

## **Restriction Strategies for Knowability: Some Lessons in False Hope**

Jonathan L Kvanvig

The knowability paradox derives from a proof by Frederic Fitch in 1963. The proof purportedly shows that if all truths are knowable, it follows that all truths are known.

Antirealists, wed as they are to the idea that truth is epistemic, feel threatened by the proof. For what better way to express the epistemic character of truth than to insist that all truths are knowable? Yet, if that insistence logically compels similar assent to some omniscience--like claim, antirealism is in jeopardy.

Response to the paradox has drifted toward a common theme, a theme I will argue is a non--starter in resolving the paradox. Seeing this point will also make clear the philosophical inadequacy of simply viewing the paradox as a refutation of a wide range of antirealisms.

### **Responses to the paradox**

One way to respond to this problem for antirealism is to question the proof itself, and there have been a number of questions raised about the proof. Such questioning seems to lead nowhere, however. The simplest form of the proof goes as follows. Where we understand the operator K as 'it is known by someone at some time that', we begin by assuming

(1)  $K(p \ \& \ \sim Kp)$ .

If we distribute the K operator across the conjunction, we get

(2)  $Kp \ \& \ K\sim Kp$ .

Since knowledge implies truth,  $K\sim Kp$  implies  $\sim Kp$ ; hence (2) implies

(3)  $Kp \ \& \ \sim Kp$ ,

allowing us to prove by reductio

(4)  $\sim K(p \ \& \ \sim Kp)$

Since (4) is a theorem, we can derive by the Rule of Necessitation

(5)  $\Box \sim K(p \ \& \ \sim Kp)$ ,

which is equivalent to

(6)  $\sim \Diamond K(p \ \& \ \sim Kp)$ .

This claim, however, is pretty obviously inconsistent with the claim that all truths are knowable.

All that is needed is for the value for  $p$  in (6) to be a truth that nobody knows, in which case

$p \ \& \ \sim Kp$  is a truth. By the knowability principle, it must be knowable, i.e.,

(7)  $\Diamond K(p \ \& \ \sim Kp)$ .

Since (7) contradicts (6), we learn that not all truths can be known (or that all truths are known,

for those who enjoy the bizarre, half-baked, inane, and philosophically balmy<sup>1</sup>).

The options for finding a logical flaw in this proof are quite limited. The only rules of inference it employs beyond those of propositional logic are these:

(K--Dist)  $K(p \ \& \ q) \ \vdash \ Kp \ \& \ Kq$

(KIT)  $Kp \vdash p$

and the metalinguistic rule

(RN)  $(\vdash p) \Rightarrow (\vdash \Box p)$ .

Of these rules, the first is the most likely candidate to be challenged, but Timothy Williamson (1993) has shown how to generate paradox without relying on (K--Dist) at all. Given this result, the only hope for avoiding the paradox is to deny that knowledge implies truth or to deny the rule of necessitation.<sup>2</sup>

In light of the well--entrenched character of these rules, it is not surprising to find the literature on the paradox turning in a different direction in attempting to save antirealism from the dark force of the knowability paradox. The dominant strategy has become to deny that the antirealist commitment to the epistemic character of truth involves any commitment at all to the claim that all truths are knowable. Instead, the heart of antirealism is to be found in some weaker claim. Dorothy Edgington (1985) proposes that antirealists need only hold that all *actual* truths are knowable. Michael Dummett (2001) insists that knowability is required only for basic statements, and Michael Hand (2003) provides a sophisticated defense of a similar point of view. Cesare Cozzo (1994) develops an alternative to the knowability claim in terms of idealized arguments, and Neil Tennant (1997: ch.8) argues that the only truths that must be knowable are those for which the assumption that they are known is logically consistent, leading to a cottage

industry regarding whether his approach is a complete non--starter.<sup>3</sup> The body of literature pursuing such restriction strategies—strategies for coping with the paradox that deny that the idea that truth is epistemic commits the antirealist to the claim that all truths are knowable—comprises an enormous percentage of the writing on this subject, and has become the favored approach among anti--realists for disarming the paradox. When one takes into account the additional literature aimed at undermining such strategies, the dominant issue displayed by recent literature on the knowability paradox is whether any such restriction strategy can successfully disarm the paradox.

Nowhere in this body of literature is the strategy itself questioned. Instead, the questions are two: is the strategy faithful to, and theoretically sustained by, the antirealist commitment to the epistemic character of truth, and does the restriction yield a claim that avoids Fitch's result? We do not need answers to these questions, however, if such approaches to the paradox are red herrings, and that is what I claim here. To defend this point, I will first explain a different approach to the paradox that is clearly a red herring. In the process, we will learn more about the heart of the knowability paradox, enough to see clearly that restriction strategies are simply lessons in false hope.

### **Theism and knowability**

The knowability paradox is typically thought of as deriving from two assumptions. The first is that all truths are knowable and the second is that some truths are unknown. Upon generating a contradiction from these two assumptions, we are required to discharge, leaving us to conclude that the knowability claim is false (and antirealism thereby threatened).

In the presentation given above, however, a different characterization of the paradox might be given. The presentation above could be characterized as a proof that there have to be unknown truths, for it begins by assuming that a particular truth is known (the truth that  $p$  is an unknown truth), and derives from that claim the impossibility of knowing this particular truth.

For philosophers of a theistic bent, this characterization of the paradox may disturb, since it threatens the idea that there is an omniscient being. Given such a disturbance, theistic philosophers may see themselves as having a strong reason to find some flaw in the proof, hoping thereby to prevent the knowability paradox from refuting their theistic perspective.

Such a response, however, is confused. After more sober reflection, the theistic philosopher may see the flaw in this reaction to the paradox. The theistic philosopher may come to see that the above proof is no threat to theism unless it is true that there is some unknown truth, for if there is no unknown truth, then omniscience does not require that it be known that there is some claim that is both true and unknown. Yet, if there is an omniscient being, then there aren't any unknown truths! Hence, if one thinks there is an omniscient being, the above proof

can be dismissed as a challenge to that viewpoint. It is no more interesting a challenge to theism than any argument that presumes an omniscient being must know what is false. So the theistic philosopher can move on to other interesting areas of philosophy, knowing that the knowability paradox is of no concern.

So characterized, the supposed reaction by the theistic philosopher is both right and wrong. It is right in that the proof above does not threaten the claim of omniscience, but it is wrong in supposing that nothing paradoxical remains about which the theistic philosopher need be concerned.

One way to see this point is to notice that the paradox does not depend simply on whether one accepts the two assumptions in question, the knowability assumption and the non--omniscience assumption. The central perplexity involved in the paradox does not depend on the idiosyncrasies of one's favored philosophy, but rather on a perplexing lost logical distinction between what is actually the case and what might be case. It is obvious that knowledge implies the possibility of such, but what is not obvious is what the Fitch proof attempts to demonstrate: that, to put it carelessly, possible knowledge implies actual knowledge. Should that distinction disappear, it would be fitting to find ourselves in a state of perplexing philosophical stupor. How could it be that there is no logical distinction between actuality and possibility in this way?

We might try for equilibrium by reminding ourselves that there are philosophical

domains in which the distinction between actuality and possibility is lost. For example, modal logicians have long been comfortable with the idea that what is actually necessary is not logically distinct from what is possibly necessary. The comfort experienced by this thought will not last long, however. We are comfortable with the lost distinction in this domain because we have a semantical theory to which to appeal to explain why there is no logical distinction here, and we became comfortable with denying the distinction here only after the development of the semantical theory that makes intelligible the loss of such a distinction.<sup>4</sup> Merely formulating S5 systems with the distinctive axiom central to the proof that eliminates the logical distinction between actual and possible necessity is not enough. As the history of the philosophy of modal logic shows, it took a semantical explanation to motivate the contemporary orthodoxy in favor of S5. Proof rules for the modal operators do not, by themselves, yield the degree of understanding necessary to rid us of the philosophical puzzle; only something more, such as is provided by Kripke semantics for quantified modal logic will help.<sup>5</sup> Nothing similar can be said when we return to the context of the knowability paradox, however: we have no semantical basis whatsoever for being sanguine about a lost distinction between actual and possible knowledge.

So even if our theistic philosopher should dissent from the assumption that there are unknown truths, said philosopher has as much reason as anyone to view Fitch's proof as establishing a very troubling conclusion. There is nothing about theism that yields an explanation

as to why actual and possible knowledge are not logically distinct. It is for this reason that the theistic response to the paradox is a red herring, even though such a philosopher can take refuge in holding that theism itself is not at stake. The imagined theistic philosopher denies that there are unknown truths, and thereby achieves serenity in the face of the paradox. Such serenity is warranted, however, only if the theistic perspective does more. It will need to explain why, to speak again in the loose and popular vernacular, there is no logical distinction between actual and possible knowledge.

### **Antirealism and knowability**

Antirealists, I maintain, do something similar to what the imagined theist has done. The imagined theist denies the assumption of non--omniscience, thereby claiming to avoid any perturbation from the paradox. The now--dominant antirealist strategy is to deny the knowability claim, substituting for it some careful emendation with weaker implications, also thereby claiming to avoid the reach of the paradox. Yet, if the theistic response to the paradox is a red herring, one should wonder why the antirealist restriction strategy isn't as well. What reason can an antirealist give on behalf of a restriction strategy that will render respectable such a response to the paradox in contrast to the theistic response?

Here is what antirealists will need to say in response to the claim above that the heart of

the paradox concerns a lost logical distinction between actuality and possibility. They can insist on more precision; they can say that we need to speak with the sophisticates rather than the vulgar. The lost distinction is not one between actual and possible knowledge, for even false (contingent) claims are objects of possible knowledge (in worlds where they are true).

Fair enough; so let's try to be more careful. The antirealist is likely to put the careful point this way: the lost distinction is a lost distinction between actual known *truths* and possible known *truths*. That is, a careful presentation of the lost distinction is:

$$(LD) \forall p((p \ \& \ \diamond Kp) \Leftrightarrow (p \ \& \ Kp)).$$

The proof from  $p \ \& \ Kp$  to  $p \ \& \ \diamond Kp$  is trivial, depending only on the modal principle that what is actual is possible. So, the antirealist can claim, the heart of the paradox is found in demonstrating that  $p \ \& \ Kp$  follows from  $p \ \& \ \diamond Kp$ . That proof, however, requires assuming that all truths are knowable.<sup>6</sup> So (LD), the careful expression of the heart of the knowability paradox in terms of a lost logical distinction between actuality and possibility, is derivable only on the assumption that all truths are knowable. Hence, a perfectly respectable strategy in responding to the paradox is to weaken the knowability assumption in such that way that (LD) can no longer be derived. So long as the resulting restriction still expresses the idea that truth is epistemic, the anti--realist has bested the theist above, for not only can the antirealist claim that their position is not undermined by the paradox but also that the paradox has been disarmed since the lost distinction at the heart

of the paradox follows only by assuming a claim that is false and necessarily so. In case we needed reminding, we might be reminded as well that strange consequences often follow from necessarily false assumptions.

It is time for a philosophical lament here, however. This response to the theist analogy is valuable because of its demand for precision regarding the lost logical distinction at the heart of the paradox. It is mistaken, however, in claiming that (LD) is the proper formulation of that distinction. The paradox is generated from two assumptions, the assumption that all truths are knowable and the assumption that some truths are not known. The proof from the latter assumption to the former is trivial; the proof from the former to the latter is just Fitch's proof. Given these two proofs, the obvious formalization of the lost distinction is not (LD) but

$$(LD^*) \forall p(p \rightarrow \diamond Kp) \Leftrightarrow \forall p(p \rightarrow Kp).$$

Whereas (LD) is not a theorem, but instead can be proven only by assuming that all truth are knowable, (LD\*) is a theorem so long as Fitch's proof is valid. I will express (LD\*) in ordinary English by saying that there is no logical distinction between universally knowable truth and universally known truth. This more careful articulation still codifies a lost logical distinction between actuality and possibility with respect to what is known.

With this more careful formulation, the analogy with the theist above is restored. The theist is both right and wrong in being undisturbed by the paradox, and the antirealist is in the

same boat. If the idea that truth is epistemic doesn't require that all truths are known, then the paradox does not threaten to undermine antirealism anymore than it threatens to undermine theism. Neither theism nor antirealism of this restricted variety has anything to say that is relevant to the paradoxicality engendered by Fitch's proof. Each view denies a different assumption in Fitch's proof, but the paradoxicality involved in (LD\*) depends in no way whatsoever on the truth of the assumptions used to generate the paradox. Antirealists may still find comfort in undermining (LD) by pursuing a restriction strategy, but they should not pretend that undermining (LD) solves the paradox.

### **Implications**

Prior to encountering the literature on the topic, the taking of (LD) as the proper careful articulation of the threat of the paradox should strike one as surprising. The obvious careful articulation is (LD\*)--after all, Fitch's proof is a derivation of the left side of (LD\*) from its right side.

There is a larger point to note as well. Critics of antirealism, such as Williamson, view the paradox as a refutation of (most versions of) antirealism, with Fitch's proof simply a display of a surprising logical result to this effect.<sup>7</sup> Such approaches to the paradox, however, leave the paradoxicality in question unresolved. What is paradoxical here is not that Fitch has discovered a proof that threatens antirealism, but rather that Fitch has discovered a proof that threatens a

logical distinction between actuality and possibility. One way to put this point is to notice that the omniscience--like claim, though not likely to be thought true (especially when we envision the quantifiers restricted to finite minds), is not obviously impossible. Contrast this point with the fact that the knowability claim, if true, is supposed to be necessarily true: it is a purported implication of a proper understanding of the nature of truth. Yet, (LD\*) claims that the two are logically equivalent, which they cannot be without have the same modal status. A satisfactory response to the paradox cannot simply swallow this result without explanation. We have already seen one example of a satisfying response to a similar situation, where we have a semantic explanation of the lost logical distinction between actual and possible necessity. If we could have the same here, we'd have a solution to the paradox. It may be there is some alternative explanation as well that is not semantic in character, but it is hard to see at this point what such a weaker explanation might look like.

I want to consider the issue of whether there might be a non--semantical explanation of the lost distinction in a moment, but I first want to emphasize the paradoxicality of the lost distinction by contrasting it with results that are merely surprising but not paradoxical to prevent some deflationary approach to the paradox that claims that the result is merely surprising. Consider, for example, Gödel's incompleteness results. These results are surprising, and threaten important philosophical perspectives, such as Hilbert's formalism. These results themselves are

not paradoxical, however. They present no challenge to anything like the edicts of common sense or the viewpoint of received opinion. That makes these results quite different from a lost logical distinction between actuality and possibility. One might disagree here with my characterization of the Gödel results, arguing that they involve real paradox, but that point can be granted without implying that there is nothing paradoxical resulting from Fitch's proof. If I'm wrong that the Gödel results do more than threaten important philosophical perspectives, then they may be paradoxical; but if so, they join the class of things already including the results of Fitch's proof rather than showing that Fitch's proof is merely surprising.

Consider for another example Vann McGee's (1985) apparent counterexample to modus ponens. The result of McGee's arguments is not merely surprising, but paradoxical. Modus ponens is so well--entrenched a part of our ordinary view of things that our reaction to his arguments is that they must contain a mistake. Suppose, however, that we are wrong. If we are wrong, and McGee is right, some explanation is in order. We need to know how it could be that our ordinary view of things could be so mistaken. It is worth noting that McGee attempts just such an explanation: logical rules should be thought of as more akin to generalizations and law--like statements in science which can be useful and instructive even if not always completely accurate.

My intention in citing McGee's explanation is not to endorse it, nor to endorse his

arguments that modus ponens is not an exceptionless logical rule.<sup>8</sup> The point is only that when a proof conflicts with ordinary understandings, a further explanatory burden must be shouldered. So it is not enough simply to accept the surprising character of Fitch's result. One must also shoulder the philosophical burden of explaining how the proof could be correct since it implies a lost distinction between actuality and possibility.

Even worse would be to respond as follows. "We've carefully considered the rules involved in the paradox, the rules of (K--Dist) and (KIT), plus the metalinguistic rule of (RN) are so inherently plausible that the conflict they create with the intuitive logical distinction between possibility and actuality is not paradoxical at all. The results are surprising and unanticipated, but not paradoxical."<sup>9</sup>

Someone is living in logical denial. No argument can conclusively show that this approach is mistaken, since the difference between what is paradoxical and what is merely surprising is, perhaps, only a difference in degree and not in kind. Even so, there is a distinction to be drawn here between the unanticipated and the seemingly contradictory, and Fitch's proof engenders the latter experience and not simply the former. It is not merely surprising when we are told that what looks like a necessary truth is logically equivalent to what looks like a contingent truth. We can't simply affirm the rules, and say, "I guess we were wrong; non--omniscience really is impossible." That's simply not an adequate explanation of what's gone

wrong; more accurately, it is not an explanation at all.

As we have seen, the paradigm example of a satisfying explanation in this regard is a semantical one. Moreover, as already noted as well, a purely syntactic explanation in terms of axioms of a system and proof rules for it, is particularly unsatisfying. Suppose, for example, that in the modal domain, we showed multiple contexts in which the introduction of a possibility operator on a formula yielded nothing logically distinct from the original formula. Such is the case, if S5 is to be believed, for logical necessities. It is also true of other modal systems, however. For example, obligation statements and possible obligation statements cannot be distinguished logically in certain deontic systems. We learn to accept such results, if we do, by being told a semantic story. In the deontic case, if we interpret obligation statements in terms of ideal worlds, worlds where everything is done properly, then we can see actual and possible obligations collapse. An actual obligation statement takes us to what is true in an ideal world, and a possible obligation statement takes us to another world where the obligation statement is interpreted with reference to ideal worlds. On the assumption that all worlds are accessible from every world, the class of ideal worlds will be the same, whether accessed from our world or some other possible world, generating a logical equivalence between possible obligation and actual obligation. If we accept this semantical story, we understand the lost distinction here. The important point to note is that understanding is not achieved merely by multiplying contexts in

which there is a similar lost distinction. All such multiplication would do is to replace a rather specific paradoxicality with a more general one. What is needed is some understanding of why an apparent logical distinction is lost, and without such understanding, we cannot say “paradox lost”.

This approach to the paradox in terms of trying to accustom us to the loss by generalizing on the syntactic features involved in the paradox has a history going back to J.L. Mackie’s (1980) early paper on the paradox. Seeing the failure of the strategy in other contexts should make us suspicious here, and it is worth taking a look at the details to see why this suspicion is correct.

Here’s an attempt along these lines.<sup>10</sup> Consider the operator “it is written on my blackboard that” and the operator “it is true that”, and the idea that anything true might be written correctly on my blackboard. If we call these operators W and T respectively, we won’t be able to get an analogue of the knowability paradox out of them, since only the latter is factive. To get a paradox, we’d have to generate a contradiction from this assumption:

$$WT(p \ \& \ \sim WTp).$$

From this formula, we can get

$$WTp \ \& \ WT\sim WTp,$$

since, we assume for now, both operators distribute over conjunction. We also assume that the

operators can be split by the following rule:

$$W_T p \vdash W p \ \& \ T p.$$

Using this rule, we can get  $T \sim W_T p$  from the second conjunct (and  $\&$ -Elim), and then get  $\sim W_T p$  from this formula given the factive character of truth:

$$T p \vdash p.$$

We may also wish to put the two operators together into a single operator  $W_T$ .

Intuitively, this operator is supposed to mean something like “written truthfully on my blackboard,” but formally speaking, the crucial idea is that  $W_T$  is both distributive and factive—that is, it borrows distributivity from the  $W$  operator and factivity from the  $T$  operator. Because it is both factive and distributive, we can generate the analogue of the contradiction crucial to the knowability paradox a bit more quickly:

$$W_T(p \ \& \ \sim W_T p)$$

$$W_T p \ \& \ W_T \sim W_T p \text{ (by distribution)}$$

$$W_T p \ \& \ \sim W_T p \text{ (by factivity)}$$

So it appears that not everything true can be correctly written on my blackboard, and thus that  $W_T$  is an analogue of the  $K$  operator. Both are distributive and factive, and hence the knowability paradox is but a special case of a more general phenomenon: find any distributive and factive operator, and the crucial contradiction in the knowability paradox will follow.

There are some niggling problems with this attempt at generalizing so as to provide a syntactic explanation of the lost logical distinction that constitutes the heart of the knowability paradox. These problems are not my fundamental reasons for rejecting this strategy, but they place important limitations on any attempt to pursue this generalization strategy. The first problem to note is that neither of the ways above of demonstrating a contradiction is quite adequate. To see the problem, let us generalize here beyond  $W$  and  $T$  and the combined operator  $W_T$  and think in terms of  $\Phi\Psi p$  and  $\Phi_{\Psi}p$ , for any operators  $\Phi$  and  $\Psi$ .

The idea above is to allow the combined operator to inherit the preferred formal features of the individual operators themselves, but there is no guarantee that this result can be achieved. Suppose we construct a new operator that is a combination of knowledge (“it is known by someone at some time that”) and necessity (“it is necessary that”). Call this the “known-to-be-necessary” operator. Assume as well that the combined operator inherits formal features from the individual operators of which it is a combination. Since  $\&--I$  works with in the context of necessity and since knowledge implies belief, we can infer from the governing of  $p$  and the governing of  $q$  by this new operator that the conjunction of the two is believed by someone at some time. This inference is faulty, however, since the first could be known by someone and the second known by someone, but the conjunction believed by no one.

For another example, combine obligation--for--everyone and knowledge--by--someone,

and you get knowingly obligatory (known--by--someone--to--be--obligatory--for--everyone). Being obligatory preserves  $\&$ -I, so if both p and q are knowingly obligatory, then someone believes both p and q (because knowledge implies belief). This inference is obviously absurd, however. The lesson here is that one can't combine operators and expect to be able to apply the usual rules for either of the individual operators that were put together to form the combined operator.

Notice further in the example about writing truly that if we keep the operators separate, we can't prove the contradiction. Consider how to try. We begin by representing the claim that a specific truth is not truthfully written on my blackboard as:

$$Tp \ \& \ \sim(Wp \ \& \ Tp).$$

Then the reductio assumption will have to be:

$$W(Tp \ \& \ \sim(Wp \ \& \ Tp)) \ \& \ T(Tp \ \& \ \sim(Wp \ \& \ Tp)).$$

If we then distribute the W and T operators, respectively, we get:

$$WTp \ \& \ W\sim(Wp \ \& \ Tp) \ \& \ TTp \ \& \ T\sim(Wp \ \& \ Tp).$$

Since truth is factive, we get from the latter two conjuncts:

$$Tp \ \& \ \sim(Wp \ \& \ Tp),$$

from which we can get

$$\sim Wp.$$

This point is somehow supposed to contradict the first conjunct  $W\top$ , but since we can't apply the factivity claim about truth inside the  $W$  operator, we can't demonstrate the contradiction.

We can avoid this problem by eliminating the truth operator in our representation:

$$p \ \& \ \sim(Wp \ \& \ p).$$

The assumption for reductio can then be

$$W(p \ \& \ \sim(Wp \ \& \ p)) \ \& \ (p \ \& \ \sim(Wp \ \& \ p)).$$

From this claim we derive by the distributivity of  $W$ :

$$Wp \ \& \ W\sim(Wp \ \& \ p) \ \& \ p \ \& \ \sim(Wp \ \& \ p).$$

The latter two conjuncts give us

$$\sim Wp,$$

which contradicts the first conjunct of this formula.

There are two points to note about this derivation. First, there is no factive operator in this proof. Even so, it is reminiscent of Fitch's proof since we represent the claim that  $p$  is an unwritten truth as the claim that  $\sim Wp \ \& \ p$ , just as we represented the idea that  $p$  is an unknown truth in Fitch's proof as the claim that  $\sim Kp \ \& \ p$ . To claim, however, that we have an analogue of Fitch's proof that mirrors its dependence on an operator that is factive and distributive is mistaken. The only operator in this representation is the  $W$  operator, and it is not factive.

This point is not important when we are looking for non--syntactic generalizations of the

knowability paradox, for the above proof is reminiscent enough of Fitch's proof that any explanation of the lost distinction engendered by Fitch's proof will shed light on this proof as well. The syntactic strategy, however, hopes that mere duplication of syntactic form will relieve our perplexity, and there is no such duplication here.

The second point to note will take a bit longer to develop, but it too casts doubt on the idea that we have a syntactic analogue of knowability here. Any syntactic explanatory value for the above derivation depends on the interpretation assumed of the operator in question, since no one should have an issue with the idea that there are distributive operators  $\psi$  for which the formula

$$\Psi(p \ \& \ \sim\Psi p)$$

implies a contradiction. Neither should anyone think that the existence of such operators appropriately addresses the knowability paradox. Instead, if any help is found in syntactic mimicry here, it comes from the assumed interpretation of the W operator in terms of what is written. Once we begin thinking carefully about the concept of writing, however, problems appear. To see them, let's think first about asserting. To assert a claim is not just to utter a sequence of phonemes that conventionally expresses the proposition in question. Instead, to assert is to express that very proposition. When speaking only of the sequence of phonemes and the associated noises, I will term the act in question an act of uttering; when the proposition itself

is expressed by such uttering, I will term such an act of asserting. What I mean by the term ‘proposition’ here is simply a bearer of truth value. Thus, I leave open whether propositions are sentences or abstract objects of some sort, and claim that there is a speech act involving the uttering of phonemes by which bearers of truth value are expressed—namely, the speech act of asserting.

We should note that uttering has formal properties that asserting does not. For one thing, I can’t avoid uttering “it is raining” by uttering “I believe it is raining,” but I do not (always) assert *it is raining* by asserting *I believe it is raining*. Furthermore, a string of phonemes is not itself a bearer of truth value, since propositions have that property exclusively. As a result, the string of phonemes that, when uttered, express a bearer of truth value is not itself a bearer of truth value, but only a vehicle by which a bearer of truth value is expressed. The lesson is that if we wish to consider operators that are both factive and distributive, we will not be able to appeal to the utterance operator, but only to the assertion operator, since only the latter governs items capable of being true or false.

Suppose then that we focus on the assertion operator. Once we notice this difference between uttering and asserting, the claim that asserting has the distributive property is in doubt. When I utter “p and q” I clearly utter “p” and I clearly utter “q”. But when I assert *p and q*, do I assert *p*? Well, when I assert *I believe p*, I don’t assert *p*, so if you think that assertion distributes

across conjunction, what's the difference? Once the question is formulated, the answer is obvious: the difference is that I've logically committed myself to  $p$  by asserting  $p$  and  $q$ . So, in order to preserve the distributive character of assertion, we have to take the concept of asserting to include not simply what propositions one expresses by uttering a string of phonemes that make up a simple declarative sentence. We'll also have to count you as asserting at least some claims to which you are logically committed in virtue of the assertions you make by saying declarative sentences. Which ones? Trying to sort leads to enough of a mess that the road most easily traveled takes us to the point of including all of them.

That's a mistake. I can assert that Fermat's last theorem is unprovable even though everything I assert logically commits me to the truth of that theorem (since it has been proven). Anti--realists are wont to appeal to idealizations, so it wouldn't be surprising to find some appealing here to the concept of what is assertible by an ideal rational agent, substituting for the concept of assertion the concept of what a logically omniscient being is committed to in virtue of what s/he asserts. Too much idealizing, methinks. The being would have to be quite unlike us, capable of knowing an uncountably infinite number of things and propositions with uncountably infinite components. If we want to speak of God here, theists like myself will have no problem with the discourse, but to think of such a being in terms of some finite extension of our own abilities and capacities is intolerable.

These same points hold for the concept of what is written. Everything written is inscribed, but sometimes only a string of morphemes is inscribed and sometimes the writing expresses a proposition as well. In my terminology to inscribe a sentence is the scribal form of uttering a string of phonemes, and writing relates to a proposition in scribal form in the way asserting relates in a vocal form to a proposition. As before, we'll have the same reasons to focus on the concept of writing rather than inscribing, since what is inscribed is not itself a bearer of truth value; but when we consider the writing operator  $W$ , we find that it is distributive only if the operator includes a reference to the logical consequences of what is written, and then the operator is not that of writing. For it is one thing to write down a claim, and it is another thing for what one has written to commit one logically to some further claim.

This problem about the  $W$  operator is not likely to detain the proof--theoretician for long. For one thing, there is no reason we can't interpret the  $W$  operator as "logically implied by what is written". Such an operator would purportedly show the falsity of the intuitive idea that anything true can be logically implied by something written truthfully on my blackboard.

Even so, there are costs to the syntactic generalization strategy. The  $W$  operator, on this understanding, is now logically complex, requiring reference to some correct logic for its interpretation. By contrast, in the knowability paradox, the rules of inference are intrinsic to the formal shorthand for the ordinary concept of knowledge. The more complex the operator, the

more tempting it is to attribute the perplexing result of the proof to the complexity of the operator and the difficulty in processing this complexity. That is, the temptation is to treat it like we do barber sentences (“there is a barber who shaves all and only them who do not shave themselves”): once we see the implications, we relieve our perplexity by simply reminding ourselves of the logical complexity of the sentence, so that the appearance of possibility is misleading. Interpreting the W operator in this complex way suggests the same kind of response, but it is a response that is not appropriate for the knowability paradox. Given these disparate reactions to two proofs that can be formally represented in the same way, it is not clear that sameness of formal representation has any power here to relieve our perplexity at the particular lost logical distinction resulting from Fitch’s proof.

This last claim raises a more general point about such a syntactic generalization strategy of finding analogues of knowability by searching for operators that mimic the distributivity and factivity features upon which Fitch’s proof relies. To see the issue, note that we can generalize the form of Fitch’s proof with a number of operators, such as the “true belief” operator, the “truly wished for” operator, the “truly imagined” operator, the “truly desired” operator, etc. In each case, the existence of thing not truly X--ed will purportedly be incompatible with the idea that any truth can be truly X--ed. Note here that we don’t need to idealize to some logically omniscient

X--er nor do we need to talk of the logical implications of what is X--ed in order to find an analogue of the knowability results, so no distinction between the arguments is naturally brought to mind by the degree of complexity of one operator over another. Does this variety of operators somehow explain the paradoxicality of a lost logical distinction between possible and actual universally known truth? I can't see why. A more plausible response to such generalizing is simply to characterize the more general paradoxicality in question. Instead of saying that there's a knowability paradox, we'd say instead that there is a paradox about any mental state operator that is factive and distributive. Syntax generalized is just paradox generalized, not paradox lost.

To generate paradox lost, my preferred explanation would be semantic. Since I've given no argument that no other explanation is possible, and it is worth considering what other non-semantic approaches might look like besides the syntactic generalization strategy just rejected. A pragmatic explanation might appeal to the notion of structural interference, claiming that the process of X--ing when applied to a conjunction can cause problems since in X--ing the first conjunct, one may affect the truth--value of the second.<sup>11</sup> Well, not quite, since the claim in question is an eternal truth if a truth at all (it quantifies over all individuals who X and all times), so a more careful claim would be that the X--ing of the first conjunct entails the falsity of the second.

These claims may be correct, but I have some questions about it. In order to carry the

explanation through, one will have to distinguish, in the case of the knowability paradox, between the executability of the basic steps of a procedure for generating knowledge (say, for knowing a conjunction) and the executability of the entire procedure itself, holding that only the basic steps are required to be executable, leaving unanswered why the basic steps are assumed to be executable.<sup>12</sup>

This point doesn't by itself cast doubt on this pragmatic approach to the paradox, however, so let's assume that this problem can be overcome. Deeper problems can be seen, though, if we simply attend to what the account is good at explaining and what it is not. Put pithily, it is good at explaining why antirealists shouldn't say that all truths are knowable, and it is not good at explaining the lost logical distinction expressed by (LD\*). It is clearly designed for the former purpose. Given this approach, one should deny that all truths are knowable by noting that those with antirealist sympathies should only go so far as to say that all truths for which structural interference is not an issue are knowable. To go further is to risk refutation by Fitch's proof.

The notion of structural interference is not designed to explain the lost logical distinction at the heart of the knowability paradox, however, and it is not very successful when turned to that purpose. To see the problem, I want to compare the omniscience--like conclusion of Fitch's proof with a more famous claim of the same sort, the claim that God exists. One's intuitive, pre-

philosophical attitude toward this claim should be that it is a contingent matter whether there is a God (just as it should be contingent whether all truths are known). There is a plausible path of reasoning to the denial of the contingency claim, however. It begins by claiming that God is the most perfect being, that He exemplifies maximal greatness. We thus identify the claim that God exists with the claim that maximal greatness is exemplified. The final steps toward a denial of the contingency assumption about God's existence is to clarify what maximal greatness involves (it is to display the maximal amount of any great-making property that has an intrinsic maximum) and argue that modal stability is itself a great-making property whose intrinsic maximum is existence in all possible worlds.

There are two quite natural responses to this threat to the contingency of the theistic claim. The first is to question the proof itself, to question the implications of the concept of maximal greatness, especially to doubt whether maximal modal stability is itself a great-making property. In doing so, one may look for analogues of the property, or one may simply construct formal notions that are claimed to have the modal stability property of being necessarily instantiated if instantiated. This strategy has a long history of threatening the ontological argument, from Gaunilo's perfect island to Arnauld's existent lion.

This approach is like the syntactic generalization strategy rejected earlier. It is, however, a more promising approach here, since the examples used do not simply mimic the problematic

proof, but constitute reductions of it. What they show is that the proof contains a mistake, even if we cannot identify exactly where the mistake occurs, thereby reaffirming our intuitive sense of the contingency of the theistic hypothesis.

This first response to the ontological argument thus could be used only to try to resurrect the syntactic generalization strategy, not to make sense of the pragmatic explanation relying on the notion of structural interference. Since we are now only considering the latter issue, we should move past this first response to look at the second kind of response. This response questions the account of the theistic claim itself. Why should we think that the claim that God exists is logically equivalent to the claim that maximal greatness is exemplified? After all, it's not as if the meaning (sense) of the term 'God' is the same as the meaning (sense) of the term 'maximally great being'.

Seeing what defenders of the argument do at this point shows why the pragmatic approach to the paradox is unsatisfying. Defenders of the ontological argument sometimes simply stipulate an understanding of 'God' in terms of maximal greatness. Such an approach leaves untouched the intuitive sense of the contingency of the theistic hypothesis, and thus provides no useful model for the pragmatic approach to the knowability paradox to emulate. What is needed instead is some way of explaining away some apparent contingency.

An alternative to this stipulation approach to the issue tries to argue against the

contingency claim. The argument takes the form of a reductio of the denial of contingency, beginning with the supposition that there is a God and also that there is a distinct being greater than God. The argument then proceeds by asking what understanding of God one might have that would call for allegiance, or worship, or religious commitment to the lesser being.

This line of argument could be resisted by insisting that a proper conception of God has no religious significance whatsoever,<sup>13</sup> but that escape route will strike most as fairly extreme. My point here, however, is not to defend this approach, but to show it can be used to explain away the apparent contingency of the theistic hypothesis. The apparent contingency is resisted by pointing out that we think in these terms because there is nothing about the meaning or sense of the claim that God exists that yields a denial of contingency, and yet there is an argument for this conclusion. Put in the language of the analytic/synthetic distinction, the claim is not necessary because analytic, but it is necessary nonetheless.

The important point to note is that the argument for this claim is not simply the original argument for the necessity of a maximally great being. Applying this point to the context of the knowability claim, we can see that the pragmatic approach to Fitch's proof in terms of structural interference is not plausibly taken at all as providing an analogue of this way of defending the necessity of the claim that God exists. To function in the same way, the pragmatic approach would need to provide a basis to argue, independently of Fitch's proof, that the omniscience--like

conclusion of that proof is necessarily false. Even the most superficial understanding of this approach shows that it would be complete pretense to assert that it can provide such an argument.

Once we appreciate the design plan of the structural interference approach, we can see why that approach seems so irrelevant to the lost logical distinction in question. The reason is that it wasn't designed to answer the question of how there could be such a loss. It was designed to answer to the question of whether an epistemic conception of truth requires affirming that all truths are knowable. Looked at from this perspective, it is a powerfully promising idea. It holds out the promise of being able to explain why "I don't exist" can't be verified even though it might be true, and how "no thinkers exist" can't be confirmed even though possible. Such conclusions are essential to an adequate defense of semantic antirealism, and it is a mark in favor of the philosophical fecundity of the idea of structural interference that it blocks these problems for antirealism while at the same time showing why Fitch's proof fails to refute the view. The lesson to learn here is that it is sometimes best to let tools be used for what they were intended, rather than to try to force them to be accomplish a job for which they were not intended. So independently of any use to which antirealists might put the concept of structural interference, we first need a solution to the paradox itself.

## **Conclusion**

In short, the paradox should disturb us all, antirealists and realists alike. It is true that the difference between a paradox and a merely surprising logical result is often not a difference in kind but only a difference in degree. Even so, there are distinctive marks of each that we look for when assessing what kind of a result we have achieved. If, for example, the result is merely one that we had no reason to think was true, we should classify such a result as a surprise. Or, again, if the result is merely one that threatens a particular philosophical perspective, such as antirealism, we should still classify the result as a merely surprising. But when the result threatens some aspect of received opinion, especially received opinion on logical matters themselves, we should not classify the result as merely surprising. In the present case, the lost logical distinction is part of a firmly entrenched understanding of the nature of the modalities of necessity, possibility, and actuality. It is not a partisan distinction that only certain philosophical perspectives could endorse, and in this way, it is paradoxical to face a derivation that undermines the distinction, in the same way it is paradoxical to be told that two grains of sand constitute a heap or that motion is impossible. The perplexity engendered by Fitch's proof is paradoxical, and the paradox cannot be addressed either by embracing Fitch's proof as a refutation of antirealism or by finding a version of antirealism that involves no commitment to the knowability claim itself. What we need is either an explanation of the failure of Fitch's proof or an explanation of the lost logical distinction between actuality and possibility that it implies—nothing short of that

constitutes a proper philosophical response to the paradox.<sup>14</sup>

## Notes

---

In thinking about the issues discussed here, I have been helped immensely in a number of conversations and blog interchanges with the following, whom I would like to thank: Mike Beaty, Bryan Frances, Michael Hand, Stephen Hetherington, Carrie Jenkins, Robert Johnson, Matt McGrath, Julien Murzi, Joe Salerno, Fritz Warfield, Jonathan Weinberg, and Tim Williamson.

<sup>1</sup> Do *not* say here: if you understand what *we* mean by the claim that all truths are known, it is not a bizarre or balmy claim. It's a sentence of English; we all speak the language; you don't get to reinterpret into your favored alternative idiolect or dialect.

<sup>2</sup> This is not to say that other strategies have not been tried. One common attempt is to reinterpret the disappointing results in intuitionistic language, i.e., to hold that, in that language, it is not so bad a thing to have to deny that there are unknown truths. I shall not comment here on this strategy beyond pointing out how philosophically strained it is.

<sup>3</sup> See, e.g., Hand and Kvanvig (1999); Williamson (2000a); DeVidi and Kenyon (2003). Replies and further discussion can be found in Tennant (2001a), (2001b) and (2002); and Brogaard and Salerno (2002).

---

<sup>4</sup> Note here how rare it is to find a defender of S5 prior to a development of the semantics in question by Kripke, and the corresponding paucity of deniers of S5 after this development.

<sup>5</sup> There is a way that the remarks in the text are a bit misleading, for it is not the mere fact of having a formal semantics that does the trick here. What is important is that the pure semantics connect up with ordinary meaning, which then gives us the explanation we seek. For more on the distinction between these two kinds of semantics, see Alvin Plantinga's (1979) distinction between pure and depraved semantics.

<sup>6</sup> Once we get to  $\sim\Diamond K(p \ \& \ \sim Kp)$ , as above, we can only get a contradiction by noting that  $p \ \& \ \sim Kp$  is true and hence knowable by the knowability principle.

<sup>7</sup> A point he made most recently to me at the *Modalism and Mentalism* conference in Copenhagen at the end of January, and most clearly made in print in Williamson (1987a) and (1993).

<sup>8</sup> My own view of the matter is that it is preferable to abandon importation/exportation in response to his arguments. If his example is put in counterfactual form, this response becomes obvious: to say that if Reagan were to lose, then if Anderson were to lose, Carter would win, is to say something false; whereas to say that if both Reagan and Anderson were to lose, Carter would win, is to say something true. Only the former, however, is of any use to McGee's argument.

McGee's argument, of course, involves indicative conditionals rather than counterfactual

---

ones. As a result, more argument is needed to get around his claims. The extra arguments needed, I believe, involve refusing to adopt the assertibility condition semantics he employs, but since that is beyond the scope of the present essay, I will leave that topic for another time and place.

<sup>9</sup> This response is in the spirit of Williamson's approach to the paradox both in print and in conversation, as well as that of Carrie Jenkins' contribution to this volume.

<sup>10</sup> I owe a great deal to Michael Hand regarding this approach. In fact, I think it fair to say that I simply would not have seen the possibility or significance of proof-theoretic insight without the long discussions we have had together.

<sup>11</sup> For the latter perspective on the paradox, Michael Hand (2003).

<sup>12</sup> See Hand (2003). Hand uses the analogy of recursion theory to show how a formal system can be developed by informal reference to procedures that might be executed, when the formal construction itself should not be thought to require such. Carrying the analogy through would lead to the conclusion that not even the basic steps of a verification procedure need to be thought of as executable.

<sup>13</sup> See, e.g., Richard Taylor (1982). Taylor endorses arguments for the existence of God, but takes this result to have no religious significance at all.

---