To appear in Ejerhed & Lindström (eds.) *Action, Language and Cognition* (Proceedings of UmLLI-93, the Umeå Colloquium on Dynamic approaches in Logic, Language, and Information, Umeå 1993.)

SITUATIONS, TRUTH AND KNOWABILITY

— A Situation-Theoretic Analysis of a Paradox by Fitch*

Sten Lindström
Department of Philosophy and Philosophy of Science
Umeå University

Abstract

According to a non-realist conception, the notion of truth is epistemically constrained: the anti-realist accepts one version or another of the *Knowability Principle* ("Any true proposition is knowable"). There is, however, a well-known argument, first published by Frederic Fitch (1963), which seems to threaten the anti-realist position. Starting out from seemingly innocuous assumptions, Fitch claims to prove: *if there is some true proposition which nobody knows to be true, then there is a true proposition which nobody can know to be true.*

Fitch's argument is simple. Suppose that q is a true proposition that is not known to be true. Consider then the proposition (p): q and it is not known that q. This is obviously a true proposition. And it cannot be known. For suppose that p is known in some counterfactual situation s. Then, the following proposition is true in s: it is known that (q and it is not known that q). Since knowledge distributes over conjunction, it would then also be true in s that: it is known that q and it is known that it is not known that q. Since knowledge implies truth, it would then follow that in s: it is known that q and it is not known that q. That is, the proposition (p) cannot be known.

I shall argue that Fitch's argument is based on an equivocation: there is a subtle reference shift as we speak of *the* proposition: *q and it is not known that q*, with respect to various contexts. In each context of utterance c, a *sentence* of the form: q and it is not known that q, refers to a certain situation s_c, which is provided by the context c, and makes the following claim about s_c:

(1) q is true in s_c and it is not known of s_c that q.

As is clear from Fitch's argument, it cannot be known in s_c itself that this claim about s_c is true. That is, it is not possible to have:

(2) true in s_c: it is known of s_c that (q and it is not known of s_c that q).

However, it might very well be known in *another situation* s that the claim (1) is true in s_c. That is, it may be the case for some situation $s \neq s_c$:

(3) true in s: it is known of s_c that (q and it is not known of s_c that q).

That is, in order to avoid Fitch's paradox, it is necessary to keep track of the situation in which a certain state of knowledge obtains as well as the situation which that state is about.

Writing T(p, s) for the proposition p being true in the situation s and K(p, s) for it being known of the situation s that p is true in s, we interpret the Knowability Principle as follows:

(KP) $\forall p \forall s(T(p, s) \rightarrow T(\Diamond K(p, s), s))$.

As a special case, we get the schema:

(K@) $A\varphi \rightarrow A\Diamond K(\varphi, @)$,

where @ is an individual constant (rigidly) designating the actual world and $A\varphi$ ("Actually φ") is an abbreviation of $T(\varphi, @)$. (KP) should be compared to Fitch's "paradoxical" principle:

(F) $\forall p \forall s(T(p, s) \rightarrow T(\Diamond K(p), s))$.

I am also going to compare my analysis of Fitch's paradox with the approaches of Edgington (1985) and Rabinowicz & Segerberg (1994) that formalize the Knowability Principle as:

(E) $A\varphi \rightarrow \Diamond K(A\varphi)$.

1. *The Paradox of Knowability*

According to a realist conception, truth is a radically non-epistemic notion: there is no absurdity in the idea of a proposition being true although we may lack any capacity to verify it or to gather evidence in its favor. A proposition, that is graspable by the mind, may be true without being knowable. For the anti-realist in contrast, the notion of truth is epistemically constrained. Although there are many true propositions that in fact are never known, the anti-realist finds the idea of an, in principle, unknowable proposition unintelligible. That is, the anti-realist accepts one version or another of the following principle:

The Knowability Principle: Any true proposition is knowable.

We leave it open, for the time being, whether the anti-realist also accepts the converse of this principle, namely, that any knowable proposition is true.

Consider, for example, various verificationist or pragmatist definitions of truth:

p is true if and only if

— it is possible, in principle, to conclusively verify p *(Verificationism)*

— it would be justified to believe p at an *ideal limit* of inquiry *(Peirce)*

— it would be justified to believe p under ideal epistemic circumstances *(Putnam)*

— there exists a state of information S, which is in principle accessible, such that p is justified by S, and p would remain justified no matter how S might be enlarged upon or improved. *(Truth as superassertability, Crispin Wright)*

On all these conceptions, it is plausible to claim that the Knowability Principle holds: truth implies possibility in principle to know.

There is, however, a well-known argument, first published by Frederic Fitch (1963), which seems to threaten the anti-realist position. Starting out from seemingly innocuous assumptions about knowledge, truth and possibility, Fitch proves:

If there is some true proposition which nobody knows (or has known or will know) to be true, then there is a true proposition which nobody can know to be true.

The argument is quite simple. Suppose that q is a true proposition that is not known to be true. Consider then the proposition:

(1) q and it is not known that q.

This is obviously a true proposition. And it cannot be known. For suppose that in some counterfactual situation:

(2) it is known that: q and it is not known that q.

Since knowledge distributes over conjunction, it would then also be true that:

(3) it is known that q

(4) it is known that it is not known that q.

But since knowledge implies truth, it would follow from (4) that:

(5) it is not known that q,

contrary to (3). The imagined situation is therefore impossible. That is, the proposition (1) cannot be known.

Suppose now that there are true propositions that are not known. It then follows from Fitch's argument that the Knowability Principle in the form:

Knowability Principle:

For any proposition p, if p is true, then it is possible in principle to know that p.

is false. Hence, if Fitch's argument is correct, we could infer from the Knowability Principle the absurd consequence that all true propositions are actually known. Since, obviously all known propositions are true, we could also infer that truth and knowledge are coextensive notions. If we took the Knowability Principle to be a necessary truth, we would be forced to conclude that truth and knowledge are necessarily coextensive, i.e., that they have the same intension.

Some philosophers, among them W. D. Hart (1979), have viewed Fitch's argument as a conclusive refutation of anti-realism concerning truth. Others view the argument as a valid refutation of the Knowability Principle, but think that there are forms of anti-realism or verificationism that are unaffected by the argument. In our view, it is quite implausible that it should be possible to establish Fitch's strong conclusion:

For each agent who is not omniscient, there is a true proposition which the agent cannot know.

by purely logical means. We would, therefore, rather like to find some fault with the reasoning leading up to it.

In the following we are going to reformulate Fitch's paradoxical argument within a situation-theoretic framework in which propositions can take the truth-values True or False relative to (possible) situations. Due to the partial nature of situations, a proposition p may also lack a truth-value relative to a situation s. Or to be more precise, neither p nor its negation ¬p may be true relative to s. We notice that admitting truth-value gaps is one way of blocking Fitch's argument. This is not our solution to the paradox, however, since counterintuitive consequences are still forthcoming. But the groundwork for a solution is laid.

According to our preferred analysis, the argument is based on an equivocation: there is a subtle reference shift as we speak of *the proposition:* q and it is not known that q, with respect to various contexts. In each context of utterance c, a *sentence* of the form:

q and it is not known that q,

refers to a certain situation s_c, which is provided by the context c, and makes the following claim about s_c:

(1) q and it is not known of s_c that q.

As we shall see, it cannot be known in s_c itself that this claim about s_c is true. That is, it is not possible to have:

(2) true in s_c: it is known of s_c that: (q and it is not known of s_c that q).

However, it might very well be known in *another situation* s that the claim (1) is true in s_c. That is, it may be the case for some situation $s \neq s_c$:

(3) true in s: it is known of s_c that: (q and it is not known of s_c that q).

That is, in order to avoid Fitch's paradox, it is necessary to keep track of the situation in which a certain state of knowledge obtains as well as the situation referred to by that state.

Our analysis has certain points in common with that of Dorothy Edgington (1985). According to Edgington, one needs to distinguish between the situation x in which a certain proposition is true and the situation y in which it is known about x that p is true in it.[1] To capture this idea she proposes the following formalization of the Knowability Principle:

(E-Knowability) $A\varphi \rightarrow \Diamond KA\varphi$,

where A is the actuality operator. In our view, however, Edgington's formalization is inadequate. It fails to distinguish between the following two principles:

(1) if actually φ, then it is possibly known *(de dicto)* that actually φ.

(2) if actually φ, then it is possible that it is known *(de re)* of the actual situation that φ is true in it.

Edgington's formalization is adequate to capture (1). But it fails to capture (2), which in our opinion is the more reasonable interpretation of the Knowability Principle. To capture (2), we would instead need something like:

$$A\varphi \rightarrow \Diamond K(\varphi, @),$$

where '@' is a constant that rigidly refers to the actual situation and K(φ, @) means that it is known of @ that φ. By the actual situation we here simply mean a designated situation supplied by context. These ideas, having been indicated here, will be developed in detail below.

2. *Knowability and Partiality*

Let us now take a closer look at Fitch's argument, in order to see which assumptions are involved. The argument is usually formulated as a formal derivation in a propositional logic with additional intensional operators ◇ and K for possibility and knowledge, respectively. Instead of this "logical" format, we shall choose a more mathematical one for presenting the argument. Our goal is to bring out in the open such assumptions that are hidden in the standard presentation of the argument.

We consider a framework with the following ingredients:

(i) A non-empty set Sit of *(possible) situations.*

(ii) A non-empty set Prop of *propositions*.

(iii) Three primitive operations on Prop, two unary ones, $\neg$ *(negation)* and K (the *knowledge operation*), and a binary one, $\wedge$ *(conjunction)*.

(iv) A binary relation $\vDash$ between situations and propositions. We read $x \vDash p$ as: the proposition p is *true in* the situation x, or alternatively as: x *supports* p.

(v) A binary relation R between possible situations. Intuitively, xRy means that the situation y is *possible* relative to the situation x.

We assume that the operations $\neg$ and $\wedge$ satisfy the following weak semantic principles:

(C) Not: $x \vDash p \wedge \neg p$ *(Law of Contradiction)*

($\wedge$) $x \vDash p \wedge q$ iff $x \vDash p$ and $x \vDash q$.

And, for the knowledge operation, we assume:

(S) If $x \vDash K(p)$, then $x \vDash p$ *(Success)*

(D) if $x \vDash K(p \wedge q)$, then $x \vDash K(p)$ and $x \vDash K(q)$. *(Distributivity)*

We are now ready to prove:

Lemma 2.1. There is no situation x and proposition p such that $K(p \wedge \neg K(p))$ is true in x.

Proof: Suppose that

(1)	$x \vDash K(p \wedge \neg K(p))$	
(2)	$x \vDash K(p)$	from (1) by Distributivity
(3)	$x \vDash K(\neg K(p))$	from (1) by Distributivity
(4)	$x \vDash \neg K(p)$	from (3) using Success
(5)	$x \vDash K(p) \wedge \neg K(p)$	from (2) and (4) using ($\wedge$).

However, (5) contradicts (LC). □

Consider now:

The Knowability Principle:

if $x \vDash p$, then $\exists y[xRy$ and $y \vDash K(p)]$.

That is, if p is true in x, then there is a situation y that is possible relative to x such that p is known in y.

If we add to the above framework a propositional operation $\Diamond$ with the truth-clause:

$$x \vDash \Diamond p \text{ iff } \exists y[xRy \text{ and } y \vDash p],$$

then we may rewrite the Knowability Principle in the following more familiar form:

$$\text{if } x \vDash p, \text{ then } x \vDash \Diamond K(p).$$

We now prove:

Theorem 2.2. The Knowability Principle implies that there is no situation x and proposition p such that:

$$x \vDash p \wedge \neg K(p).$$

Proof: Suppose that:

(1) $x \vDash p \wedge \neg K(p)$.

The Knowability Principle, then yields that for some situation y such that xRy,

(2) $y \vDash K(p \wedge \neg K(p))$.

But this contradicts Lemma 1. □

Let us say that a situation x (*logically*) *complete*, i.e., if it satisfies:

$$\forall p[x \vDash p \text{ or } x \vDash \neg p].$$

We say that the agent is *omniscient* in x, if

$$\forall p[\text{if } x \vDash p, \text{ then } x \vDash K(p)].$$

Theorem 2.3. For any complete situation x, if x satisfies the Knowability Principle, i.e., if

$$\forall p[\text{if } x \vDash p, \text{ then } x \vDash \Diamond K(p)],$$

then the agent is omniscient in x.

Proof: Suppose that x is a complete situation satisfying the Knowability Principle. Suppose also that the agent is not omniscient in x. Then for some proposition p,

$$x \vDash p \text{ and not } x \vDash K(p).$$

Since x is complete, not $x \vDash K(p)$ implies that $x \vDash \neg K(p)$. Hence, there is a proposition p such that $x \vDash p \wedge \neg K(p)$. But this is contrary to Theorem 2. □

From Theorem 3, we immediately get:

Corollary 2.4. If in a complete situation x the agent is not omniscient, then there exists a proposition p such that: $x \vDash p \wedge \neg\Diamond K(p)$.

Theorem 3 and Corollary 4 can be proved only for situations that are complete — not for partial situations. Let x be a partial situation and suppose that the agent is not omniscient in x. Then there is a proposition p such that $x \models p$ and not: $x \models K(p)$. However, in the absence of completeness, we cannot infer: $x \models \neg K(p)$. Hence, we don't get: $x \models p \wedge \neg K(p)$ either. So the Knowability Principle cannot be applied to the proposition $p \wedge \neg K(p)$. It does not follow that there exists a true proposition which the agent cannot know.

Now, if we assume that 'truth' means 'truth in the actual world' and that the actual world is a designated complete situation, we get "Fitch's Theorem":[2]

> For each agent who is not omniscient, there is a true proposition which the agent cannot know.

Fitch's Theorem can be avoided by denying the existence of a complete actual situation. In a given context, the actual situation is the — real or imagined — situation that we refer to. In general, there are good reasons not to suppose this situation to be complete. Fitch's theorem might be taken as one such reason.

Considering Theorem 3 and Corollary 4, we may even want to deny the existence of complete possible situations altogether. That is, one way of avoiding Fitch's Knowability Paradox is to deny the existence of complete situations: possible situations are always partial.

However, it seems that partiality does not solve the problem altogether. Remember Theorem 2:

> The Knowability Principle implies that there is no situation x and no proposition p such that $p \wedge \neg K(p)$ is true in x.

But a situation x in which $p \wedge \neg K(p)$ is true seems clearly possible. Consider, for instance, the situation in my room right now. Suppose it is true that I have an even number of books in my bookcase. It is also true in this situation that I know that the number of books is either odd or even. I also know that there are fewer than two hundred books. Since I haven't counted the books, it is also true that I don't know that the number of books is even. Hence, it is true in this situation: that the number of books is even and I don't know that it is. So it seems, after all, that we should have to abandon the Knowability Principle. Before doing so, however, we shall consider another possible reaction to Fitch's Paradox.

3. *Diagnosis*

In order to get what we think is a satisfactory solution to Fitch's paradox, we shall modify the framework. So far, we have treated knowledge as an operation on propositions. In such a framework, we can express knowledge claims of the following kind:

$$x \models K(p), \qquad x \models \neg K(p).$$

That is, we can say that a proposition is known (or not known) in a given situation. However, we cannot express the claim that it is known in one situation x, *about* another situation y, that p holds in y. We might want to say, for instance:

> Today, John knows that there was an even number of books in his bookcase, yesterday, and that he didn't know it then.

Here, we have two situations, today (T) and yesterday (Y), and a proposition p such that:

(1) true today: it is known of yesterday that: (q and it is not known of yesterday that q).

Writing:

$x \vDash K(p, y)$,

for:

true in situation x: it is known of the situation y that p holds in y,

we can express (1) as:

(2) $T \vDash K(q \wedge \neg K(q, Y), Y)$.

Since knowledge distributes over conjunction, we may infer from (2):

(3) $T \vDash K(q, Y)$

and

(4) $T \vDash K(\neg K(q, Y), Y)$.

But, knowledge implies truth, so we have:

(5) if $T\vDash K(\neg K(q, Y), Y)$, then $Y \vDash \neg K(q, Y)$.

Hence, by modus ponens:

(6) $Y \vDash \neg K(q, Y)$.

From (3) and (6) we get:

(8) $T \vDash K(q, Y)$ and $Y \vDash \neg K(q, Y)$.

But since T and Y are *different situations*, we don't get a contradiction. That is, in order to avoid Fitch's paradox, it is necessary to keep track both of the situation where a given knowledge state obtains and the situation which that state is about.

The idea behind this solution is the following. Rather than being an intrinsic property of a proposition, truth is a relative notion: one and the same proposition may be true relative to one situation and fail to be true relative to another. When we say that a proposition is true, simpliciter, we don't mean that it is true relative to one situation or another but that it is true relative to a particular situation which is provided by context. The concept of truth is an indexical notion.[3]

However, knowledge involves truth. Knowledge is always knowledge of a proposition being true in a situation. Therefore, the proper logical form of a knowledge claim is not K(p), but rather K(p, x). That is, it is known of the situation x that p holds in it. We may also express this as: it is known true of x that p. It follows that knowledge should not be represented as a unary operation on propositions, but rather as an function that takes a situation and a proposition as arguments and yields a proposition as value. For a given knowledge claim, we have to keep track of both the situation *in* which the claim is supposed to be true and the situation that it is about. The notation:

$$x \vDash K(p, y)$$

accomplishes just that. It says that in the situation x it is known of the situation y that p holds in it. In many contexts, x and y are the same situation. But this is far from always the case.

Hence, we modify our framework by letting knowledge be represented by a function K that takes propositions and situations as arguments and yields propositions as values. We assume that K satisfies the following conditions:

(S)	If $x \vDash K(p, y)$, then $y \vDash p$	*(Success)*
(D)	if $x \vDash K(p \wedge q, y)$, then $x \vDash K(p, y)$ and $x \vDash K(q, y)$.	*(Distributivity)*

The *Knowability Principle* now takes the following form:

$$\text{if } x \vDash p, \text{ then } \exists y[xRy \text{ and } y \vDash K(p, x)].$$

That is, if p is true in the situation x, then there is a possible situation y which is possible relative to x such that in y it is known true of x that p.

Expressed in terms of ◇, the Knowability Principle becomes:

$$\text{if } x \vDash p, \text{ then } x \vDash \Diamond K(p, x),$$

i.e., if p is true in the situation x, then it is true in x that possibly it is known true of x that p.

The latter formulation of the principle may be preferred to the former for the reason that it avoids explicit quantification over a totality that contains not only actually existing situations but also merely possible ones.

Let us see what happens to Fitch's paradox after this modification of our framework. Instead of Lemma 1, we now have:

Lemma 3.1. There is no situation x and proposition p such that:

$$x \vDash K(p \wedge \neg K(p, x), x).$$

Proof: The proof is exactly like the one for Lemma 1.

However, the Knowability Principle does not exclude situations x such that:

(1) $x \models p \wedge \neg K(p, x)$.

From (1), the Knowability Principle implies that there is a situation y such that xRy and

(2) $y \models K(p \wedge \neg K(p, x), x)$.

As we have seen, from this we can only infer:

(3) $y \models K(p, x)$ and $x \models \neg K(p, x)$,

which would yield a contradiction only if x and y were the same situation.

Now, what was it that went wrong in the derivation of Fitch's Paradox? We started with the assumption:

(1) q and it is not known that q.

Here, we are obviously speaking of some situation x that is supplied by the context. That is, the claim made is:

(2) $x \models q \wedge \neg K(q, x)$.

From (1) we inferred, using the Knowability Principle that for some counterfactual situation y:

(3) it is known that: q and it is not known that q.

However, the natural interpretation of (3) is:

(3') it is known of y that: q and it is not known of y that q,

rather than:

(3'') it is known of x that: q and it is not known of x that q.

Hence, by not keeping track of the situation referred to in (1), we infer from it that there is a possible situation y satisfying:

(4) $y \models K(q \wedge \neg K(q, y), y)$,

rather than:

(5) $y \models K(q \wedge \neg K(q, x), x)$.

(4) but not (5) is impossible. However, the Knowability Principle — once it is properly formulated — yields (5) but not (4). According to our diagnosis then, Fitch's Paradox is based on an equivocation.

The following argument against the new version of the Knowability Principle is based on suggestions made by Wlodek Rabinowicz. It is natural to think that the collection Sit of all situations comes equipped with a part-whole relation $\sqsubseteq$ that satisfies the requirements of a partial ordering. Intuitively, $x \sqsubseteq y$ means that x is a *part of* y. Possible worlds are presumably maximally inclusive situations. Hence, we define:

$$x \text{ is a possible world iff } \forall y(x \sqsubseteq y \rightarrow x = y).$$

Now, the following principle seems plausible:

(R) if $x \vDash K(p, y)$, then $\exists z(x \sqsubseteq z$ and $y \sqsubseteq z)$.

In particular, this principle yields:

(i) if w is a possible world and $w \vDash K(p, x)$, then $x \sqsubseteq w$.

(ii) if w is a possible world and $x \vDash K(p, w)$, then $x \sqsubseteq w$.

(iii) if w, w' are possible worlds, then not: $w \vDash K(p, w')$.

Suppose now that p is a proposition that is true in the actual world w_0. According to the Knowability Principle, there is some possible situation x such that:

$$x \vDash K(p, w_0).$$

The principle (R), then yields that $x \sqsubseteq w_0$. We have proved that every actually true proposition is known in some subsituation of the actual world. However, this conclusion is absurd. Hence, if we accept (R), we must give up the Knowability Principle — or so it seems. There is, however, another alternative: we may instead give up the assumption that there are maximally inclusive situations, or possible worlds.

Now, how is it possible in one situation to have knowledge of another situation. Let me just hint at an answer. One way of obtaining knowledge is by perception. The agent, in one situation x (the perceptual situation), may gain knowledge of another situation y (the *scene* observed) simply by observing that situation. By observing a situation y, I might even learn that a proposition p holds in y and that it is not known in y that p. For example, by observing a card game I may learn that one of the players, John say, has the best hand and that none of the participants of the game knows this fact. And even if there is no actual outside observer of the card game, it is still possible that such an observer could have existed. This is so, as long as we assume that all situations are partial and possible to extend.

4. A Logico-Semantical Framework

Dorothy Edgington (1985) has argued that a solution to Fitch's Paradox can be given in terms of a modal actuality operator A. She considers a propositional language with the usual connectives, modal operators $\Box$ and $\Diamond$, a knowledge operator K and the actuality operator. According to her proposal, the Knowability Principle should be formulated as:

(E) $\varphi \rightarrow \Diamond KA\varphi$.

How does this proposal compare to our own? The answer to this question is not immediately obvious, since Edgington does not provide her object language with an explicit semantics. In order to remedy this situation, we shall present a formal language $\mathcal{L}$ which is an extension of Edgington's modal language and provide $\mathcal{L}$ with a model-theoretic semantics. Within this semantic framework, we shall then compare our resolution of Fitch's paradox to that of Edgington.

We are also going to compare our treatment of the paradox to one that has been recently proposed by Rabinowicz and Segerberg (1994).

The syntax of $\mathcal{L}$. The formal language $\mathcal{L}$ has *symbols* of the following kinds:

(i) A denumerable sequence $s_1, s_2,\ldots$ of situation variables.

(ii) Two situation constants @ and #.

(iii) A denumerable sequence P_1, P_2, of proposition variables.

(iv) The sentential connectives $\neg, \wedge, \vee, \rightarrow$

(v) The modal operators $\Box$ and $\Diamond$.

(vi) The epistemic operator K.

(vii) The truth-operator T.

(viii) Punctuation symbols: parentheses and the comma sign.

The *well-formed expressions* of $\mathcal{L}$ are of two kinds: situation terms and formulas.

A *situation term* is either a situation variable or a situation constant. We use the letters σ, τ, with or without indices, as syntactic variables for situation terms.

The set of *formulas* of $\mathcal{L}$ is defined as the smallest set satisfying the following clauses:

(i) Proposition variables are formulas.

(ii) If φ and ψ are formulas, then $\neg\varphi$, $(\varphi \wedge \psi)$, $(\varphi \vee \psi)$, $(\varphi \rightarrow \psi)$, $\Box\varphi$ and $\Diamond\varphi$ are also formulas.

(iii) If σ is a situation term and φ is a formula, then $K(\varphi, \sigma)$ and $T(\varphi, \sigma)$ are formulas. The intended readings of $K(\varphi, \sigma)$ and $T(\varphi, \sigma)$ are: *it is known of* σ *that* φ; and: *it is true in* σ *that* φ, respectively.

The semantics for $\mathcal{L}$. We are now going to provide the language $\mathcal{L}$ with a model-theoretic semantics. A semantic interpretation should assign to every well-formed expression of $\mathcal{L}$ an appropriate semantic value, which we, using David Kaplan's (1989) terminology, refer to as its *content.* The content of a situation term is a function that assigns a *denotation* to the term relative to each situation. The content of a formula is the proposition that it expresses.

However, some expressions — the indexical ones — are assigned a content only relative to a context of utterance. It follows that any well-formed expression that contains indexicals, is assigned a content only relative to a context of utterance. A semantic interpretation assigns to such an expression not a determinate content but rather what Kaplan calls a *character*, i.e., a function from contexts to contents.

From an abstract point of view, a semantic evaluation for a language containing indexical expressions, can be viewed as a three-step process: first, each expression is assigned a character, i.e., a function from contexts to contents. *(semantic interpretation).* We use the notation: $[\![E]\!]$, for the character of an expression E. Next, the character together with the context of utterance c determines a content, $[\![E]\!]^c$ *(contextual interpretation)*. Finally, the content $[\![E]\!]^c$ determines the denotation (or extension) of E, taken in the context c, relative to various possible situations x. This last step we might call *evaluation.*

Situations have two distinct roles in the semantics. First, they play the role of *context situations,* i.e., situations relative to which the indexical expressions are interpreted. Secondly, situations play the role of *described situation*, i.e., situations relative to which propositions are evaluated as true or false. $\mathcal{L}$ contains only one primitive indexical expression, namely, the situation constant @ that functions as a (rigid) designator that always refers to the context situation. The content of @, relative to a context situation c, is the constant function $[\![@]\!]^c$ which for every situation x assigns to @ the situation c as its denotation $[\![@]\!]^c_x$. The constant @ should be compared to the non-indexical situation constant # which, independently of the context situation, refers to the described situation. The content of #, relative to a context situation c, is the function $[\![\#]\!]^c$ which for every situation x assigns to # the described situation x itself as its denotation $[\![\#]\!]^c_x$.

When developing the semantics for $\mathcal{L}$, there are several options to choose from. First, we need to decide on a formal representation of propositions. Should we think of propositions as structured sentence-like entities, or should we identify propositions that are true in the same situations? Here, we have opted for the latter, less fine-grained alternative, even if the former presumably is to be preferred in epistemic logic.

Hence, we represent propositions as sets of situations. This choice does not however automatically commit us to classical propositional logic. That is, we shall not assume in general that the set Prop of all propositions forms a Boolean algebra. Instead, we let Prop be closed under unions and intersections of arbitrary sets, i.e., the propositions form a complete set-lattice with Sit as its universal set.[4]

It is a well-known fact that in any complete lattice it is possible to define the operations → and ¬ of a Heyting algebra.[5] For any propositions p, q, p → q is defined as the logically weakest proposition which together with p implies q. And, ¬p is the defined as the logically weakest proposition which is incompatible with p. If we want to, we can make the logic classical by assuming that Prop is closed under set-theoretic complementation. ¬p will then be the complement of p.

The epistemic operator K is interpreted in terms of a ternary alternativeness relation E between situations. The intuitive idea behind the semantic clause for K(φ, σ) is the following:

> It is true in x: K(p, y) iff p is true in all situations z that are compatible with everything the agent knows in x about y.

Writing xE_yz for: z is compatible with everything the agent knows in x about y, we get:

> It is true in x: K(p, y) iff ∀z(if xE_yz, then p is true in z).

Now, after having explained the general ideas, it is time to get down to the technical details. By a (semantic) *interpretation* for $\mathcal{L}$ we understand a structure $\mathcal{I}$ = <Sit, Prop, R, E, V> such that:

(i) Sit is a non-empty set, the members of which are called *situations*.

(ii) Prop is a complete set-lattice with Sit as its universal set, i.e., it is a family of subsets of Sit satisfying the following conditions:

(a) Prop is closed under arbitrary unions and intersections; that is, for any set X ⊆ Prop, ∪X ∈ Prop and ∩X ∈ Prop;

(b) both ∅ and Sit are members of Prop.

The members of P are called the *propositions* of $\mathcal{I}$. We say that the proposition p is *true* in the situation x, if x ∈ p.

(iii) R is a binary relation in Sit. We assume that R is reflexive in Sit, i.e., for all x in Sit, xRx. Intuitively, xRy means that y is possible relative to x.

(iv) E is a ternary relation in Sit. Intuitively, yE_xz means that the situation z is compatible with everything that is known in the situation y about the situation x. If nothing is known in y about x, then yE_xz obtains for every z. We assume that for every y and x, yE_xx. That is, for any y, x is always compatible with what is known of x in y.

(v) If p ∈ Prop, then $\Box p =_{df}$ {x: ∀y(if xRy, then y ∈ p)} and $\Diamond p =_{df}$ {x: ∃y(xRy and y ∈ p)} are also members of Prop.

(vi) If p ∈ Prop, then $K(p, x) =_{df}$ {y: ∀z(if yE_xz, then z ∈ p)} belongs to Prop.

(vii) V is a valuation function that assigns an appropriate semantic value to every variable in $\mathcal{L}$. That is, if s is a situation variable, then V(s) ∈ Sit; and if P is a proposition variable, then V(P) ∈ Prop.

For any interpretation I, we define the notions of relative pseudo-complement → and pseudo-complement ¬. These are algebraic (or set-theoretic) counterparts of the notions of intuitionistic implication and intuitionistic negation, respectively.

For any p, q ∈ Prop, the *relative pseudo-complement of* p *with respect to* q is the proposition:

$$p \to q =_{df} \cup\{r: p \cap r \subseteq q\}.$$

Hence:

$$x \in p \to q \text{ iff } \exists r[x \in r \text{ and } p \cap r \subseteq q].$$

That is, p → q is true in a situation x if, and only if, there exists a proposition r such that (i) r is true in x; and (ii) p ∩ r entails q.

The operation → satisfies the conditions:

(i) $p \cap (p \to q) \subseteq q$

(ii) for any r, if $p \cap r \subseteq q$, then $r \subseteq p \to q$.

In other words, p → q is the weakest proposition which in conjunction with p entails q.

In terms of →, we define for each proposition p its pseudo-complement ¬p:

$$\neg p =_{df} (p \to \varnothing).$$

We also have, $\neg p = \cup\{q: p \cap q = \varnothing\}$. Hence, pseudo-complements satisfy the conditions:

(i) $p \cap \neg p = \varnothing$

(ii) for any q, if $p \cap q = \varnothing$, then $q \subseteq \neg p$.

That is, ¬p is the weakest proposition which is incompatible with p.

The algebra <Prop, ∪, ∩, →, ¬, ∅, Sit>, where ∪, ∩ are the binary set-theoretic operations of union and intersection, is a *Heyting set-algebra.* It is a *Boolean set-algebra,* if Prop is closed under set-theoretic complementation, in which case:

$$\neg p = (\text{Sit} - p),$$

i.e., for any situation x, $x \in \neg p$ iff $x \notin p$.

We say that I = <Sit, Prop, R, E, V> is a *classical* interpretation, if Prop is a closed under complements, i.e., if it is a complete set-algebra. . A *standard interpretation* is an interpretation, where Prop = $\wp$(Sit). That is, a standard interpretation is an interpretation in which every subset of Sit represents a proposition. Such interpretations can be written on the form I = <Sit, R, E, V>. Standard interpretations are, of course, classical.

Suppose that an interpretation I = <Sit, Prop, R, E, V> is given. We define for every situation term τ in $\mathcal{L}$, and every c and x in Sit, the *denotation of* τ, when considered in the context c, with respect to the described situation x (in symbols, $[\![\tau]\!]_x^{I,c}$):

(i) If τ is a situation variable, then $[\![\tau]\!]_x^{I,c} = V(\tau)$.

(ii) If τ is @, then $[\![\tau]\!]_x^{I,c} = c$.

(iii) If τ is #, then $[\![\tau]\!]_x^{I,c} = x$.

Next, we define, for each context c ∈ Sit and each formula φ of $\mathcal{L}$, the *content* $[\![\varphi]\!]^{I,c}$ of φ in the context c (relative to the interpretation I). Intuitively, $[\![\varphi]\!]^{I,c}$ is the proposition expressed by φ in the context situation c. It is easy to see that $[\![\varphi]\!]^{I,c}$ is always a member of Prop.

(i) If P is a proposition variable, then $[\![P]\!]^{I,c} = V(P)$.

(ii) $[\![\neg\varphi]\!]^{I,c} = \neg[\![\varphi]\!]^{I,c}$.

(iii) $[\![\varphi \wedge \psi]\!]^{I,c} = [\![\varphi]\!]^{I,c} \cap [\![\psi]\!]^{I,c}$.

(iv) $[\![\varphi \vee \psi]\!]^{I,c} = [\![\varphi]\!]^{I,c} \cup [\![\psi]\!]^{I,c}$.

(v) $[\![\varphi \rightarrow \psi]\!]^{I,c} = [\![\varphi]\!]^{I,c} \rightarrow [\![\psi]\!]^{I,c}$.

(vi) $[\![\Box\varphi]\!]^{I,c} = \Box([\![\varphi]\!]^{I,c}) = \{x: \forall y(\text{if } xRy, \text{ then } y \in [\![\varphi]\!]^{I,c})\}$.

(vii) $[\![\Diamond\varphi]\!]^{I,c} = \Diamond([\![\varphi]\!]^{I,c}) = \{x: \exists y(xRy \text{ and } y \in [\![\varphi]\!]^{I,c})\}$.

(viii) $[\![K(\varphi, \tau)]\!]^{I,c} = \{x: \forall y \forall z(\text{if } y = [\![\tau]\!]_x^{I,c} \text{ and } xE_yz, \text{ then } z \in [\![\varphi]\!]^{I,c})\}$.

(ix) $[\![T(\varphi, \tau)]\!]^{I,c} = \{x: \forall y(\text{if } y = [\![\tau]\!]_x^{I,c}, \text{ then } y \in [\![\varphi]\!]^{I,c})\}$.

We say that a formula φ, when considered in the context c (and relative to the interpretation I), is *true with respect to* the situation x (in symbols, $I \models_x^c \varphi$) if, and only if, x ∈ $[\![\varphi]\!]^{I,c}$. Hence, we have:[6]

(i) If P is a proposition variable, then $\models_x^c P$ iff $x \in V(P)$.

(ii) $\models_x^c \neg\varphi$ iff $\exists p(x \in p \text{ and } p \cap [\![\varphi]\!]^c = \varnothing)$.

(iii) $\models_x^c (\varphi \wedge \psi)$ iff $\models_x^c \varphi$ and $\models_x^c \psi$.

(iv) $\models_x^c (\varphi \vee \psi)$ iff $\models_x^c \varphi$ or $\models_x^c \psi$.

(v) $\models_x^c (\varphi \rightarrow \psi)$ iff $\exists p(x \in p \text{ and } p \cap [\![\varphi]\!]^c \subseteq [\![\psi]\!]^c)$.

(vi) $\models_x^c \Box\varphi$ iff $\forall y(\text{if } xRy, \text{ then } \models_y^c \varphi)$.

(vii) $\models_x^c \Diamond\varphi$ iff $\exists y(xRy \text{ and } \models_y^c \varphi)$.

(viii) $\models_x^c K(\varphi, \tau)$ iff $\forall y \forall z(\text{if } y = [\![\tau]\!]_x^{I,c} \text{ and } xE_yz, \text{ then } \models_z^c \varphi)$.

(ix) $\models_x^c T(\varphi, \tau)$ iff $\forall y(\text{if } y = [\![\tau]\!]_x^{I,c}, \text{ then } \models_y^c \varphi)$.

If the interpretation is classical, then we can simplify (ii) and (v) to:

(ii') $\models^{c}_{x} \neg\varphi$ iff not $\models^{c}_{x} \varphi$.

(v') $\models^{c}_{x} (\varphi \rightarrow \psi)$ iff not($\models^{c}_{x} \varphi$ and not $\models^{c}_{x} \psi$).

The formula φ, when interpreted in I, is *true in the context* c if and only if $\models^{c}_{c} \varphi$. In other words, φ is true in the context c if and only if $c \in [\![\varphi]\!]^{I,\,c}$, i.e., if and only if the proposition expressed by φ in c is true in the situation c itself.

By a *model* for $\mathcal{L}$, we understand an ordered pair $\mathcal{M} = \langle I, c\rangle$ of an interpretation I and a situation c of I. We say that a formula φ is *true in the model* $\mathcal{M}$ (in symbols, $\mathcal{M} \models \varphi$) if and only if $I \models^{c}_{c} \varphi$. That is, φ is true in the model $\mathcal{M} = \langle I, c\rangle$ if and only if φ, when interpreted in I, is true in the context c. We also say that $\mathcal{M}$ *satisfies* φ, if $\mathcal{M} \models \varphi$. The formula φ *entails* the formula ψ in the model $\mathcal{M} = \langle I, c\rangle$, iff $[\![\varphi]\!]^{I,\,c} \subseteq [\![\psi]\!]^{I,\,c}$. A set of formulas Γ *entails* a formula φ in the model $\mathcal{M} = \langle I, c\rangle$, iff $\cap\{[\![\psi]\!]^{I,\,c}: \psi \in \Gamma\} \subseteq [\![\varphi]\!]^{I,\,c}$.

Lemma 4.1. For all formulas φ, ψ, χ and every model $\mathcal{M} = \langle I, c\rangle$:

(i) $\varphi \wedge \psi$ entails χ in $\mathcal{M}$ iff φ entails $\psi \rightarrow \chi$ in $\mathcal{M}$.

(ii) φ entails χ in $\mathcal{M}$ iff $[\![\varphi \rightarrow \psi]\!]^{I,\,c} = \text{Sit}$.

Let K be a class of models. We say that a formula is (weakly) K-*valid*, if it is true in every model in K. A formula φ is a (weak) K-*consequence* of a set of formulas Γ (in symbols, $\Gamma \models_K \varphi$), iff for any model $\mathcal{M}$ in K, if all the formulas in Γ are true in $\mathcal{M}$, then φ is true in $\mathcal{M}$.

We say that φ is *strongly* K-*valid* if for every model $\mathcal{M} = \langle I, c\rangle$ in K, $[\![\varphi]\!]^{I,\,c} = \text{Sit}$. A formula φ is a strong K-*consequence* of a set of formulas Γ iff for any model $\mathcal{M}$ in K, Γ *entails* φ in the model $\mathcal{M}$.

We define *(logical) validity* and *(logical) consequence* as K-validity and K-consequence, where K is the set of all models. We speak of *classical validity [classical consequence]* and *standard validity [standard consequence]*, when K is the set of all classical models and all standard models, respectively. The corresponding notions with the prefix 'strong' are defined analogously. Of course, the strong notions imply the corresponding weak notions (for instance, if φ is strongly valid, then it is valid). We may also speak of K-validity, K-consequence etc., where K is a class of *interpretations.* This just means validity, consequence etc. relative to the class of all models $\mathcal{M} = \langle I, c\rangle$ in which $I \in$ K.

Lemma 4.2. Let K be any class of models for $\mathcal{L}$. Then, we can define the notion of (weak) K-validity in terms of strong K-validity:

(i) φ is weakly K-valid iff Aφ is strongly K-valid.

Moreover, the notion of (weak) K-consequence is definable in terms of strong K-consequence:

(ii) $\Gamma \models_K \varphi$ iff $A\varphi$ is a strong K-consequence of $\{A\psi: \psi \in \Gamma\}$.

Proof: φ is weakly K-valid iff for any model $\mathcal{M} = <I, c>$ in K, $\mathcal{M} \models^c_c \varphi$ iff for any model $\mathcal{M} = <I, c>$ in K and any situation x in I, $\mathcal{M} \models^c_x A\varphi$ iff φ is strongly K-valid. We omit the proof of (ii). □

We introduce the following metalinguistic abbreviations:

$K\varphi =_{df} K(\varphi, \#)$

$K^{@}\varphi =_{df} K(\varphi, @)$

$T\varphi =_{df} T(\varphi, \#)$

$A\varphi =_{df} T(\varphi, @)$.

It is easily verified that these defined operators have the following truth-clauses:

(1) $\models^c_x K\varphi$ iff $\forall y$(if xE_xy, then $\models^c_y \varphi$)

(2) $\models^c_x K^{@}\varphi$ iff $\forall y$(if xE_cy, then $\models^c_y \varphi$)

(3) $\models^c_x T\varphi$ iff $\models^c_x \varphi$

(4) $\models^c_x A\varphi$ iff $\models^c_c \varphi$.

Consider the operators K and $K^{@}$. Relative to a given context, $K\varphi$ is true in a situation x if and only if φ is true in all situations y that are compatible with what is known in x *about* x *itself*. $K^{@}\varphi$, on the other hand, is true in a situation x if and only if φ is true in all situations y that are compatible with what is known in x *about the context situation.* $K\varphi$ and $K^{@}\varphi$ may be read:

it is known that φ

it is known of the actual situation that φ,

respectively.

The difference between T and A is analogous to that between K and $K^{@}$. $T\varphi$ can be read simply as: it is true that φ. A is the actuality operator: relative to a given context situation c, $A\varphi$ is true in a situation x iff φ is true in the context situation c (the "actual" situation).

There is a subtle, but important, distinction between $K^{@}\varphi$ and $KA\varphi$. The informal readings of the two constructions are different:

(1) It is known of the actual situation c that φ is true in it.

(2) The proposition: Actually φ, is known to be true.

Intuitively, neither condition entails the other. Suppose that in the situation x it is known of the situation c that φ is true in it. From this it does not follow that the proposition Actually φ, needs to be known in x. The agent need not know in x that c is the actual situation. Conversely, suppose that it is known that φ is true in the actual situation, whichever situation that may be. That is, the agent knows (de dicto) that Actually φ. From this it does not follow that the agent knows (de re) of the situation that happens to be the actual one that φ holds in it. The distinction between $K^{@}\varphi$ and $KA\varphi$ is important when we, in the next section, discuss Edgington's proposal for a resolution of Fitch's paradox.

If we look to our formal semantics, we see that the $K^{@}\varphi$ and $KA\varphi$ are, indeed, associated with different truth-clauses:

(i) $\models^{c}_{x} K^{@}\varphi$ iff $\forall y$(if $xE_{c}y$, then $\models^{c}_{y} \varphi$)

(ii) $\models^{c}_{x} KA\varphi$ iff $\forall y$(if $xE_{x}y$, then $\models^{c}_{y} A\varphi$) iff $\forall y$(if $xE_{x}y$, then $\models^{c}_{c} \varphi$) iff $\models^{c}_{c} \varphi$,

However, the semantics for $\mathcal{L}$ has some unintuitive features.

(iii) for any model $\mathcal{M}$, if $\mathcal{M} \models \varphi$, then $\mathcal{M} \models \Box KA\varphi$ (and, $\mathcal{M} \models KA\varphi$).

(iv) for any classical model $\mathcal{M}$, $\mathcal{M} \models \varphi \rightarrow \Box KA\varphi$, i.e., every instance of the schema $\varphi \rightarrow \Box KA\varphi$ is classically valid. The same holds, of course, for: $\varphi \rightarrow KA\varphi$.

(v) The schemata $A\varphi \rightarrow \Box KA\varphi$ and $A\varphi \rightarrow KA\varphi$ are strongly valid.

(vi) The schema $K^{@}\varphi \rightarrow KA\varphi$ is strongly valid.

Proof. (iii) Consider any model $\mathcal{M} = \langle I, c\rangle$ and suppose that $\mathcal{M} \models \varphi$. This means that $\models^{c}_{c} \varphi$. Hence, by (ii) above, for any situation x, $\models^{c}_{x} KA\varphi$. It follows that, for any x such that cRx, $\models^{c}_{x} KA\varphi$. Hence, $\models^{c}_{c} \Box KA\varphi$. That is, $\mathcal{M} \models \Box KA\varphi$.

(iv) If the model is classical, then we have:

$$\models^{c}_{x} \varphi \rightarrow \psi \text{ iff (if } \models^{c}_{x} \varphi\text{, then } \models^{c}_{x} \varphi\text{).}$$

So, (iii) implies (iv).

(v) We prove that $A\varphi \rightarrow \Box KA\varphi$ is strongly valid. Suppose that for some model $\mathcal{M} = \langle I, c\rangle$ and some situation x in I, $\models^{c}_{x} A\varphi$. Then, $\models^{c}_{c} \varphi$. Hence, by (ii), for any y, $\models^{c}_{y} KA\varphi$. It follows that $\models^{c}_{x} \Box KA\varphi$. We have shown that $A\varphi$ entails $\Box KA\varphi$ in $\mathcal{M}$. By lemma 4.1, it follows that $\llbracket A\,\varphi \rightarrow \Box KA\varphi \rrbracket^{I,\, c} = \text{Sit}$. □

(vi) Consider any model $\mathcal{M} = \langle I, c\rangle$. Suppose that for any $x \in$ Sit, $\models^{c}_{x} K^{@}\varphi$. Then, $\forall y$(if $xE_{c}y$, then $\models^{c}_{y} \varphi$). Since $xE_{c}c$, it follows that $\models^{c}_{c} \varphi$. But, by (ii), this implies $\models^{c}_{x} KA\varphi$. Hence, $\llbracket K^{@}\varphi \rrbracket^{c} \subseteq \llbracket KA\varphi \rrbracket^{c}$. It follows that $\llbracket K^{@}\varphi \rightarrow KA\varphi \rrbracket^{c} = \text{Sit}$. Hence, $K^{@}\varphi \rightarrow KA\varphi$ is strongly valid. □

Now, Rabinowicz and Segerberg (1994), p. 7-8, has claimed that something must be seriously wrong with a semantic framework that yields results like (iii) - (v). In particular, the fact that, in any model, $A\varphi$ entails $KA\varphi$ is clearly unintuitive. And as we have already pointed out, neither does it seem right that $K^{@}\varphi$ should entail $KA\varphi$. Rabinowicz and Segerberg write:

> Surely, something is seriously wrong. The standard truth-conditions for the actuality operator and the epistemic operator do not mix: when we try to combine them, they yield absurdities.

What are the semantic assumptions about K that are responsible for the unintuitive consequences?[7] Let us make an intuitive distinction between the *content* $[\![\varphi]\!]^c$ of a sentence φ in a context c and its *intension* in the context in question. The content is the proposition expressed by the sentence, while the intension is the set $[\varphi]^c$ of all situations x such that $\models^c_x \varphi$. Consider now the following features of our semantics:

(1) K, like any other sentential operator in $\mathcal{L}$, is an *operator on content*: for every interpretation $\mathcal{I}$, there exists a function F such that for every context c in $\mathcal{I}$ and every formula φ:

$$[\![K\varphi]\!]^c = F([\![\varphi]\!]^c).$$

This means that the proposition that Kφ expresses, when it is considered in the context c, is functionally determined by the proposition that the embedded sentence φ expresses, relative to the same context c.

(2) *Content determines intension*, i.e., if two formulas φ and ψ have the same content with respect to a given context c (and interpretation $\mathcal{I}$), then they have the same *intension* with respect to c.

(3) *Intension determines content,* i.e., if two formulas φ and ψ have the same intension with respect to a given context c, then they also have the same content with respect to c.

Conditions (2) and (3) makes it possible to identify the proposition expressed by a sentence in a context with its intension relative to that context. However, in conjunction with (1) these conditions also imply that the operator K satisfies:

(4) For any class X of models, if $\varphi \leftrightarrow \psi$ is strongly X-valid, then $K\varphi \leftrightarrow K\psi$ is also strongly X-valid.

However, our semantics for the actuality operator A yields:

(5) For any class X of models that satisfy φ, $A\varphi \leftrightarrow T$ is strongly X-valid (where $T =_{df} \neg(P_1 \wedge \neg P_1)$.).

(4) and (5) yield:

(6) For any class X of models that satisfy φ, $KT \leftrightarrow KA\varphi$ is strongly X-valid.

From (6) we get:

(7) $A\varphi \rightarrow (KT \leftrightarrow KA\varphi)$ is strongly X-valid

(8) $\varphi, KT \models KA\varphi$.

If KT is strongly X-valid, we may conclude from (6) that $A\varphi \rightarrow KA\varphi$ is strongly X-valid. This in turn yields the rest of the conditions (iii) - (vi).

Since we don't want to abandon the principle that content determines intension, it seems that we should either have to live with (iii) - (vi) or give up one of the assumptions (1) and (3). However, abandoning the latter condition seems to be a more natural and less radical step than introducing operators that violate principle (1).

It appears to be nothing wrong with principles like:

$$A\varphi \rightarrow KA\varphi$$

and

$$K^{@}\varphi \rightarrow KA\varphi$$

as long as we give $K\varphi$ and $K^{@}\varphi$ the respective readings:

It is strictly implied by what is known that φ

It is strictly implied by what is known about the actual situation @ that φ.

Suppose, namely, that $A\varphi$, when considered in the context c, is true in the situation x. Then, $A\varphi$, when considered in c, is true in every situation y. Thus, the proposition $[\![A\varphi]\!]^c$ is universally true.[8] Then, of course, $[\![A\varphi]\!]^c$ is strictly implied by the set of all propositions that are known by the agent in the situation x. Hence, given the above reading of K, the sentence $KA\varphi$, when considered in the context c, is true in the situation x. We have shown that, for any situation x, if $[\![A\varphi]\!]^c$ is true in x, then $[\![KA\varphi]\!]^c$ is also true in x. Thus, $A\varphi \rightarrow KA\varphi$ is strictly valid.

Suppose now that $K^{@}\varphi$, when considered in the context c, is true in situation x. Then, $[\![\varphi]\!]^c$ must be true in the actual situation c. Hence, $[\![A\varphi]\!]^c$ is the universally true proposition. Again it follows that $KA\varphi$, when considered in the context c, is true in the situation x. Hence, $K^{@}\varphi \rightarrow KA\varphi$ is strictly valid.

However, if we give the formulas $K\varphi$ and $K^{@}\varphi$ their *intended readings*: "*It is known that* φ" and "*It is known of the actual world that* φ", then the above schemata are intuitively valid, only if we assume that the agent under consideration is *logically omniscient.* The reason why logical omniscience fails is presumably that the propositions that serves as objects of propositional attitudes are more finely individuated than Carnapian intensions. This indicates that it is the assumption (3) rather than (1) that should be abandoned.

5. *Fitch's Paradox Revisited*

Let us now see what happens to Fitch's paradox, when we formulate it with respect to the formal language $\mathcal{L}$. Let us understand by a *Fitch-model* any model $\mathcal{M}$ for $\mathcal{L}$ which satisfies the principle:

(Weak F-Knowability) If $\mathcal{M} \models \varphi$, then $\mathcal{M} \models \Diamond K\varphi$.

Observe that this principle is implied by but does not imply:

(F-Knowability) $\mathcal{M} \models \varphi \rightarrow \Diamond K\varphi$.

For classical Fitch-models, the two principles are, of course, equivalent.

Theorem 6.

(a) For any Fitch-model $\mathcal{M}$ and any formula φ:

$$\mathcal{M} \not\models \varphi \wedge \neg K\varphi.$$

(b) For classical Fitch-models, knowledge collapses into truth, i.e.,

If $\mathcal{M} \models \varphi$, then $\mathcal{M} \models K\varphi$.

Dorothy Edgington (1985) has proposed the following formalization of the Knowability Principle:

(E) $\varphi \rightarrow \Diamond KA\varphi$.

However, in $\mathcal{L}$ we have:

Lemma 7. Every model for $\mathcal{L}$ satisfies (E).

In order to discuss our own analysis of the paradox we introduce a second-order extension $\mathcal{L}^2$ of the language $\mathcal{L}$. We let $\mathcal{L}^2$ be the language obtained from $\mathcal{L}$ by adding universal and existential quantifiers that range over Sit and Prop. The semantics for $\mathcal{L}^2$ is obtained from that of $\mathcal{L}$ by adding semantic clauses for the quantifiers:

$\mathcal{I} \models^c_x \forall s\varphi$ [$\mathcal{I} \models^c_x \forall P\varphi$] iff for any interpretation $\mathcal{J}$ which is exactly like $\mathcal{I}$ except for the value that it assigns to the situation variable s [the proposition variable P], $\mathcal{J} \models^c_x \varphi$.

$\mathcal{I} \models^c_x \exists s\varphi$ [$\mathcal{I} \models^c_x \exists P\varphi$] iff for some interpretation $\mathcal{J}$ which is exactly like $\mathcal{I}$ except for the value that it assigns to the situation variable s [the proposition variable P], $\mathcal{J} \models^c_x \varphi$.

In $\mathcal{L}^2$ we formalize our version of the *Knowability Principle* as:

(KP) $\forall P \forall s[T(P, s) \rightarrow T(\Diamond K(P, s), s)]$.

According to (KP), if the proposition P is true in a situation s, then it possible in s to know of s that P is true there. By a *knowability model*, we understand a model that satisfies this principle.

(KP) does not logically imply Fitch's paradoxical principle:

(F) $\forall P \forall s(T(p, s) \rightarrow T(\Diamond K(p), s))$,

which written in primitive notation becomes:

$$\forall P \forall s(T(p, s) \rightarrow T(\Diamond K(p, \#), s)).$$

At first sight one might think that (KP) would imply:

(F#) $\forall P[T(P, \#) \rightarrow T(\Diamond K(P, \#), \#)]$,

which is equivalent to the following version of Fitch's principle:

$$\forall P[P \rightarrow \Diamond K(P)].$$

However, (F#) is not a logical consequence of (KP). The reason is that # is not a *rigid designator*: it denotes different situations relative to different described situations. In the context 'T(P, ...)', # functions as a name for the actual situation. But this is not true of the modal context '$\Diamond K(P, \#)$', in which # does not have a purely referential occurrence. We might say that the occurrence of # in '$\Diamond K(P, \#)$' is bound by the operator $\Diamond$. Hence, the following inference pattern is invalid:

from $\forall s\varphi(s)$, infer $\varphi(\#)$.

Instead, we only have:

from $\forall s\varphi(s)$, infer $\varphi(\sigma)$,

where σ is a situation term different from #.

Instead of *(F#)*,we have in every Knowability model:

$$\forall P[T(P, @) \rightarrow T(\Diamond K(P,@),@)],$$

which may be expressed as:

$$\forall P[AP \rightarrow A\Diamond K^{@}P].$$

Since, for any model $\mathcal{M}$:

$$\mathcal{M} \models \forall P[AP \leftrightarrow P],$$

we conclude that the following principle is true in all knowability models:

(K@) $\forall P[P \rightarrow \Diamond K^{@}P]$.

(K@) should be compared to Edgington's principle:

(E) $A\varphi \rightarrow \Diamond KA\varphi$,

that although it blocks the paradox, doesn't seem to capture the intended meaning of the knowability principle. In our semantics, Edgington's principle comes out as true in all models.

Finally, we point out that the converse of our version of the Knowability Principle,

$$\forall P \forall s(T(\Diamond K(P, s), s) \rightarrow T(P, s)),$$

is true in every model. And this, we think, is as it should be.

Notes

* The present paper has benefited a lot from inspiring and enjoyable discussions with Wlodek Rabinowicz. Earlier versions have been presented at *UmLLI-93* (*the Umeå Colloquium on Dynamic Approaches in Logic, Language and Information*), September 24-26, 1993 and at a seminar in theoretical philosophy at Uppsala University. I am grateful to the members of these audiences, especially to Peter Pagin, Dag Prawitz, Krister Segerberg, Sören Stenlund and Dag Westerståhl, for stimulating comments and discussions.

1 Cf. also Williamson (1987).

2 Fitch (1963), Theorem 4 on p. 138.

3 This is a theme that is developed in Barwise & Etchemendy (1987).

4 From an abstract point of view, the structure $\langle Sit, Prop, \subseteq, \in \rangle$ is what Barwise and Etchemendy (1990) calls an *infon algebra*, in which the elements of Prop play the role of pieces of information or *infons*. To be more specific, it is a complete and strongly balanced infon algebra. The strongly balanced infon algebras are just the ones that are isomorphic to set-algebras of the kind above. Barwise and Etchemendy make a sharp distinction between *infons*, that are made true or false by situations, and *propositions* that are true or false simpliciter. A proposition, then, has as its constituents, both an infon p and a situation x against which p is evaluated. On this conception, truth and falsity are intrinsic properties of propositions. In this paper, I adopt a more traditional view of propositions and think of them just as pieces of information, i.e., as Barwise and Etchemendy's infons.

5 For this purpose, we only need to assume that Prop forms a topology, i. e., that it contains $\varnothing$ and Sit and is closed under arbitrary unions and *finite* intersections. The assumption of closure under *arbitrary* intersections is only used in section 5, where we extend the language $\mathcal{L}$ to a language $\mathcal{L}^2$ with quantifiers ranging over propositions and over situations. The assumption that Prop is a complete set-algebra guarantees that, for every sentence of $\mathcal{L}^2$, there exists a proposition in Prop that it expresses.

6 From now on, we usually suppress the reference to the interpretation $\mathcal{I}$.

7 Here, we speak of the unary K-operator defined in the following way from the binary one: $K\varphi =_{df} K(\varphi, \#)$.

8 We do not assume in this argument that a proposition is identified with the set of situations in which it is true. Hence, we allow for there being many distinct propositions that are universally true.

References

Barwise, J. and Etchemendy, J. (1987) *The Liar — An Essay on Truth and Circularity,* Oxford University Press.

Barwise, J. and Etchemendy, J. (1990) "Information, Infons, and Inference", in Cooper, Mukai, and Perry (eds.), *Situation Theory and Its Applications,* Vol. 1, CSLI, Lecture Notes, Number 22.

Edgington, D. (1985), "The Paradox of Knowability", *Mind* **93**, pp. 557-68.

Fitch, F. B. (1963), "A Logical Analysis of Some Value Concepts", *The Journal of Symbolic Logic* **28**, pp. 135-42.

Hart, W. D. (1979), "The Epistemology of Abstract Objects", *Proceedings of the Aristotelian Society* **53**, pp. 164-165, n. 3.

Kaplan, D. (1989) "Demonstratives. An Essay on the Semantics, Logic, Metaphysics, and Epistemology of Demonstratives and Other Indexicals", in Almog, Perry and Wettstein (eds.), *Themes From Kaplan,* Oxford University Press.

Rabinowicz, W. and Segerberg, K. (1994) "Actual Truth, Possible Knowledge", *Topoi* **13**, pp. 101-115.

Segerberg, K. (1973), "Two-dimensional Modal Logic", *Journal of Philosophical Logic* **2**, pp. 77-96.

Williamson, T. (1987) "On the Paradox of Knowability", *Mind* 1987, pp. 255-261.