Alternative Logics and Applied Mathematics

[draft of 5 March 2018]

Timothy Williamson

1. Internal and external motivations for deviance

The hardest test of deviant logic is mathematics, which constitutes by far the most sustained and successful deductive enterprise in human history. With only minor exceptions, mathematicians have freely relied on *classical logic*, including principles such as the law of excluded middle, $A \vee \neg A$. They unquestioningly accept classical reasoning in proofs. When deviant logicians reject a classical principle, they face an obvious challenge: what does that mean for mathematics? Where does it leave theorems whose proofs rely on the principle? Absent a good response, the deviant logic has not earned the right to be taken seriously.

How the dialectic goes from there depends on what motivates the rejection of the classical principle. Some motivations are *internal* to mathematics. The standard example is the intuitionistic rejection of excluded middle, motivated by a constructivist conception of mathematics in general and infinity in particular. The rejection of classical mathematics is the *point* of intuitionism, not a grudgingly accepted incidental cost. Intuitionists willingly accepted the burden of building mathematics anew using only intuitionistically acceptable reasoning. The result may be no adequate substitute for classical mathematics, but that is a further issue.

A more recent example of logical deviance with a motivation at least partly internal to mathematics is paraconsistent logic. Dialetheists such as Graham Priest endorse an unrestricted comprehension principle for sets, including the inconsistent instance that generates Russell's paradox, while revising the logic to avoid the total collapse classically implicit in the claim that the Russell set is self-membered if and only if it is not self-membered (Priest 1995: 123-94). Here the motivation is overtly to simplify and clarify the axioms of set theory: some rebuilding of mathematics is intended, not just accepted as a by-product.

Even in these internally motivated cases, the usual strategy is not usually to construct a new mathematics from scratch, with no eye to the old. Often, it is to make the new as close as possible to the old, while starting with the deviant logic and respecting the motivation for reform. It is a project of *reconstruction*, not pure construction *ab initio*. That is hardly surprising, given the multitudinous successful applications of classical mathematics throughout science.

For other rejections of classical logic, the motivation is *external* to mathematics. Salient examples include various proposals motivated by phenomena of vagueness, such as sorites paradoxes. Typical cases involve non-mathematical words like 'heap', 'bald', 'red', and 'rich', which lack clear boundaries. On the assumption that the unclarity does not merely manifest our characteristic ignorance of the boundary's location, vagueness is supposed to involve some sort of non-epistemic indeterminacy that undermines the bivalent dichotomy of truth and falsity. Supervaluationism offers at least a quasi-classical treatment of such indeterminacy (§5 briefly discusses its non-classical aspect). As for more obviously non-classical treatments of vagueness, there is no consensus as to which is most promising. Many reject excluded middle and postulate some form of many-valued logic, typically either the three-valued strong Kleene logic K3 or a continuum-valued fuzzy logic. Another alternative is to reject Cut, the structural rule that allows a series of short arguments to be chained together as in a sorites paradox, so that if *C* follows from *B* and *B* from *A* then *C* follows from *A*. In any case, such proposals are usually not intended to disrupt classical mathematics. Core mathematical vocabulary is assumed to be precise, and so not to generate sorites paradoxes and other phenomena of vagueness. Thus classically valid reasoning is still supposed to preserve truth within a purely mathematical language.

One might contest the assumption that vagueness does not infect mathematics. For instance, even the language of pure set theory has been thought to be vague as to exactly what counts as a 'set', so that Cantor's continuum hypothesis comes out true on one legitimate sharpening but false on another (Feferman 2011). However, even if such vagueness is granted, it is usually assumed to be irrelevant for purposes of the ordinary working mathematician. Admittedly, some standard notation of ordinary working mathematics is vague by intention: for instance, $x \approx y$ is read '*x* is approximately equal to *y*', and $x << y$ is read '*x* is much smaller than *y*'. However, for the sake of argument we may concede that in principle such uses of vague notation in mathematical proofs can always be eliminated in favour of something more precise.

A further motivation external to mathematics for rejecting classical logic comes from the Liar and other semantic paradoxes. Metalinguistic terms like 'true' and 'false' do not figure in the core language of mathematical theories. Consequently, those who take the semantic paradoxes to motivate a retreat from classical logic to K3 or some other non-classical logic usually assume that their logical reform leaves classical mathematics itself intact.

Yet another example of logical deviance motivated externally to mathematics is quantum logic, which rejects the classical distribution principle that $A \wedge (B \vee C)$ entails $(A \wedge B) \vee (A \wedge C)$ because it supposedly fails in the quantum world. This too is usually assumed to do no damage to classical mathematics.

This chapter concerns the idea that, for all externally motivated logical deviance shows, classical logic is fine for mathematics. It has been voiced by various theorists. For example, Hartry Field takes excluded middle to fail in the presence of semantic and sorites paradoxes, but suggests (Field 2008: 101):

> Presumably excluded middle holds throughout ordinary mathematics, and indeed whenever vagueness or indeterminacy is not at issue.

Recently, Ole Thomassen Hjortland has articulated the idea in more detail (2017: 652-3):

> Mathematical proofs do contain an abundance of instances of classical principles: applications of classical *reductio ad absurdum*, conditional proof, disjunctive syllogism, the law of absorption, etc. The emphasis, however, should be on the fact that these are *instances* of classical principles. The mathematical proofs do not rely on any of these principles being unrestricted generalizations […]. They do at most rely on the principles holding restrictedly for mathematical discourse, which does not entail that the principles of reasoning hold universally. Put differently, mathematical practice is consistent with these reasoning steps being instances of *mathematical* principles of reasoning, not generalizable to all other discourses. *A fortiori*, they may very well be principles of reasoning that are permissible for mathematics, but not for theorizing about truth.

I will argue that such claims are much too optimistic about the prospects of isolating mathematics from logical deviance in non-mathematical discourse. They overlook the capacity of pure mathematics to be *applied*.

2. *The generality of mathematics*

The simplest way of applying pure mathematics is just by substituting what may loosely be called 'empirical terms' for the variables in its theorems, when the latter are universal generalizations. Of course, many actual applications to science involve something less direct and more elaborate than that, but if pure mathematics cannot be kept free of externally motivated logical deviance in the simplest applications, then the Field-Hjortland strategy is already in serious trouble.

We can start with a rather extreme external motivation for logical deviance: ontological vagueness. Consider the idea that there are vague objects, such as mountains, whose identity is vague in ways that make them counterexamples to the law of excluded middle for identity:

EMI $\forall x \ \forall y \ (x = y \vee x \neq y)$

A proponent of this view may hold that there is a mountain Everest and a mountain Chomolungma such that it is vague whether they are identical. The vagueness is not in the names but in the things named. Under the assignment of Everest and Chomolungma to the variables '*x*' and '*y*' respectively, the open formula $x = y$ is neither true nor false. Given the many-valued semantics which this theorist adopts, the disjunction $x = y \vee x \neq y$ is correspondingly neither true nor false under this assignment, so the universal generalization EMI is not true. Of course, this is a deeply contentious view, and very far from my own. Nevertheless, such views have been defended by serious, technically competent philosophers of vagueness (Parsons 2000).

For present purposes, the point is that EMI is a sentence in the pure language of first-order logic with identity, without non-logical predicates or individual constants. It therefore seems to belong to the core language of mathematics. Although Everest and Chomolungma are not mathematical objects, they are still in the range of the first-order quantifiers. Thus the logical deviance, although motivated externally, still infects mathematics. One cannot rely on excluded middle even for formulas in its core language.

For proponents of the isolationist strategy, the problem may seem to have an obvious solution. They can restrict their defence of classical logic to *pure* mathematics, where by definition the quantifiers are restricted to *purely mathematical* objects. In particular, the quantifiers in set theory as a branch of mathematics are restricted to *pure* sets, sets to which only sets bear the ancestral of the membership relation. A set is pure if and only if all its members are. Thus Everest, Chomolungma, their singletons, pair set, and so on, are excluded from the quantifiers of pure mathematics. On this reinterpretation, EMI is immune to the proposed counterexample.

But where does the proposed solution leave applied mathematics? At an elementary level, consider the statement that if there are exactly *m* apples and *n* oranges, and no apple is an orange, then there are exactly $m + n$ apples and oranges. It quantifies over non-mathematical objects, apples and oranges, and so falls outside pure mathematics, on the proposed restriction. It is not even clear how it might be derived from a theorem of pure mathematics, as just construed. Of course, Peano arithmetic itself quantifies only over natural numbers, but one might still expect pure mathematics as a whole to deliver the result that for any pluralities *xx* and *yy*, if there are exactly *m xx* and *n yy*, and none of the *xx* is one of the *yy*, then there are exactly $m + n$ *xx* and *yy*.

At a more sophisticated level, consider the application of group theory to the group of symmetries of a physical system. The symmetries are rotations, reflections, translations, or other mappings of the elements of the system onto each other. They are not purely mathematical objects. Thus group theory as a branch of pure mathematics on the proposed restriction does not deal with such symmetries. In some cases, the chief interest of a result in pure mathematics derives from its applications: Arrow's Theorem may be an example, with its application to voting rules.

At first sight, the problem of applied mathematics may in turn appear to have an easy solution. For it may be argued that any system of objects some of which are not purely mathematical is *isomorphic* to a system of purely mathematical objects. On this view, each set of non-sets is in a one-one correspondence with a set of pure sets, which induces a mapping of any properties, relations, or operations defined over the former to corresponding properties, relations, or operations defined over the latter.[1] Thus one can prove a result in pure mathematics and then by isomorphism transfer it to impure applied mathematics.

The isomorphism argument appeals to standard habits of mathematical thinking. However, they are invalidated by the very restriction of mathematics at issue. For starters, the claim 'Every set of non-sets corresponds one-one with a set of pure sets' falls outside pure mathematics on the proposed restriction, since it quantifies over objects that are not purely mathematical: sets of non-sets. It would have to be derived from some impure theory, which on the envisaged view is not entitled to rely on classical logic. Indeed, the alleged counterexample to EMI shows that the claim is quite dubious in this setting. For suppose that the set {Everest, Chomolungma} of non-sets corresponds one-one with a set X of pure sets. By hypothesis, pure set theory obeys classical logic, so either X has at most one member or X has at least two members. Therefore, since it corresponds one-one with X, either {Everest, Chomolungma} has at most one member, in which case Everest and Chomolungma are identical, or it has at least two members, in which case they are distinct. Thus the isomorphism argument conflicts with the assumption that excluded middle fails in the non-mathematical case.

One moral of that argument is that if set theory is to provide a foundation for mathematics, it must be *impure* set theory, for example ZFCU, Zemelo-Fraenkel set thery with the axiom of choice and ur-elements (non-sets), rather than just ZFC. For a key feature of pure mathematics is that it can be applied outside itself, to natural and social science. Without impure sets, set theory fails to enable such applications. An account of isomorphisms between pure and impure sets is itself part of impure set theory, not of pure set theory.

A similar moral goes for other putative foundations of mathematics: one condition of adequacy is that they enable pure mathematics to be applied outside itself. A topical case is homotopy type theory, the standard text on which advertises it as 'univalent foundations of mathematics' (Univalent Foundations Project 2013). Unfortunately, the text appears to miss this point, and presents homotopy type theory in a way which fails to enable applications outside itself. In effect, it presents homotopy type theory as failing to meet the condition of adequacy on foundations of mathematics.[2] Whether this defect can somehow be repaired is an open question.

In any case, even if we do find a sense in which every set of non-sets corresponds one-one with a set of pure sets, the strategy of transferring results about the latter to the former is undermined by the very hypothesis at issue, that classical logic holds when the quantifiers are restricted to purely mathematical objects but not when the restriction is lifted. For that hypothesis entails that what holds under the restriction is *not* structurally representative of what holds without the restriction. A parallel argument will apply to other putative foundational frameworks. Thus, given the hypothesis, the isomorphism strategy fails as a way of applying mathematics.

For all that has just been argued, some classical principles may hold under the restriction to purely mathematical objects but not more generally. But the point is that once you treat that as a live option, you cannot freely apply standard pure mathematics to science. Instead, you must develop an alternative mathematics for applications, based on your preferred non-classical logic. That is the hard work which proponents of the isolationist strategy were hoping to avoid.

Might the case of EMI be somehow misleading? After all, the hypothesis of ontologically vague identity is notoriously problematic. We may therefore consider a deviant logician who nevertheless accepts EMI. The next question is whether the standard, classical rule of universal instantiation can be applied to EMI. If so, we can derive an instance EMIi of EMI for any closed singular terms ‘*a*’ and ‘*b*’:

EMIi $\quad a = b \vee a \neq b$

For instance, ‘*a*’ and ‘*b*’ may abbreviate ‘Everest’ and ‘Chomolungma’ respectively. But a non-classical logician who regards vagueness as semantic rather than ontological may still reject EMIi, on the grounds that indeterminacy in how names refer can invalidate an instance of excluded middle.

The envisaged theorist accepts EMI but rejects its instance EMIi, and so must also restrict the standard rule of universal instantiation. Of course, that is not unprecedented: it happens in free logic. A similar restriction might be adopted here: to derive EMIi from EMI, one needs the auxiliary premises $\exists x\, x = a$ and $\exists y\, y = b$. The theorist rejects those premises, because ‘*a*’ and ‘*b*’ are (allegedly) indeterminate in reference. The trouble with this strategy is that in effect it treats semantic indeterminacy as reference failure: vague names become virtually unusable in science, because they fall outside the reach of universal generalizations, including mathematical theorems. That applies not just to radically vague names but even to slightly vague ones, since the slightest vagueness in ‘*a*’ suffices to make EMIi fail, for a suitable choice of ‘*b*’, on the envisaged view of vagueness. In practice, virtually any term for a physical quantity is at least slightly vague.

Such problems for universal instantiation arise even when the quantifiers are restricted to purely mathematical objects. For instance, in a suitable context, ‘*a*’ may abbreviate ‘the velocity of that cloud in kilometres per hour’, or ‘the current number of oxygen molecules in the Pacific ocean’, or ‘the largest small natural number’. One can talk vaguely about mathematical objects, just as one can about non-mathematical objects. On some precise, purely mathematical reading of ‘*b*’, it will then be vague whether $a = b$. Nor is this problem specific to the identity predicate. It arises for other mathematical predicates too. The same examples show that, on a non-classical semantic view of vagueness, it is problematic to reason from the classical principle TO that real or natural numbers are totally ordered in the standard way to a vague instance TOi:

TO $\quad \forall x\ \forall y\ \ (x \leq y \vee\ y \leq x)$

TOi $\quad a \leq b \vee\ b \leq a$

To avoid such problems on the semantic view, the relevant restriction is not to quantify only over purely mathematical objects but rather to use only purely mathematical vocabulary. However, that restriction merely exacerbates the problem of applied mathematics.

Is the difficulty specific to vague singular terms? If so, it should be resolved by the Quinean move of eliminating singular terms in favour of predicates. But, on reflection, such a move turns out to make little difference.

If we are allowed to quantify into predicate position, we have the second-order version of the problem of universal instantiation. For instance, on its intended interpretation the second-order universal generalization EM2 is precise:

EM2 $\quad \forall X\ \forall x\ \ (Xx \vee\ \neg Xx)$

Informally, EM2 says that for every way, everything either is or is not that way. On a suitably unrestricted understanding of the quantifiers, EM2 will be precise. Thus, if the problem for excluded middle is vagueness, then EM2 should hold. But if the second-order rule of universal instantiation is valid, then EM2 entails its instance EM2i:

EM2i $\quad \forall x\ \ (Fx \vee\ \neg Fx)$

But when the predicate *F* in EM2i is vague, those who take excluded middle to fail in case of vagueness will reject EM2i.[3] A similar problem arises if some classical principle other than excluded middle is taken to fail in case of vagueness. If mathematical theorems are formulated as precise higher-order generalizations, and classical logic holds for precise languages but not vague ones, then we cannot expect to instantiate those theorems with vague predicates, so their applications are blocked.

If one tries to approximate the generality of a second-order universal generalization (such as EM2) with that of the corresponding first-order schema, the problem is more immediate. On the envisaged view of vagueness, the unrestricted first-order schema is invalid, because instances such as EM2i fail when *F* is vague.

The point generalizes to other motivations for non-classical logic external to mathematics, such as the semantic paradoxes. For instance, EM2i is often alleged to fail when *Fx* abbreviates '*x* is true', even though EM2 remains a classical theorem in the language of pure second-order mathematics. Similarly, for proponents of quantum logic, instances of the distributive law fail for predicates ascribing properties to a particle, even though the law itself may be formulated as a classical theorem in the language of pure second-order mathematics. In such cases, pure mathematics cannot be applied by means of second-order universal instantiation.

It will not do simply to claim that we can recover the instances of the second-order universal generalization when we need them in unproblematic cases. For *Fx* may be problematic, generating a paradox, for some values of '*x*' and not for others, but to restrict second-order universal instantiation we must classify the general predicate *F* as problematic or unproblematic, since what gets substituted for the second-order predicate variable *X* is *F*, not the open formula *Fx* for some particular value of '*x*'.

Of course, in many circumstances (though not all) we can use first-order generalizations over sets instead of second-order generalizations into predicate position. Instead of instantiating the second-order predicate variable *X* with *F*, we then instantiate a first-order set variable with a set term such as '{*x*: *Fx*}'. For example, instead of EM2 and EM2i we have excluded middle for membership and its instances respectively:

EMM $\quad \forall s\ \forall x\ (x \in s \lor x \notin s)$

EMMi $\quad \forall x\ (x \in \{x\text{: } Fx\} \lor x \notin \{x\text{: } Fx\})$

But that is no help. For if there is such a set as {*x*: *Fx*}, then $x \in \{x\text{: } Fx\}$ should reduce to *Fx*, and $x \notin \{x\text{: } Fx\}$ to ¬*Fx*, so EMMi should reduce to EM2i, which is just what the envisaged non-classical logician rejects, while accepting the first-order universal generalization EMM, since it is a classical theorem in the language of pure mathematics. On the other hand, if there is no such set as {*x*: *Fx*}, then EM2i should not yield EMMi anyway. Either way, the envisaged non-classical logician refuses to go from EMM to EMMi, and so faces the first-order version of the obstacle to applying pure mathematics, already discussed above.

3. Are there false classical theorems?

For a deviant logician who rejects bivalence, a more modest claim is that, even if not every theorem of classical mathematics is true, none of them is false. Thus one will never go too badly wrong if one relies on classical mathematics. Some initial results point that way. For instance, if *A* is a theorem of classical logic, then no valuation in the Kleene weak or strong three-valued semantics makes *A* false.[4] One might hope to extend that result to all theorems of classical mathematics.

However, the result depends on peculiarities of the Kleene semantics. It does not apply to all forms of many-valued semantics. For proponents of many-valued logic, a natural thought about perfect borderline cases for vague expressions and Liar cases is that we are facing an aporia because there is an underlying semantic equality between a paradox-inducing sentence *A* and its negation ¬*A*, so they should get the very same value (neutral). But, on the analogy between many-valued and two-valued semantics, it is natural to allow a biconditional $\leftrightarrow$ that does for sentences what '=' does for singular terms: a symbol for extensional identity. That is what $\leftrightarrow$ does in two-valued semantics. Thus if the sentences *B* and *C* get the very same value, $B \leftrightarrow C$ should be true. In particular, $A \leftrightarrow \neg A$ should be true in case of aporia. Since the negation of a true sentence should be false, $\neg(A \leftrightarrow \neg A)$ should be false. But $\neg(A \leftrightarrow \neg A)$ is a theorem of classical logic; it is a two-valued tautology. In such ways, many forms of many-valued semantics allow some classical theorems to be false.

In the three-valued semantic framework, the absence of a connective behaving as $\leftrightarrow$ above looks like an expressive limitation of the Kleene approach. On the latter, one is forced instead to define $A \leftrightarrow B$ as $(A \wedge B) \vee (\neg A \wedge \neg B)$, which is neutral rather than true when both *A* and *B* are neutral; it entails both $A \vee \neg A$ and $B \vee \neg B$. This has major repercussions for the development of set theory on the Kleene approach. For instance, consider the axiom of extensionality:

AE $\quad \forall s\ \forall t\ (s = t \leftrightarrow \forall x\,(x \in s \leftrightarrow x \in t))$

Since $s = s$, from AE we can trivially obtain this from AE:

$$\forall s\ \forall x\ (x \in s \leftrightarrow x \in s)$$

But on the Kleene definition of $\leftrightarrow$, this entails EMI, excluded middle for membership. In a setting where excluded middle is not generally valid, that is a significant restriction: membership is never indefinite. Similarly, consider the existence of singletons:

ES $\forall x\ \exists s\ \forall y\ (y \in s \leftrightarrow x = y)$

On the Kleene definition of $\leftrightarrow$, ES entails EMI, excluded middle for identity.[5] In a setting where excluded middle is not generally valid, that too is a significant restriction: identity is never indefinite. To avoid such restrictions, many many-valued theorists will want a semantics on which $B \leftrightarrow C$ is true when B and C get the same value.[6] As noted above, they are then liable to falsify some theorems of classical logic and mathematics.

Proponents of such an approach might hope that classical theorems only get falsified when bivalence fails 'badly'. But that is not in general so. For instance, on a standard fuzzy logical approach, the values are real numbers in the interval [0, 1], where 1 represents perfect truth and 0 perfect falsity. The connectives are typically given the continuum-valued Łukasiewicz semantics, on which $B \leftrightarrow C$ gets value 1 whenever B and C get the same value. One might expect that if some atomic formulas are evaluated 'almost classically', all getting values very close to 1 or 0, then any classical theorem built up out of them will also be evaluated almost classically. But it is not so. Pick any positive real number δ, however small. Then on some valuation which assigns an atomic formula P a value within δ of a classical value, 0 or 1, some classical tautology built up only from P, $\leftrightarrow$, and ¬ comes out perfectly false (the appendix gives details). On such an approach, the slightest degree of vagueness can falsify some classical theorems.

That is far from the only case where a non-classical approach falsifies some classical theorems. For example, in intuitionistic mathematics, the theory of choice sequences has theorems of the form $\neg \forall x\ (Fx \vee \neg Fx)$ (Dummett 1977: 84; 2000: 61-2). In intuitionistic logic, although it is inconsistent to deny any specific instance of excluded middle, it is not always inconsistent to deny a universal generalization over infinitely many instances of excluded middle.

Although not all non-classical logics diverge so dramatically from classical logic, claims of approximately classical behaviour must be established, not assumed.

4. *The example of probability theory*

On the evidence so far, attempts to maintain classical logic within pure mathematics while rejecting it elsewhere create severe obstacles to the application of mathematics. We can test this conclusion against the case of probability theory, a branch of mathematics in which theorems are rigorously proved but applications supply much of the motivating force.

From a purely mathematical point of view, a *probability space* is any triple $\langle \Omega, \mathcal{F}, \mathrm{P} \rangle$ where Ω is a nonempty set, $\mathcal{F}$ is a collection of subsets of Ω containing the empty set and closed under countable unions and complements in Ω (a σ-field), P is a function mapping each member of F to a nonnegative real number, $P(\Omega) = 1$, and P satisfies countable additivity: if X is the union of the pairwise disjoint members X_0, X_1, X_2, … of F then $P(X)$ is the sum of $P(X_0)$, $P(X_1)$, $P(X_2)$, … . Thus what it is to be a probability space is defined in purely mathematical terms; the conditions amount to the Kolmogorov axioms for probability and entail other standard features of probability. The restriction of P to a σ-field is to avoid technical problems which otherwise occur when Ω is infinite. Such a definition enables the theory of probability spaces to be a branch of pure mathematics.

Informally, however, a more specific interpretation motivates the development. A standard textbook opens with this (Grimmett and Stirzaker 2001: 1):

> **Definition**. The set of all possible outcomes of an experiment is called the **sample space** and is denoted by Ω.

Another textbook begins ‘Let us consider an experiment of which all possible results are included in a finite number of outcomes’ (Shiryaev 2016: 1). A third just speaks of ‘outcomes’ in inverted commas, without saying what they are outcomes of (Durrett 2010: 1). Although the informal idea of a possible outcome of an experiment plays no essential role in the proofs of theorems, it helps give point to the enterprise.[7]

The informal understanding of probability spaces goes further (Grimmett and Stirzaker 2001: 2):

> we think of *events* as subsets of the sample space Ω. Whenever *A* and *B* are events in which we are interested, then we can reasonably concern ourselves also with the events $A \cup B$, $A \cap B$, and A^c, representing ‘*A* or *B*’, ‘*A* and *B*’, and ‘not *A*’ respectively.

Of course, the standard definitions of the set-theoretic operations of union, intersection, and complementation already link them to the respective sentential connectives: $A \cup B$ is $\{x: x \in A$ or $x \in B\}$, $A \cap B$ is $\{x: x \in A$ and $x \in B\}$, and A^c is $\{x: x \in \Omega$ but not $x \in A\}$. More generally, when we apply the mathematical apparatus, we wish to assign probabilities to events in the informal sense, which we can specify by sentences such as ‘There was an explosion’. To do so, we must first interpret those sentences by mapping them to events in the technical sense, subsets of some suitable Ω, sets of possible outcomes of an experiment (often in a very broad sense).

To be more explicit about the relation between sentences and events as subsets of Ω, let the sentence *A* be interpreted by the event [*A*]. We can also use standard symbols in place of the natural language connectives ‘or’, ‘and’, and ‘not’. Thus:

$$[A \vee B] = [A] \cup [B]$$

$[A \wedge B] = [A] \cap [B]$
$[\neg A] = [A]^c$

Consequently, in classical mathematics, $[A \vee \neg A] = [A] \cup [A]^c = \Omega$, so $P([A \vee \neg A]) = P(\Omega) = 1$ in any probability space. Thus, if classical logic is valid for pure mathematics, applying standard probability theory involves assigning probability 1 to the law of excluded middle, which opponents of the law will resist. A similar argument obviously goes for any other classical tautology in disjunction, conjunction, and negation, and for the distributivity principle that fails in 'quantum logic'.

Of course, we can conceive alternative rules for assigning events to formulas of the forms $A \vee B$, $A \wedge B$, and $\neg A$. But they all amount to deviant interpretations of the connectives, and so merely change the subject.

Thus if deviant logicians accept a classical tautology involving disjunction, conjunction, and negation within mathematics but not outside, it is quite unclear how they can apply the results of standard probability theory to natural science.

5. *Mathematical reasoning*

So far, applying pure mathematics has been treated simply as a process of instantiating its theorems. However, mathematics is used in science far more pervasively than that. For example, given a scientific law or hypothesis in the form of an equation, one uses ordinary mathematical reasoning to deduce its consequences, typically with statements of initial conditions or other auxiliary assumptions as extra premises. Rejections of classical logic often have the effect of calling such reasoning into question, even when motivated externally to mathematics.

To illustrate the variety of inference rules rejected by one non-classical logician or another, we need look no further than the problem of vagueness, which has been taken to motivate the rejection of almost any given rule:

Some fuzzy logicians define a valid argument as one such that on every interpretation on which every premise reaches a fixed threshold of degree of truth, say 0.9, so does the conclusion. By that standard, the rule of modus ponens comes out invalid: even if its premises are 'true enough', its conclusion may not be. For instance, in a sorites series for the vague predicate *F*, if *Fa* has degree of truth 0.91 while *Fb* has degree of truth only 0.89, then $Fa \rightarrow Fb$ has degree of truth 0.98 by one standard semantic rule (see the Appendix), so *Fa* and $Fa \rightarrow Fb$ reach the threshold of 0.9 while the conclusion *Fb* falls short.[8]

By contrast, some paraconsistent accounts of vagueness invalidate the rule of disjunctive syllogism, on the grounds that it is not truth-preserving. For if borderline statements are treated as truth-value gluts, both true and false, and *A* is borderline while *B* is plain false, then $A \vee B$ is true (because *A* is true), and ¬*A* is true (because *A* is also false), while the conclusion *B* is plain false.

Supervaluationist approaches to vagueness typically validate all theorems of classical logic and all classical rules of inference (including modus ponens and disjunctive syllogism) from premises in the object-language to a conclusion in the object-language. However, one of the main supervaluationist approaches invalidates some classical meta-rules according to which the validity of one or more object-language arguments implies the validity of another object-language argument. More specifically, it invalidates standard classical meta-rules which discharge premises, such as conditional proof (if *B* follows from Γ and *A* then $A \rightarrow B$ follows from Γ alone), reductio ad absurdum (if both *B* and ¬*B* follow from Γ and *A* then ¬*A* follows from Γ alone), and proof by cases (if *C* follows from Γ and *A*, and also from Δ and *B*, then *C* follows from Γ, Δ, and $A \vee B$).[9]

Others have rejected standard structural meta-rules, which are not specific to a particular connective, such as the Cut rule (if *A* follows from Γ, and *B* follows from Δ and *A*, then *B* follows from Γ and Δ), in order to block sorites paradoxes. For example, let *S* be the tolerance principle that the successor of any small natural number is small, where 'small' is vague; then even if our other rules tell us that '1 is small' follows from '0 is small' and *S*, and that '2 is small' follows from '1 is small' and *S*, we typically need Cut to conclude that '2 is small' follows from '0 is small' and *S*.[10]

The rejection of any of the rules just mentioned could also have been illustrated from non-classical treatments of the semantic paradoxes.

The difficulty of being classical inside mathematics but non-classical outside generalizes from theorems to rules and meta-rules. For suppose that a deviant logician accepts classical reasoning within mathematics but rejects a classical rule or meta-rule *R* as invalid outside, and so rejects some classically valid reasoning from premises Γ to a conclusion *A* because one of its steps applies *R* to a non-mathematical atomic expression *E*, such as a vague or semantic predicate. Now replace *E* by a new variable *V* of the same grammatical type throughout the reasoning. Unlike *E*, *V* lacks a distinctively non-mathematical meaning. Repeat the process until all the non-mathematical atomic expressions in the original reasoning have been replaced by distinct new variables. The result is some analogous reasoning from premises Γ* to a conclusion *A**, which is still classically valid because it has the same relevant form as the original classically valid reasoning. Our deviant logician accepts the reasoning from Γ* to *A** because it is classically valid and it contains no extra-mathematical elements. But we can recover the original reasoning from Γ to *A* simply by substituting *E* back for *V*, or interpreting *V* in terms of *E*, and likewise for any other substitutions made.

The deviant logician will doubtless refuse to give up, and will instead reject the substitution of *E* for *V* as illicit. This rejection is either sensitive or insensitive to the details of the reasoning from Γ* to *A**. If it is insensitive, then it amounts to a blanket rejection of the application of all mathematics and logic to expressions such as *E*, which puts the latter beyond the reach of reason. But terms such as 'heap' and 'small', 'true' and 'false', are *not* beyond the reach of reason. On the other hand, if the rejection is sensitive to the details of the reasoning, then the deviant logician is in effect already engaged in the project of reconstructing mathematics to determine, for purposes of applying it, how much can survive the retreat to the deviant logic.

Sometimes the variable *V* will be of a higher grammatical type, not usual in ordinary mathematical notation. The reasoning from Γ* to *A** may then be more logical than mathematical in flavour. For present purposes, however, that is a superficial feature. For example, if '*x* is red' in the reasoning from Γ to *A* became *Rx*, for a predicate variable *R*, in the reasoning from Γ* to *A**, we could instead have used $x \in r$ (with an eye to the set of all red things) or '$r(x) = 1$' (with an eye to the characteristic function for 'red'). That would have given the substitute argument a more characteristically mathematical flavour, while raising the same key issues. If there is no set of all red things, no characteristic function for 'red', and so on, then how can we apply standard set-theoretic or function-theoretic mathematical reasoning to the distinction between red and non-red at all?[11]

For mathematical reasoning, the conclusion is the same as for mathematical theorems: if classical logic fails outside pure mathematics, and pure mathematics is fit to be applied outside itself, then classical logic fails inside pure mathematics too.

6. The closure role of logic

Another way to think of the background role of logic and mathematics in natural and social science is as a *closure operator*. We want to test scientific hypotheses by their logico-mathematical consequences, so a system of logico-mathematical reasoning should induce a mapping from each set of hypotheses Γ to its set of consequences Cn(Γ) according to that system, the closure of Γ under its consequence relation. For simplicity, we can treat the hypotheses and the consequences as sentences of the same language. Indeed, for many purposes we can identify the system of reasoning with this closure operation.

Closure operations have some standard structural features:[12]

(i) If $\Gamma \subseteq \Delta$ then $Cn(\Gamma) \subseteq Cn(\Delta)$

(ii) $\Gamma \subseteq Cn(\Gamma)$

(iii) $Cn(Cn(\Gamma)) \subseteq Cn(\Gamma)$

Arguably, all three features are assumed in the ordinary testing of scientific theories. There is no limit to the length of permitted chains of reasoning, so consequences of consequences of Γ are consequences of Γ, as (iii) says. Even a zero-length chain of reasoning counts, so members of Γ are consequences of Γ, as (ii) says. Finally, if $\Gamma \subseteq \Delta$, then we can trivially reason from Δ to each member of Γ; since we can reason from Γ to each member of Cn(Γ), $\text{Cn}(\Gamma) \subseteq \text{Cn}(\Delta)$, as (i) says.[13]

Together, (i)-(iii) imply all the standard structural rules for a consequence relation, including Cut, which in this notation says: if $A \in \text{Cn}(\Gamma)$ and $B \in \text{Cn}(\Delta \cup \{A\})$ then $B \in \text{Cn}(\Gamma \cup \Delta)$.[14] This is a problem for those who reject Cut in an attempt to solve semantic paradoxes or the paradoxes of vagueness. For example, suppose that we start by accepting that the successor of any small natural number is small and that 0 is small. We then deduce that 1 is small, and accept that too. Since we now accept that the successor of any small natural number is small and that 1 is small, we deduce that 2 is small, and accept that too, even though it does not follow from our original two postulates in a logic where Cut fails. We continue indefinitely, at each stage accepting only consequences in the Cut-free system of what we already accept. Evidently, we have not escaped the sorites paradox. In response to such difficulties, David Ripley, a friend of logics where Cut fails, admits that they are ill-suited to playing the closure role; he has other purposes in mind for them.[15]

Another substructural strategy involves replacing sets of premises by *multisets*, in which the same sentence may occur many times. The idea is that using a premise many times may require many occurrences of that premise. This involves rejecting the rule of Contraction, on which one occurrence of a premise is as good as two. The paradoxes may then be blamed on illicit uses of Contraction. However, when we are investigating the consequences of a scientific hypothesis, 'How many occurrences of the hypothesis are we assuming?' is not a sensible question.[16]

In general, substructural logics are ill-suited to acting as background logics for science.

7. Non-classical mathematics

When non-classical logics are introduced to treat paradoxes of vagueness, semantic paradoxes, or the like, the aim is to provide a systematic framework for correct deductive reasoning with vague, semantic, or other vocabulary for which classical reasoning is supposed to break down. This should extend to a systematic framework for correct *mathematical* reasoning with the same vocabulary. As already argued, if the non-classical logic is needed at all, it is also needed for mathematical reasoning with those terms. If we could simply substitute them for variables in classical mathematical reasoning, there would have been no need for the non-classical logic in the first place. The lazy strategy does not work: mathematics must be redeveloped from scratch within the non-classical framework. Although there may be pockets within mathematics where everything behaves classically, on the non-classical view they cannot include mathematics for general applications.

What will non-classical mathematics be like? Obviously, that depends on the non-classical logic. As a case study, let us consider the prospects for the classical *least number principle* in a setting where excluded middle is supposed to fail for vague or semantic vocabulary.

A set-theoretic version of the least number principle says that every nonempty set of natural numbers has a least member. In practice, we often need to apply the principle to a complex predicate *F*, in the form:

LNP $\exists n\, Fn \rightarrow \exists n\, (Fn \;\&\; \forall k\, (k<n \rightarrow \neg Fk))$

To obtain LNP from the set-theoretic version, we need a set $\{n: Fn\}$, where $k \in \{n: Fn\}$ is equivalent to *Fk*. For present purposes, we can ignore the set-theoretic digression, and focus directly on the predicate version. Classically, LNP can be simply proved from the principle of mathematical induction. One shows by induction on *m* that this holds for all natural numbers *m*:

$$\exists n\, (n<m \;\&\; Fn) \rightarrow \exists n\, (Fn \;\&\; \forall k\, (k<n \rightarrow \neg Fk))$$

But if excluded middle fails in case of vagueness, LNP is implausible for many vague predicates *F*. For instance, if *Fn* is read as '*n* is large', where 'large' is vague, LNP goes from the obvious premise that there is a large natural number to the highly problematic conclusion that there is a least large natural number—what is it?

In response to this problem, Hartry Field (2008: 100-1) proposes a watered-down version of LNP:

GLNP $\exists n\, (Fn \;\&\; \forall k\, (k<n \rightarrow (Fk \vee \neg Fk))$ implies $\exists n\, (Fn \;\&\; \forall k\, (k<n \rightarrow \neg Fk))$

The point of the restriction in the premise is that $Fn \;\&\; \forall k\, (k<n \rightarrow (Fk \vee \neg Fk))$ in Field's framework implies that *F* behaves like a precise predicate from *n* down, and *Fn* implies that there is no need to look above *n* for the least satisfier of *F*.

The trouble with GLNP is that it is a one-off postulate in place of a mathematical *result*, proved from more basic principles. Consider any other theorem-schema of classical mathematics, formulated with predicate variables like *F*. Does the non-classical logician simply look at it, and make an educated guess at the minimal mutilation of it to escape counterexamples with vague or semantic predicates in place of the variables? Such an unsystematic, conjectural approach would fall far short of the standards of contemporary mathematics, which provides an enormous accumulating body of theorems ultimately derived from a very small group of first principles, such as the axioms of ZFCU. If mathematics were done in such an ad hoc spirit, one should have much less confidence in its results.

Alternatively, the non-classical logician might try to derive, within the favoured non-classical logic, some watering-down of all classical mathematics from a small group of first principles, presumably similar to those of classical mathematics but rendered invulnerable to counterexamples with vague or semantic predicates. That looks like an enormous undertaking, if done properly.

One tiny step would be to derive GLNP from a suitable version of mathematical induction. In the setting of strong Kleene logic, it is best formulated as an inference rule (Halbach and Horsten 2006: 692), where *t* can be any term for a natural number:

MI If from given side premises Γ (in which the variable '*n*' does not occur free) one can prove both *F*0 and, also given *Fn*, *Fn*+1, then from Γ one can prove *Ft*.

One can then universally generalize the conclusion over all natural numbers. With an appropriate background logic, one should then be able to derive GLNP from MI.

The point of requiring a proof of *Fn*+1 from the additional premise *Fn*, rather than simply a proof of $Fn \rightarrow Fn{+}1$, is that the logic affords modus ponens but not the deduction theorem, so the former requirement is weaker, making MI more general. For instance, if an inductive definition reduces *Fn*+1 to *Fn*, but does not exclude the case when both take the intermediate value, one may be able to prove *Fn*+1 from *Fn* without being able to prove $Fn \rightarrow Fn{+}1$, which itself takes the intermediate value in that case. In strong Kleene logic, the standard axiom schema for mathematical induction is unsatisfactory for the opposite sort of reason:

MI_{Ax} $(F0 \wedge \forall n\, (Fn \rightarrow Fn{+}1)) \rightarrow \forall n\, Fn$

MI_{Ax} is too strong in this setting, since not all its instances are true. To take an extreme case, if *Fn* takes the intermediate value whatever natural number is assigned to the variable '*n*', then so does the corresponding instance of MI_{Ax}. The inferential forms of mathematical induction avoid that problem.[17]

Some non-classical logicians may even reject MI, because they reject its instance when *Fn* is read as '*n* is small'; they interpret the derivation of *Fn*+1 from *Fn* as a sort of tolerance principle for 'small'. They may accept a watering-down of MI with the extra premise $\forall n\, (Fn \vee \neg Fn)$, but if excluded middle fails for vague predicates, the question is whether any watering-down of mathematical induction is applicable to vague predicates of natural numbers. If not, we are drastically limited in our mathematical reasoning with such predicates.

Those who endorse strong Kleene logic may accept MI. But then they will reject the classical derivation of LNP from MI, for on their view LNP fails when *Fn* is read as '*n* is small'.

Treatments of the semantic paradoxes typically accept mathematical induction in the form of the rule MI, even for semantic predicates *F* (Halbach and Horsten 2006). However, standard mathematical principles of transfinite induction can be unavailable in non-classical settings. For instance, induction is provable in full generality only up to the ordinal ω^{ω} in the partial Kripke-Feferman system PKF, a formal theory of truth based on strong Kleene logic, motivated by the semantic paradoxes, although it can be proved in PKF up to ε_0 for formulas without the truth predicate (Halbach 2015, corollary 16.7).

The underivability of LNP from the rule MI also demonstrates that the common practice of measuring the mathematical strength of theories of arithmetic by which instances of mathematical induction they prove is inadequate when the background logic is non-classical. For a theory may do very well by that standard, yet lack many classically trivial consequences of those instances, and so be near-useless for applications. In the non-classical setting, we must check that standard results over a much wider range are derivable before we can start to regard the theory as seriously applicable.

These difficulties with induction and the least number principle are just a foretaste of the challenges for the reconstruction of mathematics within a non-classical logic. One can expect the non-classical logician to respond wherever possible with local 'recovery' results to the effect that classical logic can be recovered for a specified class of formulas on condition that their constituents are 'well-behaved', for example by obeying excluded middle. For this strategy to be cogent, at least the initial recovery results must themselves be proved within the corresponding non-classical metalogic. For metalogic is itself a branch of mathematics; to use a classical metalogic at this stage would be to assume something of the very kind that was to be proved. That applies to the whole meta-logical background, including the numerical coding of formulas. Jack Woods has raised serious doubts as to whether the required non-classical proofs of recovery results will be available to non-classical logicians. More specifically, even if there is a classical proof that there is a proof in the relevant non-classical system of ϕ, that classical meta-theorem does not entitle non-classical logicians by their lights to rely on ϕ in their own inquiries, as opposed to using it *ad hominem* against their classical opponent. They need to display the non-classical proof itself, or at least to display a proof in their non-classical meta-logic that there is such a ground-level non-classical proof. That may be humanly impossible, for instance if the length of the corresponding non-classical proof grows exponentially with the length of the classical growth (Woods 2017).

Suppose, however, that a recovery result can somehow be non-circularly obtained, and we use it to justify relying on classical mathematical reasoning in a given scientific explanation.

This will involve adding an auxiliary premise to the explanation, to ensure that the relevant terms are well-behaved, for instance by obeying excluded middle. That might be assumed directly, just by listing the relevant instances of excluded middle on a case-by-case basis. However, the more extra premises one adds, the more the explanatory cost, just as *avoiding* reliance on a premise *adds* explanatory value, by making the explanation more elegant and economical. Alternatively, one might add both an extra major premise, to the effect that excluded middle holds for all precise, non-semantic terms, and an extra minor premise, to the effect that the terms at issue are precise and non-semantic. But such metalinguistic premises seem even more out of place in an ordinary natural scientific explanation of a non-linguistic state of affairs. They are not even partly why the state of affairs obtains. Thus the recovery strategy has a tendency to degrade ordinary explanations in natural science.

8. The easy-going attitude

One sometimes encounters the easy-going attitude that classical logic is a good enough approximation for some purposes and not for others, just as physicists who accept Einsteinian special relativity theory may still treat Newtonian mechanics as sometimes but not always a good enough approximation. An advantage of this attitude is that it enables one to use classical logic when one judges it to be good enough, with no great pressure to specify the precise conditions under which it is good enough. One simply judges it to be good enough in ordinary mathematical contexts, but not good enough in the presence of paradoxes of vagueness or semantic paradoxes. That is compatible with several contrasting attitudes to the non-classical logic. Some may regard it as, unlike classical logic, correct without exceptions. Others may regard it as a better approximation than classical logic, but still only an approximation to genuine logical consequence. Still others may be undecided between those two attitudes.

The easy-going attitude is often used as an excuse for not developing an alternative mathematics on the basis of the preferred non-classical logic. But the excuse presupposes that the contexts in which one needs mathematics are disjoint from the contexts in which one needs non-classical logic. That presupposition is hard to motivate. Suppose, for example, that one takes the semantic paradoxes to require a non-classical logic. Versions of those paradoxes arise when one adds a truth predicate to Peano arithmetic, by the standard devices of Gödel numbering and diagonalization. Thus, in order to give a rigorous theory of truth in mathematics as a whole, the theorist's commitments require a development of mathematics on the basis of the preferred non-classical logic. Again, since quantum mechanics applies mathematics ubiquitously, those who propose non-distributive quantum logic as a serious alternative to classical logic are not excused from the need to reconstruct mathematics on the basis of their quantum logic.

Theorists who motivate a non-classical logic by appeal to paradoxes of vagueness may take themselves to escape the requirement, on the grounds that when applying mathematics in science one can *pretend* that the relevant non-mathematical terms are precise, because the gap between this pretence and reality is irrelevant for scientific purposes. But that argument is simplistic. Here are some reasons why:

(I): Physics and other natural sciences are often in the business of making accurate quantitative predictions, sometimes with an expected error much less than one in a thousand. Vast sums of money may be spent on the basis of those predictions, as when a probe is sent to another planet. Theories are also tested by such predictions, as with a particle accelerator: then accuracy is needed for the value the quantity *would* take if the theory *were* true. Such predictions rest on complex mathematical reasoning and calculations, using standard classical mathematics. If there is vagueness in the relevant terms, and the true mathematics for vague terms is non-classical (because the correct logic for them is), then one would expect 'vagueness errors' in the predictions. Moreover, it would not be reasonable to expect the vagueness errors to take the form of random noise, normally distributed about the true value. Since the true non-classical mathematics will differ structurally from standard classical mathematics, the difference might easily result in errors of a more systematic kind, with some sort of structural bias. When scientists spend so much effort in meticulously controlling for other sources of error, should they not also make some attempt to control for vagueness errors? That would require a careful study of the nature of the divergences between classical mathematics and the true non-classical mathematics, which would in turn require the latter to have been developed. Here the easy-going attitude is not good enough.

(II): Consider the application of probability theory (§4) to the beliefs and reasoning of uncertain agents who specify the relevant events in vague terms, as real agents do. Such vagueness is highly relevant to doxastic and evidential probability, because borderline cases result in characteristic uncertainty effects. If the true logic and mathematics of vagueness is non-classical, then any psychologically realistic account must take that into account. Probabilities are being assigned to 'vague events'. For reasons already explained, in such a setting non-classical theorists cannot simply help themselves to standard classical mathematics in the background. From their perspective, they must carefully justify any distinctive use of classical mathematics, on pain of distorting the very phenomena they are trying to understand. Here too the easy-going attitude is out of place.

(III): consider the use of mathematics in legal cases, for instance to analyse statistical evidence, as with DNA testing. In legal reasoning, such mathematical considerations interact with the imprecise terms in which the law is often cast and evidence is often given. If the true mathematics for such legal discourse is non-classical, that needs to be taken into account. When a verdict of innocence or guilt is at stake, the easy-going attitude is hardly appropriate.

Of course, in practice we seem to get by well enough in (I)-(III) without taking into account the possibility of vagueness errors in classical mathematics. But friends of a non-classical logic of vagueness should take no comfort from that fact, since it is some evidence that the logic of vagueness is classical after all (Williamson 1994). It does not support the easy-going attitude.

Friends of non-classical logic sometimes assume that classical logic is simply refuted by paradoxes of one sort or another, and is no longer even a candidate for non-approximate correctness. But such crude falsificationism is methodologically naïve. Classical treatments are available for all the paradoxes used to motivate non-classical logics. Although some people find those treatments implausible, the issues are too subtle, abstract, and theoretical to be settled either way by such pre-theoretical judgments. Rather, in the long run they will be decided by an abductive comparison of the rival accounts (Williamson 2013: 423-9; 2017). Since the accounts differ in their component logics, a major aspect of the comparison will naturally be the serviceability of those logics for mathematics and its applications in science, which is just where the easy-going attitude undermines non-classical logics. Since it invokes them only when the need is supposed to be most urgent, in the immediate presence of paradoxes, they lack the massive abductive support that accrues to classical logic in the vast range of non-paradoxical cases where explanations rely on standard classical mathematics. The effect of the easy-going attitude is to concede the abductive comparison to classical logic.

9. Conclusion

Pure mathematics is applicable to the world outside pure mathematics. As a result, in a non-classical world, pure mathematics is no safe haven for classical logic. Advocates of non-classical logic motivated by non-mathematical considerations have often grossly underestimated the challenge of reproducing the success of classical mathematics within their framework.[18]

Appendix

On the standard continuum-valued Łukasiewicz semantics for fuzzy logic, a valuation v maps each formula A to a real number $v(A)$ in the interval [0, 1] (see Priest 2008: 224-231 for an introduction). Informally, 1 is understood as perfect truth and 0 as perfect falsity. For the usual connectives, the compositional evaluation rules are these:

$v(\neg A) = 1 - v(A)$

$v(A \wedge B) = \text{minimum}\{v(A), v(B)\}$

$v(A \vee B) = \text{maximum}\{v(A), v(B)\}$

$v(A \rightarrow B) = \text{minimum}\{1, 1 - (v(A) - v(B))\}$

For the biconditional, $A \leftrightarrow B$ is taken to abbreviate $(A \rightarrow B) \wedge (B \rightarrow A)$, which induces this evaluation rule:

$v(A \leftrightarrow B) = 1 - |v(A) - v(B)|$

Here $|v(A) - v(B)|$ is the positive difference between $v(A)$ and $v(B)$; thus the more A and B differ in value from each other, the more the value of $A \leftrightarrow B$ falls short of perfect truth.

For any formula A, the formula $\neg(A \leftrightarrow \neg A)$ is a classical tautology. A quick calculation shows that on the continuum-valued semantics, $v(\neg(A \leftrightarrow \neg A)) = |2v(A) - 1|$. Take any real number $\delta > 0$. For some large enough natural number $k \geq 1$, $2^{-k} < \delta$. Let P be an atomic formula. Inductively define:

$\#^0(P) = P$

$\#^{n+1}(P) = \neg(\#^n(P) \leftrightarrow \neg\#^n(P))$

Consider a valuation v such that $v(P) = 1 - 2^{-k}$. We show by induction that for $n \leq k$, $v(\#^n(P)) = 1 - 2^{n-k}$. The basis ($n = 0$) is trivial. For the induction step, suppose that $v(\#^n(P)) = 1 - 2^{n-k}$ and $n+1 \leq k$, so $1 - 2^{n-k} \geq ½$. Then:

$v(\#^{n+1}(P)) = v(\neg(\#^n(P) \leftrightarrow \neg\#^n(P))) = |2v(\#^n(P)) - 1| = |2(1 - 2^{n-k}) - 1| = 1 - 2^{n+1-k}$

This completes the induction. Hence, putting $n = k$, $v(\#^k(P)) = 1 - 2^{k-k} = 0$. Since $k \geq 1$ by hypothesis, $\#^k(P)$ is a classical tautology, and it is built up by logical connectives out of the formula P whose value under v differs by less than δ from the 'classical' value 1, but v assigns $\#^k(P)$ the lowest value, 0.

Notes

1 See Uzquiano 2015: 15 for a case (which he does not quite endorse) that the non-sets can be mapped 1-1 into the ordinals, which are purely mathematical objects (such a mapping implies a choice principle for non-sets). The mapping would then induce a 1-1 mapping from sets of non-sets to sets of ordinals, which are also purely mathematical objects, and so on for all impure sets. He also provides a useful account of relations between various choice and maximality principles for proper classes in the setting of impure set theory (op. cit.: 17-19).

2 Three examples, which do not detract from the intrinsic mathematical interest of homotopy type theory, only from its capacity to enable applications: (a) The univalence axiom is informally paraphrased as saying 'that isomorphic things can be identified' in the strong sense that 'every property or construction involving one also applies to the other' (Univalent Foundations Program 2013: 5); but if the parts of two snowflakes constitute isomorphic physical systems with respect to given spatial relations, that does not make those physical systems have the same properties; for example, they differ in spatial location. (b) Homotopy type theory is supposed to subsume set theory because 'we can *define* a class of types which behave like sets' (*ibid.*: 6), but such structural analogies are no substitute for impure sets as a bridge between pure mathematics and its applications. (c) Homotopy type theory has a constructivist aspect: 'we can regard a term $a : A$ as both an element of the type A (or in homotopy type theory, a point of the space A), and at the same time, a proof of the proposition A' (*ibid*.: 8). A particle may be in some sense an element of a type or a point of an abstract space, but it is not itself a mathematical proof (though propositions about it have various entailments).

3 If we analyse the vague name '*a*' by the vague definite description 'the *F*', treated Russell's way, the assumption $\exists x\, a = x$ becomes equivalent to $\exists x\ \forall y\ (Fy \leftrightarrow y = x)$, given which the definiteness of *F*, EM2i, should follow from the definiteness of identity, EMI. Thus if one wants the definiteness of identity but not of *F*, one must give up the assumption $\exists x\, a = x$, and so cannot expect to substitute the name in universal generalizations.

4 Proof: Let the three values be T (true), N (neutral), and F (false). Define a partial order ≤ on {T, N, F} thus: X ≤ Y iff either X = N or X = Y. A valuation V *refines* a valuation U iff for every atomic formula *P*, U(*P*) ≤ V(P); V is *bivalent* iff for every

atomic formula *P*, V(*P*) ∈ {T, F}. One easily shows by induction on the complexity of the formula *A* that, for both the weak and strong Kleene semantics, if V refines U then U(*A*) ≤ V(*A*) and if V is bivalent then V(*A*) is the same as the value of *A* on a classical valuation coinciding with V on atomic formulas. But every valuation U is refined by some bivalent valuation V; thus U(*A*) ≤ V(*A*), so if U(*A*) = F then V(*A*) = F and *A* is not classically valid. The induction step works for the usual quantifiers as well as the usual sentential connectives.

5 One can also use ES to obtain the definiteness of identity, EMI, from the definiteness of membership, on any reasonable account of ↔.

6 For a many-valued treatment of the semantic paradoxes, one drawback of such a deviation from the Kleene approach is that it can generate revenge paradoxes. This applies to both one-way and two-way conditionals, which are naturally interdefinable in the presence of conjunction: $A \rightarrow B$ as $A \leftrightarrow (A \wedge B)$ and $A \leftrightarrow B$ as $(A \rightarrow B) \wedge (B \rightarrow A)$. Hartry Field has struggled with the challenge of introducing a good conditional into such a treatment of the paradoxes; for a recent version see Field 2016.

7 For the modal aspect of a probability space of possible outcomes see Williamson 2016.

8 For details of various ways of rejecting modus ponens in many-valued logic see Williamson 1994: 103-24.

9 Γ and Δ are sets of premises. For details on the failures of the meta-rules under supervaluationism see Williamson 1994: 151-2. See also Williamson 2018 for further extensions of these results, including to a standard elimination rule for the existential quantifier and to vague mathematical notation (the cases of ≈ and << mentioned in §1). Not all theorists count such meta-rules as part of ‘classical logic’, but standard mathematical reasoning relies on them.

10 See Ripley 2013. He does not count the structural meta-rules as part of ‘classical logic’, but standard mathematical reasoning relies on them.

11 For purposes of applications, the sets and functions at issue can typically be restricted to a set-sized domain of discourse, and so are not ‘too large to exist’.

12 Tarski's five axioms for such a closure operation Cn imply (i)-(iii) (Tarski 1983: 31). They also have other implications of less interest here. His axiom 1 says that the set of sentences of the language is at most countably infinite. Axiom 4 postulates compactness: a sentence follows from a set if and only if it follows from a finite subset of that set; this fails for second-order logic under its standard semantics. Axiom 5 says that there is a sentence from which every sentence follows; this fails for some versions of relevance logic. The axioms for a closure operator in Ripley 2015 are equivalent to (i)-(iii).

13 Of course, the monotonicity principle (i) fails for defeasible reasoning, but that is not at issue here.

14 Proof: If $A \in \mathrm{Cn}(\Gamma)$ then $A \in \mathrm{Cn}(\Gamma \cup \Delta)$ by (i); but $\Delta \subseteq \mathrm{Cn}(\Gamma \cup \Delta)$ by (i) and (ii), so $\Delta \cup \{A\} \subseteq \mathrm{Cn}(\Gamma \cup \Delta)$, so by (i) $\mathrm{Cn}(\Delta \cup \{A\}) \subseteq \mathrm{Cn}(\mathrm{Cn}(\Gamma \cup \Delta))$; hence $\mathrm{Cn}(\Delta \cup \{A\}) \subseteq \mathrm{Cn}(\Gamma \cup \Delta)$ by (iii); thus if $B \in \mathrm{Cn}(\Delta \cup \{A\})$ then $B \in \mathrm{Cn}(\Gamma \cup \Delta)$.

15 See Ripley 2013: 9. He considers, but does not pursue, an alternative strategy which treats the problem as based on an illicit application of Cut in the metalanguage. However, if the metalinguistic argument is mathematical, this strategy is not available to those who endorse classical reasoning (including Cut) within mathematics.

16 See Ripley 2015 for a discussion, from a starting-point more sympathetic to the use of multisets, of difficulties in playing a closure role for logics without Contraction.

17 An advantage of formulating mathematical induction axiomatically is that in the setting of second-order logic with the standard semantics one gains the full intended content of mathematical induction by prefixing $\mathrm{MI}_{\mathrm{Ax}}$ with the second-order quantifier $\forall F$, which is needed for the categoricity of arithmetic (Shapiro 1991). By contrast, the theory of arithmetic obtained with the inferential versions of mathematical induction such as MI do not attain such generality. Of course, one may announce one's intention to accept all instances of MI in extensions of the language, but that does not amount to expressing the generality in the theory itself.

18 Many of the ideas in this chapter were presented at conferences on Metaphysics and Semantics at Yale University and on the Normativity of Logic at the University of Bergen, the 2016 Logic Colloquium at the University of Leeds, the 11th Panhellenic Logic Symposium at Delphi, and classes and colloquia at Brown University, the

University of Salzburg, and the University of Oxford. Many thanks are owed to participants for valuable comments and questions, and especially to Beau Mount, Hartry Field, Volker Halbach, and Leon Horsten for discussion and correspondence. This article is supported by the research project 17ZDA024 funded by the National Foundation of Social Science, China.

Bibliography

Dummett, Michael. 1977: *Elements of Intuitionism*. Oxford: Clarendon Press. 2nd ed. 2000.
Durrett, Rick. 2010: *Probability: Theory and Examples*, 4th ed. Cambridge: Cambridge University Press.
Feferman, Solomon. 2011: 'Is the Continuum Hypothesis a definite mathematical problem?'. *Exploring the Frontiers of Incompleteness.* Harvard: http://logic.harvard.edu/EFI_Feferman_IsCHdefinite.pdf
Field, Hartry. 2008: *Saving Truth from Paradox*. Oxford: OUP.
Field, Hartry. 2016: 'Indicative conditionals, restricted quantification, and naïve truth', *Review of Symbolic Logic*, 9, 181-208.
Grimmett, Geoffrey, and Stirzaker, David. 2001: *Probability and Random Processes*, 3rd ed. Oxford: Oxford University Press.
Halbach, Volker. 2015: *Axiomatic Theories of Truth*, 2nd ed. Cambridge: Cambridge University Press.
Halbach, Volker, and Horsten, Leon. 2006: 'Axiomatizing Kripke's theory of truth', *The Journal of Symbolic Logic*, 71: 677-712.
Hjortland, Ole Thomassen. 2017: 'Anti-exceptionalism about logic', *Philosophical Studies*, 174: 631-658.
Parsons, Terence. 2000: *Indeterminate Identity: Metaphysics and Semantics*. Oxford: Oxford University Press.
Priest, Graham. 1995: *Beyond the Limits of Thought*. Cambridge: Cambridge University Press.
Priest, Graham. 2008: *An Introduction to Non-Classical Logic: From Ifs to Is*, 2nd ed. Cambridge: Cambridge University Press.
Ripley, David. 2013: 'Revising up: strengthening classical logic in the face of paradox', *Philosophers' Imprint*, 13, 5: 1-13.
Ripley, David. 2015: 'Contraction and closure', *Thought*, 4: 131-138.
Shapiro, Stewart. 1991: *Foundations without Foundationalism: A Case for Second-Order Logic*. Oxford: Clarendon Press.
Shiryaev, Albert. 2016: *Probability-1*, 3rd ed. New York: Springer.
Tarski, Alfred. 1983: *Logic, Semantics, and Metamathematics*, 2nd ed., tr. J.H. Woodger, ed. J. Corcoran.
Univalent Foundations Program. 2013: *Homotopy Type Theory: Univalent Foundations of Mathematics*. Princeton: Institute for Advanced Study.
Uzquiano, Gabriel. 2015: 'Recombination and paradox', *Philosophers' Imprint*, 15, 19: 1-20.
Williamson, Timothy. 1994: *Vagueness*. London: Routledge.
Williamson, Timothy. 2013: *Modal Logic as Metaphysics*. Oxford: Oxford University Press.
Williamson, Timothy. 2016: 'Modal science', *Canadian Journal of Philosophy*, 46: 453-492.
Williamson, Timothy. 2017: 'Semantic paradoxes and abductive methodology', in Bradley

Armour-Garb (ed.), *Reflections on the Liar*. Oxford: Oxford University Press. 325-346.

Williamson, Timothy. 2018: 'Supervaluationism and good reasoning'. *Theoria* (Spain), forthcoming.

Woods, Jack. 2017: 'Logical partisans'. MS.