CrossMark

Ignorance of ignorance

Kit Fine[1]

Abstract I discuss the question of when knowledge of higher order ignorance is possible and show in particular that, under quite plausible assumptions, knowledge of second order ignorance is impossible.

Keywords Knowledge · Ignorance · Indeterminacy · Contingency · Modal logic · Fitch's paradox

> *Not ignorance, but ignorance of ignorance, is the death of knowledge.*
> Alfred North Whitehead

> *But there are also unknown unknowns. These are things we don't know we don't know.*
> Donald Rumsfeld

It is good to know when one is ignorant. But sometimes one is ignorant of one's ignorance. One might think that, in that case, it would be good to know of one's second order ignorance. But it can be shown, under quite plausible assumptions, that it is impossible to know of one's second order ignorance. Second order ignorance is inevitably third-order ignorance. Likewise, third-order ignorance is inevitably fourth order ignorance and so on, indefinitely, through all the orders of ignorance. When one is second order ignorant one enters a black hole from which there is no epistemic escape.[1]

✉ Kit Fine
kit.fine@nyu.edu

[1] Department of Philosophy, New York University, New York, NY, USA

What goes for ignorance also goes for indeterminacy or being borderline. In a sorites series, there may be borderline cases. It is often supposed that between the borderline and the definite cases, there may be borderline borderline cases. But again, it can be shown that, under plausible assumptions, it can never be correct to assert that a given case is borderline borderline. Any second order borderline case is inevitably third-order. Likewise, any third-order borderline case is inevitably fourth order and so on, indefinitely, through all the orders of being borderline.

These result are surprising and, although they are implicit in the earlier literature on contingency, they have, as far as I know, never been made explicit or further developed.[2]

Let me give an informal proof of the result, leaving a formal treatment of this and related results to the appendix. We presuppose the modal system S4, in which whatever is necessarily the case is both the case and is necessarily necessarily the case. Thus under the epistemic interpretation, we presuppose that whatever is known to be the case is both in fact the case and is known to be known to be the case (positive introspection) and, under the indeterminacy interpretation, we presuppose that whatever is definitely the case is both in fact the case and is definitely definitely the case.

First, let us fix on terminology. Say that:

one is *ignorant that* p if one does not know that p ($\neg Kp$);

it is (*epistemically*) *possible* that p if one is ignorant that not-p ($Mp =_{df} \neg K\neg p$)

one is *ignorant of the fact that* p if p is the case and one is ignorant that p ($p \wedge \neg Kp$);

one is (*first-order*) *ignorant whether* p if one is both ignorant that p and ignorant that not-p ($Ip =_{df} \neg Kp \wedge \neg K\neg p$);

one is *Rumsfeld ignorant of* p if one is ignorant of the fact that one is ignorant whether p ($Ip \wedge \neg KIp$); and

one is *second-order ignorant whether* p if one is ignorant whether one is ignorant whether p (IIp).

Note that first-order ignorance whether p amounts to its being epistemically possible that p and epistemically possible that not-p ($Mp \wedge M\neg p$).

We now establish various claims (all within the context of the system S4):

(1) Second-order ignorance implies first-order ignorance.

For suppose that one is not first-order ignorant. Then either one knows p (Kp) or ones know not-p ($K\neg p$). Suppose the former (the other case is similar). Then since one knows that p, one knows that one knows that p (KKp) and hence one knows that one is not ignorant whether p (Lemma 3).

(2) Second-order ignorance implies Rumsfeld ignorance.

For second order ignorance implies first order ignorance by (1) and ignorance *whether* one is ignorant whether p implies ignorance *that* one is ignorant whether p.

Let us note that the converse also holds:

(2*) Rumsfeld ignorance implies second-order ignorance.

For suppose one is Rumsfeld ignorant. Then one does not know that one is first-order ignorant. But given that one *is* first-order ignorant, one also does not know that one is *not* first-order ignorant.

(3) One does not know that one is Rumsfeld ignorant.

For suppose one knows that one is Rumsfeld ignorant (K(Ip ∧ ¬KIp)). Then one knows, in particular, that one is first-order ignorant (KIp) and, it follows from the truth of what one knows ((K(Ip ∧ ¬KIp)), that one does not know that one is first-order ignorant (¬KIp). A contradiction.

From (2) and (3) it follows that:

(4) One does not know that one is second-order ignorant.

For if one knew one were second-order ignorant whether p, one would know that one was Rumsfeld ignorant of p by (2), which is impossible by (3).

(4) is a *theorem* of the modal system S4. We have therefore established the logical impossibility—at least relative to system S4—of knowing that one is second-order ignorant (Theorem 4).

Let us make a number of observations.

Why surprising?

We should perhaps first consider why this result is so surprising even though, from a technical point of view, it is relatively trivial. Suppose one does not know certain propositions. Assuming negative introspection (that one knows what one does not know), it follows that one knows that one does not know each of these propositions. As is well-known, negative introspection does not follow from positive introspection. But there would appear to be no good reason why positive introspection should *preclude* one's knowing that one does not know each of these proposition.

Indeed, suppose one were to list all of the propositions that one knows:

$$p_1, p_2, \cdots$$

This is something one should in principle be able to do, whilst knowing that one knows each of the propositions on the list. Given that the list is exhaustive, then even though it is not guaranteed that one would know that the list is exhaustive, there would appear to be nothing to prevent one from knowing that it is exhaustive, that there was nothing else left to know; and so if one failed to know any proposition, one would know that it was not known since one could see that it did not appear on the list.

Our result shows that this line of reasoning is mistaken. For if one is second order ignorant whether p, then the fact that one does not know that p and the fact that one does not know that not-p will not appear on the list. But one could not know that they do not appear on the list, since this would amount to knowledge of second order ignorance.

Thus the result shows that what might appear to be a plausible model for achieving knowledge of what one does not know is not something that will always work. The content of everything that one knows might prevent one from seeing that it is the content of everything that one knows.

Orders of ignorance

If one suffers from second-order ignorance then one does not know that one suffers from second-order ignorance and, of course, if one suffers from second-order ignorance then one does not know that one does *not* suffer from second-order ignorance, since it is impossible to know what is false. So second-order ignorance implies third-order ignorance (ignorance whether one is second-order ignorant). Likewise third-order ignorance implies fourth order ignorance and so on, indefinitely, through all the orders of ignorance.

Moreover, as we have seen, second-order ignorance implies first-order ignorance. Consequently, third-order ignorance implies second-order ignorance and so on, indefinitely, through all the orders of ignorance. Thus any form of higher order ignorance (beyond the first) will imply all orders of ignorance (including the first).

Although knowledge of second-order ignorance is impossible, knowledge of first-order ignorance certainly is possible and, indeed, would appear to be commonplace. I am ignorant of whether is it raining in London right now, for example, and I certainly know that I am ignorant since I know that I do not have the evidence required to settle the question.

Second-order ignorance would also appear to be possible, even though knowledge of second-order ignorance is not. Indeed, let us suppose that there is a counter-example to the characteristic axiom 5 of S5, i.e. a case in which one is ignorant that p ($\neg Kp$) but ignorant that one is ignorant that p ($\neg K\neg Kp$). Given that one is ignorant that one is ignorant that p ($\neg K\neg Kp$), it follows that one is ignorant that not-p ($\neg K\neg p$) and hence that one is ignorant whether p (Ip). But also one is ignorant that one is ignorant *that* p ($\neg K\neg Kp$) and hence ignorant that is one ignorant *whether* p ($\neg KIp$). (The converse also holds: any case of second-order ignorance will be a counter-example to axiom 5. For second-order ignorance implies Rumseldian ignorance—($\neg Kp \wedge \neg K\neg Kp$) $\vee$ ($\neg K\neg p \wedge \neg K\neg K\neg p$)—and either disjunct will then serve as a counter-example to axiom 5).

We therefore see, when we look at the series of ignorance claims:

(I1) one is ignorant (i.e. first-order ignorant) whether p
(I2) one is second-order ignorant whether p
(I3) one is third-order ignorant whether p
⋮

that there are only three possible distributions of truth-value (Theorem 6). *Either* all of the ignorance claims are false—one is not ignorant whether p and consequently one does not suffer from any form of higher order ignorance (since that would require that one be first-order ignorant); *or* the first claim is true and all the others are false—one knows one is ignorant whether p and consequently one does not suffer from any higher order of ignorance since that would imply that one suffers from second-order ignorance; *or* all of the ignorance claims are true—one is second-order ignorant and hence ignorant at every order. Thus either one is *not* ignorant or one is *knowingly* ignorant or one is *thoroughly* ignorant. Ignorance of ignorance, rather than ignorance, is indeed the death of knowledge, though perhaps not in the sense that Whitehead intended.

I have so far only looked at the modalities which can be formed from the ignorance operator: ignorance whether p; ignorance whether ignorance whether p; and so on.

One might also bring in the operators for knowledge and epistemic possibility, thereby allowing ourselves to talk of ignorance whether one knows that p, for example, or ignorance whether one knows that it is epistemically possible that p. But even if one allows the use of these other epistemic operators in the formation of the prefix to p, one's 'epistemic profile' will peter out at the third level: once told how things are under a threefold iteration of these operators, there will be nothing else of this sort to be told (Theorem 12).[3]

Fitch The present result is closely related to the famous result of Fitch (1963).[4] He pointed out that ignorance of the fact that q (q ∧¬Kq) could not itself be known. Although second order ignorance is not itself a form of Fitchean ignorance, it is equivalent (within S4) to Rumsfeld ignorance, which *is* a form of Fitchean ignorance, in which the base state q is constituted by ignorance whether p for some given state p.

The present case of Fitch-style unknowability is of interest for a number of different reasons. In the first place, there is perhaps special interest attaching to the state of second-order ignorance and to how, in particular, it may differ from first-order ignorance (with the former being unknowable and the latter not).

In the second place, second-order ignorance may be regarded as a purely mental state, not consisting in a connection between something non-mental (the base state q) and something mental (¬Kq). Indeed, it might even be thought that ignorance of ignorance could be an *internal* mental state in that it could be grounded in something completely internal to the mind (as given, let us say, by the lack of evidence under an internalist conception of evidence).

Fitchean states may, of course, be purely mental or internal in this sense. Indeed, Rumsfeld ignorance may well be a mental state of this sort. But second order ignorance has a quite special status. For it may be shown to be the *weakest* purely mental state that is incapable of being known. If a purely mental state concerning a given proposition p is incapable of being known then it will imply a state of second order ignorance concerning p (Theorem 10). In other words, the unknowability of a purely mental state is always attributable to second order ignorance; given that the purely mental state requires second order ignorance, we can always see its unknowability as arising from the unknowability of second order ignorance.

This is in marked contrast to the unknowability of states which may not be purely mental, as with one's not knowing p, when p itself is not a mental state . In this case, without the restriction to purely mental states, it may be shown that there is no weakest unknowable state and hence no common source for the unknowability (Theorem 11).

Other systems The proofs of the above results all presuppose the principles of the system S4; and the characteristic 4 axiom of S4 is, of course, somewhat controversial—both for knowledge and for definiteness. Suppose we drop the characteristic 4 axiom of S4 ($\Box A \supset \Box\Box A$), thereby obtaining the system T. Then it can no longer be shown that knowledge of second-order ignorance is impossible.

---

[3] I have talked in this section on 'orders' in an informal way. See Williamson (1999) for a more formal discussion of the notion.

[4] See Brogaard and Salerno (2013) for a survey of recent work on the topic.

Indeed, we can establish a strong result to the contrary. For consider the series of ignorance claims (I1), (I2), (I3), . . . above and any possible assignment of truth-values to each claim in the series. Thus it might be supposed that one is n-th order ignorant for each even n and not n-th order ignorant for each even n or that one is n-th order ignorant for prime n and not n-th order ignorant for composite n. It can then be shown that each such supposition is consistent within the system T, i.e. that there is no logical basis, within T, for ruling it out (Theorem 15).

A related result can also be established. Say that one *absolutely* knows the proposition p if one knows that p, knows that one knows that p, knows that one knows that p, and so on (Kp, KKp, KKKp, . . .). Of course, within S4 one absolutely knows whatever one knows but, within T, this need not be so. It can then be shown that each order of ignorance is absolutely knowable. In other words, for each order n, one may consistently suppose that one has absolute knowledge of one's n-th order ignorance (corollary 16). Thus, somewhat paradoxically, the failure to know certain kinds of things (in particular, what one knows) may enable one to know other kinds of things (such as second order ignorance) which one would otherwise be incapable of knowing.

Moving in the other direction, stronger systems than S4 may narrow the possibility of knowledgeable ignorance. We previously noted that S4 left open three consistent assignments of truth values to the series of ignorance claims (I1), (I2), (I3), . . .: no ignorance; knowledgeable first-order ignorance; and thorough-going ignorance. Within the system S5 (whose additional axiom is $A \supset \Box\Diamond A$), all ignorance will be known, thereby excluding the possibility of thorough-going ignorance. On the other hand in the system S4M (whose additional axiom is $\Diamond\Box A \vee \Diamond\Box\neg A$), all ignorance will be thorough-going, thereby excluding the possibility of knowledgeable ignorance.[5]

Finally, note that the negative result on second order ignorance does not depend upon having axiom T. The result can also be establishing using axiom $D(= \Diamond\top)$ in place of T (observation to Theorem 4). This means that there are also plausible tense-logical, doxastic and proof-theoretic analogues of the result, where contingency corresponds to A sometimes but not always being the case under a tense logical reading, to suspension of belief under a doxastic reading, and to undecidability under a proof-theoretic reading.

General ignorance The ignorance we have so far dealt with concerns *specific* ignorance, i.e. ignorance concerning a specific proposition. But one may also be ignorant about some subject matter. One may be ignorant, for example, about which candidates will be elected to the senate.

We might think of the subject matter as given by a set of propositions $p_1, p_2, \ldots, p_n$. Thus in the case of the senatorial contest, the propositions might be that John McCain is elected, that Elizabeth Warren is elected, and so on.

Ignorance about a given subject matter can then take two forms. It can be *complete* ignorance or *partial*. In the case of complete ignorance, one is ignorant whether p for each of the propositions constituting the subject matter and, in the case of partial

---

[5] The system S4M, in which all continency is contingent, features prominently in Suzanne Bobzien's work on vagueness, as in Bobzien (2010) for example.

ignorance, one is ignorant whether p for at least one of the propositions constituting the subject matter.

One might think that one could escape the barrier to knowledgeable second order ignorance when the ignorance in question concerns some subject matter rather than a specific proposition. But this is not so. It is impossible to know of one's ignorance of one's ignorance concerning some subject matter or to know of one's second order ignorance concerning some subject-matter. And these results hold whether the ignorance be complete or partial (Theorem 7).

Indeterminacy

The notion of being borderline is the analogue of the notion of contingency in the modal sphere and of the notion of ignorance in the epistemic sphere. Where D is used for the definitely operator (it is definitely the case that), we may define it being borderline that A (BA) in terms of its being neither definitely the case that A nor definitely the case that not-A ($\neg DA \wedge \neg D\neg A$), in analogy to the definition of contingency in terms of necessity and to the definition of ignorance in terms of knowledge.

The above results are relevant to the phenomenon of vagueness, when characterized in terms of the property of being borderline, in a number of different ways (indeed, it was my interest in vagueness which originally motivated my interest in these results). In the first place, it is often supposed that just as one may be in a position to assert that a case is borderline without being in a position to assert that it is an actual case of the given predicate, so one may be in a position to assert that a case is borderline borderline without being in a position to assert that it is a borderline case of the predicate.

But this line of thought is misguided for two separate reasons. For even though a borderline case of a predicate need not be an actual case of a predicate, a borderline borderline case of a predicate must be a borderline case of the predicate (or, at least, within the context of S4). Moreover, one never is in a position to assert that a case is borderline borderline (again, in the context of S4) if one can only properly assert what is definitely true. For even though a case can be borderline borderline, it can never be definitely borderline borderline.

In the second place, it is often supposed that the familiar cases of vagueness are thoroughgoing in the sense of not simply requiring the presence of borderline cases but also the presence of borderline borderline cases. But if vagueness is thoroughgoing, one is never in a position to assert its existence. For its existence would require that one of the cases to which the predicate applies be borderline borderline; and, as we have seen, this is not something that can definitely be the case. This, in my opinion, is one reason among others for being skeptical as to whether one can even characterize vagueness in terms of the notion of borderline case (Fine 2008).

Formal appendix

We adopt the usual symbolism of propositional modal logic ▽ and standard results and terminology will be taken for granted. ▽A (it is contingent that A) is an abbreviation for ($\Diamond A \wedge \Diamond \neg A$). Under an epistemic interpretation of the modalities, □A means that ▽one knows that A, ◇A that it is possible for all one knows that A, and ▽A that one is ignorant whether or not A. Under the vagueness-theoretic interpretation, □A means that it is definitely the case that A, ◇A that it is indefinite that A, and ▽A that it is

indefinite whether or not A. We use numerical superscripts in an obvious way for iterations of a modality. Thus $\Box^2$ is short for $\Box\Box$ and $\triangledown^3$ for $\triangledown\triangledown\triangledown$.

We show how to simplify formulas of the form $\triangledown^2 A$ within the modal systems T and S4. But first:

Lemma 1 *For* $n = 1, 2, \ldots,$

(i) $\Box(\Box A \vee \Box \neg A)$ *is provably equivalent in T to* $\Box^2 A \vee \Box^2 \neg A$; *and*
(ii) $\Diamond(\Diamond A \wedge \Diamond \neg A)$ *is provably equivalent in T to* $\Diamond^2 A \wedge \Diamond^2 \neg A$.

Proof (ii) follows from (i) given the duality of $\Box$ and $\Diamond$.

As for the right-to-left direction of (i), $\Box A \vdash \Box A \vee \Box \neg A$ and so $\Box^2 A = \Box\Box A \vdash \Box(\Box A \vee \Box \neg A)$. Similarly, $\Box^2 \neg A \vdash \Box(\Box A \vee \Box \neg A)$; and so, putting together these two results, $\Box^2 A \vee \Box^2 \neg A \vdash \Box(\Box A \vee \Box \neg A)$.

In proving the left-to-right direction of (i), it will be convenient to employ a natural deduction style of reasoning. Suppose $\Box(\Box A \vee \Box \neg A)$; and make the temporary supposition of A. Since $\Box A \vee \Box \neg A$ (given $\Box(\Box A \vee \Box \neg A)$) and since A is inconsistent with $\Box \neg A$, we may infer $\Box A$. But then from $\Box(\Box A \vee \Box \neg A)$ and $\Box A$, we may infer $\Box(A \wedge (\Box A \vee \Box \neg A))$; and since A is inconsistent with $\Box \neg A$, we may infer $\Box\Box A$. Similarly, under the supposition of $\neg A$, we may infer $\Box^2 \neg A$; and so from the main supposition, we may infer $\Box^2 A \vee \Box^2 \neg A$, as required.   □

We now have:

Lemma 2 (i) $\triangledown^2 A$ *is provably equivalent in T to* $(\Diamond^2 A \wedge \Diamond\Box\neg A) \vee (\Diamond^2 \neg A \wedge \Diamond\Box A)$
(ii) $\neg\triangledown^2 A$ *is provably equivalent in T to* $\Box^2 A \vee \Box^2 \neg A \vee \Box\triangledown A$
(iii) $\triangledown^2 A$ *is provably equivalent in S4 to* $(\Diamond A \wedge \Diamond\Box\neg A) \vee (\Diamond\neg A \wedge \Diamond\Box A)$
(iv) $\neg\triangledown^2 A$ *is provably equivalent in S4 to* $\Box A \vee \Box\neg A \vee \Box\triangledown A$

Proof (i) $\triangledown^2 A$
-||- $\Diamond\triangledown A \wedge \Diamond\neg\triangledown A$ by definition of $\triangledown$
-||- $\Diamond(\Diamond A \wedge \Diamond\neg A) \wedge \Diamond\neg(\Diamond A \wedge \Diamond\neg A)$ by definition of $\triangledown$
-||- $\Diamond^2 A \wedge \Diamond^2 \neg A \wedge \Diamond\neg(\Diamond A \wedge \Diamond\neg A)$ by Lemma 1(ii)
-||- $\Diamond^2 A \wedge \Diamond^2 \neg A \wedge \Diamond(\Box\neg A \vee \Box A)$ by simplification in T
-||- $\Diamond^2 A \wedge \Diamond^2 \neg A \wedge (\Diamond\Box\neg A \vee \Diamond\Box A)$ by simplification in T
-||- $(\Diamond^2 A \wedge \Diamond^2 \neg A \wedge \Diamond\Box\neg A) \vee (\Diamond^2 A \wedge \Diamond^2 \neg A \wedge \Diamond\Box A)$ by truth-functional (tf) logic
-||- $(\Diamond^2 A \wedge \Diamond\Box\neg A) \vee (\Diamond^2 \neg A \wedge \Diamond\Box A)$ since $\Diamond\Box\neg A$ implies $\Diamond\Diamond\neg A$ and $\Diamond\Box A$ implies $\Diamond\Diamond\neg A$.
(ii) $\neg\triangledown^2 A$
-||- $\neg[(\Diamond^2 A \wedge \Diamond\Box\neg A) \vee (\Diamond^2 \neg A \wedge \Diamond\Box A)]$ by (i)
-||- $(\neg\Diamond^2 A \vee \neg\Diamond\Box\neg A) \wedge (\neg\Diamond^2 \neg A \vee \neg\Diamond\Box A)$ by tf logic
-||- $(\Box^2 \neg A \vee \Box\Diamond A) \wedge (\Box^2 A \vee \Box\Diamond\neg A)$ by simplification in T
-||- $(\Box^2 \neg A \wedge \Box^2 A) \vee (\Box^2 \neg A \wedge \Box\Diamond\neg A) \vee (\Box\Diamond A \wedge \Box^2 A) \vee (\Box\Diamond A \wedge \Box\Diamond\neg A)$ by tf logic
-||- $\Box^2 A \vee \Box^2 \neg A \vee \Box\triangledown A$

since

(a) $(\Box^2 \neg A \wedge \Box^2 A)$ -||- $\Box^2(\neg A \wedge A)$ -||- $\bot$,
(b) $(\Box^2 \neg A \wedge \Box \Diamond \neg A)$ -||- $\Box^2 \neg A$ given that $\Box^2 \neg A \vdash \Box \Diamond \neg A$,
(c) $(\Box^2 A \wedge \Box \Diamond A)$ -||- $\Box^2 A$ similarly, and
(d) $(\Box \Diamond A \wedge \Box \Diamond \neg A)$ -||- $\Box(\Diamond A \wedge \Diamond \neg A)$ -||- $\Box \nabla A$.
(iii) & (iv). From (i) and (ii) by the properties of S4.   □

Lemma 3 $\nabla^2 A$ *provably implies* ∇A *in S4.*

Proof $\nabla^2 A$ -||- $(\Diamond A \wedge \Diamond \Box \neg A) \vee (\Diamond \neg A \wedge \Diamond \Box A)$ by Lemma 2(i). But $\Diamond \Box \neg A \vdash \Diamond \Diamond \neg A \vdash \Diamond \neg A$; and so $(\Diamond A \wedge \Diamond \Box \neg A) \vdash \Diamond A \wedge \Diamond \neg A = \nabla A$. Similarly, $(\Diamond \neg A \wedge \Diamond \Box A) \vdash \nabla A$; and so $\nabla^2 A \vdash \nabla A$.   □

Say that the formula A is *boxable in* the system S if $\Box A$ is consistent in S and that otherwise A is *unboxable*. (Any inconsistent formula is trivially unboxable and so our interest is only in the consistent unboxable formulas).

Theorem 4 *The formula* $\nabla^2 A$ *is unboxable in S4.*

Proof By Lemma 3, $\Box \nabla^2 A \vdash \Box \nabla A$. Also:

$\Box \nabla^2 A$
$\vdash \nabla^2 A$
$\vdash (\Diamond A \wedge \Diamond \Box \neg A) \vee (\Diamond \neg A \wedge \Diamond \Box A)$ by Lemma 2(i)
$\vdash \Diamond \Box \neg A \vee \Diamond \Box A$.

So $\Box \nabla^2 A \vdash (\Diamond \Box \neg A \wedge \Box \nabla A) \vee (\Diamond \Box A \wedge \Box \nabla A)$. But $(\Diamond \Box \neg A \wedge \Box \nabla A) \vdash \Diamond(\Box \neg A \wedge \nabla A) \vdash \bot$ and $(\Diamond \Box A \wedge \Box \nabla A) \vdash \Diamond(\Box A \wedge \nabla A) \vdash \bot$; and so $\Box \nabla^2 A \vdash \bot$.   □

Let us observe that it can also be shown that $\nabla^2 A$ is unboxable in K4[D], where K4[D] is the non-normal extension of K4 with the axiom D $= \Diamond \top$. So this result includes certain unimodal tense logics, certain logics of belief, and the logic of provability (with the understanding that its theorems should be true but not necessarily provable).[6]

Corollary 5 *For* $m, n > 1$, $\nabla^m A$ -||- $\nabla^n A$ *in S4.*

Proof $\nabla^2 A \vdash \nabla A$ by Lemma 3; and so, substituting ∇A for A, $\nabla^3 A \vdash \nabla^2 A$. By Theorem 4, $\vdash \neg \Box \nabla^2 A \vdash \Diamond \neg \nabla^2 A$; and so $\nabla^2 A \vdash \nabla^2 A \wedge \Diamond \neg \nabla^2 A \vdash \nabla^3 A$. Hence $\nabla^2 A$ -||- $\nabla^3 A$. Appropriately substituting for A, $\nabla^k A$ -||- ∇^{k+1}, $k > 1$, from which the result follows.

By a *contingency profile* α is meant an infinite sequence $\alpha_1 \alpha_2 \alpha_3 \ldots$ of T's and F's. Corresponding to a contingency profile α is the set $\nabla_\alpha = \{A_1, A_2, A_3, \ldots\}$, where A_k is $\nabla^k p$ when α_k is T and A_k is $\neg \nabla^k p$ when α_k is F. A contingency profile α is said to be *consistent in* the system S if the corresponding set ∇_α is consistent in S.   □

---

[6] Thanks to Martin Pleitz for pointing me in the direction of this observation.

Theorem 6 *A contingency profile* α *is consistent in S4 iff it is of the form FFF . . . or TFF . . . or TTT . . ..*

Proof By Corollary 5, a consistent contingency profile must be of the form α_1 FF . . . or α_1 TT . . .. But it is readily shown that if α_1 = F then each α_k = F.

The consistency of FFF . . . is established by the model:

$$\circ \quad p$$

the consistency of TFF . . . by the model:

$$\circ \quad \leftrightarrow \quad \circ p$$

and the consistency of TTT . . . by the model:

$$\circ \quad \rightarrow \quad \circ p \qquad\qquad \square$$

We may establish some somewhat more general results:

Theorem 7 *For* $n > 0$*, the formulas* $\nabla(\nabla p_1 \wedge \nabla p_2 \wedge \ldots \wedge \nabla p_n)$*,* $\nabla(\nabla p_1 \vee \nabla p_2 \vee \ldots \vee \nabla p_n)$*,* $\nabla^2 p_1 \wedge \nabla^2 p_2 \wedge \ldots \wedge \nabla^2 p_n$ *and* $\nabla^2 p_1 \vee \nabla^2 p_2 \vee \ldots \vee \nabla^2 p_n$ *are unboxable in S4.*

Proof Let us deal with the formula $\nabla(\nabla p_1 \wedge \nabla p_2 \wedge \ldots \wedge \nabla p_n)$ (the proof for the second formula is similar).

$\nabla(\nabla p_1 \wedge \nabla p_2 \wedge \ldots \wedge \nabla p_n)$
$\vdash \lozenge(\lozenge p_1 \wedge \lozenge\neg p_1 \wedge \lozenge p_2 \wedge \lozenge\neg p_2 \wedge \ldots \wedge \lozenge p_n \wedge \lozenge\neg p_n)$ by def. of ∇ and the system T
$\vdash \lozenge\lozenge p_1 \wedge \lozenge\lozenge\neg p_1 \wedge \lozenge\lozenge p_2 \wedge \lozenge\lozenge\neg p_2 \wedge \ldots \wedge \lozenge\lozenge p_n \wedge \lozenge\lozenge\neg p_n$ by T
$\vdash \lozenge p_1 \wedge \lozenge\neg p_1 \wedge \lozenge p_2 \wedge \lozenge\neg p_2 \wedge \ldots \wedge \lozenge p_n \wedge \lozenge\neg p_n)$ by S4.

Hence:
(*)$\square\nabla(\nabla p_1 \wedge \nabla p_2 \wedge \ldots \wedge \nabla p_n) \vdash \square[\lozenge p_1 \wedge \lozenge\neg p_1 \wedge \lozenge p_2 \wedge \lozenge\neg p_2 \wedge \ldots \wedge \lozenge p_n \wedge \lozenge\neg p_n]$.
Also:

$\square\nabla(\nabla p_1 \wedge \nabla p_2 \wedge \ldots \wedge \nabla p_n)$
$\vdash \nabla(\nabla p_1 \wedge \nabla p_2 \wedge \ldots \wedge \nabla p_n)$
$\vdash \lozenge\neg(\nabla p_1 \wedge \nabla p_2 \wedge \ldots \wedge \nabla p_n)$
$\vdash \lozenge(\neg\nabla p_1 \vee \neg\nabla p_2 \vee \ldots \vee \neg\nabla p_n)$
$\vdash \lozenge(\square p_1 \vee \square\neg p_1 \vee \square p_2 \vee \square\neg p_2 \vee \ldots \vee \square p_n \vee \square\neg p_n)$.

But then, given (*) above, it follows that:

$\square\nabla(\nabla p_1 \wedge \nabla p_2 \wedge \ldots \wedge \nabla p_n)$
$\vdash \lozenge[(\square p_1 \vee \square\neg p_1 \vee \square p_2 \vee \square\neg p_2 \vee \ldots \vee \square p_n \vee \square\neg p_n) \wedge (\lozenge p_1 \wedge \lozenge\neg p_1 \wedge \lozenge p_2 \wedge \lozenge\neg p_2 \wedge \ldots \wedge \lozenge p_n \wedge \lozenge\neg p_n)]$,

which is readily shown to be contradictory.

The proof for the formula $\nabla^2 p_1 \wedge \nabla^2 p_2 \wedge \ldots \wedge \nabla^2 p_n$ is straightforward since if $\nabla^2 p_1 \wedge \nabla^2 p_2 \wedge \ldots \wedge \nabla^2 p_n$ were boxable then $\nabla^2 p_1$ would also be boxable.

The final case is the formula $\nabla^2 p_1 \vee \nabla^2 p_2 \vee \ldots \vee \nabla^2 p_n$. We prove $\square(\nabla^2 p_1 \vee \nabla^2 p_2 \vee \ldots \vee \nabla^2 p_n)$ inconsistent in S4 by induction on n. The case $n = 1$ is Theorem 4.

Consider the case n = k +1. Suppose, for reduction, $\Box(\nabla^2 p_1 \vee \nabla^2 p_2 \vee \ldots \vee \nabla^2 p_n)$. Then $\nabla^2 p_1 \vee \nabla^2 p_2 \vee \ldots \vee \nabla^2 p_n$ and so $(\Diamond p_1 \wedge \Diamond\Box\neg p_1) \vee (\Diamond\neg p_1 \wedge \Diamond\Box p_1) \vee (\Diamond p_2 \wedge \Diamond\Box\neg p_2) \vee (\Diamond\neg p_2 \wedge \Diamond\Box p_2) \ldots \vee (\Diamond p_n \wedge \Diamond\Box\neg p_n) \vee (\Diamond\neg p_n \wedge \Diamond\Box p_n)$. Without loss of generality, suppose $(\Diamond p_1 \wedge \Diamond\Box\neg p_1)$ and so, in particular, $\Diamond\Box\neg p_1$. Given $\Box(\nabla^2 p_1 \vee \nabla^2 p_2 \vee \ldots \vee \nabla^2 p_n)$, $\Diamond(\Box\Box\neg p_1 \wedge \Box(\nabla^2 p_1 \vee \nabla^2 p_2 \vee \ldots \vee \nabla^2 p_n))$ and so $\Diamond\Box(\Box\neg p_1 \wedge (\nabla^2 p_1 \vee \nabla^2 p_2 \vee \ldots \vee \nabla^2 p_n))$. But $\Box\neg p_1$ is inconsistent with $\nabla^2 p_1$ and so $\Diamond\Box(\nabla^2 p_2 \vee \nabla^2 p_3 \vee \ldots \vee \nabla^2 p_n))$, which is inconsistent by the inductive hypothesis. $\square$

A p-*formula* is any formula whose sole sentence letter is p. The formula A is *modalized* if every sentence letter of A occurs within the scope of a model operator.

Lemma 8 *Any p-formula $\Box B$, for B non-modal, is equivalent in T to: $\top$, $\bot$, $\Box p$, or $\Box\neg p$.*

Proof Any non-modal p-formula B is equivalent to $\top$, $\bot$, p, or $\neg$p; and from this it is readily shown that any p-formula $\Box B$, for B non-modal, is equivalent to $\top$, $\bot$, $\Box p$, or $\Box\neg p$. $\square$

Say that the formula A is *complete for* a class of formulas Δ *in* the modal system S if for any formula B of Δ either $A \vdash_S B$ or $A \vdash_S \neg B$.

Lemma 9 *In S4, each of the formulas $\Box p$, $\Box\neg p$ and $\Box\nabla A$ is complete for the class of modalized p-formulas.*

Proof We first show by inspection that each of the formulas $\Box p$, $\Box\neg p$ and $\Box\nabla p$ is complete for the class of formulas $\{\top, \bot, \Box p, \Box\neg p\}$ and hence, by the previous lemma, is complete for the class of all p-formulas of the form $\Box B$, for B non-modal.

It is easy to show that completeness is preserved under negation and conjunction, i.e. that if A is complete for Δ then it is complete for $\{\neg C : C \in \Delta\}$ and for $\{C \wedge D : C, D \in \Delta\}$. It therefore suffices to show that completeness is preserved under necessitation, i.e. that if A is complete for Δ then it is complete for $\{\Box C : C \in \Delta\}$. But it is readily shown that completeness is preserved under necessitation for any formula of the form $\Box B$. For suppose $\Box B$ is complete for the class of formulas Δ. Take now any formula of the form $\Box C$ for $C \in \Delta$. Then either $\Box B \vdash C$ or $\Box B \vdash \neg C$. In the former case, $\Box\Box B \vdash \Box C$ and so $\Box B \vdash \Box C$ given that, in S4, $\Box B \vdash \Box\Box B$. In the latter case, $\Box\Box B \vdash \Box\neg C$ and so $\Box B \vdash \Box\Box B \vdash \Box\neg C \vdash \neg\Box C$. $\square$

It can be shown that, in contrast to the case for S4, there is no consistent p-formula of T that is complete for the class of modalized p-formulas.

Theorem 10 *In the system S4, the modalized p-formula A is unboxable iff it implies $\nabla\nabla p$.*

Proof The right to left direction follows from Theorem 4. For the other direction, suppose that A does not imply $\nabla\nabla p$. Then A is consistent with $\neg\nabla\nabla p$ and hence, by Lemma 2(iv), consistent with $\Box p \vee \Box\neg p \vee \Box\nabla p$. But then A is consistent with $\Box p$ or $\Box\neg p$ or $\Box\nabla p$. So by Lemma 9, one of these three formulas of the form $\Box B$ implies A. But then $\Box B$ implies $\Box A$ and, since $\Box B$ is consistent then so is $\Box A$. $\square$

This result does not hold when the requirement that the formula A be modalized is dropped. For consider the Fitch formula $p \wedge \Diamond\neg p$. As is well known, this is unboxable

(in T and not just S4). However, it does not imply $\triangledown\triangledown p$ (even in S4) since it is compatible with $\Box\triangledown p$.

Indeed, it may be shown that:

Theorem 11 *In S4, there is no weakest unboxable p-formula, i.e. no unboxable p-formula implied by all unboxable formulas.*

Proof $p \wedge \Diamond\neg p$ and $\neg p \wedge \Diamond p$ are both unboxable. Now if there were a weakest unboxable formula A, it would be implied by both $p \wedge \Diamond\neg p$ and $\neg p \wedge \Diamond p$ and hence implied by their disjunction $(p \wedge \Diamond\neg p) \vee (\neg p \wedge \Diamond p)$, which is equivalent in T to $\triangledown p$. But $\Box\triangledown p$ is consistent in S4 and hence so is $\Box A$ and A would be boxable after all. $\square$

It should be noted that any unboxable formula A (in T) will imply a Fitch formula of the form $B \wedge \Diamond\neg B$ for, since $\Diamond\neg A$ is a theorem, A will imply and, indeed, be equivalent to $A \wedge \Diamond\neg A$. However, this is not a case in which we can attribute the unboxability of A to that of $A \wedge \Diamond\neg A$, since the unboxability of A is already presupposed. I do not know if there is a meaningful result in which we are able to attribute the unboxability of p-formulas in T to a common source.

An (*affirmative*) *modality* is a (possibly empty) string of $\Box$'s and $\Diamond$'s; and an *extended* (*affirmative*) modality is a (possibly empty) string of $\Box$'s, $\Diamond$'s and $\triangledown$'s. Two modalities φ and ψ are equivalent in a given system if φp is provably equivalent to ψp in the system; and similarly, for when one modality implies another.

It is well known that are are exactly 7 non–equivalent modalities in S4. Using $\wedge$ for the empty string, any modality is equivalent in S4 to one of: $\wedge$, $\Box$, $\Diamond$, $\Diamond\Box$, $\Box\Diamond$, $\Diamond\Box\Diamond$ and $\Box\Diamond\Box$. Using the easy direction of Theorem 10 and some of our earlier results, we can extend this result to the extended modalities.

Theorem 12 *Any extended modality is equivalent in S4 either to an unextended modality or to one of the following:* $\triangledown, \Box\triangledown, \triangledown\Diamond, \triangledown\Box, \triangledown\triangledown, \Diamond\Box\triangledown, \Box\triangledown\triangledown, \triangledown\Box\triangledown, \triangledown\Box\Diamond, \triangledown\Box\triangledown.$

Proof (*Sketch*) Show that the result of prefixing $\Box$, $\Diamond$ or $\triangledown$ to any modality on the list is equivalent to a modality on the list. We do not go into details but let us note that Theorem 10 is useful in that if we can show that a modality of the form $\triangledown\ldots$ implies $\triangledown\triangledown$ then it follows that $\Diamond\triangledown\ldots$ is equivalent to $\triangledown\ldots$ (since $\Diamond\triangledown$ is equivalent to $\triangledown$), that $\Box\triangledown\ldots$ is equivalent to $\Box\triangledown\triangledown$, which is equivalent to $\bot$, and that $\overline{\triangledown}\triangledown\ldots$ is equivalent to $\triangledown\ldots$. $\square$

Thus, as with unextended modalities, any extended modality will be equivalent to one of length at most 3. It may be verified that all of the modalities listed under the corollary are non-equivalent. I do not know if there are any normal modal systems that have only finitely many non-equivalent modalities but infinitely many non-equivalent extended modalities.

We turn finally to the possibilities for contingency within T and other logics. But first we review a familiar result on modal degree. The *modal degree* md(A) *of* a formula A is the maximum number of nested occurrences of $\Box$ within the formula. It may be recursively defined by:

md(p) = 0
md($\neg$A) = md(A)

md$((A \lor B)) = \max(\mathrm{md}(A), \mathrm{md}(B))$

md$(\square A) = \mathrm{md}(A) + 1$.

Given a binary relation R on a set W, we take $wR^n v$ to hold, for $w, v \in W$ and $n \geq 0$, if there is a sequence of elements $w_0, w_1, \ldots, w_n$ for $w_0 = w$, $w_n = v$ and $w_k R w_{k+1}$ for $k = 0, 1, \ldots, n - 1$; and given a Kripke model $\boldsymbol{M} = (W, R, \varphi, w)$, we let $\boldsymbol{M}^n$, for $n \geq 0$, be the model (W^n, R^n, φ^n, w), where $W^n = \{v \in W : wR^k v$ for $k \leq n\}$ and R^n and φ^n are the restriction of R and φ to W^n.

Lemma 13 *Given any formula A of modal degree $\leq n$ and any model $\boldsymbol{M} = (W, R, \varphi, w)$:*

A is true at $\boldsymbol{M}$ iff A is true at $\boldsymbol{M}^n$.

Proof By a straightforward induction on n.　　　　　　　　　　　　□

Use of the lemma will often be implicit.

By a *strip model of order* n, $n \geq 0$, is meant a Kripke-model of the form $\boldsymbol{M} = (W, R, \varphi, 0)$, where:

(i) $W = \{0, 1, 2, \ldots, n\}$;

(ii) $R = \{< k, k + 1 >: k = 0, 1, 2, \ldots, n\text{-}1\} \cup \{< k, k >: k \in W\}$

Note that the valuation φ is left undetermined; and, since our only interest is in the single sentence letter p, φ will be determined by an assignment of truth-values to each of the worlds 0, 1, 2, …, n. We may therefore represent φ by a sequence $p_0 p_1 \ldots p_n$ of length $n + 1$, where p_k is p should p be true at k and p_k is should p be false at k. Thus a strip model will look like so:

$$p_0 \quad p_1 \quad p_2 \cdots \qquad p_n$$
$$0 \rightarrow 1 \rightarrow 2 \ldots \rightarrow n.$$

We say that the strip model $\boldsymbol{M} = (W^+, R^+, \varphi^+, 0\}$ of order n+1 *extends* the strip model $\boldsymbol{M}^+ = (W, R, \varphi, 0)$ of order n if φ is the restriction of φ^+ to W, i.e. if the truth-values of p at the worlds 0, 1, 2, …, n are the same in each model. Let us use $\pm A$ ambiguously for A or $\neg A$.　　　　　　　　　　　　□

Lemma 14 *For each strip model $\boldsymbol{M} = (W, R, \varphi, 0)$ of order n and each formula $\pm \nabla^{n+1} p$, $n \geq 0$, there is an extension $\boldsymbol{M}^+ = (W^+, R^+, \varphi^+, 0)$ of $\boldsymbol{M} = (W, R, \varphi, 0)$ at which $\pm \nabla^{n+1} p$ is true.*

Proof By induction on n. For each n, there are two cases to consider: when $\pm \nabla^{n+1} p = \nabla^{n+1} p$ and when $\pm \nabla^{n+1} p = \neg \nabla^{n+1} p$.

<u>n = 0</u> The given model $\boldsymbol{M} = (W, R, \varphi, 0)$ is of the form $(\{0\}, R, \varphi, 0)$ and the associated sequence is either p or $\bar{p}$.

Suppose first that $\pm \nabla^1 p = \nabla p$. If the associated sequence is p, we may let the sequence associated with $\boldsymbol{M}^+$ be $p\bar{p}$; and if the associated sequence is, we may let the sequence associated with $\boldsymbol{M}^+$ be $\bar{p}p$.

Now suppose that $\pm \nabla^1 p = \neg \nabla p$. If the associated sequence is p, we may let the sequence associated with $\boldsymbol{M}^+$ be pp; and if the associated sequence is ($\bar{p}$), we may let the sequence associated with $\boldsymbol{M}^+$ be $\bar{p}\bar{p}$.

<u>n = k + 1</u> The given model $M = (W, R, \varphi, 0)$ is of the form $(\{0, 1, \ldots, k+1\}, R, \varphi, 0)$.

Suppose first that $\pm\nabla^{k+1}p = \nabla^{k+1}p = \Diamond\nabla^{k}p \wedge \Diamond\neg\nabla^{k}p$. There are two subcases. (a) $\nabla^{k}p$ is true at M. Consider the model $N = (W, R, \varphi, 1)$. By IH, there is an extension N^+ of N^k (which no longer contains the point 0) at which $\neg\nabla^{k}p$ is true. Let M^+ be the corresponding extension of M (in which the point 0 is restored). Then $\nabla^{k}p$ will be true at 0 in M^+ and $\nabla^{k}p$ false at 1 in M^+; and so $\nabla^{k+1}p$ will be true at M^+. (b) $\nabla^{k}p$ is false at M. The proof is similar to that for (a) but with the role of $\nabla^{k}p$ and $\neg\nabla^{k}p$ reversed.

Now suppose $\pm\nabla^{k+1}p = \neg\nabla^{k+1}p$. Again, there are two subcases. (a) $\nabla^{k}p$ is true at M. By IH, there is an extension N^+ of N^k at which $\nabla^{k}p$ is true (with N as before). Let M^+ be the corresponding extension of M. Then $\nabla^{k}p$ will be true at 0 in M^+ and $\nabla^{k}p$ true at 1 in M^+; and so $\Box\nabla^{k}p$ and hence $\neg\nabla^{k+1}p$ will be true at M^+. (b) $\nabla^{k}p$ is false at M. By IH, there is an extension N^+ of N^k at which $\neg\nabla^{k}p$ is true. Let M^+ be the corresponding extension of M. Then $\neg\nabla^{k}p$ will be true at 0 in M^+ and also true at 1 in M^+; and so $\Box\nabla^{k}p$ and hence $\neg\nabla^{k+1}p$ will be true at M^+. □

Theorem 15 *Each set of formulas* $\{\pm\nabla^{1}p, \pm\nabla^{2}p, \pm\nabla^{3}p, \ldots\}$ *is consistent in the system T.*

Proof By Lemma 14, there will be a strip model M_1 (of order 1) at which $\pm\nabla^{1}p$, an extension M_2 of M_1 at which $\pm\nabla^{1}p$ and $\pm\nabla^{2}p$ are true; and so on. The union M of the models $M_1, M_2, \ldots$ will then be a model for $\{\pm\nabla^{1}p, \pm\nabla^{2}p, \pm\nabla^{3}p, \ldots\}$ since each of $M_1, M_2, \ldots$ are generated submodels of M and hence will agree with M. □

For each formula A, let $\Box^{\infty}A$ be the set of formulas $\{A, \Box A, \Box^{2}A, \Box^{3}A, \ldots\}$.

Corollary 16 *For each* $n = 1, 2, \ldots,$ $\Box^{\infty}\nabla^{n}p$ *is consistent in T.*

Proof By the theorem, $\{\nabla^{n}p, \neg\nabla^{n+1}p, \neg\nabla^{n+2}p, \ldots\}$ is consistent in T. But it is then readily shown that each of the formulas $\nabla^{n}p, \Box\nabla^{n}p, \Box\Box\nabla^{n}p, \ldots$ is a consequence of $\{\nabla^{n}p, \neg\nabla^{n+1}p, \neg\nabla^{n+2}p, \ldots\}$. □

The strip models (and their unions) satisfy the condition that wRv and wRu implies $w = v$ or $w = u$ or $u = v$. The various results will therefore also hold in the system obtained by adding the 'linearity' axiom $A \vee \Box(A \supset B) \vee \Box(A \supset \neg B)$ to T. The corollary can also be established for the system B.

References

Bobzien, S. (2010). Higher-order vagueness, radical unclarity and absolute agnosticism. *Philosophers' Imprint*, *10*(10), 1–30.

Brogaard, B., & Salerno, J. (2013) Fitch's paradox of unknowability. In *Stanford encyclopedia of philosophy*. https://plato.stanford.edu/entries/fitch-paradox/.

Cresswell, M. J. (1988). Necessity and contingency. *Studia Logica*, *47*, 145–149.

Fine, K. (2008). The impossibility of vagueness. *Philosophical Perspectives*, *22*(1), 111–136.

Fitch, F. (1963). A logical analysis of some value concepts. *The Journal of Symbolic Logic*, *28*, 135–142.

Humberstone, I. L. (1995). The logic of non-contingency. *Notre Dame Journal of Formal Logic*, *36*(2), 214–229.

Kuhn, T. S. (1995). Minimal non-contingency logics. *Notre Dame Journal of Formal Logic*, *36*, 230–234.

Montgomery, H., & Routley, R. (1966). Contingency and non-contingency bases for normal modal logics. *Logique et Analyse*, *9*, 318–328.

Montgomery, H., & Routley, R. (1968). Non-contingency axioms for S4 and S5. *Logique et Analyse*, *11*, 422–424.

Montgomery, H., & Routley, R. (1969). Modalities in a sequence of normal non-contingency modal systems. *Logique et Analyse*, *12*, 225–227.

Mortensen, C. (1976). A sequence of normal modal systems with non-contingency bases. *Logique et Analyse*, *19*, 341–344.

Van der Hoek, W., & Lomuscio, A. (2004). A logic for ignorance. *Electronic Notes in Computer Science*, *85*(2), 1–17.

Williamson, Timothy. (1999). On the structure of higher-order vagueness. *Mind*, *108*, 127–142.