Fitch's paradox and *ceteris paribus* modalities

Carlo Proietti and Gabriel Sandu

March 20, 2008

One way to represent the so-called Verification Thesis (VT) *every truth is knowable* into a formal language, is using the following schema which combines the possibility and the knowledge operator:

(1) $\phi \rightarrow \Diamond K\phi$

Fitch's paradox [6] consists in using (1) in order to derive

(2) $\phi \rightarrow K\phi$

which means, according to the given reading, that *every truth is known.* Clearly this is too much, at least for a non *very-hard* antirealist about truth.[1] The structure of Fitch's derivation is roughly the following: first we assume *non omniscience*, i.e. the existence of a particular ψ such that

(3) $\psi \wedge \neg K\psi$

and then, by replacing ϕ with (3) in (1), we obtain a contradiction, from which we infer classically the negation of (3), (i.e. (2)) and we discharge the assumption.

We shall focus in this paper on two kinds of proposals for blocking the derivation of (2)[2]. Each questions one of the fundamental assumptions underlying (1):

(a) Restrictions on the *substitution formulas* in schema (1)

(b) Reformulation of the Verification Thesis

[1] This term is inspired by H. Rueckert's [11] adaptation of N. Tennant's terminology, used in [12] (chap. 8), and his distinction between various forms of antirealism, i.e. *soft*, *hard* and *very hard*, each of them inspiring different types of responses to Fitch's paradox.

[2] Actually, many different strategies for solving (or *dissolving*) the paradox have been displayed during the last three decades. Nearly every inferential step of Fitch's argument was brought into question. But we will focus only on two of them, which have the advantage of letting untouched some fundamental features of classical normal modal logic, thus seeming reasonable both from a realist and a *softly* antirealist point of view.

(a) has been advocated by M. Dummett [4], N. Tennant [12] and, from the perspective of dynamic epistemic logic, by J. van Benthem and others.[3] Perhaps the most basic but important observation in this context is due to Tennant: even if $\psi \wedge \neg K\psi$ is consistent, the formula $K(\psi \wedge \neg K\psi)$ is not, and that is the source of the contradiction. Thus, it is argued, we should not substitute in (1) formulas ϕ for which $K\phi$ leads to inconsistencies. We think this is a correct point to make. But we also think that the most plausible strategy for avoiding the paradox is pursuing (b).

The most discussed proposal for reformulating (VT) is due to D. Edgington [5]. She considers a scenario in which ϕ stand for "something wrong is happening in L.A. tonight" (and I don't know it). This is precisely one of those propositions (or statements) that I may come to know, maybe by reading tomorrow's newspapers. But if this happens, I will then know that "something wrong *was* happening in L.A. *yesterday* night", or that "something wrong is happening in L.A. tonight" *was* true. This suggests that, independently of whether our formulas represent *sentences* or *propositions*, a *temporal or situational reference* matters in the reformulation of (1), which is not taken into account by, say

(4) $\phi \rightarrow FK\phi$.

The standard way to interpret (4) in a Kripke-style semantics says that if something is true at a certain point t, it will become known at a future moment t', that ϕ is true *then* (i.e. at t'). In the same way (1) says that if ϕ is true, then there is another situation so that ϕ is known *there*.

Edgington proposes to reformulate (1) in a way that is more sensitive to the temporal or situational aspect just mentioned. In order to do this, she adopts a hybrid modal language with a *satisfaction operator* @.[4] The operator @ can be interpreted either as a "metaphysical" *actuality* operator A or as a temporal operator *now*, N. Then (1) and (4) are replaced with

(5) $A\phi \rightarrow \Diamond KA\phi$

and respectively

(6) $N\phi \rightarrow FKN\phi$.

[3] A fundamental connection between Fitch's paradox and self-refuting statements in dynamic epistemic logic was highlighted by van Benthem in [2] and further explored in recent collective works like [1].

[4] For every ϕ, $@\phi$ fixes the evaluation of ϕ to a predetermined world of reference, regardless of the modal scope in which $@\phi$ appears. Basic modal logic is clearly not expressive enough to do that. For example, in the temporal language, a better reformulation of (1) would be

$$\phi \rightarrow FKP\phi$$

but, again, this is not what we meant.

Edgington's suggestion has two problems to deal with. The first one, raised by T. Williamson in [16], is that the semantics for hybrid logic renders the following principle valid:

$$@\phi \leftrightarrow \Box @\phi.$$

In other words, on Edgington's variant of the verification thesis, this last would concern only *necessary* (or *a priori*) truths of the form $@\phi$. But this is clearly inadequate, for one would want also some *contingent* truths to fall in its range.[5] This objection has been amplified by W. Rabinowicz and K. Segerberg [15] who pointed out that, when combined with the satisfaction operator, *every universal modal operator* collapses. In particular, we have also:

$$@\phi \leftrightarrow K@\phi.$$

We notice that the left to right direction is precisely a revised version of (2), a principle that the strategy of reformulation was calculated to avoid.

One should note at this point that this kind of collapse of universal modal operators with the operator @ arises only within the standard semantics for basic modal logic with satisfaction operators. In such a framework, where models are typically triples $M = (W, R, V)$, the interpretation of the actuality operator @, or rather $@_i$ ('i' is the nominal functioning as the name of a world in W) is given by

$$(M, w) \models @_i\varphi \text{ iff } (M, v) \models \varphi, \text{ where } v \text{ is the denotation of '}i\text{' under } V$$

and the interpretation of the necessity operator is given by the standard clause

$$(M, w) \models \Box\varphi \text{ iff } (M, v) \models \varphi, \text{ for every } v \text{ such that } Rwv.$$

Now it is obvious that whenever we have $(M, w) \models @_i\varphi$ we do have also $(M, w) \models \Box @_i\varphi$. The same reasoning is reproduced for $@\phi \leftrightarrow K@\phi$.

Thus it seems that if we want to avoid the collapse of the universal operators, we have to look for a more fine-grained interpretation. Rabinowicz and Segerberg followed this path (in [15]). They defined a two-dimensional

[5] Williamson's criticism is formally correct but seems a little bit odd, for it depends on what notion of necessity is at stake. The fact that $\Box\phi$ is true in the whole model does not imply that ϕ is *necessary* in a *strong* sense: actually there are cases of a *contingent necessity*, even if this sounds like a *contradictio in adjecto*. We can illustrate it via an example made by Williamson himself. Let ϕ stand for "Italy actually won the 1982 World Cup". ϕ is true and, given a certain reading of $\Box$ also $\Box\phi$ is, for past events are settled and may not be changed (read $\Box$ as representing the notion of historical necessity given by Prior in [10]). Nevertheless we wouldn't say that ϕ is necessary in a *strong* sense, nor that it is *a priori*, because things *could* have been different and Brazil could have won. As we will see later, the Priorian reading of $\Box$ plays an important role in our approach.

semantics in which the actuality operator can denote different worlds of *reference* in the same model, depending on the *perspective* which is taken into account. Our proposal in this paper will be similar in spirit.

A second objection to Edgington's approach, raised also by Williamson, concerns the dubious notion of a *non actual knowledge of actual truths*. It is hard to make sense of a non actual knower referring to the actual situation. We will not discuss this point in detail, we content ourselves to observe that this possibility seems to be less problematic in a temporal setting (a possible future knower may very well be able to refer to a specified past time)[6]

1 A temporal approach

We set ourselves the task to formulate notions of "metaphysical" necessity and knowledge for which it is consistent to reformulate the Verification Thesis along the lines proposed by Edgington while still avoiding the collapse of metaphysical or epistemic modalities. We will present a uniform semantics for both these notions. We do not claim to have a solution to Fitch's paradox but rather to model it in a different way.

Our language will consist of *temporal indexes* $t, t', \ldots$ and temporal satisfaction operators $@_t, @_{t'} \cdots$ which operate in a way analogous to that of

[6]Actually things are more complicated: Williamson put forward an argument for showing that, in order to refer to a situation, the agent should be able to specify it via some *necessary and sufficient* conditions (or also, as an alternative, by some other devices, namely by *counterfactual*, by *space-time coordinates* or by *ostension*) but that, again, would trivialise the agent's knowledge to knowledge of some logical truths. There is no place here to discuss in detail Williamson's argument for that conclusion. We only hint at the fact that also in this argument Williamson makes a rather ambiguous use of the notion of necessity. See for example in the following passage:

> Assume first that, in knowing that in *s, p* , one can specify *s* in way (i) [i.e. by necessary and sufficient conditions]. Thus, for some value of '*q*', necessarily, *s* obtains if and only if *q*: moreover, the knowledge that, necessarily, if *q* then *p*, counts as knowledge that in *s, p*. Now it is easy to show that, necessarily, *s* obtains if and only if both *p* and *q*. Thus the condition that both *p* and *q* specifies *s* in way (i) just as well as the condition that *q* does. [,...] In particular, the knowledge that, necessarily, if both *p* and *q* then *p* counts as knowledge that in *s, p*. Thus [,...] the knowledge that in *s,p* requires no more than the knowledge of a trivial logical truth. ([16] p. 259)

A key passage in this argument seems to be the assumption that q specifies the situation s as well as $q \wedge p$ does. In other words, if the situation of the world in 1982 could be specified by, say, q this should be necessarily equivalent, according to Williamson, with specifying it by "q and Italy won the World Cup". But this seems too strong: even if it is in some sense necessary (settled) that Italy won the world cup in 1982 we cannot claim that knowledge of this fact is knowledge of a plain logical truth.

This remark could also be regarded as a reductio of Williamson's initial claim that in order to know that in *s, p* we should have specified *s* by necessary and sufficient conditions: we never specify alternative situations this way and, nevertheless, we constantly have epistemic claims on them.

nominals and actuality operators in hybrid logics[7].

Let $(T, <)$ be a *strict linear order*, which we also assume to be *dense*, *right and left unbounded*[8]. Our language $\mathcal{L}$ is built over a set of propositional variables $\Phi = \{p, q, r..\}$ and a set $\mathcal{T} = \{\bar{t} \mid t \in T\}$[9], disjoint from Φ, called the set of *temporal indexes* (or *t-indexes*). Formulas are given by the clauses:

$$\mathcal{L} ::= p, q, ... \mid \bot \mid \neg\phi \mid \phi \wedge \psi \mid P\phi \mid F\phi \mid @_{\bar{t}}\phi$$

where $@_t\phi$ is taken to mean "ϕ holds at time t".

$\mathcal{L}$ is interpreted on temporal frames $\mathcal{F} = (W \times T, <^F)$ where W is a nonempty set and $(w, t) <^F (w', t')$ if and only if $w = w'$ and $t < t'$. A temporal model $\mathcal{M}$, based on a frame $\mathcal{F}$, is a pair $(\mathcal{F}, V)$, with V a valuation function from the set of propositional letters into $\mathcal{P}(W)$. Truth at world-time pairs is defined in the usual way:

- $(w, t) \models p$ iff $w \in V(p)$
- $(w, t) \models \phi \wedge \psi$ iff $(w, t) \models \phi$ and $(w, t) \models \psi$
- $(w, t) \models \neg\phi$ iff $(w, t) \not\models \phi$
- $(w, t) \models F\phi$ iff there is a t' such that $t < t'$ and $(w, t') \models \phi$
- $(w, t) \models P\phi$ iff there is a t' such that $t' < t$ and $(w, t') \models \phi$
- $(w, t) \models @_{\bar{t'}}\phi$ iff $(w, t') \models \phi$

Obviously such models can be regarded as collections of strictly ordered *flows of time* (or *runs*), where a *flow of time* is a subset of $W \times T$ consisting of pairs having the same first coordinate. The satisfaction clause for $@_{\bar{t}}\phi$ presents a main difference with respect to the usual semantics for hybrid modal logics: given a moment t of time, there is not a single world of reference for it, with respect to which $@_t\phi$ should be evaluated, but one for every single flow of time.

We take the propositional variables in Φ to express some kind of *contingent instantaneous sentences*, like "David is sitting" or "Mary is washing her car". They may be true or false at two different moments. What about *propositions*? They are commonly expressed by "David is sitting at time t" or "Mary is washing her car at time t'"[10]. Formulas like $@_tp$ are supposed

[7]The frameworks underlying our models are actually inspired by bidimensional semantics of kind $W \times T$, where points are couples (w, t), (w', t') made of a *world coordinate* and a *temporal coordinate* (with a total ordering on T). For more detail see R. H. Thomason's presentation of those models in [13].

[8]We are presupposing, for the sake of simplicity, a particular kind of temporal structure, but other assumptions can be made.

[9]The $\bar{t}$'s are names for the elements in $t \in T$ but we will simply write them as t where no confusion is possible.

[10]We are taking into account only temporal indexes.

to achieve the kind of specification of the temporal parameter t specific of "eternal" sentences. We will take them to express propositions. And indeed we notice that

$$@_t\phi \rightarrow H@_t\phi \wedge @_t\phi \wedge G@_t\phi$$

is valid.

2 Universal operators

When defining necessity and knowledge, we have to be more careful. Moving to bidimensional frames $\mathcal{F} = (W \times T, <^F)$ does not automatically help us to avoid the validity of $@_t\phi \leftrightarrow \Box@_t\phi$ and $@_t\phi \leftrightarrow K@_t\phi$.

Usually, when defining these notions[11], one introduces an equivalence relation $w \approx_t w'$ which holds between the worlds w and w' whenever they share the same past up to and including t. Finally one stipulates:

$$(w, t) \models \Box\phi \text{ iff } (w^*, t) \models \phi \text{ for any } w^* \text{ such that } (w, t) \approx (w^*, t)$$

But now very much depends on what "sharing the same past" is taken to mean. If it is defined by:

> w and w' satisfy the same formulas at each moment in the past
> up to and including t

then we are in trouble. For in that case, supposing that $@_i\phi$ is true at (w, t), it is going to be true at any (w^*, t) such that $(w, t) \approx (w^*, t)$.

So this way to define "sharing the same past" is not going to help us. In fact the main problem with it is that the formula φ or $@_t\varphi$ may contain some elements that are extraneous to the past of a given world, on our normal understanding of this notion.

So what are the alternatives that we have here?

The crucial point here is to specify the past in such a way that extraneous elements (pertaining to the future) do not enter into its description. There are several ways to implement this idea, all variants of each others. Let us mention a few.

Freddoso [7] discusses a relevant example. Suppose ϕ is the sentence "Mary is sitting" and suppose it is true at some moment of time t. Then at a much earlier moment of time $t' < t$ it is true that Mary will be sitting (at t). ($F\varphi$ is true at t'.) And at each moment t'' later than t' but earlier than t, it is true that it was the case (at t') that Mary will be sitting (at t). ($PF\varphi$ is true at each t''.) But now it follows that any other world which shares the same history (in the sense just mentioned) with the one being scrutinized is such that $PF\varphi$ is true in it between t' and t. Thereby $PF\varphi$ is necessary between t' and t. Freddoso shows that this later claim, together with a couple of

[11] See Thomason [13], p. 209.

undeniable logical principles, leads to the conclusion that φ is unavoidable at t. To avoid this undesirable conclusion, he formulates a notion of "sharing the same past" according to which statements like $PF\varphi$, which bears on future events, are not taken into account. His "Ockhamist solution" is to give preference to pure present tense sentences which express "time-indifferent" propositions. These are atomic sentences like "Mary is sitting" together with sentences which are reducible to them through various logical constructions (which exclude future tense operators and certain propositional attitudes.) The logical outcome is the restricted validity of the schema $\varphi \to \Box\varphi$ which now holds only for time-indifferent propositions φ.

One finds a variant of this idea in Prior's notion of *historical necessity*. In addition to standard propositional symbols $p, q, r \ldots$, Prior introduces special letters $a, b, c \ldots$ "with the special property that their truth is independent of the future course of history". Like in the previous case, the schema $a \to \Box a$ is valid on the condition that only formulas not involving future tense operators can be substituted for the special letters. One cannot derive, for instance, $Fa \to \Box Fa$.

For the purpose of the present paper, we shall be satisfied with a strategy based on Thomason ([13]), inspired by Prior, which, like the other two variants, gives a special treatment to atomic sentences.

Let us fix a set of sentences Γ, *the smallest set containing atomic sentences and closed under conjunctions, and tautological inferences*[12]

Next we define the needed notion of equivalence relative to Γ :

$$(w, t) \approx_\Gamma (w^*, t) \text{ if and only if } \forall t' < t : (w, t') \models \varphi \Longleftrightarrow (w^*, t') \models \varphi, \text{ for all } \varphi \in \Gamma$$

This definition says, informally, that two worlds w and w' are Γ-*equivalent* at time t only when at every previous moment t' they verify the very same formulas in the set Γ; but they can diverge, and this is also fundamental for our proposal, concerning formulas not in Γ (such as future-tensed formulas).

Finally we form a new modal operator, $[\Gamma]\phi$, interpreted by

$$(w, t) \models [\Gamma]\phi \text{ if and only if for all } w^* : \text{if } (w, t) \approx_\Gamma (w^*, t), \text{ then } (w^*, t) \models \phi$$

In other words we say that ϕ is *necessary relative to* Γ in w at a moment t when ϕ is true in all the relevant alternatives w', i.e. those which are Γ-equivalent with w up to (but not including) t. The clause for its dual $\langle\Gamma\rangle$ is standard.

The new notion has the virtue to block, as wanted by Freddoso and Prior, the unrestricted validity of the schema $\varphi \to [\Gamma]\varphi$. For instance, for $t < t'$, we can very well have $(w, t) \models @_{t'}p$, but also $(w, t) \not\models [\Gamma]@_{t'}p$.

[12] Actually this closure condition is superfluous in this case.

Notice however, that all tautological future formulas, like $Fp \vee \neg Fp$ are necessary, i.e., for any Γ, $[\Gamma](Fp \vee \neg Fp)$ is valid.[13]

The operator $[\Gamma]$ expresses a kind of *ceteris paribus* modality: all things being equal with respect to Γ. Below we shall identify both $\Box\varphi$ and $K\varphi$ to $[\Gamma]\varphi$. This move should come as no surprise. Freddoso and Prior have blocked the unrestricted validity of the schema $\varphi \to \Box\varphi$ in order to avoid logical (metaphysical) determinism. Here we shall use $[\Gamma]\varphi$ to avoid the epistemic determinism ($\phi \to K\phi$) associated with Fitch's paradox.

3 $[\Gamma]$ as an epistemic operator

Let us now interpret $[\Gamma]$ as a knowledge operator K. Thus whenever we write $K\phi$ we have in mind $[\Gamma]\varphi$. We may even modify some of the clauses for $(w,t) \approx_\Gamma (w^*,t)$ so that sentences do not need to become known in the same temporal order. But independently of that, the main idea is that $(w,t) \approx_\Gamma (w^*,t)$ reflects the fact that, for a given agent, the two histories are "observationally equivalent" up to and including the time t. Thus the notion of "observational equivalence" replaces in the present context the earlier notion of "time-indifferent" proposition. Typically Γ would include contingent instantaneous sentences and those sentences reducible to them through various logical operators. Variations are allowed here, as in previous cases. For instance, we may allow closure under knowledge but not under future operators. Given that $\approx_\Gamma$ is an equivalence relation, $[\Gamma]\varphi$ is, from a technical point of view, a *bona fide* S5 epistemic operator.

Now back to Fitch's paradox.

Recall the example discussed in the first section, where ϕ stands for "something wrong is happening in L.A. tonight" (and I don't know it). There, following Edgington, we observed that what will become known is that "something wrong *was* happening in L.A. *yesterday* night", or that "something wrong is happening in L.A. tonight" *was* true. In this light, it is useful to look at several reformulations of the Verification Principle in our temporal setting:

(a) $@_t\phi \to \Diamond K @_t\phi$

(b) $@_t\phi \to FK@_t\phi$

(c) $@_t\phi \to FKP@_t\phi$

(d) $\phi \to FKP\phi$.

Let us look at the following scenario.

[13] This last fact comes as no surprise because, given the linearity condition on *flows*, formulas like $Fp \vee \neg Fp$ are satisfied everywhere, then also in every point sharing the same story as the point of evaluation. Moreover, given the definition of the $@_i$'s, we have also that $(w,t) \models @_{t'}\phi \vee \neg @_{t'}\phi$, even if $t < t'$.

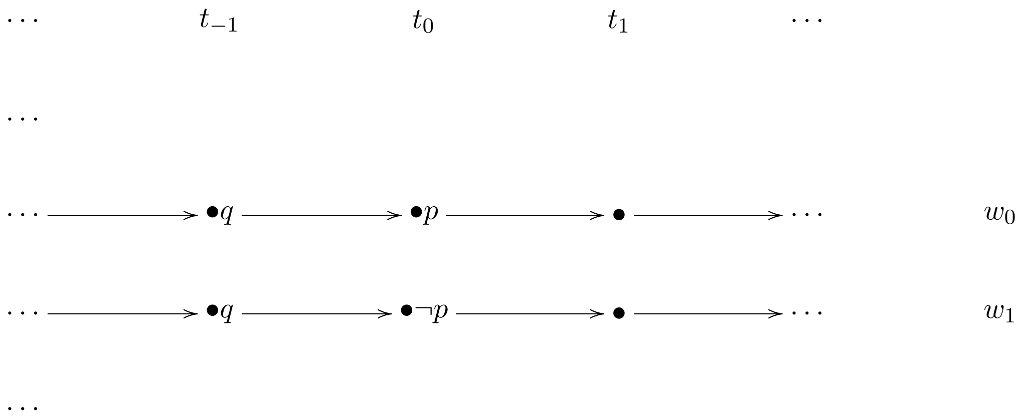

Figure 1:

We notice that (w_0, t_0) and (w_1, t_0) are epistemic alternatives to each other, but not (w_0, t_1) and (w_1, t_1). As time flows, and new contingencies p are observed, the agent comes to know that they happened immediately afterwards, for any scenario which is an alternative to the actual one must have the same contingencies true at the same moments. But then, no matter what Γ is, (w_0, t_1) and (w_1, t_1) cannot be equivalent. Thus $@_{t_0}p \wedge \neg K @_{t_0}p$ and $p \wedge \neg Kp$ are both true at (w_0, t_0) and the same holds of all the schemata (a)-(d), when p is substituted for ϕ and t_0 is substituted for t.

Nevertheless the problem pointed out by Tennant and van Benthem remains here: as in its alethic counterpart, even the modified version of the Verification Principle does not hold in general (we will see below what restrictions are required).

Clearly some instances of non omniscience like $@_{t_0}p \wedge \neg K @_{t_0}p$ do not give rise to Fitch's paradoxical conclusion when we perform the substitution into the revised versions of the verification schema. For instance, it may be checked that when we substitute it for ϕ in (b), the result is a true sentence. By our remarks a few lines earlier, the antecedent

$$@_{t_0}(@_{t_0}p \wedge \neg K @_{t_0}p)$$

is true at (w_0, t_0). Actually $@_{t_0}(@_{t_0}p \wedge \neg K @_{t_0}p)$ is satisfied all along the flow of w_0, because w_0 is indistinguishable from w_1 up to t_0 (and $(w_1, t_0) \models \neg p$).

In addition, the consequent

$$FK @_{t_0}(@_{t_0}p \wedge \neg K @_{t_0}p)$$

is also true at (w_0, t_0), given that $K @_{t_0}(@_{t_0}p \wedge \neg K @_{t_0}p)$ is true at (w_0, t_1).

The reader may check that all the other versions of the Verification Thesis are true.

As we said before, this version of VT is not absolutely substitution free, even if our interpretation of $[\Gamma]$ endows the epistemic operator with some exceptionally strong properties. It makes the agent have *unbounded memory* and *perfect recall*: every time that some atomic fact p is true at a given time t, then we also have $FK@_t p$, because the evaluation of the epistemic operator excludes all the histories in which p is not true at time t. But this doesn't force the agent to know at $t' > t$ every formula which is true at t. For example Kp (where we don't specify any temporal index) is not forced at t' [14]. Unrestricted substitution fails also for (b): take $\phi = Gp$ and observe that $@_t Gp$ is not known to be true at t'. (Think about a model with a point (w, t) in which p is satisfied for every (w, t') with $t < t'$, but which contains an infinity of temporal flows $w_1, w_2, w_3 \ldots$. Every w_i is Γ-equivalent with w up to a certain t_i but diverges afterwards. Suppose that the series of the t_i is right unbounded in $(T, <)$. Then there is no $t' > t$ such that $(w, t') \models K@_t Gp$.) There is nothing surprising in all this. The agent can achieve knowledge of only those future events which are either "tautological" or whose truth becomes "settled" after a moment of time, i.e., they become historical necessities.

4 Burgess' Discovery principle

In his paper *Can Truth Out?* (in [3]) J. Burgess also considers Fitch's paradox in a temporal framework. He proposes a reformulation of what he calls the Discovery Principle (4) ($\phi \to FK\phi$) as:

(i) $G\phi \to FK\phi$

He shows that (i) together with the usual axioms for epistemic and linear temporal logic (call it system T) proves the following:

(ii) $P\phi \to FKP\phi$

(iii) $\phi \to FKP\phi$

(iv) $F\phi \to FKP\phi$

(v) $G\phi \to FKG\phi$

but not the undesired

(3) $\phi \to K\phi$

[14] This makes perfectly good sense given our reading of atomic formulas $p, q \ldots$ as *contingent instantaneous sentences.*

Burgess also mentions that, if one adds actuality operators $@_t$ to the language[15] $\mathcal{L}$ (call it $\mathcal{L} \cup \{@\}$), then the resulting system is stronger than Edgington's proposal, since it proves her version of the verification principle, i.e.,

(b) $@_t\phi \rightarrow FK@_t\phi$.

One of Burgess' motivations for adopting such a strong version of the discovery principle is to avoid its applicability to *ephemeral* truths.[16] Such truths are represented in our language by atomic sentences $p, q \ldots$. A reason to adopt a more expressive language like $\mathcal{L} \cup \{@\}$ and to maintain (b) as a reformulation of the Verification Thesis is to hold to the standard view according to which *truth* is a property of *propositions.* Propositions as they are usually understood, are *atemporally true or false.* Thus neither ephemeral truths nor formulas such as Gp represent propositions, and for this reason they cannot, strictly speaking, be true or false. In our semantics we took $@_t\phi$ to express propositions. It is then consistent to talk about the truth of say (b).

Burgess' system T contains all the axioms of *linear temporal logic* and S5 axioms for the epistemic operator K. The semantics we introduced in the preceding section validates all theorems of linear temporal logic and the axioms of S5 for epistemic logic (but contrary to Burgess, we operate with a $\mathcal{L} \cup \{@\}$ language). Given this, it is then straightforward to check that, as in the case of Burgess' system T, (i) $G\phi \rightarrow FK\phi$ and $@_t\phi \rightarrow G@_t\phi$ (which is valid in our semantics) entail (b). Also, (i) is clearly stronger than (b), for there are formulas ϕ, models M and points (w, t^*) in M in which (b) is true but (i) is not. The picture below, which represents the model outlined in the last paragraph of the preceding section, is a case in point.

Take $\phi = p$ and observe that $(w_0, t_0) \models @_{t_0}\phi \rightarrow FK@_{t_0}\phi$, but $w_0, t_0) \not\models G\phi \rightarrow FK\phi$ (notice also that the corresponding instance of (v) is false at (w_0, t_0)).

Actually we have seen in the last section that (b) is not generally valid (Gp is one of the counter-instances.) Thus Burgess' temporal system is considerably stronger that any logical system that is sound for the semantics given in this paper. One could of course exclude models like the ones we have used to invalidate instances of (i) and (b) (that is, models in which the truth of a given ϕ is never "settled"), but to borrow a quote from Burgess, that would be in our opinion more like an act of faith in Verificationism than a decision about which logical principles to include in our logic.

[15]Burgess actually is against using actuality operators for reasons which shall not be discussed here.

[16]Ephemeral truths are what we have called contingent instantaneous sentences, like "Smith is murdering Jones".

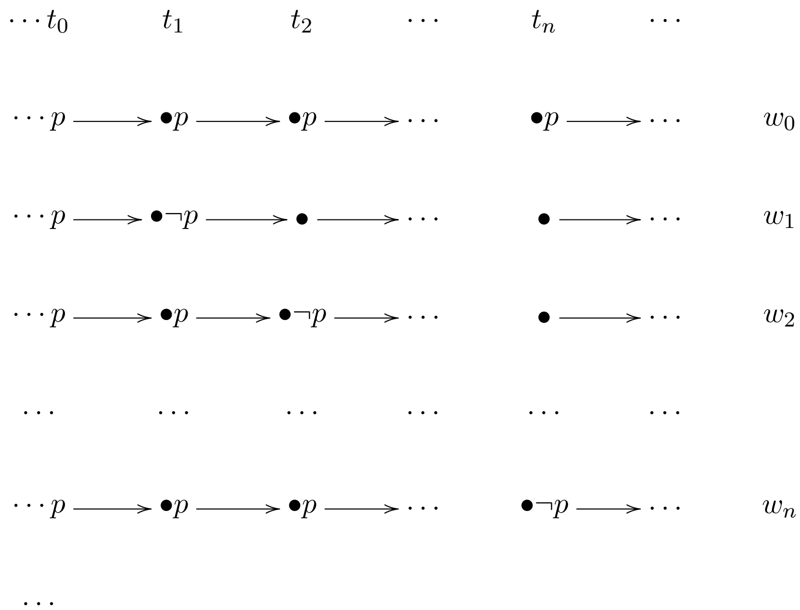

Figure 2:

5 [Γ] as a metaphysical operator

There is a vantage point from which the temporal variants of the verification thesis, like $\phi \rightarrow FK\phi$, or (b) $@_t\phi \rightarrow FK@_t\phi$ are inadequate. There is no need to be a realist about truth in order to acknowledge, as Dummett does, that there are unknown truths that will perhaps remain forever unknown (because the means of verifying them are no longer at our disposal).[17] Nevertheless, the arguments continues, these truths are *knowable* in a wider sense, one that our version of the Verification Thesis is insensitive to.

In order to take into account such a possibility, we could try to reformulate (VT) as

$$@_t\phi \rightarrow \Diamond FK@_t\phi^{18}.$$

while still wanting to avoid the collapsing principle $@_t\phi \rightarrow FK@_t\phi$.

---

[17]This case could not be better illustrated than Williamson does by citing Thomas Gray's *Elegy Written in a Country Churchyard*

Full many a gem of purest ray serene
The dark unfathom'd caves of ocean bear:
Full many a flower is born to blush unseen,
And waste its sweetness on the desert air.

See [17] p. 273.

[18]Where $\Diamond$ should be regarded as ranging over many different possible future times and not only "the actual future".

Such a move was actually considered by Rueckert in [11], but he dismissed it on the ground that it is coherent only by assuming that future tense propositions have no truth-value.[19]. If that were so, that would be a major drawback.

It doesn't seem to us that we are we forced into this conclusion. It is enough to remind ourselves that the models considered here are based on (linear) flows of time and thus there is no indeterminacy about the future, i.e. future tense statements have a fixed value at every point. In the same time, the operators $[\Gamma]$ and $\langle\Gamma\rangle$ can take us from one flow of time to another.

Let us interpret $\langle\Gamma\rangle$ as a "metaphysical" operator $\Diamond$ and let K be primitive. The picture below represents a situation where p is true at (w_0, t_0), but the agent never comes to know it, i.e. $(w_0, t_n) \models \neg K@_{t_0}p$ for every $t_n > t_0$. Then $(w_0, t_0) \models \neg FK@_{t_0}p$. Nevertheless there is an alternative flow w_1 in which "all things are equal" up to t_0 such that $@_{t_0}p$ becomes known at a later time, i.e. $(w_1, t_2) \models K@_{t_0}p$. Thus we have $(w_0, t_0) \models \Diamond FK@_{t_0}p$, a counter-example to $@_t\phi \to FK@_t\phi$.

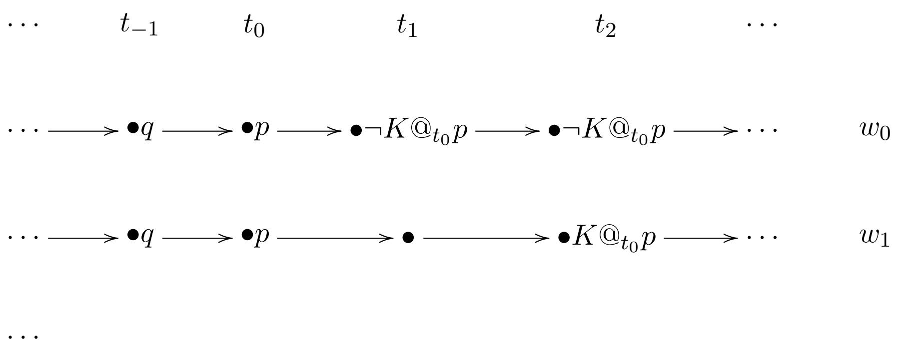

Figure 3:

6 $[\Gamma]$ operator as a general *ceteris paribus* operator.

As we mentioned in section 2 the operator $[\Gamma]$ has a long history. Our source of inspiration is *ceteris paribus preference* logic. The first modal treatment of ceteris paribus preference goes back to Von Wright's seminal work [14] which is a logical study of "everything else being equal p is preferred to q". A natural way to interpret the ceteris paribus notion in Kripke models consists in saying that, given two points w and w', w' can be a ceteris paribus alternative to w only if w' agrees with w on a certain set Γ of sentences of

[19] This is the standard strategy for avoiding logical determinism.

the modal language. This is precisely the way how ceteris paribus preference operators $[\Gamma]^<$ and $[\Gamma]^\leq$ are interpreted in [9] and in [8]. The operator $[\Gamma]$ we used in the present paper belongs to the same family, for $[\Gamma]\phi$ says, roughly speaking, "everything else in the past being equal: ϕ".

In [9] Γ is taken to be an arbitrary set of sentences. There are actually some good motivations to use this more general option also in temporal reasoning: if we give up some of the closure conditions previously associated with Γ, the operators $[\Gamma]$ and $\langle\Gamma\rangle$ can be given many interesting interpretations. Consider for example the following ϕ: "even if John had been injuried, John would have played today". Suppose that p stands for "John plays" q for "John has an injury" and t is today time. It is then natural to interpret ϕ by considering all the alternatives which share the same history with the actual one, except for q. Assuming Γ does not contain q we can represent ϕ by $@_t p \wedge [\Gamma]@_t p$.

References

[1] Philippe Balbiani, Hans van Ditmarsch, Andreas Herzig, Tomohiro Hoshi, and Tiago de Lima. What can we achieve by arbitrary announcements? a dynamic take on fitch's knowability. In *Proceedings of the eleventh Theoretical Aspects of Rationality and Knowledge conference (TARK)*, 2007. (to appear).

[2] J. Van Benthem. What one may come to know. *Analysis*, 64:95–105, 2004.

[3] J. Burgess. Can truth out? In J. Salerno, editor, *New Essays on the Knowability Paradox*. Oxford University Press, Oxford, forthcoming.

[4] M. Dummett. Victor's error. *Analysis*, 61:1–2, 2001.

[5] D. Edgington. The paradox of knowability. *Mind*, 94:557–568, 1985.

[6] F. B. Fitch. A logical analysis of some value concepts. *Journal of Symbolic Logic*, 28:135–142, 1963.

[7] A. J. Freddoso. Accidental necessity and logical determinism. *The Journal of Philosophy*, 80:257–278, 1983.

[8] P. Girard. *Modal Logic for Belief and Preference Change*. PhD in Philosophy, Stanford University, 2008.

[9] J. Van Benthem P. Girard, O. Roy. Everyhing else being equal. a modal logic approach to *ceteris paribus* preferences.

[10] A. N. Prior. *Past, Present and Future*. Oxford University Press, Oxford, 1967.

[11] H. Rueckert. A solution to fitch's paradox of knowability. In J. Symons S. Rahman, editor, *Logic, Epistemology and the Unit of Science*, pages 351–380. Kluwer Academic Publishers, Amsterdam, 2004.

[12] N. Tennant. *The Taming of the True*. Clarendon Press, Oxford, 1997.

[13] R. H. Thomason. Combinations of tense and modality. In F. Guenthner D. Gabbay, editor, *Handbook of Philosophical Logic*, pages 205–234. Kluwer Academic Publisher, Dordrecht, 2002.

[14] G. H. von Wright. *The Logic of Preference*. University of Edinburgh Press, Edinburgh, 1963.

[15] K. Segerberg W. Rabinowicz. Actual truth, possible knowledge. In Ronald Fagin, editor, *Proceedings of the 5th Conference on Theoretical Aspects of Reasoning about Knowledge, Pacific Grove, CA, USA, March 1994*, pages 122–137. Morgan Kaufmann, 1994.

[16] T. Williamson. On the paradox of knowability. *Mind*, 96:256–261, 1987.

[17] T. Williamson. *Knowledge and its Limits*. Oxford University Press, Oxford, 2000.