On a new tentative solution to Fitch's paradox

Abstract

In a recent paper, Alexander argues that relaxing the requirement that sound knowers know their own soundness might provide a solution to Fitch's paradox and introduces a suitable axiomatic system where the paradox is avoided. In this paper an analysis of this solution is proposed according to which the effective move for solving the paradox depends on the axiomatic treatment of the ontic modality rather than the limitations imposed on the epistemic one. It is then shown that, once the ontic modality is standardly introduced, the paradox still follows and, in addition, some puzzling consequences arise.

Keywords: knowledge; knowability; Fitch's paradox; epistemic modality; ontic modality.

1. Introduction

Fitch's paradox is the thesis that, if all truths are knowable in principle, then all truths are in fact known. To be sure, if *p* is an unknown truth, then that *p* is both true and unknown is an unknowable truth. For if it were possible to know that *p* is both true and unknown, then it would be possible to know that *p* and to know that *p* is unknown and, since knowledge implies truth, it would be possible for *p* to be both known and unknown, contradiction. Hence, there is no unknown truth, and so all truths are in fact known.[1]

It is well-known that one can carry out the proof of this thesis in any classical bi-modal system where the rule of ontic necessitation is assumed (if $\vdash \varphi$, then $\vdash \Box\varphi$), and the epistemic operator K satisfies some minimal conditions.[2] These conditions are to the effect that (I) knowing a conjunction entails knowing its conjuncts and (II) knowledge entails truth:

(I) $\vdash K(\varphi \wedge \psi) \rightarrow K\varphi \;/\; K\psi$

[1] See [16] for a synthetic presentation and [23] ch. 12 for a more in-depth discussion.

[2] In order to obtain the conclusion, *K* is to be interpreted as: it is known by someone at some time that.

(II) $\vdash K\varphi \rightarrow \varphi$

The first part of the proof runs as follows:

(1)	$\vdash K(\varphi \wedge \neg K\varphi) \rightarrow K\varphi$	by (I)
(2)	$\vdash K(\varphi \wedge \neg K\varphi) \rightarrow K\neg K\varphi$	by (I)
(3)	$\vdash K\neg K\varphi \rightarrow \neg K\varphi$	by (II)
(4)	$\vdash K(\varphi \wedge \neg K\varphi) \rightarrow \neg K\varphi$	by (2), (3) and logic
(5)	$\vdash \neg K(\varphi \wedge \neg K\varphi)$	by (1), (4) and logic

Then, given the rule of ontic necessitation and the principle of knowability, we conclude:

(6)	$\vdash \Box\neg K(\varphi \wedge \neg K\varphi)$	by (5) and $\Box$-necessitation
(7)	$\vdash \varphi \wedge \neg K\varphi \rightarrow \neg\Box\neg K(\varphi \wedge \neg K\varphi)$	by knowability (ontic)
(8)	$\vdash \neg(\varphi \wedge \neg K\varphi)$	by (6), (7) and logic
(9)	$\vdash \varphi \rightarrow K\varphi$	by (8) and classical logic

If we substitute the epistemic modality for the ontic modality, a similar proof can be carried out in any modal system where every derivable truth is known to be true and the knowledge operator satisfies the foregoing minimal conditions. Actually, it suffices to proceed as follows:

(6K)	$\vdash K\neg K(\varphi \wedge \neg K\varphi)$	by (5) and *K*-necessitation
(7K)	$\vdash \varphi \wedge \neg K\varphi \rightarrow \neg K\neg K(\varphi \wedge \neg K\varphi)$	by knowability (epistemic)
(8K)	$\vdash \neg(\varphi \wedge \neg K\varphi)$	by (6K), (7K) and logic
(9K)	$\vdash \varphi \rightarrow K\varphi$	by (8K) and classical logic

It is worth noting that the reason why the first conclusion, but not the second, is considered paradoxical is that the negation of the ontic version of the knowability principle, i.e. φ

is true but necessarily unknown, can strike us as counterintuitive, whereas the negation of the epistemic version of the knowability principle, i.e. φ is true but we know that it is unknown, seems to be perfectly intuitive.[3]
It is possible, following Tennant [17], to identify three main general strategies for solving the paradox.

(1) A *semantic strategy*, consisting in giving a new interpretation of the principle of knowability, so to change the most important premise of the argument.[4]

(2) A *scope strategy*, consisting in limiting the class of propositions that are assumed to be knowable, so to dismiss a premise that is essential for the argument.[5]

(3) a *logical strategy*, consisting in limiting the logical framework in which the argument is settled, so to hinder a step that is essential for obtaining the conclusion.[6]

In this paper I focus on a new solution of the paradox based on the limitation of the axio-

---

[3] To be more accurate, it is perfectly intuitive *from a classical point of view*, where negation is not conceived of as entailment of inconsistency. Indeed, every time we know that we wonder about p, i.e. every time when it is true that $K(\neg Kp \wedge \neg K\neg p)$, we are in a condition where $\neg K\varphi$ for some true φ. This is why no principle like $\varphi \rightarrow \neg K\neg K\varphi$ is assumed in a classical framework. In addition, Alexander was able to show the interesting result that $\varphi \rightarrow \neg K\neg K\varphi$ implies $\varphi \rightarrow KK\varphi$ even without assuming (II).
[4] This strategy was adopted by Edgington in [8], and developed in [9]. According to Edgington, what the principle of knowability actually states is that all *actual* truths are knowable. A logical characterization of Edgington's interpretation of the paradox is given in [14], where the costs of the original proposal are highlighted and an improved solution is given in terms of a novel semantic characterization of the actuality operator. Edgington's interpretation has been strongly criticized by Williamson, in [21] and [23], 12.5, but see [5] and [12] for a different implementation of the basic idea. Recently, a new interesting solution of the paradox, based on a definition of knowability in terms of existence of a proof, has been advanced in [6].
[5] This strategy was adopted by Tennant in [17] and originated a decennial debate with Williamson. See [18], [19], and [24], [25]. The idea proposed by Tennant, i.e., to limit the scope of the knowability principle to propositions whose knowledge does not give rise to contradiction, is still a promising way out of the paradox. In the same vein, in [3] it is proposed to limit the principle of knowability with respect to stable propositions, i.e., propositions that, once true, cannot change their truth value. This is an elegant and intuitive solution, whose only problem seems to be that it prevents the general possibility of knowing that we don't know some true proposition.
[6] This is the strategy preferred by constructivists. The central intuition is that the last step in the argument is only classically valid and that the conclusion that can be achieved by using intuitionistic rather than classical logic, i.e., $\varphi \rightarrow \neg\neg K\varphi$, can be accepted by a constructivist once the connectives are construed in the right way. A standard way in which the paradox can be interpreted from a constructivist point of view is proposed by Dummett in [7]. See [10], [20], [22] for a discussion of the problems that constructivists have to take into account. See also the general proof-theoretic approach developed in [13], where it is shown that the adoption of intuitionistic logic allow us to block all the possible derivation of the paradox.

matic system in which the knowability principle is introduced. The novelty of the approach, which follows a logical strategy, consists in the fact that the limitations imposed on the set of logical principles do not concern principles of classical logic, but modal principles only. *In particular, as we will see in the next section, the closure axioms characterizing the epistemic operator are explicitly restricted, while the set of axioms characterizing the ontic operator is empty.*

In section 2 I present the general logical framework within which the solution of the paradox is developed. Since the systems to be analysed are new in the panorama of modal logic, the presentation is rather comprehensive. In section 3 a detailed exposition of all the steps of the solution is provided. In section 4 it is shown why the solution is to be rejected and an interesting variant of it is discussed, and eventually rejected as well.

2. The logical framework

The systems we are going to work with are defined relative to a standard propositional modal language $\mathcal{L}$ containing the ontic and the epistemic modal operators. In particular, they are extensions of the same basic system. In this section a general framework for treating modal logics that are less powerful than standard systems of normal modal logic is developed in some details, so to provide a self contained presentation of both Alexander's results and some extensions thereof. Let us introduce some key definitions[7].

Definition: kinds of modal logics.
Let $\mathcal{L}$ be a propositional modal language and $\Box$ a modality.
The following definition is based on [4].

A *general modal logic* **L** is a set of formulas of $\mathcal{L}$ such that:

P: $\varphi \in \mathbf{L}$ provided φ is a classical tautology.

MP: **L** is closed under *modus ponens*.

[7] See [4], ch 4, for a general introduction to modal systems. Note that instances of classical tautologies are allowed to contain modal operators. Hence, e.g., $\Box p \rightarrow \Box p$ is to be considered a classical tautology.

A *pre-normal modal logic* **L** is a general modal logic such that:

$\Box$**1**: $\Box(\varphi \rightarrow \psi) \rightarrow (\Box\varphi \rightarrow \Box\psi) \in \mathbf{L}$.

$\Box$**2**: $\Box\varphi \in \mathbf{L}$, provided φ is a classical tautology.

A *normal modal logic* **L** is a general modal logic such that:

$\Box$**1**: $\Box(\varphi \rightarrow \psi) \rightarrow (\Box\varphi \rightarrow \Box\psi) \in \mathbf{L}$.

$\Box$**R**: $\Box\varphi \in \mathbf{L}$, provided $\varphi \in \mathbf{L}$

Definition: derivability.

Let $\varphi_1, \ldots, \varphi_n, \varphi$ be in $\mathcal{L}$. Say that φ is propositionally derivable from $\varphi_1, \ldots, \varphi_n$ if and only if $\varphi_1 \wedge \ldots \wedge \varphi_n \rightarrow \varphi$ is a tautology. In this case, we will write: $\varphi_1, \ldots, \varphi_n \vdash_{\mathbf{PL}} \varphi$. In general, I will write $\varphi_1, \ldots, \varphi_n \vdash_{\mathbf{L}} \varphi$ whenever $\varphi_1 \wedge \ldots \wedge \varphi_n \rightarrow \varphi \in \mathbf{L}$. The following propositions are direct consequences of the previous definition.

Proposition 1: Let **L** be a *general modal logic*, then

$\varphi_1, \ldots, \varphi_n \vdash_{\mathbf{PL}} \varphi \Rightarrow \varphi_1, \ldots, \varphi_n \vdash_{\mathbf{L}} \varphi$, by def. **L**

Proposition 2: Let **L** be a *pre-normal modal logic*, then

$\varphi_1, \ldots, \varphi_n \vdash_{\mathbf{PL}} \varphi \Rightarrow \Box\varphi_1, \ldots, \Box\varphi_n \vdash_{\mathbf{L}} \Box\varphi$, by $\Box$**1** and $\Box$**2**

Proposition 3: Let **L** be a *normal modal logic*, then

$\varphi_1, \ldots, \varphi_n \vdash_{\mathbf{L}} \varphi \Rightarrow \Box\varphi_1, \ldots, \Box\varphi_n \vdash_{\mathbf{L}} \Box\varphi$, by $\Box$**1** and $\Box$**R**

Hence, the main difference between a pre-normal modal logic and a normal modal logic concerns the conditions of application of the general rule of necessitation.

2.1. Systems and semantics

Let $\mathcal{L}$ be specified by the following grammar:

$$\varphi := p \mid \neg\varphi \mid \varphi \wedge \psi \mid K\varphi \mid \Box\varphi$$

where $p \in P$ and P is a set of propositional variables. The other propositional connectives are defined as usual. On the intended interpretation, K is the epistemic modality of knowledge, while $\Box$ is the ontic modality of necessity.

In what follows we mainly will work with bi-modal logical systems, which are obtained by extending a basic system **S** with axioms from the following table.

TABLE 1:

Axioms on K		Axioms on $\Box$	
K1:	$K\varphi$, if φ is propositionally valid	**$\Box$1**:	$\Box\varphi$, if φ is propositionally valid
K2:	$K(\varphi \to \psi) \to (K\varphi \to K\psi)$	**$\Box$2**:	$\Box(\varphi \to \psi) \to (\Box\varphi \to \Box\psi)$
K3:	$K\varphi \to \varphi$	**$\Box$3**:	$\Box\varphi \to \varphi$
K4:	$K\varphi$, if φ is in **K1**$\cup$**K2**	**$\Box$4**:	$\Box\varphi$, if φ is in $\Box$**1**$\cup\Box$**2**
K5:	$K\varphi$, if φ is in **K1**$\cup$**K2**$\cup$**K3**	**$\Box$5**:	$\Box\varphi$, if φ is in $\Box$**1**$\cup\Box$**2**$\cup\Box$**3**
Bridge Principles		Knowability Principle	
K$\Box$1:	$K\varphi \to \Box\varphi$		
K$\Box$2:	$\Box\varphi$, if φ is in **K1**$\cup$**K2**	**KP**:	$\varphi \to \neg\Box\neg K\varphi$
K$\Box$3:	$\Box\varphi$, if φ is in **K1**$\cup$**K2**$\cup$**K3**		

Definition: basic system **S**.

S is the smallest system which is closed under *modus ponens* and contains φ, whenever φ is a classical tautology, and **K3**.[8] Axiom **K3** is indeed necessary for characterizing the epistemic modality as a *factive* modality.

The semantics we are going to introduce is based on the notion of frame for $\mathcal{L}$, which is a triple $\langle W, \mathcal{K}, \mathcal{E}\rangle$, where W is a set of states, and $\mathcal{K}$ and $\mathcal{E}$ are functions assigning sets of formulas to elements of W:

[8] Since it is a straightforward consequence of the definition of general modal logic that the intersection of any set of general modal logics is in turn a general modal logic, **S** turns out to be the intersection of all the general modal logics over $\mathcal{L}$.

(i) $\mathcal{K}$: W –> $\wp(\mathcal{L})$

(ii) $\mathcal{E}$: W –> $\wp(\mathcal{L})$

Intuitively, $\mathcal{K}$ specifies the set of formulas that are known in a state, while $\mathcal{E}$ specifies the set of formulas that are necessary in a state. The sense in which $\mathcal{E}$ can be considered a correct specification of what is necessary in a state will be explored below. The notion of model is defined accordingly: a model for $\mathcal{L}$ based on $\langle W, \mathcal{K}, \mathcal{E}\rangle$ is a tuple $\mathcal{M} = \langle W, \mathcal{K}, \mathcal{E}, V\rangle$, where V is a function assigning sets of propositional variables to elements of W.

The notion of truth of a formula at a state in a model is recursively defined as follows:

$$\mathcal{M},w \mid= p \iff p \in V(w)$$

$$\mathcal{M},w \mid= \neg\varphi \iff \text{not } \mathcal{M},w \mid= \varphi$$

$$\mathcal{M},w \mid= \varphi \wedge \psi \iff \mathcal{M},w \mid= \varphi \text{ and } \mathcal{M},w \mid= \psi$$

$$\mathcal{M},w \mid= K\varphi \iff \mathcal{M},w \mid= \varphi \text{ and } \varphi \in \mathcal{K}(w)$$

$$\mathcal{M},w \mid= \Box\varphi \iff \varphi \in \mathcal{E}(w)$$

Note that (i) the ontic modality is not required to be *factive* and (ii) the semantic condition corresponding to a factive ontic modality is

$$\mathcal{M},w \mid= \Box\varphi \iff \mathcal{M},w \mid= \varphi \text{ and } \varphi \in \mathcal{E}(w)$$

The notion of logical consequence and logical truth are defined as usual. Hence

$$X \mid\mid\!- \varphi \iff \mathcal{M},w \mid= X \Rightarrow \mathcal{M},w \mid= \varphi, \text{ for all } \mathcal{M} \text{ and } w$$

where X is a set of formulas and $\mathcal{M},w \mid= X$ precisely when $\mathcal{M},w \mid= \varphi$ for every $\varphi \in X$.

Finally, it is not difficult to see that the following conditions on functions $\mathcal{K}$ and $\mathcal{E}$ are the ones we need in order to ensure the validity of axioms in table 1.

TABLE 2:

Conditions on $\mathcal{K}(w)$, for all w		Conditions on $\mathcal{E}(w)$, for all w	
$\mathcal{K}1$:	$\varphi\in\mathcal{K}(w)$, if φ is a classical tautology	$\mathcal{E}1$:	$\varphi\in\mathcal{E}(w)$, if φ is a classical tautology
$\mathcal{K}2$:	$(\varphi \rightarrow \psi),\varphi\in\mathcal{K}(w) \Rightarrow \psi\in\mathcal{K}(w)$	$\mathcal{E}2$:	$(\varphi \rightarrow \psi),\varphi\in\mathcal{E}(w) \Rightarrow \psi\in\mathcal{E}(w)$
$\mathcal{K}3$:	–	$\mathcal{E}3$:	$\mathcal{M},w \models \Box\varphi \Leftrightarrow \mathcal{M},w \models \varphi$ and $\varphi\in\mathcal{E}(w)$
$\mathcal{K}4$:	$\mathbf{K1}\cup\mathbf{K2} \subseteq \mathcal{K}(w)$	$\mathcal{E}4$:	$\Box\mathbf{1}\cup\Box\mathbf{2} \subseteq \mathcal{E}(w)$
$\mathcal{K}5$:	$\mathbf{K1}\cup\mathbf{K2}\cup\mathbf{K3} \subseteq \mathcal{K}(w)$	$\mathcal{E}5$:	$\Box\mathbf{1}\cup\Box\mathbf{2}\cup\Box\mathbf{3} \subseteq \mathcal{E}(w)$
Conditions on $\mathcal{K}(w)$ and $\mathcal{E}(w)$, for all w			
$\mathcal{KE}1$:	$\mathcal{K}(w) \subseteq \mathcal{E}(w)$		
$\mathcal{KE}2$:	$\mathbf{K1}\cup\mathbf{K2} \subseteq \mathcal{E}(w)$		
$\mathcal{KE}3$:	$\mathbf{K1}\cup\mathbf{K2}\cup\mathbf{K3} \subseteq \mathcal{E}(w)$		

As we can see, any model is characterized by a syntactic component, encoded by functions $\mathcal{K}$ and $\mathcal{E}$. In this way, we get the most general modal semantics for systems of modal logic.[9]

2.2. Soundness and completeness

Let us prove soundness and completeness of system **S**.

Soundness: **S** is sound with respect to the class of all models for $\mathcal{L}$.

The proof that all the classical tautologies are true at every world in every model is straightforward. Hence, we only have to prove that $\models K\varphi \rightarrow \varphi$.

K3: $\models K\varphi \rightarrow \varphi$

$\mathcal{M},w \models K\varphi$ implies $\mathcal{M},w \models \varphi$, by the definition of $\models$.

Completeness: **S** is complete with respect to the class of all models for $\mathcal{L}$.

[9] See [11], ch. 9 for an overall introduction to this kind of semantics.

The proof is by construction of a canonical model for $\mathcal{L}$.[10]

The canonical model for $\mathcal{L}$ is the tuple $\mathcal{M} = \langle W, \mathcal{K}, \mathcal{E}, V \rangle$, where

- $W = \{w \mid w \text{ is } \mathbf{S}\text{-maximally consistent}\}$
- $\mathcal{K}$ is such that $\mathcal{K}(w) = \{\varphi \mid K\varphi \in w\}$
- $\mathcal{E}$ is such that $\mathcal{E}(w) = \{\varphi \mid \Box\varphi \in w\}$
- V is such that $V(w) = w$

Lemma 1: $\mathcal{M}$ is a model for $\mathcal{L}$.

Straightforward, since, for all $w \in W$, $\mathcal{K}(w)$ and $\mathcal{E}(w)$ are sets of formulas.

Truth Lemma: $\mathcal{M},w \models \varphi \iff \varphi \in w$

The interesting cases are the ones concerning the modal operators.

(1) $\mathcal{M},w \models K\varphi \iff K\varphi \in w$.

Suppose $\mathcal{M},w \models K\varphi$; then $\varphi \in \mathcal{K}(w)$, by def. $\models$, and so $K\varphi \in w$, by the definition of $\mathcal{K}(w)$. Suppose now $K\varphi \in w$; then $\varphi \in w$ and $\varphi \in \mathcal{K}(w)$, by **K1** and the definition of $\mathcal{K}(w)$; therefore $\mathcal{M},w \models \varphi$ and $\varphi \in \mathcal{K}(w)$, by inductive hypothesis, so that $\mathcal{M},w \models K\varphi$, by def. $\models$.

(2) $\mathcal{M},w \models \Box\varphi \iff \varphi \in \mathcal{E}(w) \iff \Box\varphi \in w$, by definition of $\mathcal{E}(w)$.

It is not difficult now to extend the previous proof to systems obtained by adding to **S** axioms from table 1. To wit, let us prove soundness and completeness for two of the systems we will use in the following, call them **S1** and **S2**.

S1 is constituted by **S** plus axioms **K1**, **K2**, **K4**, **K□1**.

[10] See [4], ch. 4 for an introduction to this procedure and proofs of the standard steps.

Note that **K□2** is derivable in **S1**.

S2 is constituted by **S** plus axioms **K1**, **K2**, **K5**, **K□1**.
Note that both **K□2** and **K□3** are derivable in **S2**.

Soundness: **S1**/**S2** is sound with respect to the class of models for $\mathcal{L}$ satisfying conditions $\mathcal{K}1$, $\mathcal{K}2$, $\mathcal{K}4$/$\mathcal{K}5$, $\mathcal{KE}1$.

K1: $||- K\varphi$, if φ is propositionally valid.
Suppose φ is propositionally valid. Then, for all $w \in W$, $\mathcal{M},w \models \varphi$, since every propositionally valid formula is true at every state, and $\varphi \in \mathcal{K}(w)$, by condition $\mathcal{K}1$. Thus, for all $w \in W$, $\mathcal{M},w \models K\varphi$, by the definition of $\models$, and so $||- K\varphi$.

K2: $||- K(\varphi \rightarrow \psi) \rightarrow (K\varphi \rightarrow K\psi)$
Suppose $\mathcal{M},w \models K(\varphi \rightarrow \psi)$ and $\mathcal{M},w \models K\varphi$. Then, $\mathcal{M},w \models (\varphi \rightarrow \psi)$ and $\mathcal{M},w \models \varphi$, by the definition of $\models$, and so $\mathcal{M},w \models \psi$. In addition, $(\varphi \rightarrow \psi) \in \mathcal{K}(w)$ and $\varphi \in \mathcal{K}(w)$, again by the definition of $\models$, and so $\psi \in \mathcal{K}(w)$, by condition $\mathcal{K}2$. Thus, $\mathcal{M},w \models K\psi$.

K4: $||- K\varphi$, if φ is in **K1**∪**K2**.
Since φ is in **K1**∪**K2**, $\mathcal{M},w \models \varphi$, by the previous points. Hence, $\mathcal{M},w \models K\varphi$, by $\mathcal{K}4$.

K5: $||- K\varphi$, if φ is in **K1**∪**K2**∪**K3**.
Since φ is in **K1**∪**K2**∪**K3**, $\mathcal{M},w \models \varphi$, by the previous points. Hence, $\mathcal{M},w \models K\varphi$, by $\mathcal{K}5$.

K□1: $||- K\varphi \rightarrow \Box\varphi$
Straightforward, since $\mathcal{K}(w) \subseteq \mathcal{E}(w)$, by condition $\mathcal{KE}1$.

Completeness: **S1**/**S2** is complete with respect to the class of models for $\mathcal{L}$ satisfying conditions $\mathcal{K}1$, $\mathcal{K}2$, $\mathcal{K}4/\mathcal{K}5$, $\mathcal{KE}1$.

The proof is similar to the one relative to system **S**.

It is worth noting that by structurally similar proofs we can show that systems obtained from the basic one by adding axioms concerning the ontic modality are both sound and complete with respect to models satisfying the corresponding conditions. In addition, the following important proposition is provable.

Proposition 4: let φ be **S***-consistent, where **S*** is the system obtained by adding to **S** all the axioms in table 1 except **KP**, then $K\varphi$ is **S***-consistent.

Proof. Suppose φ is **S***-consistent, i.e. $\not\vdash_{\mathbf{S1}} \neg\varphi$. Then, by completeness, there is some model $\mathcal{M} = \langle W,\mathcal{K},\mathcal{E},V\rangle$ and a state w such that $\mathcal{M},w \models \varphi$. Let $\mathcal{M}^* = \langle W^*,\mathcal{K}^*,\mathcal{E}^*,V^*\rangle$ be such that

(i) $W^* = \{w\}$
(ii) $\mathcal{K}^*(w) = \mathcal{E}^*(w) = \mathcal{L}$
(iii) $V^* = V$ restricted to W^*.

It is then straightforward that $\mathcal{M}^*$ is a model for $\mathcal{L}$ satisfying all the conditions in table 2 and that $\mathcal{M},w \models K\varphi$. Hence, $\not\vdash_{\mathbf{S1}} \neg K\varphi$, by soundness, and so $K\varphi$ is **S***-consistent.

3. A tentative solution

In his recent [1], Alexander attempts at solving the paradox by limiting the epistemic power of the agent. The systems Alexander works with are **S1** plus **KP** and **S2** plus **KP** and the solution he proposes is justified on the basis of the following two claims, that I comment in order to make explicit what is actually at stake.

Claim 1 [1], p. 2015: "Fitch's paradox is the fact that certain set of assumptions imply that

$\varphi \rightarrow K\varphi$". More precisely: Fitch's paradox is the fact that certain set of assumptions, which are introduced in order to capture our basic intuitions about modal and epistemic facts, imply that $\varphi \rightarrow K\varphi$. Indeed, what is paradoxical in a paradox is precisely the fact that some firm intuitions clash.

Claim 2: [1], p. 2016: "to resolve Fitch's paradox is to weaken the assumptions so that the conclusion no longer follows". More precisely: to resolve Fitch's paradox is to alter the assumptions so that (i) the basic intuitions are preserved or better captured and (ii) the conclusion no longer follows. To see the significance of (i), just consider that we can obtain a simple solution to Fitch's paradox by adopting no axioms on K except **KP**.

What Alexander shows is, firstly, that Fitch's paradox can be derived in **S2**, but not in **S1** and, secondly, that the move from **S2** to **S1** is intuitive, since what is eliminated is the schema $K(K\varphi \rightarrow \varphi)$[11], and Alexander argues that $K(K\varphi \rightarrow \varphi)$ may to be too strong for characterizing actual knowledge (see [1], p. 2016). Hence, assuming **S1** as our basic system allows us to provide a straightforward solution to Fitch's paradox.

3.1. The first step: the derivation of the paradox in S2+KP

A complete version of the proof of the paradox is as follows.

Proposition 5: $\vdash_{\mathbf{S2+KP}} \varphi \rightarrow K\varphi$.

Proof: set

$H = K(\varphi \wedge \neg K\varphi)$

$H_1 = K(\varphi \wedge \neg K\varphi \rightarrow \varphi)$

$H_2 = K(\varphi \wedge \neg K\varphi \rightarrow \neg K\varphi)$

$H_3 = K(\varphi \wedge \neg K\varphi \rightarrow \varphi) \rightarrow (K(\varphi \wedge \neg K\varphi) \rightarrow K\varphi)$

$H_4 = K(\varphi \wedge \neg K\varphi \rightarrow \neg K\varphi) \rightarrow (K(\varphi \wedge \neg K\varphi) \rightarrow K\neg K\varphi)$

[11] More precisely the schema $K^N(K\varphi \rightarrow \varphi)$, where K^N is a string of N > 0 epistemic operators.

$H_5 = K\neg K\varphi \to \neg K\varphi$

It is then immediate that:

$H_1, H_2 \in$ **K1**; $H_3, H_4 \in$ **K2**; $H_5 \in$ **K3**, and so $KH_1, KH_2, KH_3, KH_4, KH_5 \in$ **K5**.

(1) $H_1 \wedge H_2 \wedge H_3 \wedge H_4 \wedge H_5 \wedge H \to K\varphi \wedge \neg K\varphi$ is a classical tautology

(2) $H_1 \wedge H_2 \wedge H_3 \wedge H_4 \wedge H_5 \to \neg H$ is a classical tautology

(3) $H_1, H_2, H_3, H_4, H_5 \vdash_{\mathbf{PL}} \neg H$, by the definition of $\vdash_{\mathbf{PL}}$

(4) $KH_1, KH_2, KH_3, KH_4, KH_5 \vdash_{\mathbf{S2}} K\neg H$, by proposition 2

(5) $\vdash_{\mathbf{S2}} K\neg H$, since $KH_1, KH_2, KH_3, KH_4, KH_5 \in$ **K5**

(6) $\vdash_{\mathbf{S2}} \Box\neg H$, by **K□1** and *modus ponens*

(7) $\varphi \wedge \neg K\varphi \vdash_{\mathbf{S2+KP}} \neg\Box\neg H$, by **KP**

(8) $\vdash_{\mathbf{S2+KP}} \neg(\varphi \wedge \neg K\varphi)$, by propositional logic

(9) $\vdash_{\mathbf{S2+KP}} \varphi \to K\varphi$, by propositional logic

3.2. The second step: the identification of the culprit

According to Alexander, the key step is (5), which allows us to obtain $K\neg K(\varphi \wedge \neg K\varphi)$, and then $\Box\neg K(\varphi \wedge \neg K\varphi)$, i.e. the impossibility of knowing an unknown proposition. Still, this step is admissible only because the epistemic agent is allowed to know that knowledge is factive and, in Alexander's opinion, this assumption is to be dropped. Indeed, Alexander proposes two main philosophical reasons for dropping the axiom: the first one related to Gödel's second incompleteness theorem; the second one related to how explicit knowledge is modelled in some systems of explicit modal logic. Let us consider these reasons in turn.

(1) **K5** and Gödel's second incompleteness theorem.

Let K_{PA} be the operator of knowledge interpreted as provability in PA, and assume that PA is consistent. Then, while $K_{PA}\varphi \to \varphi$ is true, $K_{PA}(K_{PA}\varphi \to \varphi)$ fails. To be sure, $K_{PA}(K_{PA}\varphi \to \varphi)$ says that it is possible to prove in PA that what is provable in PA is true, which contradicts Gödel's second incompleteness theorem.

(2) **K5** in specific systems of explicit epistemic logic.

Let K be the operator of knowledge interpreted as explicit knowledge, i.e. knowledge which is currently immediately available to an epistemic agent. Then, while $K\varphi \rightarrow \varphi$ is true, by definition, $K(K\varphi \rightarrow \varphi)$ fails, since the epistemic agent might be in a position where she is not aware of the fact that knowledge is factive.

3.3. The final step: the paradox is not derivable in S1

Axiom **K5** is the source of the paradox. The intuitive move is then to drop it, which amounts to move from system **S2** to system **S1**. Hence, in order to give a complete solution to the paradox, we have to show that the paradox is not derivable in **S1**. In particular, in view of the results of the previous section, it suffices to show that there is a model $\mathcal{M}$ for $\mathcal{L}$ such that $\mathcal{M}$ satisfies conditions $\mathcal{K}1$, $\mathcal{K}2$, $\mathcal{K}4$, $\mathcal{KE}1$ and

(1) $\varphi \rightarrow K\varphi$ is not valid in $\mathcal{M}$.

(2) $\varphi \rightarrow \neg\Box\neg K\varphi$ is valid in $\mathcal{M}$;

Proposition 6: $\not\vdash_{\mathbf{S1+KP}} \varphi \rightarrow K\varphi$.

Let $\mathcal{M} = \langle W, \mathcal{K}, \mathcal{E}, V\rangle$, where

- $W = \{w\}$
- $\mathcal{K}$ is such that $\mathcal{K}(w) = \{\varphi \mid \vdash_{S1} \varphi\}$
- $\mathcal{E}$ is such that $\mathcal{E}(w) = \{\varphi \mid \vdash_{S1} \varphi\}$
- V is such that $V(w) = \{p\}$

First, given the definition of $\mathcal{K}$ and $\mathcal{E}$, it is not difficult to see that $\mathcal{M}$ satisfies conditions $\mathcal{K}1$, $\mathcal{K}2$, $\mathcal{K}4$, $\mathcal{KE}1$, given the soundness theorem. Furthermore

(1) $\mathcal{M},w \not\models Kp$. Indeed, $\not\vdash_{S1} p$, since it is well known that any proof of p based on instances

of axiom schemata can be transformed in a proof of $\neg p$, against the soundness theorem.

(2) $\mathcal{M},w \models \varphi \rightarrow \neg\Box\neg K\varphi$. Suppose $\mathcal{M},w \models \varphi$. Then $\not\models_{S1} \neg\varphi$, by the soundness theorem. Hence, $\not\models_{S1} \neg K\varphi$, by proposition 4, and so $\neg K\varphi$ is not in $\mathcal{E}(w)$. Therefore, $\mathcal{M},w \not\models \Box\neg K\varphi$, so that $\mathcal{M},w \models \neg\Box\neg K\varphi$, whence the conclusion.

4. Shortcomings

Is Alexander's strategy successful in solving the paradox? What I'm willing to show is that the success of Alexander's solution is not based on the limitations imposed on the epistemic operator, but on neglecting our basic intuitions about the behaviour of the ontic modality. In particular, I will focus on three problems that emerge in relation to the three steps presented above.

4.1. On the derivation of the paradox in S2

The derivation of the paradox in **S2** is crucially dependent on the assumption of axiom **K□1**. This observation draws our attention both to the reasons for introducing **K□1** and to its credibility[12]. Actually, the unique reason why **K□1** is introduced is to allow for the derivation of the paradox without exploiting the □-necessitation rule. However, is **K□1** a credible bridge principle with respect to knowledge and necessity? In a sense, it is. Indeed, one might argue that knowledge concerns *stable propositions* only, i.e. propositions that possess a truth value independently of the circumstance in which they are uttered, thus being necessarily true, if true.[13] Still, the principle of knowability, limited to stable propositions, would become $\Box\varphi \rightarrow \neg\Box\neg K\varphi$, and it is well-known that this principle is not subjected to any paradox, being consistent with the negation of $\varphi \rightarrow K\varphi$, even in strong systems where both the necessity operator and the knowledge operator are modelled as **KT5** modalities.[14]

[12] Alexander attributes **K□1** to Salerno. However, as far as I can see, no such axiom is proposed. Actually, what we find there the assumption that what is necessarily false is impossible: $\Box\neg\varphi \rightarrow \neg\Diamond\varphi$.

[13] See [3] for an analysis of this position. Limiting knowledge to necessary propositions is a consequence of Edgington's approach in [8]. Williamson, in [23], criticizes this move by highlighting that Fitch's paradox displays a limit to possible knowledge of contingent truths, so that it cannot be solved by limiting knowledge to necessary propositions.

[14] To see that, consider a system where both □ and K are **KT5** modalities and the following axiom is added: $K\varphi \leftrightarrow \Box\varphi$. In this system □ and K are in fact the same modality, so that $\Box\varphi \rightarrow \neg\Box\neg K\varphi$ turns out to coincide with axiom **5**: $\Box\varphi \rightarrow$

Therefore, no new solution should be required.

4.2. On the identification of the culprit

Let us consider now the main reasons for discarding $K(K\varphi \to \varphi)$. The first reason is that, if the operator of knowledge interpreted as provability in PA, then $K(K\varphi \to \varphi)$ is in contradiction with Gödel's second incompleteness theorem. That is true, but the interpretation of K as provability in PA is not the intended one. As it is known, in the context of Fitch's paradox, $K\varphi$ is interpreted as stating that φ it is known by some agent at some time. Since an agent is not constrained to work within a fixed axiomatic system, $K(K\varphi \to \varphi)$ can be also interpreted as $K_T(K_{PA}\varphi \to \varphi)$, where T is any theory. Still, $K_T(K_{PA}\varphi \to \varphi)$ is not in contradiction with Gödel's second incompleteness theorem, provided T is sufficiently powerful (indeed more powerful than PA). Hence, the first reason is not compelling. The second reason is that, if K is construed as an operator of explicit knowledge, then the epistemic agent might be in a position where she is not aware of the fact that knowledge is factive, so that $K(K\varphi \to \varphi)$. However, K cannot possibly be interpreted as an operator of explicit knowledge in systems that contains axioms like **K1** and **K2**, since in these systems the epistemic agent is assumed to know all the valid propositions and to be able to apply *modus ponens* to all known propositions, while no agent is able to explicitly know an infinite number of propositions and to perform infinitely many applications of a rule. Hence, also the second reason is not compelling. In sum, the reasons for discarding $K(K\varphi \to \varphi)$ seem to be ineffective.

4.3. On the impossibility of deriving the paradox in S1

Let us come now to the crucial problem: once the necessity operator is standardly introduced, the derivation of the paradox cannot be blocked by dropping **K5**. Let us consider the system **S1** plus any standard axiomatisation of the notion of necessity that includes the rule of necessitation[15]. Then, Fitch's paradox follows, since principle (II), i.e., $K\varphi \to \varphi$, is

---

$\neg\Box\neg\Box\varphi$, whereas the triviality axiom $\varphi \leftrightarrow \Box\varphi$ is not derivable.

[15] Let us remind that the axiomatic systems that are usually considered appropriate for capturing the notion of necessity are **KT**, **KT4**, and **KT5**, defined according to the following axioms:

assumed and principle (I), i.e., $K(\varphi \wedge \psi) \rightarrow K\varphi$ / $K\psi$ is derivable:

(1)	$\vdash (\varphi \wedge \psi) \rightarrow \varphi$	by **P**
(2)	$\vdash K((\varphi \wedge \psi) \rightarrow \varphi)$	by (1) and **K1**
(3)	$\vdash K(\varphi \wedge \psi) \rightarrow K\varphi$	by (2) and **K2**

Hence, **S1** + $\Box$-necessitation rule suffices to get the paradox. This is a terrific problem, especially in view of the fact that both Alexander's systems are silent with respect to how necessity has to be axiomatically characterized. Indeed, both **S1** and **S2** contain no axiom on $\Box$, and without any axiomatic characterization of this operator *it is difficult even to say that* **KP** *is a principle of knowability*. To be sure, to say that φ is knowable amounts to saying that φ can be known, i.e. that it is possible for φ to be known. Still, it seems impossible to say that the notion captured by $\neg\Box\neg$ in systems **S1** and **S2** is a notion of possibility, because it obeys none of the principles we intuitively acknowledge as correct relative to this notion.

4.4. A variant of the solution

In this last section I present a variant of Alexander's solution that seems to succeed in meeting all the problems raised above.[16] Let us consider the following pre-normal systems:

SN1 is constituted by **S** plus axioms **K1**, **K2**, **□1**, **□2**, **K□2**.
SN2 is constituted by **S** plus axioms **K1**, **K2**, **□1**, **□2**, **K□2**, **K□3**.

It is now possible to prove that, while Fitch's paradox is derivable in **SN2+KP**, it is not de-

---

K: $\Box(\varphi \rightarrow \psi) \rightarrow (\Box\varphi \rightarrow \Box\psi)$
T: $\Box\varphi \rightarrow \varphi$
4: $\Box\varphi \rightarrow \Box\Box\varphi$
5: $\neg\Box\varphi \rightarrow \Box\neg\Box\varphi$
Rule of necessitation: $\vdash \varphi \Rightarrow \vdash \Box\varphi$
KT5 is commonly considered the right system to capture both the notion of logical necessity and the notion of metaphysical necessity. As far as I can see, no scholar has questioned the interpretation of the necessity operator, and so I will assume that one of the abovementioned notions is at work in the formalization of Fitch's paradox. Nevertheless, as we will see, the full strength of **KT5** is not needed in order to obtain the paradox.

16 This interesting variant and its fundamental problem has been suggested by an anonymous referee.

rivable in **SN1+KP**.

Proposition 7: $\vdash_{\mathbf{SN2+KP}} \varphi \rightarrow K\varphi$.

Let H, H_1, H_2, H_3, H_4, H_5 be as in proposition 5.

Then $H_1, H_2 \in$ **K1**; $H_3, H_4 \in$ **K2**; $H_5 \in$ **K3**, and so $\Box H_1, \Box H_2, \Box H_3, \Box H_4, \Box H_5 \in$ **K□3**.

Proof: completely analogous to the proof of proposition 5.

Proposition 8: $\nvdash_{\mathbf{S1+KP}} \varphi \rightarrow K\varphi$.

Let $\mathcal{M} = \langle W, \mathcal{K}, \mathcal{E}, V \rangle$ be as in proposition 6. Then $\mathcal{M}$ satisfies all the conditions corresponding to the axioms of **SN1**, as can be checked without any difficulty, and

(1) $\varphi \rightarrow K\varphi$ is not valid in $\mathcal{M}$.

(2) $\varphi \rightarrow \neg\Box\neg K\varphi$ is valid in $\mathcal{M}$.

Proof: completely analogous to the proof of proposition 6.

The advantage of the present strategy is threefold: (1) the use of a dubious principle like **K□1** is completely avoided; (2) it is no more necessary to discard $K(K\varphi \rightarrow \varphi)$ in order to block the paradox; (3) the ontic operator is precisely, although minimally, characterized in both **SN1** and **SN2** as a factive operator which can be applied to any formula which is propositionally valid. It seems then to be possible to provide a simple solution to all the shortcomings previously highlighted. Unfortunately, even this new solution is unsuccessful. In this case, the problem is due to the fact that the principle that has to be rejected in order to conclude is

$$\Box(K\varphi \rightarrow \varphi)$$

which is the principle that characterizes knowledge as necessarily factive. However, the factivity of knowledge is hardly disputable, since it is the very property that distinguish knowledge from weaker epistemic conditions such as belief or justified belief, and is universally acknowledge as a non-negotiable premise.

References

[1] Alexander, S. (2013). An Axiomatic Version of Fitch's paradox. Synthese, 190: 2015-2020.

[2] Alexander, S. (2012). A purely epistemological version of Fitch's paradox. The Reasoner, 6: 59-60.

[3] Artemov, S. & Protopopescu, T. (2013). Discovering Knowability: a Semantic Analysis, Synthese, 190: 16, 3349-3376.

[4] Blackburn, P., de Rijke, M., Venema, Y. (2001). Modal Logic. Cambridge: CUP.

[5] Burgess, J. (2009). Can Truth Out? In [15], 147-162.

[6] Dean, W. and Kurokawa, H. (2010). From the Knowability Paradox to the Existence of Proofs, Synthese 176: 177 225.

[7] Dummett, M. (2009). Fitch's Paradox of Knowability. In [15], 51-52.

[8] Edgington, D. (1985). The Paradox of Knowability, Mind, 94: 557-568.

[9] Edgington, D. (2010). Possible Knowledge of Unknown Truth, Synthese, 173: 41-52.

[10] Fagin, R., Halpern, Y., Moses, Y., Vardi, M. (1995). Reasoning about Knowledge, Cambridge: MIT Press.

[11] Hand, M. (2010). Antirealism and Universal Knowability, Synthese, 173: 25-39.

[12] Kvanvig, J.L. (2006). The Knowability Paradox. Oxford: OUP.

[13] Maffezioli, P., Naibo, A. and Negri, S. (2013). The Church-Fitch knowability paradox in the light of structural proof theory, Synthese, 190, pp. 2677-2716.

[14] Rabinowicz, W. and Segerberg, K. (1994). Actual Truth, Possible Knowledge, Topoi, 13: 101-115.

[15] Salerno, J. (2009). New Essays on the Knowability paradox. Oxford: OUP.

[16] Salerno, J. (2010). Introduction to Knowability and Beyond. Synthese, 173, 1-8.

[17] Tennant, N. (1997). The Taming of the True, Oxford: OUP, Chapter 8.

[18] Tennant, N. (2009). Revamping the Restriction Strategy, in [15], 223-238.

[19] Tennant, N. (2010). Williamson's Woes, Synthese, 173: 9-23.

[20] Williamson, T. (1982). Intuitionism Disproved?, Analysis, 42: 203-207.

[21] Williamson, T. (1987). On the Paradox of Knowability, Mind, 96: 256-61.

[22] Williamson, T. (1988). Knowability and Constructivism, Philosophical Quarterly, 38: 422-432.

[23] Williamson, T. (2000). Knowledge and its Limits. Oxford: OUP.

[24] Williamson, T., (2000). Tennant on Knowable Truth, Ratio, 13: 99-114.

[25] Williamson, T., (2009). Tennant's Troubles, in [15], 183-204.