The Paradox of Idealisation

Salvatore Florio* and Julien Murzi†

October 1, 2008

A well-known proof by Alonzo Church, first published in 1963 by Frederic Fitch, shows that all truths are knowable only if all truths are known.[1] This is the Paradox of Knowability. If we take it, quite plausibly, that we are not omniscient, the proof appears to undermine metaphysical doctrines committed to the knowability of truth, such as semantic anti-realism. Since its rediscovery by W. D. Hart and Colin McGinn (1976), many solutions to the paradox have been offered. In this paper, we present a new proof to the effect that not all truths are knowable, which rests on different assumptions from those of the original argument published by Fitch. We highlight the general form of the knowability paradoxes, and argue that anti-realists who favour either a hierarchical or an intuitionistic approach to the Paradox of Knowability are confronted with a dilemma: they must either give up anti-realism or opt for a highly controversial interpretation of the principle that every truth is knowable.

1 The Church-Fitch Paradox

The proof of the Church-Fitch Paradox requires only that knowledge be factive and that it distribute over conjunction. Let '$\Diamond$' and '$\Box$' denote some notion of possibility and some correlative notion of necessity respectively. Then, one can prove that what Williamson (2000) calls *weak verificationism*:

$$\text{(WVER)}\ \ \forall\varphi(\varphi \to \Diamond\mathcal{K}\varphi),$$

collapses into *strong verificationism*:

$$\text{(SVER)}\ \ \forall\varphi(\varphi \to \mathcal{K}\varphi),^{2}$$

*The Ohio State University, USA, florio.2@osu.edu

†University of Sheffield, S10 2TN, UK, J.Murzi@sheffield.ac.uk

[1] See Fitch (1963) and Church (forthcoming).

[2] These principles are usually meant to apply only to propositions expressed by sentences we understand, and the quantifiers are interpreted substitutionally. Neither complication affects our points below.

where $\mathcal{K}\varphi$ reads 'someone knows at some time that φ'.[3] One first shows that for any particular proposition p,

$$(1)\ \neg\Diamond\mathcal{K}(p \wedge \neg\mathcal{K}p)$$

is provable. One then proceeds to show that all truths are knowable only if all truths are known. In a nutshell, if $\mathcal{K}$ is factive and distributes over conjunction, truths of the form $\varphi \wedge \neg\mathcal{K}\varphi$ are provably unknowable. Yet on the anti-realist assumption that all truths are knowable, unknowable propositions are to be regarded as false. By an elementary, though exclusively classical step, it follows that all truths are known. Since this latter claim is clearly false—indeed, we are not omniscient—anti-realism is under threat.

2 Intuitionistic and Hierarchical Treatments

According to Timothy Williamson (1982), the Paradox of Knowabilty is no straightforward *reductio* of semantic anti-realism. As he points out, within intuitionistic logic WVER only implies

$$(\mathsf{WVER}^*)\ \forall\varphi(\varphi \rightarrow \neg\neg\mathcal{K}\varphi).$$

But unlike SVER, WVER* is not obviously problematic. As Williamson puts it:

> it forbids intuitionists to produce claimed *instances* of truths that will never be known: but why should they attempt something so foolish? (Williamson, 1982, p. 206)

Furthermore, given the intuitionistic invalidity of the step from $\neg\forall x\phi$ to $\exists x\neg\phi$, intuitionists can deny that all truths will be known at some time without thereby being committed to the existence of any forever unknown truth. The paradox constrains anti-realism, Williamson concludes, but does not necessarily undermine it:

> That a little logic should short circuit an intensely difficult and obscure issue was perhaps too much to hope, or fear (Williamson, 1982, p. 207).[4]

[3] For the sake of clarity, we shall occasionally make the quantifiers explicit and write $\exists x\mathcal{K}_x\varphi$ and $\forall x\mathcal{K}_x\varphi$.

[4] In a recent response to the paradox, Michael Dummett endorses an intuitionist strategy similar to the one outlined above (see Dummett (2007, pp. 348–50)). On Dummett's view, intuitionists can escape the Paradox as long as they can avoid commitment to the existence of forever unknown truths (notice that in light of the Paradox, asserting the existence of any such truth is intuitionistically inconsistent with WVER). Dummett claims that intuitionists do not incur such a commitment since the Law of Bivalence can only be legitimately applied to decidable *mathematical* statements, and not to empirical statements that we could have known but no longer can. He writes: "[the realist] relies on assuming bivalence in order to provide an example of a true statement that will never be known to be true—more exactly, of a pair of

A second, quite natural way to block the paradox had already been suggested by Church in 1945:

> Of course the foregoing refutation [...] is strongly suggestive of the paradox of the liar and other epistemological paradoxes. It may therefore be that Fitch can meet this particular objection by incorporating into the system of his paper one of the standard devices for avoiding the epistemological paradoxes. (Church, forthcoming)

Bernard Linsky and Alexander Paseau have recently developed this thought.[5] Though the Church-Fitch proof makes no use of self-referential sentences, they observe, it is nevertheless invalid on a logical account of knowledge reminiscent of Russell's theory of types. The intuitive idea is that each formula is assigned a *logical type*, which reflects the nesting of occurrences of $\mathcal{K}$ within that formula. Formally, one introduces infinitely many knowledge operators $\mathcal{K}_n$, one for each natural number n. The type of any formula φ is defined by the greatest index of the knowledge operators occurring in φ. A formula of the form $\mathcal{K}_n\varphi$ is well-formed just in case n is strictly greater than the type of φ. In this framework, only $\Diamond(\mathcal{K}_{n+2}\varphi \wedge \neg\mathcal{K}_{n+1}\varphi)$ follows from WVER. But unless it is assumed that $\mathcal{K}_{n+1}\varphi$ entails $\mathcal{K}_n\varphi$ for every index n and formula φ, that is not a formal contradiction.

Does the hierarchical treatment represent a viable answer to the Church-Fitch Paradox? And can a simple appeal to intuitionistic logic salvage semantic anti-realism from its paradoxical consequences?

3 The Paradox of Idealisation

There is a dispute among anti-realists over whether or not knowability requires idealisation. Strict Finitists think that idealisation is not required: the word 'knowable', for them, is to be interpreted as 'possibly known by agents just like us'. Strict Finitism has highly revisionary consequences. On that view, any decidable proposition that cannot be known for mere 'medical' limitations, e.g. some arithmetical propositions involving very large numbers, turns out to be meaningless, if not false. But this result is hardly acceptable. As Dummett puts it:

> The intuitionist sanctions the assertion, for any natural number, however large, that it is either prime or composite, since we have a

statements one of which is true. He has to. If he could instance a specific true statement, he would know that it was true. This illustrates how important the principle of bivalence is in the controversy between supporters and opponents of realism" (Dummett, 2007, p. 350). Our new paradox circumvents the problem raised by Dummett. The first version of it, which we present in Section 3, implies the existence of forever unknown truths, but we argue that it does so consistently with Dummett's take on the Law of Bivalence. As for the modal version we give in Section 4, it does not imply the existence of forever unknown truths.

[5] See Linsky (forthcoming) and Paseau (forthcoming).

> method that will, at least in principle, decide the question. But suppose that we do not, and perhaps in practice cannot apply that method: is there nevertheless a fact of the matter concerning whether the number is prime or not? There is a strong impulse that there must be. (Dummett, 1994, pp. 296–7)

Dummett's reason for rejecting Strict Finitism is in line with his discussion of apodictic numbers in "Wang's Paradox". A number n is apodictic "if it is possible for a proof (which we are capable of taking in, i.e. of recognizing as such) to contain as many as n steps" (Dummett, 1975, p. 253). Dummett assumes the existence a number M "sufficiently large that it is plainly not a member of the totality [of apodictic numbers]" (*Ibid.*) in his *reductio* of Strict Finitism. Consider now some decidable mathematical proposition s whose proof has at least as many as M steps. In Dummett's view, anti-realists can legitimately say that either s or its negation is true: although neither s nor its negation is feasibly knowable, at least (and at most) one of them is nevertheless knowable in an idealised sense.

Following Dummett, most anti-realists concede that 'knowable' in WVER is to be read as 'knowable in principle', i.e. knowable by agents endowed with cognitive capacities like ours or that finitely exceed ours.[6] Here is Neil Tennant:

> The truth does not have to be knowable by all and sundry, regardless of their competence to judge. [...] This would be to hostage too much of what is true to individual misfortune. At the very least, we have to abstract or idealize away from the limitations of actual individuals. [...] At the very least, then, we have to imagine that we can appeal to an ideal cognitive representative of our species. (Tennant, 1997, p. 144)

Call such anti-realists *moderate*. In spite of its initial plausibility, this move runs the risk of becoming a Trojan horse.

Our argument starts from the moderate anti-realist's concession that there are feasibly unknowable truths, i.e. truths that, because of their complexity or of the complexity of their proofs, can only be known by agents whose cognitive capacities finitely exceed ours. In symbols:

$$(2)\ \exists\varphi(\varphi \wedge \Box\forall x(\mathcal{K}_x\varphi \rightarrow Ix)).$$

Let q be one such truth and let 'Ix' read 'x is an idealised agent', where an agent counts as idealised if and only if her cognitive capacities—perceptual discrimination, memory, working memory etc.—finitely exceed ours.[7] Now let us assume that there are no idealised agents:

$$(3)\ \neg\exists x Ix.$$

It can be proved that the conjunction

[6] See especially, Tennant (1997, Chapter 5).

[7] We shall consider an alternative definition of an idealised agent in Section 4.

$$(4)\ q \wedge \neg\exists x I x$$

is unknowable:

> *Proof*: Assume that $q \wedge \neg\exists x I x$ is knowable. Then there is a world w where some agent knows $q \wedge \neg\exists x I x$. Call this agent a. By (2), every agent who knows q in w is idealised. Therefore, a is idealised. However, since a knows $q \wedge \neg\exists x I x$, by distributivity and factivity, $\neg\exists x I x$ is true at w. Hence, a cannot be an idealised agent. Contradiction. Therefore, $q \wedge \neg\exists x I x$ is unknowable. ■

We call this the Paradox of Idealisation.

The argument generalizes. Similar proofs can be constructed for every formula φ and $\mathcal{P}(x, \varphi)$ such that the following holds:

$$(5)\ \exists\varphi(\varphi \wedge \Box\forall x(\mathcal{K}_x\varphi \rightarrow \mathcal{P}(x,\varphi)) \wedge \neg\exists x\mathcal{P}(x,\varphi)).$$

Relevant instances of $\mathcal{P}(x, \varphi)$ may include traditional necessary conditions for knowledge, such as justification or belief. The Paradox of Knowability itself may be thought of as a trivial instance of (5), with $\mathcal{P}(x, \varphi) \equiv \mathcal{K}_x\varphi$:

$$(5')\ \exists\varphi(\varphi \wedge \Box\forall x(\mathcal{K}_x\varphi \rightarrow \mathcal{K}_x\varphi) \wedge \neg\exists x\mathcal{K}_x\varphi).$$

The argument poses a problem for anti-realists who appeal to intuitionistic logic to block the Church-Fitch Paradox. If it is not to be regarded as a *reductio* of WVER, anti-realists have no choice but to deny either (2) or (3). We argue below that neither option seems viable, regardless of whether intuitionistic logic is adopted. The new paradox equally threatens to undermine hierarchical approaches to the Paradox of Knowability.[8] Although the definition of 'Ix' involves reference to cognitive capacities, it does not involve reference to knowledge of any particular proposition. Hence, typing '$\mathcal{K}$' would be uneffective here.[9] We now turn to some potential concerns about the soundness of our proof.

4 Objections and Replies

Let us begin with (2), i.e. the claim that there are feasibly unknowable truths. In light of the Paradox of Idealisation, anti-realists might reconsider their moderation and argue that for any true proposition φ, it is possible that φ be known by a non-idealised agent:

[8] Thanks to Tim Williamson for pointing this out.

[9] It might be objected that anti-realists could still block the Paradox of Idealisation by typing the predicate 'Ix'. It is however unclear whether they would have any independent reason for doing so. As Paseau (forthcoming) points out, the main motivation for typing $\mathcal{K}$ is to avoid other paradoxes, such as the Paradox of the Knower. Yet, no analogous motivation seems to be available in the case of 'Ix'. Moreover, it is worth reminding that merely typing 'Ix' will not do: anti-realists would also need to type any other predicate one could substitute in (5).

$$(6)\ \forall\varphi(\varphi \to \Diamond\exists x(\mathcal{K}_x\varphi \wedge \neg Ix)).$$

Since (6) intuitionistically entails the falsity of (2), our paradox would be blocked. This thought might be motivated in different ways. For instance, anti-realists might claim that, if there is a method to verify φ, then there is a possible world whose space-time structure is such that agents with cognitive capacities just like ours know that φ. Alternatively, they might claim that for any true φ, there is a possible world at which φ itself, or a proof of it, are expressed in a language that renders them cognitively accessible.[10]

This objection does not work. Let S be a description of the space-time structure of the actual world or a description of which languages are actually used. Now consider the modified premise:

$$(2^*)\ \exists\varphi((\varphi \wedge S) \wedge \Box\forall x(\mathcal{K}_x(\varphi \wedge S) \to Ix)).$$

In perfect analogy with the Paradox of Idealisation, we can argue as follows:

> *Proof*: Assume that $(q \wedge S) \wedge \neg\exists x Ix$ is knowable. Then there is a world w where some agent a knows $(q \wedge S) \wedge \neg\exists x Ix$. This forces w to have the space-time structure described by S, or a to speak an actual language. It also follows that $\neg\exists x Ix$ is true in w. Therefore, a is a non-idealised knower of q in a world whose space-time structure is S or where no non-actual language is used. Contradiction, since we are assuming that, necessarily, $\forall x(\mathcal{K}_x(q \wedge S) \to Ix))$. Thus, $(q \wedge S) \wedge \neg\exists x Ix$ is unknowable. ■

Anti-realists might reply by exploiting the characteristic weakness of intuitionistic logic. They may deny (7), on the one hand, and express their moderation by claiming that not every truth is feasibly knowable, on the other:

$$(7)\ \neg\forall\varphi(\varphi \to \Diamond\exists x(\mathcal{K}_x\varphi \wedge \neg Ix)).$$

Classically, (7) is inconsistent with the denial of (2), but not intuitionistically. The problem with this move, though, is that intuitionists seem to be in a position to *prove* the existence of feasibly unknowable truths. Let q be some decidable yet undecided mathematical statement whose decision procedure is feasibly unperformable. Then, q satisfies both of the following:

$$(8)\ \Box\forall x(\mathcal{K}_x q \to Ix);$$
$$(9)\ \Box\forall x(\mathcal{K}_x\neg q \to Ix).$$

Since q is *ex hypothesi* decidable, even the intuitionist should be willing to assert that either q or its negation is true. The existence of a feasibly unknowable truth can then be easily derived from $q \vee \neg q$, (8), and (9).

Intuitionists might object that one can never rule out that a sentence that is now feasibly unknowable will turn out to be feasibly knowable. However, on the same grounds, one would be prevented from asserting empirical generalisations, as Dummett himself observes:

[10] We thank Cesare Cozzo and Luca Incurvati for pressing this point.

> there may be some point in saying that, for any statement not known to be false, we can never absolutely rule out the possibility that some indirect evidence for its truth may turn up; but if we are ever to be credited with knowing the truth of a universal empirical statement other than one that follows from scientific laws, this possibility may be so remote that we are sometimes entitled to say—as we often do—that it will be never be known whether p. (Dummett, 2001, p. 1)

Moderate anti-realists might bite the bullet and, instead, deny (3), i.e. the claim that there are no idealised agents. But would this be advisable? We see two possibilities, depending on how anti-realists define the notion of an idealised agent. If an agent counts as idealised just in case her cognitive capacities finitely exceed those of *any actual epistemic agent*, then (3) is indeed an a priori truth. It would say that there are no (actual) epistemic agents whose cognitive capacities finitely exceed those of any (actual) epistemic agent, which is of course a truism. One might object that, on this reading, the claim that there is a decidable proposition satisfying (8) and (9) would be hardly acceptable. For how do we know that in the actual world there will never be agents so clever that they will be able to decide q? However, the existence of a decidable proposition satisfying (8) and (9) is only problematic if one assumes that there is no bound to the cognitive capacities of *actual* epistemic agents. If, as we think plausible, there is a bound, then it would seem difficult to maintain that there is no decidable and yet feasibly unknowable proposition. On the other hand, anti-realists might take (3) to be an empirical claim, for example following Tennant in defining 'Ix' in terms of human cognitive capacities. The worry would then be that a principle such as WVER, thought to be necessary and a priori, would carry a commitment, $\neg\neg\exists x Ix$, that is open to empirical refutation.

Be that as it may, if anti-realists went as far denying $\neg\exists x Ix$, this would not help them with another variant of our paradox, that rests on the following weaker assumption:

$$(10)\ \exists\varphi(\Diamond(\varphi \wedge \neg\exists x Ix) \wedge \Box\forall x(\mathcal{K}_x\varphi \rightarrow Ix)).$$

Presumably, even for an anti-realist there is some feasibly unknowable proposition φ, such that φ and $\neg\exists x Ix$ are co-possible. Provided that the relation of accessibility is transitive, we can now run a version of the Paradox of Idealisation via (10) and the necessitated formulation of WVER:

$$(\mathsf{WVER}^{**})\ \Box\forall\varphi(\varphi \rightarrow \Diamond\mathcal{K}\varphi).$$

Anti-realists could reply by rejecting WVER**, thereby sticking to WVER. This, however, would be a desperate move: it would leave them with a contingent version of their core metaphysical tenet. They might still maintain that WVER is a priori, though contingent. But this does not seem to square with the modal profile of WVER as supported by the standard anti-realist arguments: semantic anti-realists like Dummett would find it problematic to give up the thought that, as a matter of *conceptual necessity*, truth cannot outstrip our capacity to know. Then, provided that the logic of conceptual necessity obeys the minimal modal principles required for our proof, the problem would still remain. Anti-realists would thus seem to have only one option left: giving up transitivity. But this would be a surprising consequence of accepting WVER.

Conclusion

The Paradox of Idealisation threatens the viability of intuitionist and hierarchical defences of semantic anti-realism. Hierarchical approaches might block the original Paradox of Knowability, but fail to block the cognate Paradox of Idealisation. As for the appeal to intuitionistic logic, it does not help the anti-realist avoid the inconsistency among the three assumptions on which our paradox depends. Denying (3) does not seem an option, independently of whether classical logic is admitted. Rejecting (2), on the other hand, is tantamount to abandoning moderate anti-realism. Anti-realists who favour either an intuitionist or a hierarchical approach to the Paradox of Knowability appear to be confronted with a dilemma: they must either negate WVER or give up their moderation. Several other solutions to the paradox have been proposed so far.[11] Although they are all controversial, our result suggests that a viable defense of anti-realism may turn on whether or not they are acceptable. We leave to anti-realists the hard task of providing an adequate defence of their metaphysical views.[12]

References

Church, A.: forthcoming, Anonymous referee reports, *in* J. Salerno (ed.), *New Essays on the Knowability Paradox*, Oxford University Press, Oxford.

Dummett, M.: 1975, Wang's paradox, *Synthese* **30**, 301–304. Page references are to reprint in Dummett, 1978, pp. 248–268.

Dummett, M.: 1978, *Truth and Other Enigmas*, Duckworth, London.

Dummett, M.: 1994, Reply to Prawitz, *in* D. Prawitz and G. Olivieri (eds), *The Philosophy of Michael Dummett*, Kluwer, Dordrecht, pp. 292–98.

Dummett, M.: 2001, Victor's error, *Analysis* **61**(269), 1.2.

[11]See, e.g., Edgington (1985), Tennant (1997) and Tennant (forthcoming)

[12]Many thanks to Cesare Cozzo, Dominic Gregory, Bob Hale, Luca Incurvati, Steve Laurence, Alexander Paseau, Stephen Read, Joe Salerno, George Schumm, Neil Tennant, Tim Williamson, Elia Zardini, and an anonymous referee for helpful comments on previous drafts of this paper. Earlier versions of this material were presented at the Universities of Geneva, Leeds, Bristol, Norwich, Oxford, Cambridge and the Eastern Division of the American Philosophical Association. We are grateful to the members of these audiences for their valuable feedback.

Dummett, M.: 2007, Reply to Wolfgang Künne, *in* R. E. Auxier and L. E. Hahn (eds), *The Philosophy of Michael Dummett*, Open Court, Chicago, pp. 345–350.

Edgington, D.: 1985, The paradox of knowability, *Mind* **94**, 557–68.

Fitch, F.: 1963, A logical analysis of some value concepts, *Journal of Philosophical Logic* **28**, 135–42.

Hart, W. and McGinn, C.: 1976, Knowledge and necessity, *Journal of Philosophical Logic* **5**, 205–208.

Linsky, B.: forthcoming, Logical types in some arguments about knowability and belief, *in* J. Salerno (ed.), *New Essays on the Knowability Paradox*, Oxford: Oxford University Press.

Paseau, A.: forthcoming, Fitch's paradox and typing knowledge, *Notre Dame Journal of Formal Logic* .

Tennant, N.: 1997, *The Taming of the True*, Oxford University Press, Oxford.

Tennant, N.: forthcoming, Revamping the restriction strategy, *Synthese* .

Williamson, T.: 1982, Intuitionism disproved?, *Analysis* **42**, 203–7.

Williamson, T.: 2000, *Knowledge and its Limits*, Oxford University Press, Oxford.