single statement serving as premiss to a single conclusion, but it may also be formalized as characterizing a relation $X : A$ between a finite set of premisses X and a single conclusion A. That formalization includes the usual structural rules of Reflexivity, Dilution and Cut, and the following rules for the operators:

$\wedge$-intro $\quad$ If $X : A$ and $Y : B$ then $X, Y : A \wedge B$

$\wedge$-elim $\quad$ If $X, A, B: C$ then $X, A \wedge B : C$

$\vee$-intro $\quad$ If $X : A$ then $X : A \vee B$ and $X : B$ then $X : A \vee B$

$\vee$-elim $\quad$ If $A : C$ and $B : C$ then $A \vee B : C$

DN $A : \neg\neg A$ and $\neg\neg A : A$

Con If $A : B$ then $\neg B : \neg A$

Red If $X : A$ then $X, \neg A : \bot$

DS If $B : A$ then $A \wedge (\neg A \vee B) : B$.

Quantum logic does not contain the unrestricted version of $\vee$-elimination (with side-premisses): if $X, A : C$ and $Y, B : C$ then $X, Y, A \vee B : C$. The form of Disjunctive Syllogism that I have called DS, and which is a 'mixed rule' in that involves all three operators, partly compensates for the loss of deductive power that the restriction on $\vee$-elimination engenders. (One may prove that these rules are sound and complete with respect to the recommended account of consequence by adapting the completeness proof for quantum logic in Goldblatt 1974, §6.)

9. The problems resolved

How does this help to diagnose the flaw in the *Sorites* argument of §7?

Where A_n is the statement 'Tube a_n is red', the first phase of that argument was a demonstration that a trio of plausible premisses was inconsistent. We formulated the first of these premisses—which says that the red tubes have no sharp boundary—as the negation of a disjunction, viz., $\ulcorner\neg[(A_1 \wedge \neg A_2) \vee \ldots \vee (A_{99} \wedge \neg A_{100})]\urcorner$. But since all of De Morgan's Laws are valid in quantum logic, I shall now take the equivalent formula below to be the canonical formulation of premiss (1):

(1) $\neg(A_1 \wedge \neg A_2) \wedge \ldots \wedge \neg(A_{99} \wedge \neg A_{100})$.

As before, premisses (2) and (3) of the *Sorites* argument are A_1 and $\neg A_{100}$. Let us call premiss (1), 'B'

The first part of our soritical deduction was directed towards showing that premisses (1), (2) and (3) are an inconsistent trio—in other words, that $B, A_1, \neg A_{100} : \bot$. Our rules enable us to derive this sequent as follows. From Reflexivity and $\wedge$-elim we have $B : \neg(A_1 \wedge \neg A_2)$, which gives $B : \neg A_1 \vee A_2$ by De Morgan and DN. Since $A_2 : A_1$, the rule DS yields $A_1, \neg A_1 \vee A_2 : A_2$ whence $B, A_1 : A_2$ by Cut. A similar deduction shows that $B, A_2 : A_3$, whence $B, A_1 : A_3$ by Cut. By repeating this inferential sub-routine a further 98 times, we eventually reach $B, A_1 : A_{100}$, whence $B, A_1, \neg A_{100} : \bot$ by Red. Our analysis confirms, then, that premisses (1), (2), and (3) are an inconsistent trio.

What about the next stage of the *Sorites* argument, which moves from the two apparently incontrovertible premisses—(2) and (3)—to the negation of premiss (1)? Our rules validate this stage of the argument too. We already have $A_1, B : A_{100}$, which yields $A_1 \wedge B : A_{100}$ by $\wedge$-elim. This gives $\neg A_{100} : \neg(A_1 \wedge B)$ by Con, whence $\neg A_{100} : \neg A_1 \vee \neg B$ by De Morgan. We also have $\neg B : A_1$. The formula $\ulcorner \neg B \urcorner$ can only be rendered true if some tube in the sequence is red; since a_1 is the reddest tube in the sequence, it must be counted red if any tube is, so any pds that verifies $\ulcorner \neg B \urcorner$ will verify A_1. Accordingly DS yields $A_1, \neg A_1 \vee \neg B : \neg B$. Cut then gives $A_1, \neg A_{100} : \neg B$. This validates the second stage of the original argument: A_1 and $\neg A_{100}$ are its premisses (2) and (3), and $\ulcorner \neg B \urcorner$ is the negation of premiss (1).

At this stage, it may seem as though we have made no progress in escaping from the *Sorites* paradox. Premisses (2) and (3) are highly plausible: if we can be sure of the truth-values of any statements involving vague terms, we can be sure that our paradigm red tube a_1 is red and that our paradigm orange tube a_{100} is not. So it seems that we must accept the conclusion $\ulcorner \neg B \urcorner$. But B is equivalent to the long negated disjunction $\ulcorner \neg[(A_1 \wedge \neg A_2) \vee \ldots \vee (A_{99} \wedge \neg A_{100})] \urcorner$, and hence $\ulcorner \neg B \urcorner$ is equivalent to the doubly negated disjunction $\ulcorner \neg\neg[(A_1 \wedge \neg A_2) \vee \ldots \vee (A_{99} \wedge \neg A_{100})] \urcorner$. In quantum logic double negations are eliminable, so our conclusion $\ulcorner \neg B \urcorner$ is equivalent to the long disjunction

$$(A_1 \wedge \neg A_2) \vee \ldots \vee (A_{99} \wedge \neg A_{100}),$$

a formula I shall label *C*. *C* is simply Wright's 'unpalatable' conclusion, unpalatable because it seems to say that there is a cut-off point somewhere in the sequence. Our logic, then, has taken us from two apparently certain premisses—viz., A_1 and $\neg A_{100}$—to an apparently unacceptable conclusion, viz., C. So it may seem that no substantial progress has been made in resolving the paradox.

In fact, though, we have made progress. For, when C is understood in the way our semantic theory requires, it is in truth entirely palatable. C appears to be unpalatable because it seems to assert the existence of a sharp boundary between the red and the non-red tubes. We imagine that C can be true only if one of its component disjuncts is true—say $\ulcorner A_{50} \wedge \neg A_{51} \urcorner$. In turn, that disjunct can be true only if there is a sharp boundary between the red members of the sequence—$a_1, \ldots, a_{50}$—and the non-red members: $a_{51}, \ldots, a_{100}$. Given our semantics, though, this assumption about the truth-conditions of disjunctions is not correct. The semantic axiom for 'or' does not say that a disjunction is true just when one of its disjuncts is true. Rather, it says that a partial determination of sense verifies a disjunction if it belongs to the *closure* of the union of the verifiers of the disjuncts, where closure is double orthocomplementation. It is therefore entirely possible for a pds to verify a disjunction without verifying either disjunct and this makes it possible for a disjunction to be true without either disjunct's being true.

Indeed, when we think what the present closure operator means, we can see how C may be true even though the concept *red* lacks a sharp boundary. Our semantic axiom for disjunction says that $|A \vee B| = (|A| \cup |B|)^{\perp\perp}$. A pds belongs to $U^{\perp\perp}$ if it is incompatible with every pds that is itself incompatible with all the members of U. Hence a pds verifies C if it is incompatible with every pds that falsifies the claim that there is a cut-off point in the sequence of red tubes. Now a pds falsifies that claim only when *either* the entire sequence is red *or* the entire sequence is not red. Given premisses (2) and (3)—that tube a_1 is red and that tube a_{100} is not red—neither of these possibilities obtains. So the customary current sense of the word 'red', partial as it is, is indeed incompatible with every pds in the orthocomplement of the union

$$|A_1 \wedge \neg A_2| \cup \ldots \cup |A_{99} \wedge \neg A_{100}|$$

and hence belongs to $|(A_1 \wedge \neg A_2) \vee \ldots \vee (A_{99} \wedge \neg A_{100})|$. It belongs to that set, and hence verifies C, because it verifies'a_1 is red' while falsifying 'a_{100} is red'. But because that pds verifies no particular disjunct $\ulcorner A_n \wedge \neg A_{n+1} \urcorner$, the truth of C carries no commitment to the existence of a sharp cut-off in the sequence of red tubes. As one might put it: C says that there is a shift from red to non-red tubes *somewhere* in the sequence, but there need be no determinate place where the shift occurs. It makes no sense to ask which is the last red tube, not because there is such a tube but we cannot in principle know which it is (as in Timothy Williamson's (1994) account of these matters), but because there is no such thing. The question presupposes a determinate answer, but none exists.

How, though, is such indeterminacy possible? We can, perhaps, dispel any remaining mystery by recalling the semantic freedom that is the concomitant of Peirce's account of vagueness. Suppose one is called upon to answer the successive questions 'Is tube a_1 red?', 'Is tube a_2 red?', etc. What C records is the unsurprising fact that one will have to stop answering 'yes' somewhere if one is to avoid giving an affirmative answer to the question 'Is tube a_{100} red?'. Within limits, though, one may choose where to stop answering 'yes'. One has to stop somewhere, but there is no place where one has to stop. The verifying disjunct of C, then, is indeterminate because it is, within limits, *arbitrary*. When introduced to the *Sorites*, many people's first reaction is to say: there must be a switch in colour somewhere, but there is no fact of the matter where, and the place we choose to make the switch does not mark any boundary of the concept *red*. Our analysis shows that this naïve reaction paradox is essentially right. Or better, the Peircean semantics provides a theoretical context within which this naïve thought can alleviate the sense of perplexity that the truth of C—the ineluctable conclusion of our *Sorites* argument—presents.

Where, finally, does this leave the Revised Argument for Bivalence, when this is applied to a vague statement? Unlike the supervaluationist, we can allow that a vague statement expresses just one thought: the statement A_{50} (for example) says that the tube a_{50} is red—no more, no less. We can also accept the relevant instance of Excluded Middle: we can say that the tube a_{50} is either red or not red. This too is attractive: a_{50} is problematic because it is indeterminate whether it is red; there is no impulse to deny that it is one or the other. All the same, the statement A_{50} is not bivalent. The Revised Argument to show that it is bivalent falls at the final hurdle, viz. in the application of $\vee$-elimination to reach line (16). In quantum logic, we cannot apply $\vee$-elimination in the presence of a non-logical premiss; yet the assumption that the relevant statement satisfies (D), made at line (3), is such a premiss. (The analogous argument to show that a true disjunction contains a true disjunct breaks down at the corresponding point.) In both set theory and vague discourse we find statements that are neither true nor false, but only vague statements put pressure on classical logic itself.\[20]/

[20] *Note added* (July 2015). I completed this essay in 2012 and my view of some of the matters it treats has evolved since I finished it. (1) I now prefer to justify the use of classical logic when reasoning about absolutely all sets by way of a negative translation into an intuitionistic language rather than via the supervaluational semantics proposed in §§5 and 6. This alternative approach confirms the main thesis of those sections—that there is no general reason to expect the statements of set theory to be bivalent. (2) Contrary to §§7-9, I now think that the *Sorites* Paradox, and Wright's Paradox of Sharp Boundaries, can be resolved without deviating from classical logic. For a defence of these claims, see Chapters 9 and 8 of my book, *The Boundary Stones of Thought* (Oxford: Clarendon Press, 2015).

REFERENCES

Andjelković, M. and T. Williamson. 2000. 'Truth, falsity and borderline cases'. *Philosophical Topics* **28**: 211-244.

Anscombe, G.E.M. 1956. 'Aristotle and the sea battle'. *Mind* **65**: 1-15.

Auxier, R.E. and L.E. Hahn, eds. 2007. *The Library of Living Philosophers: Michael Dummett*. Chicago, Ill.: Open Court.

Baldwin, J.M., editor-in-chief. 1901-5. *Dictionary of Philosophy and Psychology*. New York: Macmillan.

Benacerraf, P. 1965. 'What numbers could not be'. *The Philosophical Review* **74**: 495-512.

Benacerraf, P. and H. Putnam. 1983. 'Introduction'. In Benacerraf and Putnam, eds., 1983, pp.1-37.

——, eds. 1983. *Philosophy of Mathematics: Selected Readings*, 2nd ed. Cambridge: Cambridge University Press.

Brouwer, L.E.J. 1927. 'Intuitionistische Betrachtungen über den Formalismus'. *Koninklijke Akademie van wtenschappen te Amsterdam, Proceedings of the Section of Sciences* **31**: 374-79. Page references are to the English translation by Stefan Bauer-Mengelberg, entitled 'Intuitionistic reflections on formalism', in Jean van Heijenoort, ed., *From Frege to Godel: A Source Book in Mathematical Logic, 1879-1931* (Cambridge, Mass: Harvard University Press, 1967), pp.490-92.

Cohen, P.J. 1966. *Set Theory and the Continuum Hypothesis*. New York: W.A. Benjamin.

Dummett, M.A.E. 1959. 'Truth'. *Proceedings of the Aristotelian Society* **59**: 141-62. Page references are to the reprint in Dummett 1978, pp.1-19.

——. 1978. *Truth and Other Enigmas*. London: Duckworth.

——. 2000. *Elements of Intuitionism,* 2nd edition. Oxford: Clarendon Press.

——. 2007. 'Reply to Crispin Wright'. In Auxier and Hahn, eds., pp.445-54.

Ewald, W.B., ed. 1996. *From Kant to Hilbert: A Source Book in the Foundations of Mathematics*, volume II. Oxford: Clarendon Press.

Feferman, S. 2009. 'Conceptions of the continuum'. *Intellectica* **51**: 169-189.

——. Forthcoming. 'Is the Continuum Hypothesis a definite mathematical problem?'

Field, H.H. 1998. 'Which undecidable mathematical sentences have determinate truth values?' In H. Garth Dales and Gianluigi Oliveri, eds., *Truth in Mathematics* (Oxford: Clarendon Press), pp.291-310.

Fine, K. 1975. 'Vagueness, truth, and logic'. *Synthese* **30**: 265-300. Reprinted in Keefe and Smith, eds., 1996, pp.119-50.

Glanzberg, M. 2001. 'The liar in context'. *Philosophical Studies* **103**: 217-251.

Gödel, K.F. 1940. *The Consistency of the Axiom of Choice and of the Generalized Continuum-Hypothesis with the Axioms of Set Theory*. Princeton, NJ: Princeton University Press.

——. 1964. 'What is Cantor's continuum problem? (revised and expanded version)'. In Benacerraf and Putnam, eds., 1983, pp.470-85.

Goldblatt, R.I. 1974. 'Semantic analysis of orthologic'. *Journal of Philosophical Logic* **3**: 19-35.

——. 1984. *Topoi: The Categorial Analysis of Logic*. Amsterdam: Elsevier.

Heyting, A. 1934. *Mathematische Grundlagenforschung. Intuitionismus. Beweistheorie*. Berlin: Springer Verlag.

Hintikka, K.J.J. 1964. 'The once and future sea fight: Aristotle's discussion of future contingents in *De Interpretatione* IX'. *The Philosophical Review* **73**: 461-92.

——, ed. 1969. *The Philosophy of Mathematics*. Oxford: Oxford University Press.

Isaacson, D.R. 2011. 'The reality of mathematics and the case of set theory'. In Z. Novák and A. Simonyi, eds., *Truth, Reference, and Realism* (Budapest: Central European University Press, 2011), pp.1-75.

Keefe R. and P. Smith, eds. 1996. *Vagueness: a reader*. M.I.T. Press.

Kneale, W.C. 1972. 'Propositions and truth in natural languages'. *Mind* **81**: 225-43.

Kneale, W.C. and Kneale, M. 1962. *The Development of Logic*. Oxford: Clarendon Press.

Koellner, P. 2010. 'On the question of absolute undecidablity'. In S. Feferman et al., eds., *Kurt Gödel: Essays for his Centennial* (Cambridge: Cambridge University Press), pp.189-225.

Kreisel, G. 1967. 'Informal rigour and completeness proofs'. In Lakatos, ed., 1967, pp.138-171. Page references are to the partial reprint in Hintikka, ed., 1969, pp.78-94.

Lewis, D.K. 1988. 'Vague identity: Evans misunderstood'. *Analysis* **48**: 128-130.

Łukasiewicz, J. 1922. 'O determinizmie'. Printed in Lucasiewicz, ed. J. Slupecki, *Z zagadnieh logiki i filosofii* (Warsaw, 1961). Translated into English as 'On determinism' by Z. Jordan in Storrs McCall, ed., *Polish Logic 1920-1939* (Oxford: Clarendon Press).

McGee, V. 1997. 'How we learn mathematical language'. *The Philosophical Review* **106**: 35-68.

Mackie, J.L. 1973. *Truth, Probability, and Paradox: Studies in Philosophical Logic*. Oxford: Clarendon Press.

Menzel, C. 1986. 'On the iterative explanation of the paradoxes'. *Philosophical Studies* **49**: 37-61.

——. 2014. 'Wide sets, ZFCU, and the iterative conception'. *The Journal of Philosophy* **111**: 57-83.

Myhill, J. 1952. 'A derivation of number theory from ancestral theory'. *Journal of Symbolic Logic* **17**: 192-197.

Parsons, C.D. 1974. 'The Liar Paradox'. *Journal of Philosophical Logic* **3**: 381-412.

Prior, A.N. 1971. *Objects of Thought*. Oxford: Clarendon Press.

Rumfitt, I. 2011. 'Ramsey on truth and meaning'. In Benjamin Morison and Katerina Ierodiakonou, eds., *Episteme, etc: Essays in Honour of Jonathan Barnes* (Oxford: Clarendon Press), pp.213-245.

——. 2014. 'Truth and meaning'. *Proceedings of the Aristotelian Society Supplementary Volumes* **88**: 21-55.

——. 2015. *The Boundary Stones of Thought: An Essay in the Philosophy of Logic*. Oxford: Clarendon Press.

——. 2016. 'Vagueness and intuitionistic logic'. Forthcoming in Alexander Miller, ed., *Logic, Language, and Mathematics: a Festschrift for Crispin Wright* (Oxford: Clarendon Press).

Sainsbury, R.M. 1990. *Concepts without Boundaries*. London: King's College. Page references are to the reprint in Keefe and Smith, eds., 1996, pp.251-64.

Smiley, T.J. 1993. 'Can contradictions be true?' *Proceedings of the Aristotelian Society Supplementary Volumes* **67**: 17-33.

Tarski, A. 1935. 'Der Wahrheitsbegriff in den formalisierten Sprachen'. *Studia Philosophica* **I**: 261-405. Translated by J.H. Woodger as 'The concept of truth in formalized languages' in Tarski, eds. Woodger and J. Corcoran, *Logic, Semantics, Metamathematics* (Indianapolis: Hackett, 1983), pp.152-278.

——. 1944. 'The semantic conception of truth and the foundations of semantics'. *Philosophy and Phenomenological Research* **4**: 341-76.

Uzquiano, G. 1999. 'Models of Zermelo set theory'. *The Bulletin of Symbolic Logic* **5**: 289-302.

——. 2002. 'Categoricity theorems and conceptions of set'. *Journal of Philosophical Logic* **31**: 181-96.

Varzi, A.C. 2007. 'Supervaluationism and its logics'. *Mind* **116**: 633-76.

Walmsley, J.O. 2002. 'Categoricity and indefinite extensibility'. *Proceedings of the Aristotelian Society* **102**: 217-35.

Weston, T. 1976. 'Kreisel, the Continuum Hypothesis, and second-order set theory'. *Journal of Philosophical Logic* **5**: 281-298.

Williamson, T. 1994. *Vagueness*. London: Routledge.

——. 1998. 'Indefinite extensibility'. *Grazer Philosophische Studien* **55**: 1-24.

Wright, C.J.G. 2007. 'Wang's Paradox'. In Auxier and Hahn, eds., pp.415-44.

Wright, G.H. von. 1984. 'Determinism and future truth'. In his *Truth, Knowledge, and Modality: Philosophical Papers Volume III* (Oxford: Blackwell), pp.1-13.

Zermelo, E.F.F. 1908. 'Untersuchungen über dir Grundlagen der Megenlehre I'. *Mathematische Annalen* **65**: 261-81. Translated by Stefan Bauer-Mengelberg in van Heijenoort 1967, pp.199-215.

——. 1930. 'Über Grenzzahlen und Mengenbereiche: Neue Untersuchungen über die Grundlagen der Megenlehre'. *Fundamenta Mathematicae* **16**: 29-47. Page references are to the translation by Michael Hallett, 'On boundary numbers and domains of sets: new investigations in the foundations of set theory', in Ewald, ed., 1996, pp.1219-33.