Knowledge, Time, and Paradox:
Introducing Sequential Epistemic Logic*

Wesley H. Holliday

Abstract Epistemic logic in the tradition of Hintikka provides, as one of its many applications, a toolkit for the precise analysis of certain epistemological problems. In recent years, *dynamic* epistemic logic has expanded this toolkit. Dynamic epistemic logic has been used in analyses of well-known epistemic "paradoxes", such as the Paradox of the Surprise Examination and Fitch's Paradox of Knowability, and related epistemic phenomena, such as what Hintikka called the "anti-performatory effect" of Moorean announcements. In this paper, we explore a variation on basic dynamic epistemic logic—what we call *sequential epistemic logic*—and argue that it allows more faithful and fine-grained analyses of those epistemological topics.

Key words: sequential epistemic logic, dynamic epistemic logic, temporal logic, surprise exam paradox, Fitch's paradox of knowability, Moorean announcements, anti-performatory announcements, unassimilable announcements

1 Introduction

Epistemic logic in the tradition of Hintikka (1962) has found myriad applications, spanning philosophy, computer science, game theory, and linguistics, in addition to developing a set of topics and agenda of its own (see the recent *Handbook of Epistemic Logic*, van Ditmarsch et al 2015). While enjoying this wide-ranging success, epistemic logic has not forgotten its philosophical roots. To the contrary, its application to problems in epistemology has undergone a kind of renaissance in recent years (for surveys and references, see Egré 2011 and Holliday Forthcoming).

Wesley H. Holliday
Department of Philosophy and
Group in Logic and the Methodology of Science
University of California, Berkeley
e-mail: wesholliday@berkeley.edu

* Preprint of June 1, 2016. Forthcoming in a volume in honor of Jaakko Hintikka, edited by Hans van Ditmarsch and Gabriel Sandu, for Springer's *Outstanding Contributions to Logic* series.

Another important development in recent epistemic logic has been the rise of *dynamic* epistemic logic (for a survey, see Pacuit 2013, and for textbooks, see van Ditmarsch et al 2008 and van Benthem 2011). This dynamic turn has led to diverse directions of research, yet it has also overlapped with the renaissance of epistemological applications of epistemic logic. Dynamic epistemic logic has been used in analyses of well-known epistemic "paradoxes", such as the Paradox of the Surprise Examination (Gerbrandy, 1999, 2007) and Fitch's Paradox of Knowability (van Benthem, 2004; Balbiani et al, 2008; van Ditmarsch et al, 2011). In addition, it has been used to analyze related epistemic phenomena, such as what Hintikka (1962, §4.17) called the "anti-performatory effect" of Moorean announcements (van Ditmarsch and Kooi, 2006; Holliday and Icard, 2010; Holliday et al, 2013).

In this paper, we will explore a variation on basic dynamic epistemic logic—what we call *sequential epistemic logic*—and argue that it allows more faithful and fine-grained analyses of the epistemological topics just mentioned. The basic idea of sequential epistemic logic (SEL) is that it allows us to reason about the full temporal *sequence* of agents' epistemic states, including agents' changing knowledge of their own and others' *past* and *future* epistemic states, in terms of the same kinds of epistemic transformations as studied in dynamic epistemic logic.[2]

The idea of adding some kind of temporality to dynamic epistemic logic is not new. Motivations for such a move are widely acknowledged in the literature and worked out in different ways (see, e.g., Hoshi 2008, van Benthem et al 2009, Gierasimczuk 2010, van Benthem 2011, Ch. 11, references therein, and references at the end of §3.1 and §3.2 below). The novelty of sequential epistemic logic is in the specific way this move is carried out, which is motivated by the specific epistemological applications at hand. For other applications, other ways of temporalizing dynamic epistemic logic will no doubt be more appropriate.

The plan for our introduction to sequential epistemic logic is as follows.

In §2, we briefly review basic dynamic epistemic logic, in particular, what has come to be called Public Announcement Logic (PAL). This name is unfortunate for our purposes, insofar as it suggests that the framework should be of interest to operators of loudspeakers, but perhaps not philosophers. To the contrary, the basic and familiar idea of the framework—information acquisition as the elimination of possibilities—is of interest in many disciplines, including philosophy. A notable development of the idea in philosophy is the picture of inquiry in Stalnaker 1984.

In §3, we review the standard analyses of Fitch's paradox and the surprise exam paradox based on PAL. Although these analyses are a good start, we point out ways in which a richer framework is needed to capture key ideas raised by the paradoxes.

In §4, we propose a candidate for such a richer framework. As an example of the general idea of SEL, we introduce what could be called (aligning with the accepted nomenclature) Sequential Public Announcement Logic (SPAL).

In §5, we turn to applications of SPAL. First, we show how to enlarge the standard PAL taxonomy of successful, unsuccessful, and self-refuting sentences with useful new distinctions expressible with SPAL, such as a distinction between *assim-*

[2] For a different variation on dynamic epistemic logic in a similar spirit, see Cohen 2015a,b.

ilable and *unassimilable* sentences (applied in Holliday 2016a,b), and a distinction between *ascertainable* and *unascertainable* sentences. We then show that the second distinction is key to a more faithful analysis of Fitch's paradox, while the first distinction is key to a more faithful analysis of the surprise exam paradox.

2 Basic Dynamic Epistemic Logic

In this section, we review the language, semantics, and some basic results about the dynamic epistemic logic PAL (Plaza, 1989; Gerbrandy and Groeneveld, 1997).

Let At be a countably infinite set of atomic formulas and Agt a nonempty countable set of agent symbols. The language of PAL, $\mathscr{L}_{PAL}$, is given by the following grammar:

$$\varphi ::= p \mid \neg\varphi \mid (\varphi \wedge \varphi) \mid K_a\varphi \mid \langle !\varphi \rangle \varphi,$$

where $p \in \mathsf{At}$, $a \in \mathsf{Agt}$. $\mathscr{L}_{EL}$ is the fragment without formulas of the form $\langle !\varphi \rangle \psi$, i.e., the language of epistemic logic. As usual, K_a is agent a's knowledge operator.

The appropriate reading of formulas of the form $\langle !\varphi \rangle \psi$ depends on whether we are in the single-agent case where $|\mathsf{Agt}| = 1$ or the multi-agent case where $|\mathsf{Agt}| > 1$. In the single-agent case, we read $\langle !\varphi \rangle \psi$ as "φ is true, and if the agent updates her epistemic state with the proposition which φ expressed before the update, then ψ will be true." In the multi-agent case, we read $\langle !\varphi \rangle \psi$ as "φ is true, and if all agents update their epistemic states with the proposition which φ expressed before the update—with it being common knowledge that all agents perform this update—then ψ will be true." The notion of *update* will be made precise shortly.

Traditionally $\langle !\varphi \rangle \psi$ is read as something like "φ is true, and if φ is *publicly announced*, then ψ will be true." In the multi-agent case, an act of public announcement may be one typical way of triggering the relevant kind of update with the relevant common knowledge. But according to the view adopted in this paper, the real subject matter of PAL is the information updates themselves. This is especially clear in the single-agent case, where we are reasoning about any update of the agent's epistemic state, regardless of whether it is triggered by an utterance, experiment, perception, etc. In this sense, PAL is a logic of pure information update, and the term 'public announcement' suggests that PAL has a narrower significance than it in fact has. (The term 'Information Update Logic' would be preferable in the single-agent case, and the term 'Common Information Update Logic' might be preferable in the multi-agent case. But alas it seems too late to change tradition now.)

We interpret $\mathscr{L}_{PAL}$ in *epistemic models* $\mathscr{M} = \langle W, \{R_a\}_{a \in \mathsf{Agt}}, V \rangle$ where W is a nonempty set, each R_a is a binary relation on W, and $V \colon \mathsf{At} \to \wp(W)$. A *pointed model* is a pair of a model $\mathscr{M} = \langle W, \{R_a\}_{a \in \mathsf{Agt}}, V \rangle$ and a $w \in W$.

For terminology: W is the set of *worlds*; subsets of W are *propositions*; R_a is agent a's *epistemic accessibility relation*; and $R_a(w) = \{v \in W \mid wR_av\}$ is agent a's *epistemic state* at w in $\mathscr{M}$. We say that agent *a knows* a proposition $Q \subseteq W$ at a world w iff $R_a(w) \subseteq Q$. We will assume that each R_a is at least reflexive.

We evaluate formulas of $\mathcal{L}_{PAL}$ at pointed models. The clauses for $\mathcal{L}_{EL}$ are:

- $\mathcal{M}, w \vDash p$ iff $w \in V(p)$, for $p \in \mathsf{At}$;
- $\mathcal{M}, w \vDash \neg\varphi$ iff $\mathcal{M}, w \nvDash \varphi$;
- $\mathcal{M}, w \vDash \varphi \wedge \psi$ iff $\mathcal{M}, w \vDash \varphi$ and $\mathcal{M}, w \vDash \psi$;
- $\mathcal{M}, w \vDash K_a\varphi$ iff for all $v \in R_a(w)$: $\mathcal{M}, v \vDash \varphi$.

We call $[\![\varphi]\!]^{\mathcal{M}} = \{w \in W \mid \mathcal{M}, w \vDash \varphi\}$ the *proposition expressed by* φ *in* $\mathcal{M}$. The clause for $K_a\varphi$ can be reformulated as: $[\![K_a\varphi]\!]^{\mathcal{M}} = \{w \in W \mid R_a(w) \subseteq [\![\varphi]\!]^{\mathcal{M}}\}$.

There are two well-known semantics for the update formulas of $\mathcal{L}_{PAL}$: *world-elimination* semantics, where "updating" with a formula φ means deleting all states where φ was false; and *link-cutting* semantics, where "updating" with a formula φ means cutting epistemic accessibility links between states that differed on their truth values for φ. These semantics are equivalent for the language $\mathcal{L}_{PAL}$, but not necessarily for more expressive languages. The world-elimination semantics is more common, whereas the link-cutting semantics will be better for our purposes in §4.

For world-elimination, given a model $\mathcal{M} = \langle W, \{R_a\}_{a\in\mathsf{Agt}}, V\rangle$ and proposition $Q \subseteq W$, we define the update of $\mathcal{M}$ by Q, $\mathcal{M}_{\upharpoonright Q} = \langle W_{\upharpoonright Q}, \{R_{a\upharpoonright Q}\}_{a\in\mathsf{Agt}}, V_{\upharpoonright Q}\rangle$, by:

- $W_{\upharpoonright Q} = Q$; $R_{a\upharpoonright Q}$ is the restriction of R_a to $W_{\upharpoonright Q}$; $V_{\upharpoonright Q}(p) = V(p) \cap W_{\upharpoonright Q}$.

Then the semantic clause for update formulas is:

- $\mathcal{M}, w \vDash \langle !\varphi\rangle\psi$ iff $\mathcal{M}, w \vDash \varphi$ and $\mathcal{M}_{\upharpoonright [\![\varphi]\!]^{\mathcal{M}}}, w \vDash \psi$,

where $[\![\varphi]\!]^{\mathcal{M}}$ is the proposition expressed by φ in $\mathcal{M}$ as above. For cleaner notation, we can define $\mathcal{M}_{\upharpoonright\varphi} = \mathcal{M}_{\upharpoonright [\![\varphi]\!]^{\mathcal{M}}}$, but we will use the proposition notation.

Finally, *validity* is defined as usual: an $\mathcal{L}_{PAL}$ formula φ is valid (notation: $\vDash_{PAL} \varphi$) iff for every pointed epistemic model $\mathcal{M}, w$, we have $\mathcal{M}, w \vDash \varphi$.

To understand what updating the initial epistemic model $\mathcal{M}$ to the new $\mathcal{M}_{\upharpoonright [\![\varphi]\!]^{\mathcal{M}}}$ means conceptually, let us begin with the case where there is only one agent.

The equation $W_{\upharpoonright [\![\varphi]\!]^{\mathcal{M}}} = [\![\varphi]\!]^{\mathcal{M}}$ shows that the information the agent receives is the proposition which φ expressed *before the update*. We stress *before the update*, because *after the update* φ may express a different proposition, i.e., we may have

$$[\![\varphi]\!]^{\mathcal{M}_{\upharpoonright [\![\varphi]\!]^{\mathcal{M}}}} \neq [\![\varphi]\!]^{\mathcal{M}},$$

as in Example 1 below. The reason is that the knowledge operator K introduces *indexicality*: the proposition expressed by a formula containing K depends on the agent's *current epistemic state* (the relation R), and what counts as the "current epistemic state" may change as a result of an update (as we move from $\mathcal{M}$ to $\mathcal{M}_{\upharpoonright Q}$). The English sentence 'Smith does not know that George Washington was the first president of the U.S.' carries this kind of indexicality. Today (an utterance of) it may express a true proposition, while tomorrow a false proposition. (Here we are assuming the so-called eternalist view of propositions expressed by tensed sentences.)

By contrast, the definition of $V_{\upharpoonright Q}(p) = V(p) \cap W_{\upharpoonright Q}$ shows that we are assuming our atomic formulas are *non-indexical*. The truth value of an atomic formula at a world cannot change as the agent's epistemic state changes from $\mathscr{M}$ to $\mathscr{M}_{|Q}$. If we think of updates as occurring in time, then the atomic formulas correspond to Quine's (1960, §40) *eternal sentences*, whose truth values do not change with time. By contrast, formulas of the form $K\varphi$ would correspond to *present tense knowledge attributions*. In Holliday et al 2013, we used the mnemonic PKEA—for *p*resent *k*nowledge *e*ternal *a*toms—to remember how to think about formulas of PAL.

This different treatment of atomic formulas and epistemic formulas has an important logical effect in PAL. Observe that the formula $(p \wedge q) \rightarrow \langle !p\rangle q$ is valid, i.e., if p and q are true, then after update with p, q is still true. But now substitute $\neg Kp$ for q to obtain $(p \wedge \neg Kp) \rightarrow \langle !p\rangle \neg Kp$. This latter formula is *not* valid. (Quite the contrary: $p \rightarrow \langle !p\rangle Kp$ is valid.) Thus, the set of valid PAL formulas is not closed under *uniform substitution*. Put differently, say that a formula is *schematically valid* iff all of its uniform substitution instances are valid. Then the set of schematically valid PAL formulas is a *proper* subset of the set of valid PAL formulas. All of the following formulas are examples (from Holliday et al 2012) of formulas that are valid but not schematically valid, where $[!\varphi]\psi := \neg\langle !\varphi\rangle\neg\psi$:

$$
\begin{array}{ll}
[!p]p & K_a p \rightarrow [!p]K_a p \\
{[!p]K_a p} & K_a p \rightarrow [!p](p \rightarrow K_a p) \\
{[!p](p \rightarrow K_a p)} & K_a(p \rightarrow q) \rightarrow (\langle !q\rangle K_a r \rightarrow \langle !p\rangle K_a r) \\
{[!p \wedge \neg K_a p]\neg(p \wedge \neg K_a p)} & (\langle !p\rangle K_a r \wedge \langle !q\rangle K_a r) \rightarrow \langle !p \vee q\rangle K_a r.
\end{array}
$$

That these formulas are valid but not schematically valid is cautionary: it shows that general principles about information dynamics that one might have intuitively accepted are prone to falsification when we plug in epistemic formulas (see Holliday et al 2013 for discussion). We will return to this point in Example 1 below.

The set of *valid* PAL formulas can be axiomatized as follows (cf. Plaza 1989). It is the smallest set $\mathbf{L} \subseteq \mathscr{L}_{PAL}$ that contains all uniform substitution instances of

- classical propositional tautologies
- $K_a(p \rightarrow q) \rightarrow (K_a p \rightarrow K_a q)$ and $K_a p \rightarrow p$
- $\langle !p\rangle\neg q \leftrightarrow (p \wedge \neg\langle !p\rangle q)$ and $\langle !p\rangle(q \wedge r) \leftrightarrow (\langle !p\rangle q \wedge \langle !p\rangle r)$
- $\langle !p\rangle K_a q \leftrightarrow (p \wedge K_a(p \rightarrow \langle !p\rangle q))$,

for all $a \in \mathsf{Agt}$, and contains

- $\langle !\varphi\rangle p \leftrightarrow (\varphi \wedge p)$

for all $\varphi \in \mathscr{L}_{PAL}$ and $p \in \mathsf{At}$, while being closed in the following ways:

- if $\varphi \in \mathbf{L}$, then $K_a\varphi \in \mathbf{L}$
- if $\psi \leftrightarrow \chi \in \mathbf{L}$, then $\varphi[\psi/p] \leftrightarrow \varphi[\chi/p] \in \mathbf{L}$.

The set of *schematically valid* PAL formulas (for Agt infinite) was shown to be decidable in Holliday et al 2011, 2013 and was finitely axiomatized in Holliday et al 2012 with a system of Uniform Public Announcement Logic (UPAL). UPAL was also shown to axiomatize the set of validities for an alternative semantics. In that semantics, atomic formulas are treated as genuine propositional variables, standing in for arbitrary formulas; thus, the truth value of an atomic formula—like an epistemic formula—can change across the transitions associated with update operators.

Another important point to make about PAL is that the defined notion of *updating* with a proposition Q is a strong notion. Not only does the agent come to know the proposition, so $R_{a\restriction Q}(w) \subseteq Q$, but also she comes to *know that she knows* the proposition Q, so $R_{a\restriction Q}(w) \subseteq \{v \in W \mid R_{a\restriction Q}(v) \subseteq Q\}$, and so on up to every level. In light of the arguments that an agent can know a proposition without knowing that she knows it (see, e.g., Williamson 2000), there is reason to study a weaker notion of update, where an agent could come to know P without necessarily coming to know that she knows P; but we will not study such a weaker notion here.

Finally, let us suppose there is more than one agent. Then as we go from $\mathscr{M}$ to $\mathscr{M}_{\restriction Q}$, not only does each agent come to know the proposition Q, and that she knows that she knows the proposition Q, and so on up to every level, but also each agent comes to know of each other agent that she knows Q up to every level, that each other agent knows that each other agent knows Q up to every level, and so on. In short, the proposition Q becomes *common knowledge* in the sense of Lewis 1969.

Let us now see how the PAL semantics handles a well-known example.

Example 1 (The Moore Formula). In Section 4.17 of *Knowledge and Belief*, Hintikka (1962) discusses what he calls the "Analogue to Moore's paradox for the second person." Hintikka asks us to consider the sentence 'p but you do now know that p', which he labels as sentence (52). Hintikka makes several observations about announcements of (52), including the following: "If you know that I am well-informed and if I address the words (52) to you, these words have a curious effect which may perhaps be called anti-performatory. You may come to know that what I say *was* true, but saying it in so many words has the effect of making what is being said false" (pp. 68-69). Or perhaps more carefully: it has the effect that a subsequent announcements using the same words would be false. Although Hintikka's static epistemic logic could not capture this point, the framework of PAL can capture it.

First, formalize (52) as $p \wedge \neg Kp$. Then observe that whenever $p \wedge \neg Kp$ is true at a world w in a model $\mathscr{M}$, i.e., $\mathscr{M}, w \vDash p \wedge \neg Kp$, then "announcing" $p \wedge \neg Kp$ results in a model $\mathscr{M}_{\restriction [\![p\wedge\neg Kp]\!]^{\mathscr{M}}}$ in which p is *known*, i.e., $\mathscr{M}_{\restriction [\![p\wedge\neg Kp]\!]^{\mathscr{M}}}, w \vDash Kp$, which means that $p \wedge \neg Kp$ becomes *false*, i.e., $\mathscr{M}_{\restriction [\![p\wedge\neg Kp]\!]^{\mathscr{M}}}, w \vDash \neg(p \wedge \neg Kp)$.

For concreteness, consider a model $\mathscr{M}$ with just two worlds w and v, the accessibility relation R as the universal relation, and p true only at w. Hence $\mathscr{M}, w \vDash p \wedge \neg Kp$. Then the updated model $\mathscr{M}_{\restriction [\![p\wedge\neg Kp]\!]^{\mathscr{M}}}$ contains just the world w with a reflexive accessibility relation, and p is still true at w, so $\mathscr{M}_{\restriction [\![p\wedge\neg Kp]\!]^{\mathscr{M}}}, w \vDash Kp$.

Since the observation holds for any model $\mathscr{M}$ such that $\mathscr{M}, w \vDash p \wedge \neg Kp$, the principle $(p \wedge \neg Kp) \to \langle !p \wedge \neg Kp \rangle \neg(p \wedge \neg Kp)$ is a valid principle of PAL. This supports Hintikka's point about the "anti-performatory" effect of announcing (52).[3] But the support is only partial. Hintikka presumably intended 'p' in (52) to stand in for any sentence; but if so, then (52) is not always anti-performatory. As shown by the "Puzzle of the Gifts" in Holliday et al 2013, there are complex epistemic formulas φ such that after the announcement of "φ but you do not know that φ", it

[3] We will return to Hintikka's point that "You may come to know that what I say *was* true" in §5.

remains true that φ but the agent does not know that φ, so the announcement was not anti-performatory. In the context of PAL, this means that, quite surprisingly, $(p \wedge \neg Kp) \to \langle !p \wedge \neg Kp \rangle \neg (p \wedge \neg Kp)$ is not a *schematically* valid principle. This a good illustration of how we can use dynamic epistemic logic not just to formalize our preexisting intuitions, but to discover surprising counterexamples.

The final task in our review of PAL is to explain the *link-cutting* semantics mentioned above, which appears in van Benthem and Liu 2007 (cf. van Linder et al 1994; Herzig et al 2000). For this semantics, given a model $\mathscr{M} = \langle W, \{R_a\}_{a \in \mathsf{Agt}}, V \rangle$ and proposition $Q \subseteq W$, we define the update of $\mathscr{M}$ by Q, $\mathscr{M}_{|Q} = \langle W, \{R_{a|Q}\}_{a \in \mathsf{Agt}}, V \rangle$, by:

- $vR_{a|Q}u$ iff both $vR_a u$ and [$v \in Q$ iff $u \in Q$].

Thus, we cut all epistemic accessibility links between worlds in Q and worlds not in Q, but we do not throw away the latter worlds from our new model, as we did with world-elimination. Now the semantic clause for update formulas is:

- $\mathscr{M}, w \vDash \langle !\varphi \rangle \psi$ iff $\mathscr{M}, w \vDash \varphi$ and $\mathscr{M}_{|[\![\varphi]\!]^{\mathscr{M}}}, w \vDash \psi$.

We mentioned above that world-elimination and link-cutting are equivalent for $\mathscr{L}_{PAL}$. In particular, if we let $\vDash_e$ be the satisfaction relation for world-elimination semantics and $\vDash_c$ for the link-cutting semantics, then it is easy to check that for any pointed epistemic model $\mathscr{M}, w$ and $\varphi \in \mathscr{L}_{PAL}$: $\mathscr{M}, w \vDash_e \varphi$ iff $\mathscr{M}, w \vDash_c \varphi$.[4] Thus, the set of valid formulas is the same for world-elimination and link-cutting.

An important point about link-cutting is that the models $\mathscr{M}_{|Q}$ and $\mathscr{M}_{|W \setminus Q}$ are the *same*. Link-cutting represents an event that truly answers the question of *whether or not* Q from $\mathscr{M}$ to $\mathscr{M}_{|Q}$, or equivalently, from $\mathscr{M}$ to $\mathscr{M}_{|W \setminus Q}$. At each world w in Q, all agents come to know Q, i.e., $R_{a|Q}(w) \subseteq Q$. But at each world v not in Q, all agents come to know *not* Q, i.e., $R_{a|Q}(v) \subseteq W \setminus Q$. Of course, since worlds in Q and worlds in $W \setminus Q$ become disconnected in $\mathscr{M}_{|Q}$, whatever agents come to know at a world $v \in W \setminus Q$ is irrelevant to the truth values of $\mathscr{L}_{PAL}$ formulas at a world $w \in Q$ in $\mathscr{M}_{|Q}$—hence the previous paragraph. But when we move to a more expressive language in §4, what happens at worlds in $W \setminus Q$ will no longer be irrelevant.

3 Paradoxes and Problems

In this section, we explain how basic dynamic epistemic logic has been used to analyze two well-known "epistemic paradoxes": Fitch's Paradox of Knowability (§3.1) and the Paradox of the Surprise Examination (§3.2). We highlight the valuable ideas in these analyses, as well as ways in which these analyses could be improved.

[4] An analogous statement is true for a third semantics for PAL based on *arrow-elimination* (as in Gerbrandy and Groeneveld 1997 and Kooi 2007), where the new relation $R_{a \uparrow Q}$ is defined from the original relation R_a by: $vR_{a \uparrow Q}u$ iff both $vR_a u$ and $u \in Q$. Note, however, that non-trivial arrow-elimination will turn the reflexive R_a into a non-reflexive $R_{a \uparrow Q}$, so the updated model will not be an epistemic model. For this reason, it is used to model the updating of belief rather than knowledge.

3.1 Knowability

What is known as Fitch's paradox originates in the following statement of Fitch (1963): "THEOREM 5. 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" (p. 139). Contrapositively: if every true proposition is such that somebody *can* know it to be true, then every true proposition is such that somebody knows, has known, or will know it to be true—a surprising result!

The argument is often formalized in modal logic as follows. Consider a propositional modal language $\mathscr{L}(\Diamond, \mathrm{K})$ with two modal operators $\Diamond$ and K. The principle that every true proposition can be known is formalized by the schema:

$$\varphi \to \Diamond \mathrm{K}\varphi. \tag{1}$$

The principle that every truth is sometime known is formalized by the schema:

$$\psi \to \mathrm{K}\psi. \tag{2}$$

Consider a set $\mathbf{L}$ of $\mathscr{L}(\Diamond, \mathrm{K})$ formulas such that $\mathbf{L}$ contains all substitution instances of classical propositional tautologies and such that for all formulas α and β:

(i) $\mathrm{K}(\alpha \wedge \beta) \to (\mathrm{K}\alpha \wedge \mathrm{K}\beta) \in \mathbf{L}$;
(ii) $\mathrm{K}\alpha \to \alpha \in \mathbf{L}$;
(iii) if $\neg\alpha \in \mathbf{L}$ then $\neg\Diamond\alpha \in \mathbf{L}$.

Give a set $\Gamma \cup \{\varphi\}$ of formulas, let $\Gamma \vdash_{\mathbf{L}} \varphi$ iff φ belongs to the smallest set $\Sigma \supseteq \Gamma \cup \mathbf{L}$ that is closed under modus ponens: if $\alpha \in \Sigma$ and $\alpha \to \beta \in \Sigma$, then $\beta \in \Sigma$.

Then it can be shown that for every formula ψ, there is a formula φ such that:

$$\{\varphi \to \Diamond \mathrm{K}\varphi\} \vdash_{\mathbf{L}} \psi \to \mathrm{K}\psi. \tag{3}$$

Take $\varphi := \psi \wedge \neg\mathrm{K}\psi$. Using properties (i) and (ii) of $\mathbf{L}$, it is easy to see that $\neg\mathrm{K}(\psi \wedge \neg\mathrm{K}\psi) \in \mathbf{L}$, whence $\neg\Diamond\mathrm{K}(\psi \wedge \neg\mathrm{K}\psi) \in \mathbf{L}$ by (iii). Then classical propositional reasoning give us (3). Thus, if the correct logic of $\Diamond$ and K is at least as strong as $\vdash_{\mathbf{L}}$, and if every instance of (1) is true, then so is every instance of (2).

What are we to make of this formal argument? Following Fitch, let us interpret $\mathrm{K}\varphi$ as meaning that someone knows, has known, or will know φ. Then the idea that all instances of (2) are true seems absurd, so what went wrong?

Given our intended interpretation of $\mathrm{K}\varphi$, property (i) of $\mathbf{L}$ seems unimpeachable. So does property (ii)—provided we think of α as a sentence such that if it expresses a true (resp. false) propositions at one time, then it does so at all times. Note that this is not how formulas are typically understood in temporal logics with past and future operators P and F. If we take the italic $K\varphi$ to mean that someone knows φ, $PK\varphi$ to mean that someone knew φ, and $FK\varphi$ to mean that someone will know φ, then we could think of our roman $\mathrm{K}\varphi$ as $K\varphi \vee PK\varphi \vee FK\varphi$. Then we would not want to add to our temporal logic the principle that $\mathrm{K}p \to p$, or informally, "if p is ever known to express a truth, then p expresses a truth now." For further discussion of this point, see Burgess 2009.[5] So let us assume that all our sentences are such that if they express a true (resp. false) proposition at one time, they do so at all times. Note that if α and β have this property, then so do all Boolean combinations of α and β, and so does $\mathrm{K}\alpha$, given the interpretation of $\mathrm{K}\varphi$ in the previous paragraph.

What about $\Diamond$? Some commentators on Fitch have been tempted to read the $\Diamond$ as a kind of metaphysical "it could have been the case that...". Then the compound $\Diamond\mathrm{K}\varphi$ would mean that it could have been the case that someone at some time knew φ. An arguably more interesting reading is the one suggested by Fitch's original language: $\Diamond\mathrm{K}\varphi$ means that someone *can* sometime know φ. To say that someone can know φ is not the same as saying that it could have been that someone knew φ. An entailment from 'can' to 'could have been' is plausible, but the converse is not. For 'can' takes into account, at least to some extent, constraints imposed by contingent features of the actual world. It is debatable whether the notion of *can know* is factorable into some kind of *can* modality $\Diamond$ and a *know* modality K. As an approximation, we might take $\Diamond\varphi$ to mean that (at some time) it can be brought about that φ; so $\Diamond\mathrm{K}\varphi$ would mean that (at some time) it can be brought about that someone (at some time) knows φ. Under this reading of $\Diamond$, property (iii) of **L** above seems uncontroversial: one cannot bring it about that α if $\neg\alpha$ is a logical truth.

To block the derivation that all instances of (2) are true, the obvious move is to deny that all instances of (1) are true. Not all truths can be known. The principle that all truths can be known is implied by certain anti-realist or verificationist views about truth. So much the worse for these views, one might say.

Rather than dwelling on Fitch's proof as a refutation (or not) of certain anti-realist views, we can take Fitch's proof as inspiration for the study of different notions of *knowability* and the associated limits of knowledge (cf. Williamson 2000, Ch. 12).

This is where dynamic epistemic logic enters the story, with the important notion of knowability proposed by van Benthem (2004). (For simplicity, let us imagine there is just one agent, and we are asking what is knowable for that agent.) In essence, van Benthem's idea is to read $\Diamond\varphi$ as "there is a possible change from the agent's current epistemic state to a new epistemic state such that if the change occurs, then φ is true." So we read the compound $\Diamond K\varphi$ as "there is a possible change from the agent's current epistemic state to a new epistemic state such that if the change occurs, then the agent knows φ." Thus, van Benthem construes knowability as what one may come to know from one's current epistemic state—not from some counterfactual epistemic state that one could have had but doesn't.

Technically, van Benthem's proposal can be seen as extending $\mathscr{L}_{PAL}$ to a language $\mathscr{L}_{APAL}$ with an operator $\langle ! \rangle$, interpreting $\langle ! \rangle\varphi$ to mean that there is a formula ψ of $\mathscr{L}_{PAL}$ (or equivalently, of $\mathscr{L}_{EL}$) such that updating the agent's current epistemic state with ψ results in a new epistemic state relative to which φ is true:

- $\mathscr{M}, w \vDash \langle ! \rangle\varphi$ iff $\exists\psi \in \mathscr{L}_{PAL}$: $\mathscr{M}_{|[\![\psi]\!]^{\mathscr{M}}}, w \vDash \varphi$.

The logic given by PAL semantics plus this semantics for $\langle ! \rangle$ is the Arbitrary Public Announcement Logic (APAL) of Balbiani et al 2008 (cf. van Ditmarsch et al 2011).

[5] Also see Proietti and Sandu 2010 and Wansing 2015 on the role of time in knowability principles.

So far, so good. A difficulty arises, however, in trying to apply this formalism to express the idea that if a proposition is true, then it is knowable. It is suggested in the literature that this is expressed by $\varphi \to \langle ! \rangle K\varphi$ in APAL. This, however, is not quite right. The problem is similar to the one pointed out by Lewis in the following:

> I say (1) pigs fly; (2) what I just said had fewer than three syllables (true); (3) what I just said had fewer than four syllables (false). So 'less than three' does not imply 'less than four'? No! The context switched midway, the semantic value of the context-dependent phrase 'what I just said' switched with it. (Lewis, 1996, p. 564)

Here is the similarity: since the meaning of the K operator in APAL involves indexicality—$K\varphi$ means that in the agent's *current* epistemic state, she knows φ—and since the $\langle ! \rangle$ operator in APAL shifts the index for that indexical—it shifts what the current epistemic state is—the crucial formula

$$(p \wedge \neg Kp) \to \langle ! \rangle K(p \wedge \neg Kp)$$

expresses, according to the APAL interpretation: "if p is true but in the agent's current epistemic state, she doesn't know p, then after some change in the agent's epistemic state, the agent knows that [p is true but in the agent's current epistemic state, she doesn't know p]." But the second occurrence of 'current epistemic state' refers to something different than the first. The context switched midway.

Thus, the "knowability principle" $\varphi \to \langle ! \rangle K\varphi$ of APAL is not capturing the idea that if a proposition is true, then one can come to know *that* proposition; instead, it is capturing the idea that if a proposition is true, one can come to know some *perhaps different* proposition that is expressed by the same indexical sentence in a different context. That may be interesting, but it is not the central knowability principle.

Fortunately, there is a solution. Roughly, we need an operator Y such that when we interpret $\varphi \to \langle ! \rangle KY\varphi$, $Y\varphi$ will express relative to the new epistemic state induced by $\langle ! \rangle$ the same proposition that φ expressed relative to the old epistemic state (now thinking in terms of link-cutting update, where we do not eliminate worlds—for eliminating worlds could make it impossible to express the same proposition). The choice of the letter Y is no accident—it suggests the *yesterday* operator of temporal logic. Extensions of dynamic epistemic logic with devices similar to a yesterday operator are proposed in Hoshi and Yap 2009; Renne et al 2009; Sack 2010; Yap 2011; Renne et al 2016. In §4, we will implement the idea of a yesterday operator in an especially simple way within the framework of sequential epistemic logic.

What the discussion of this section shows is the *need for the past* in dynamic epistemic logic. In the next section, we shall see the *need for the future*.

3.2 Surprise

Discussions of the Surprise Exam Paradox, also known as the Prediction Paradox, date from at least the 1940s (see Sorensen 1988, §7). Since then, a number of variants of the paradox have appeared, designed to block purported solutions to the original paradox. In this paper, it will best suit our purposes to consider the variant that Sorensen (1984) calls the *designated student paradox*:

> Consider the designated student paradox. Here, only one examination is to be given to one of five students: Art, Bob, Carl, Don, Eric. The teacher lines them up alphabetically so that Eric can see the backs of each of the four students in front of him, Don can see the backs of the three students in front of him (but not Eric's since Eric stands behind him), and so on. The students are then shown four silver stars and one gold star. One star is put on the back of each student. The teacher then announces that the gold star is on the back of the designated student. He informs them that the designated student must take the examination. The examination is unexpected in the sense that the designated student will not know he is the designated student until after the students break formation. One of the students objects that the examination is impossible. 'We all know that Eric is not the designated student since, if he were, he would see four silver stars in front of him and deduce that he must have had the gold star on his back. But then he would know that he was the designated student. The designated student cannot know he is the designated student; contradiction. We all know that Don cannot be the designated student since, if he were, he would see three silver stars in front of him, and since he knows by the previous deduction that Eric has the remaining silver star, he would be able to deduce that he is the designated student. In a similar manner, Carl, Don, and Art can be eliminated. Therefore, the examination is impossible.' The teacher smiles, has them break formation, and Carl is surprised to learn that he has the gold star, and so is the designated student, and so must take the examination. (Sorensen, 1984, p. 357)

What went wrong, then, with the student's reasoning?

In Holliday 2016b, I argue for an answer to this question using an analysis with *static* multi-agent epistemic logic. Here I will review the analysis with *dynamic* epistemic logic due to Gerbrandy (2007). Gerbrandy focuses on the original version of the surprise exam paradox, where the five students Art, Bob, Carl, Don, and Eric are replaced by a single student who could have an exam on Monday, Tuesday, Wednesday, Thursday, or Friday. But if Gerbrandy's analysis gets to the heart of the original paradox, then a similar analysis should apply to the designated student paradox (and indeed, Gerbrandy (2007, p. 26) refers to the designated student paradox).

Let us apply to the designated student paradox the analogue of the analysis of the surprise exam paradox in Section 4 of Gerbrandy 2007, which uses the *semantics* of PAL.[6] For a semantic analysis, the first step is to draw an appropriate epistemic model representing what the students know just after "One star is put on the back of each student" in Sorensen's description above, but before the teacher announces that the student with the gold star will not know that he or she has it until the students break formation. For simplicity, let us consider just three students: 1, 2, and 3.

A natural candidate for the model is the $\mathscr{M}$ shown on the left of Figure 1. Each world is identified with the set of atomic formulas true at that world. The atomic formula g_i means that student i has the gold star on his or her back. Where R_i is the epistemic accessibility relation for student i, we assume that for each world w, we have wR_iw; but to reduce clutter we do not draw these reflexive loops in the diagram. Only the relations R_1 and R_2 relate distinct worlds, representing the uncertainty of students 1 and 2, as shown in the diagram. Note that in this model, it is common

---

[6] By contrast, the analysis in Section 3 of Gerbrandy 2007 involves syntactic derivations, as does the analysis in Holliday 2016b.

knowledge that someone has the gold star. Although Sorensen's description does not say explicitly that the students know of each other that they all saw the silver stars and one gold star (it just says "The students are then shown four silver stars and one gold star"), let us assume that this fact and the fact that the teacher distributed the stars on the students' backs are common knowledge. Then the model says that whichever world is actual, student 3 knows who has the gold star, which seems right. In no case does student 1 know who has the gold star, which also seems right. And it is only if the gold star is on the back of student 1 that student 2 knows who has the gold star; otherwise 2 is uncertain whether he has it or 3 has it, which also seems right. Finally, we assume that all of this is common knowledge among the students.

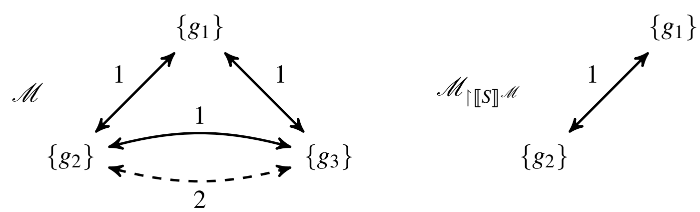

Fig. 1 Models for the designated student paradox.

The next step in the dynamic analysis is to formalize the teacher's announcement that the student with the gold star will not know that he or she has it until after the students break formation. A first try would be the formula

$$S := (g_1 \wedge \neg K_1 g_1) \vee (g_2 \wedge \neg K_2 g_2) \vee (g_3 \wedge \neg K_3 g_3).$$

Before assessing whether S is a faithful formalization, let us see what happens to our initial model when S is "announced." Observe that S is true at $\{g_1\}$ and $\{g_2\}$ but false at $\{g_3\}$ in $\mathscr{M}$ in Figure 1, since $g_3 \wedge K_3 g_3$ is true at $\{g_3\}$. Thus, the result of announcing S is the model $\mathscr{M}_{\upharpoonright [\![S]\!]^{\mathscr{M}}}$ on the right of Figure 1.

Now note what has happened. While S was *true* at $\{g_2\}$ in $\mathscr{M}$, that same S has become *false* at $\{g_2\}$ in $\mathscr{M}_{\upharpoonright [\![S]\!]^{\mathscr{M}}}$. Thus, if $\{g_2\}$ is the actual world, then S initially expressed a true proposition, but after the announcement of S, S expresses a false proposition; and whichever is the actual world, after the announcement of S, student 1 does not know the proposition expressed by S. This, Gerbrandy's analysis suggests, is the key to solving the paradox. What he says about the surprise exam paradox for one student, Marilyn, and three days, Wednesday, Thursday, and Friday, can be applied to the designated student paradox with three students, 1, 2, and 3:

> When Marilyn learns that S is true, she eliminates the world in which S is false from her information state. The state that results is $\{s_{\mathsf{we}}, s_{\mathsf{th}}\}$. Now, if the exam is actually on Thursday, it is not a surprise anymore: $\{s_{\mathsf{we}}, s_{\mathsf{th}}\}, s_{\mathsf{th}} \models \neg\mathsf{S}$. However, if the exam is given on Wednesday, it will remain to be a surprise. In either case, the sentence is not successful: Marilyn does not know whether the exam will be a surprise or not, even if she just learned that it would be.

> If S correctly paraphrases the teacher's announcement, then Marilyn's reasoning is cut short after having excluded the last day as the day of the exam. She continues her argument by reasoning that the exam cannot be on Thursday either, because that would contradict the claim of the teacher that the exam comes as a surprise. To be sure, she is correct in concluding that, now, after the announcement, it will not be a surprise if the exam is on Thursday, and she is correct in that the teacher said that it would be, but she is not correct in seeing a contradiction between these two claims. If the exam is on Thursday, then S is true before the teacher makes his announcement, but it becomes false after she learns of its truth. This may be confusing, but it is not paradoxical. (Gerbrandy, 2007, pp. 26-27)

Similarly, the analysis of the designated student paradox would be that if the gold star is on student 2's back, then the sentence S is true before the teacher makes the announcement, but it becomes false after the announcement. The mistake of the student's reasoning, then, is to assume that S is still true.

As elegantly simple as this analysis is, unfortunately there is a problem. The problem is that we have not correctly formalized the teacher's announcement. The teacher does not announce a sentence like "None of you *now* knows that you have the gold star on your back." That would indeed be a sentence such that after it is truly announced, it would express a false proposition, in the case where student 2 has the gold star. But that is not even in the ballpark of the teacher's announcement. The teacher announces something like: "None of you will know that you have the gold star *until you break formation*" which entails that they will not know *even after that very announcement*. (Similarly in the original paradox, the teacher says, "You will not know *until the time of the exam* which day the exam is on," which entails that the student will not know *even after that very announcement*.) Indeed, the teacher could redundantly add: "None of you will know that you have the gold star, *even after this very announcement*, until you break formation." Gerbrandy (2007, Section 5) recognizes that the formalization of the teacher's announcement with his sentence S does not capture the *even after this announcement* aspect; but also there seems to be no one way to capture exactly that content in the language of PAL.

We shall see that in the framework of *sequential* epistemic logic, we can formalize an announcement such as "None of you will know that you have the gold star *even after this announcement*." We will do so using a *next time* operator X, such that announcing $X\varphi$ amounts to announcing "after this announcement, φ will be true." A similar approach is sketched in the lecture slides of Baltag and Smets (2010) (see Marcoci 2010 for discussion), who analyze the surprise exam paradox using plausibility models for conditional belief, adding a next time operator to capture the teacher's announcement.[7] (As we stress in §6, the idea of sequential epistemic logic can be implemented with types of models and model transformations other than the epistemic models and updates of PAL—including belief revision models.)

Our analysis of the paradox in §5 will differ from the analysis above. While the quoted analysis from Gerbrandy supports the elimination of an exam on the last day, given the teacher's announcement, the analysis in §5 will *not* support the conclusion that student 2 can eliminate the possibility of a gold star on student 3, given

[7] Also cf. van Ditmarsch et al 2013, which uses an existential branching next-time operator, parametrized by epistemic actions, in connecting dynamic epistemic and epistemic temporal logic.

the teacher's announcement. The reason is that while the analysis above treats the teacher's announcement as an *unsuccessful* announcement, our analysis will treat the teacher's announcement as what we call an *unassimilable* announcement.

4 Sequential Epistemic Logic

In this section, we illustrate the general idea of sequential epistemic logic by introducing a sequential epistemic analogue of APAL, which we will call **SPAL**.

Let At and Agt be the same sets we used in defining $\mathscr{L}_{PAL}$. The **language of SPAL**, $\mathscr{L}_{SPAL}$, is generated by the following grammar:

$$\varphi ::= p \mid \neg\varphi \mid (\varphi \wedge \varphi) \mid X\varphi \mid Y\varphi \mid F\varphi \mid P\varphi \mid K_a\varphi$$
$$\langle\varphi\rangle\varphi \mid \langle !\varphi,\dots,\varphi\rangle\varphi \mid \langle\rangle\varphi \mid \langle !_n\rangle\varphi,$$

where $p \in \mathsf{At}$, $a \in \mathsf{Agt}$, and $n \in \mathbb{N}$. On the first line, we have familiar operators from propositional temporal logic, namely *next time* X, *previous time* Y, *future* F, and *past* P, as well as the usual knowledge operator K_a for agent a. On the second line, the operator $\langle\varphi\rangle$ is what we call a *descriptive update* operator, explained in §4.1, while the operator $\langle !\varphi_1,\dots,\varphi_n\rangle$ is what we call a *hypothetical update* operator, explained in §4.2. The operators $\langle\rangle$ and $\langle !_n\rangle$ are the "arbitrary" versions of $\langle\varphi\rangle$ and $\langle !\varphi_1,\dots,\varphi_n\rangle$, respectively, which will also be explained in §§4.1-4.2. It is important to note that we will call X, F, $\langle\varphi\rangle$, and $\langle\rangle$ the *futuristic* operators.

Other boolean connectives and temporal operators (H, G) are defined as usual. It should be stressed that the reader may plug in a more expressive temporal base language, e.g., including *since* and *until* operators, hybrid tense logic operators, etc. The interpretation of the update operators in the second line above, as described in §§4.1-4.2, will remain the same regardless of the temporal base.

Before giving the official semantics of the update operators, we can already note that syntactically we will consider PAL formulas of the form $\langle !\varphi_1\rangle\psi$ as SPAL formulas of the form $\langle !\varphi_1,\dots,\varphi_n\rangle\psi$ with $n = 1$, so we will take $\mathscr{L}_{PAL}$ to be a fragment of $\mathscr{L}_{SPAL}$. We will also consider APAL formulas of the form $\langle !\rangle\psi$ as SPAL formulas of the form $\langle !_1\rangle\psi$, so we will also take $\mathscr{L}_{APAL}$ to be a fragment of $\mathscr{L}_{SPAL}$.

A **base model** is a pair $M = \langle W, V\rangle$ where W is a nonempty set and $V\colon \mathsf{At} \to \wp(W)$. An alternative setup takes $V\colon \mathsf{At} \to \wp(W \times \mathbb{N})$, allowing the truth values of atomic formulas to vary over time, as in the semantics of UPAL (Uniform Public Announcement Logic) mentioned in §2. But in this paper, we will follow traditional PAL in treating atomic formulas as eternal sentences.

A **sequential epistemic model** is a pair $\mathscr{S} = \langle M, \sigma\rangle$ where $M = \langle W, V\rangle$ is a base model and σ is an ω-sequence $\langle R_0, R_1, R_2, \dots\rangle$ where for each $t \in \mathbb{N}$, R_t is a function assigning to each $a \in \mathsf{Agt}$ a binary relation R^a_t on W. Intuitively:

- R^a_t is agent a's epistemic accessibility relation at time t.

An alternative setup takes R^a_t to be a binary relation on $W \times \mathbb{N}$, allowing an agent to be uncertain at a time t about what time it is. But in this paper, we will assume *synchronicity*: agents know what time it is (cf. van Benthem et al 2009).

The **truth clauses** for $\mathscr{L}_{SPAL}$ formulas without update operators are standard (the reason for displaying σ on the left of the turnstile will become clear in §§4.1-4.2):

1. $M,w,t,\sigma \vDash p$ iff $w \in V(p)$;
2. $M,w,t,\sigma \vDash \neg\varphi$ iff $M,w,t,\sigma \nvDash \varphi$;
3. $M,w,t,\sigma \vDash \varphi \wedge \psi$ iff $M,w,t,\sigma \vDash \varphi$ and $M,w,t,\sigma \vDash \psi$;
4. $M,w,t,\sigma \vDash X\varphi$ iff $M,w,t+1,\sigma \vDash \varphi$;
5. $M,w,t,\sigma \vDash Y\varphi$ iff $t = 0$ or $M,w,t-1,\sigma \vDash \varphi$;
6. $M,w,t,\sigma \vDash F\varphi$ iff $\exists t' > t$: $M,v,t',\sigma \vDash \varphi$;
7. $M,w,t,\sigma \vDash P\varphi$ iff $\exists t' < t$: $M,v,t',\sigma \vDash \varphi$;
8. $M,w,t,\sigma \vDash K_a\varphi$ iff $\forall v \in W$: if $wR^a_t v$, then $M,v,t,\sigma \vDash \varphi$.

In the next two subsections, we will describe the key ideas of SPAL: the semantics for the operators $\langle\varphi\rangle$ and $\langle !\varphi_1,\dots,\varphi_n\rangle$ (and their "arbitrary" versions $\langle\rangle$ and $\langle !_n\rangle$).

Already we should say that SPAL validity will be defined as expected: φ is SPAL-valid ($\vDash_{SPAL} \varphi$) iff $M,w,t,\sigma \vDash \varphi$ for every base model $M = \langle W,V\rangle$, sequential epistemic model $\langle M,\sigma\rangle$, $w \in W$, and $t \in \mathbb{N}$; and φ is SPAL-satisfiability iff $\nvDash_{SPAL} \neg\varphi$.

Finally, for continuity with our discussion of PAL, we will continue to use the term 'proposition' for a set of worlds. Thus, in SPAL semantics, a non-atomic φ may express different propositions $[\![\varphi]\!]^{\mathscr{S},t} = \{w \in W \mid M,w,t,\sigma \vDash \varphi\}$ at different times t. We can also consider the set $[\![\varphi]\!]^{\mathscr{S}} = \{\langle w,t\rangle \in W \times \mathbb{N} \mid M,w,t,\sigma \vDash \varphi\}$ of world-time pairs at which φ is true, but we reserve 'proposition' for the set of worlds.

4.1 The Descriptive Update Operator

In SEL, there is a basic distinction between *descriptive* and *hypothetical* operators. Descriptive operators allow us to describe the *actual sequence* of epistemic changes. For example, we may wish to make the following descriptive claim about the actual sequence: (i) the agents' epistemic states at time $t+1$ are obtained from those at t by everyone publicly learning whether φ held at t. To keep the syntax of SPAL close to that of PAL, we will use formulas $\langle\varphi\rangle\psi$ to express a bit more: (ii) φ is true at $\langle w,t\rangle$, the agents' epistemic states at $t+1$ are obtained from those at t by everyone publicly learning whether φ held at t, and ψ is true at $\langle w,t+1\rangle$. Formally, we give the following semantics for our *descriptive* update operator $\langle\varphi\rangle$.

Definition 1 (Descriptive Update). $M,w,t,\sigma \vDash \langle\varphi\rangle\psi$ iff the following hold:

1. $M,w,t,\sigma \vDash \varphi$;
2. for each $a \in \mathsf{Agt}$, R^a_{t+1} is the set of all pairs $\langle v,u\rangle \in W \times W$ such that

a. $vR^a_t u$ and

b. $M,v,t,\sigma \vDash \varphi$ iff $M,u,t,\sigma \vDash \varphi$;

3. $M,w,t+1,\sigma \vDash \psi$.

Note that condition 2 says that R^a_{t+1} is obtained from R^a_t by link-cutting update with φ, as in §2. This leads to several other notes about the definition.

First, we have $M,w,t,\sigma \vDash \langle\varphi\rangle\top \vee \langle\neg\varphi\rangle\top$ iff condition 2 above holds, so the formula $\langle\varphi\rangle\top \vee \langle\neg\varphi\rangle\top$ expresses the claim (i) above that the agents' epistemic states at $t+1$ are obtained from those at t by everyone publicly learning whether φ held at t. Let us use the abbreviation $?\varphi$ for $\langle\varphi\rangle\top \vee \langle\neg\varphi\rangle\top$. Then $\langle\varphi\rangle\psi$ is equivalent to $\varphi \wedge ?\varphi \wedge X\psi$. Thus, we could have instead started with an operator ?, such that $M,w,t,\sigma \vDash ?\varphi$ iff condition 2 above holds, and then treated $\langle\varphi\rangle\psi$ as defined. But again, we start with $\langle\varphi\rangle\psi$ to stay close to the familiar syntax of PAL.

Second, if $M,w,t,\sigma \vDash ?\varphi$, then it is common knowledge at t that the question of *whether φ held at t* will be answered from t to $t+1$. Since we did not put common knowledge in our language, we will express this fact as follows.

Proposition 1 (Common Knowledge of Upcoming Updates). *If $M,w,t,\sigma \vDash ?\varphi$, then for any sequence $a_1,\dots,a_n \in \mathsf{Agt}^*$,*

$$M,w,t,\sigma \vDash K_{a_1}\dots K_{a_n}((\varphi \to \langle\varphi\rangle\top) \wedge (\neg\varphi \to \langle\neg\varphi\rangle\top)).$$

The *descriptive* character of $\langle\varphi\rangle$ leads to quite different logical behavior than that of the standard PAL operator, as shown by the following examples.

Example 2. According to PAL semantics, $p \to \langle !p\rangle\top$ is schematically valid. Whatever is true *can be* truly announced. But we do not want to say that whatever is true *is in fact* truly announced. Indeed, for the descriptive operator, $p \to \langle p\rangle\top$ is not even a SPAL validity. Just because p is true, it does not follow that the next epistemic state in the actual sequence is obtained from the present epistemic state by update with p. Note, by contrast, that $\langle p\rangle\top \to p$ is a SPAL schematic validity.

Example 3. According to PAL semantics, $\langle !p\rangle\langle !q\rangle r$ is schematically equivalent to $\langle !p \wedge \langle !p\rangle q\rangle r$. Whatever *can be* accomplished with two consecutive announcements also *can be* accomplished with one. But we would not way to say that whatever *is* accomplished with two consecutive announcement *is* accomplished with one, which almost sounds contradictory. Indeed, for the descriptive operator, none of the following formulas is equivalent to either of the others:

1. $\langle p\rangle\langle q\rangle\top$; 2. $\langle p \wedge \langle p\rangle q\rangle\top$; 3. $\langle p \wedge q\rangle\top$.

The difference between formulas 1 and 3 should be clear. The difference between formula 2 and the others should also be clear when one observes that $\neg K(p \to q) \to \neg\langle p \wedge \langle p\rangle q\rangle\top$ is valid. To see this, note that for $\langle p \wedge \langle p\rangle q\rangle\top$ to be true, $\langle p\rangle q$ must be true, so the next epistemic state must be obtained by update with p. In addition, for $\langle p \wedge \langle p\rangle q\rangle\top$ to be true, the next epistemic state must be obtained by update with $p \wedge \langle p\rangle q$. But if $\neg K(p \to q)$ is true, then update with p is not equivalent to update with $p \wedge \langle p\rangle q$. So in this case $\langle p \wedge \langle p\rangle q\rangle\top$ cannot be true.

Next, observe that versions of the standard PAL recursion axioms are SPAL valid.

Proposition 2 (Recursion Axioms). *For any* $\varphi, \psi, \chi \in \mathscr{L}_{SPAL}$ *and* $p \in \mathsf{At}$*, the following are SPAL validities:*

1. $\langle\varphi\rangle p \leftrightarrow (\langle\varphi\rangle\top \wedge p)$;
2. $\langle\varphi\rangle(\psi \wedge \chi) \leftrightarrow (\langle\varphi\rangle\psi \wedge \langle\varphi\rangle\chi)$;
3. $\langle\varphi\rangle\neg\psi \leftrightarrow (\langle\varphi\rangle\top \wedge \neg\langle\varphi\rangle\psi)$;
4. $\langle\varphi\rangle K\psi \leftrightarrow (\langle\varphi\rangle\top \wedge K(\varphi \to \langle\varphi\rangle\psi))$.

Unlike in PAL, in SPAL we cannot reduce every formula to one without update operators. What blocks the usual reduction strategy is the fact, which we saw above, that $\langle p\rangle\top \leftrightarrow p$ is not valid. The left-to-right direction is, but the right-to-left direction is not. As we also saw, $\langle\varphi\rangle\psi$ is equivalent to $\varphi \wedge ?\varphi \wedge X\psi$, so every formula can be reduced to one in which all descriptive operators are followed by $\top$, as in $\langle\alpha\rangle\top$. An important SPAL schematic validity involving such formulas is

$$\langle p\rangle\top \to XKYp.$$

But there is no formula φ without update operators such that $\varphi \to \langle p\rangle\top$ is valid.[8]

Finally, let $\mathscr{L}^-_{SPAL}$ be the fragment of $\mathscr{L}_{SPAL}$ without the arbitrary update operators $\langle\rangle$ and $\langle !_n\rangle$. Then the truth clause for $\langle\rangle$ is:

- $\mathscr{M}, w, t, \sigma \vDash \langle\rangle\psi$ iff $\exists\varphi \in \mathscr{L}^-_{SPAL}$: $\mathscr{M}, w, t, \sigma \vDash \langle\varphi\rangle\psi$.

Observe that just as $\langle\varphi\rangle\psi$ is equivalent to $\varphi \wedge ?\varphi \wedge X\psi$, the formula $\langle\rangle\psi$ is equivalent to $\langle\rangle\top \wedge X\psi$, where $\langle\rangle\top$ simply says that the next epistemic state is obtained from the current one by some $\mathscr{L}^-_{SPAL}$-definable update. A more minimalist approach would therefore start with a single formula with the same semantics as $\langle\rangle\top$ and then treat formulas of the form $\langle\rangle\psi$ as defined abbreviations. Nothing will turn on this here, and we will only briefly touch upon the operator $\langle\rangle$ in §5.

Before discussing further properties of these descriptive operators, we will introduce their hypothetical siblings in the following subsection.

4.2 The Hypothetical Update Operator

In addition to describing the actual sequence of epistemic changes, we can make claims about hypothetical sequences of epistemic changes, such as the following:

- We can suppose that, instead of whatever actually happened after time t, first everyone publicly learned that φ_0 and then everyone publicly learned that φ_1—with it being common knowledge in advance that *whether* φ_0 and *whether* φ_1 would be publicly answered in that order, and then nothing else would happen. Assuming all of this, ψ would hold after the two epistemic changes.

[8] One way to see this is to note that the truth of our formulas without update operators is preserved under taking disjoint unions of models, defined in an obvious way, whereas $\langle p\rangle\top$ concerns the model globally, not just what is reachable from the point of evaluation, so it is not preserved under taking disjoint unions of models. (If we had a universal modality, there would be more to say.)

To express claims like this, we use the *hypothetical* update operator $\langle !\varphi_0, \dots, \varphi_{n-1}\rangle$, with the following semantics, where we take $n = \{0, \dots, n-1\}$.

Definition 2 (Hypothetical Update). $M, w, t, \langle R_0, R_1, \dots\rangle \vDash \langle !\varphi_0, \dots, \varphi_{n-1}\rangle\psi$ iff for each $i \in n$, there is a function $S_i\colon \mathsf{Agt} \to \wp(W \times W)$ such that for all $i \in n$:

1. $M, w, t+i, \langle R_0, \dots, R_t, S_0, \dots, S_{n-1}, S_{n-1}, \dots\rangle \vDash \varphi_i$;[9]
2. for each $a \in \mathsf{Agt}$, S_i^a is the set of all pairs $\langle v, u\rangle \in W \times W$ such that
   a. $vS_{i-1}^a u$ (where $S_{-1}^a = R_t^a$) and
   b. $M, v, t+i, \langle R_0, \dots, R_t, S_0, \dots, S_{n-1}, S_{n-1}, \dots\rangle \vDash \varphi_i$ iff
   $M, u, t+i, \langle R_0, \dots, R_t, S_0, \dots, S_{n-1}, S_{n-1}, \dots\rangle \vDash \varphi_i$;
3. $M, w, t+n, \langle R_0, \dots, R_t, S_0, \dots, S_{n-1}, S_{n-1}, \dots\rangle \vDash \psi$.

Observe what is going on here: we are asking whether there is a hypothetical future evolution of epistemic states, given by the S_i's, such that *relative to that future*, φ_0 expresses a truth at the initial time t, the next epistemic state for $t+1$ is obtained by everyone publicly learning whether φ_0, etc. The point of stressing *relative to that future* is that φ_0 may contain future operators. Thus, we are asking whether φ_0 can make a true claim about a future brought about in part by update with φ_0 itself, where the update content of φ_0 depends on that future. This sounds like it could be problematically circular. A feature of SPAL is that it can handle these questions with the well-defined formal semantics above. Note that what we have just described is exactly what is going on in the surprise exam paradox with the teacher's announcement: "You will have an exam that comes as a surprise, i.e., relative to your future epistemic state brought about in part by this very announcement." We will describe this application of SPAL to the surprise exam paradox in §5.

There are a number of further points to make about the above definition.

First, we capture PAL semantics with SPAL semantics (notation: $\vDash_{PAL}$ and $\vDash_{SPAL}$) as in the following proposition, provable by an easy induction on φ.

Proposition 3 (From PAL to SPAL). *For any base model $M = \langle W, V\rangle$, epistemic model $\mathscr{M} = \langle W, R, V\rangle$ where $R\colon \mathsf{Agt} \to \wp(W \times W)$, $w \in W$, $t \in \mathbb{N}$, and $\varphi \in \mathscr{L}_{PAL}$:*

$$\mathscr{M}, w \vDash_{PAL} \varphi \quad \textit{iff} \quad M, w, t, \langle R, R, \dots\rangle \vDash_{SPAL} \varphi.$$

Call a sequential epistemic model $\mathscr{S} = \langle W, V, \sigma\rangle$ *constant* if $\sigma(n) = \sigma(m)$ for every $n, m \in \mathbb{N}$. It is easy to see that if a $\varphi \in \mathscr{L}_{PAL}$ is SPAL-satisfiable, then it is SPAL-satisfiable in a constant model. This fact, plus Proposition 3, gives us:

Proposition 4 (Agreement over $\mathscr{L}_{PAL}$). *For any $\varphi \in \mathscr{L}_{PAL}$: $\vDash_{PAL} \varphi$ iff $\vDash_{SPAL} \varphi$.*

Although by Proposition 4, the PAL validity $p \to \langle !p\rangle\top$ is also a SPAL validity (in contrast to $p \to \langle p\rangle\top$ in Example 2), it is not a SPAL *schematic* validity. For it may be incoherent to suppose that everyone publicly learns a certain truth about the *future*. The following is one of the key examples to remember.

[9] The '$\dots$' after S_{n-1} indicates that all coordinates of the new ω-sequence are S_{n-1} thereafter, representing the supposition that "nothing else happens" after the update with φ_{n-1}.

Example 4 (An Incoherent Supposition). The formula $p \wedge X\neg Kp$ is satisfiable, but the formula $\langle !\, p \wedge X\neg Kp\rangle\top$ is unsatisfiable. For $\langle !\, p \wedge X\neg Kp\rangle\top$ to be true at $\langle w,t\rangle$, it must be possible to find a future evolution of epistemic states, given by the S_i's, starting with an update by $p \wedge X\neg Kp$, such that *relative to that future evolution*, $p \wedge X\neg Kp$ is true at $\langle w,t\rangle$. This is clearly impossible, since update by $p \wedge X\neg Kp$ leads to Kp holding at $t+1$, in which case $X\neg Kp$ does not hold at t.

When it is coherent to suppose that a certain update takes place, what we are supposing is that everyone learns that φ was true before the update. Formally,

$$\langle !p\rangle\top \to \langle !p\rangle KYp$$

is schematically valid, a point to which we will return in §5.

Next, let us observe that although $\langle !p\rangle\top \to p$ is a SPAL validity, it is not a SPAL schematic validity (whereas recall that $\langle p\rangle\top \to p$ is a SPAL schematic validity).

Example 5. The formula $\langle !\, p \wedge XKp\rangle\top \to (p \wedge XKp)$ is not a SPAL validity. Just because we can hypothetically consider a future relative to which it would be true that $p \wedge XKp$, and in which the next epistemic state would be obtained by update with $p \wedge XKp$, it does not follow that in the *actual* sequence, $p \wedge XKp$ is true.

An announcement of $p \wedge XKp$ is a kind of *self-fulfilling* announcement, in the sense that by announcing $p \wedge XKp$, one makes it true that XKp.

Example 6 (Self-fulfilling Announcements). Even more curious is an announcement of simply XKp, as in "After I say this, you will know p." Of course, such an announcement could be true, because it could be that right after the announcement, you acquire knowledge of p in some other way. More interesting is the question of whether the announcement of "After I say this, you will know p" could be the *source* of your new knowledge of p. Since we are assuming that p is an eternal sentence, "you will know p" entails p by the factivity of knowledge, so you could plausibly reason as follows: I have been told something that entails p by an authoritative source I trust, so p is true. Perhaps you could thereby acquire not only belief but also knowledge of p. Compare this with the case, used in science fiction stories, of an agent who encounters a machine that can predict the future, and the machine predicts for her that she will come to know some important proposition (tenselessly formulated, let us suppose). Does the agent thereby come to know the proposition?

In SPAL, the formula $p \to \langle !XKp\rangle Kp$ is *valid*. To see this, consider any sequential epistemic model with $w \in W$ and $t \in \mathbb{N}$ such that $M,w,t,\sigma \vDash p$. Where R_t is the agent's actual epistemic accessibility relation at time t, consider the hypothetical epistemic accessibility relation S for time $t+1$ defined by vSu iff (i) vR_tu and (ii) $M,v,t,\sigma \vDash p$ iff $M,u,t,\sigma \vDash p$. Then observe that for every $v \in W$:

$$M,v,t,\sigma \vDash p \quad \text{iff} \quad M,v,t+1,R_0,\dots,R_t,S,S,\dots \vDash Kp,$$

which implies

$$M,v,t,\sigma \vDash p \quad \text{iff} \quad M,v,t,R_0,\dots,R_t,S,S,\dots \vDash XKp.$$

It follows that we have vSu iff both $vR_t u$ and

$$M, v, t, R_0, \ldots, R_t, S, S, \cdots \vDash XKp \quad \text{iff} \quad M, u, t, R_0, \ldots, R_t, S, S, \cdots \vDash XKp.$$

In other words, we can see S as coming from R_t by update with XKp as in Definition 2. It follows from all of the above that S is a witness for the fact that $\mathscr{M}, w, t, \sigma \vDash \langle !XKp \rangle Kp$, which completes the proof that $p \to \langle !XKp \rangle Kp$ is valid.

To be careful, what this result shows is that if p is true, then there is a set-of-worlds proposition P such that if the agent's epistemic state were updated with P, then Kp would be true at $t+1$ and P would turn out to be exactly the set of worlds w such that XKp was true at w relative to t; so the sentence XKp would turn out to express the proposition P relative to t. This does not show, of course, that some particular source uttering the words "After this announcement, you will know p" would succeed in causing an agent's epistemic state to be updated with that P. Officially SPAL says nothing about what kinds of utterances would cause agents' epistemic states to be updated by propositions. It only talks about such epistemic updates themselves (recall our discussion of "public announcement" in §2).

Next, let us observe a connection between hypothetical formulas of the form $\langle !\varphi_1, \ldots, \varphi_n \rangle \psi$ and descriptive formulas of the form $\langle \varphi_1 \rangle \ldots \langle \varphi_n \rangle \psi$. Although there is no guarantee that the truth values of $\langle !\varphi_1, \ldots, \varphi_n \rangle \psi$ and $\langle \varphi_1 \rangle \ldots \langle \varphi_n \rangle \psi$ will be the same at a particular pointed model, there is the following connection.

Proposition 5 (Equisatisfiability). *For any $\varphi_1, \ldots, \varphi_n, \psi \in \mathscr{L}_{SPAL}$ that do not contain futuristic operators:*

$$\langle !\varphi_1, \ldots, \varphi_n \rangle \psi \text{ is SPAL-satisfiable} \quad \text{iff} \quad \langle \varphi_1 \rangle \ldots \langle \varphi_n \rangle \psi \text{ is SPAL-satisfiable.}$$

Proof (Sketch). If $\langle !\varphi_1, \ldots, \varphi_n \rangle \psi$ is SPAL-satisfiable, then there is a sequential epistemic model $\mathscr{S}$, world w, and time t such that the actual history σ in $\mathscr{S}$ can be changed to a *hypothetical* history σ' that witnesses the truth of $\langle !\varphi_1, \ldots, \varphi_n \rangle \psi$ at w, t, σ as in Definition 2. Now let σ' be the *actual* history of a new sequential epistemic model $\mathscr{S}'$. Then $\langle \varphi_1 \rangle \ldots \langle \varphi_n \rangle \psi$ will be true at w, t, σ' in $\mathscr{S}'$, so it is SPAL-satisfiable. (Note that this holds for all SPAL formulas whatsoever.)

In the other direction, if $\langle \varphi_1 \rangle \ldots \langle \varphi_n \rangle \psi$ is SPAL-satisfiable, so there is a sequential epistemic model $\mathscr{S}$ whose actual history σ makes the formula true at some world w and time t, then the history σ_n that is exactly like σ up to $t+n$, but then repeats with $\sigma_n(k) = \sigma(t+n)$ for all $k \geq t+n$, witnesses the truth of the hypothetical formula $\langle !\varphi_1, \ldots, \varphi_n \rangle \psi$ at w, t, σ. The "freezing" of the future after $t+n$ in the hypothetical history σ_n (hypothetical histories being essentially finite) does not affect the truth values of $\varphi_1, \ldots, \varphi_n, \psi$, since they do not contain futuristic operators. □

However, this equisatisfiability does not extend to more complex formulas, as shown by the following example.

Example 7. The formula $\neg K(p \leftrightarrow q) \wedge \langle !p \rangle \top \wedge \langle !q \rangle \top$ is *satisfiable*, but the formula $\neg K(p \leftrightarrow q) \wedge \langle p \rangle \top \wedge \langle q \rangle \top$ is *unsatisfiable*. For if p and q do not express the same proposition at t, then it cannot be both that the epistemic state at $t+1$ was obtained by update with p and that the epistemic state at $t+1$ was obtained by update with q.

It is important to observe that putting a sequence of formulas inside the hypothetical update operator is equivalent to using a sequence of one-formula operators provided that the formulas do not contain futuristic operators.

Proposition 6 (Hypothetical Sequences). *For any base model $M = \langle W, V\rangle$, sequential epistemic model $\mathscr{S} = \langle M, \sigma\rangle$, $w \in W$, $t \in \mathbb{N}$, and $\varphi_1, \ldots, \varphi_n, \psi \in \mathscr{L}_{SPAL}$ such that $\varphi_1, \ldots, \varphi_n$ contain no futuristic operators:*

$$M, w, t, \sigma \vDash \langle !\varphi_1, \ldots, \varphi_n\rangle\psi \quad \textit{iff} \quad M, w, t, \sigma \vDash \langle !\varphi_1\rangle \ldots \langle !\varphi_n\rangle\psi.$$

Proof (Sketch). From right to left, the truth of $\langle !\varphi_1\rangle \ldots \langle !\varphi_n\rangle\psi$ at w, t, σ requires that there be a hypothetical history σ_1, in which everything is like σ up until t and then the accessibility relations at $t+1$ are obtained from those at t by update with φ_1 and then nothing happens, such that $\langle !\varphi_2\rangle \ldots \langle !\varphi_n\rangle\psi$ is true at $w, t+1, \sigma_1$; and that in turn requires that there be a hypothetical history σ_2, in which everything is like σ_1 up until $t+1$ and then the accessibility relations at $t+2$ are obtained from those at $t+1$ by update with φ_2 and then nothing happens, such that $\langle !\varphi_3\rangle \ldots \langle !\varphi_n\rangle\psi$ is true at $w, t+2, \sigma_2$; and so on. Since $\varphi_1, \ldots, \varphi_n, \psi$ contain no futuristic operators, the truth of φ_1 at $t+1$ is unaffected by changing what happens after $t+1$, the truth of φ_2 at $t+2$ is unaffected by changing what happens after $t+2$, and so on. Using this fact, one can check that σ_n witnesses the truth of $\langle !\varphi_1, \ldots, \varphi_n\rangle\psi$ at w, t, σ according to Definition 2. In the other direction, any hypothetical history σ' that witnesses the truth of $\langle !\varphi_1, \ldots, \varphi_n\rangle\psi$ at w, t, σ gives rise to a series of truncated hypothetical histories $\sigma'_1, \ldots, \sigma'_n$, with σ'_k agreeing with σ' up to $t+k$ and then making no changes after $t+k$, that witness the truth of $\langle !\varphi_1\rangle \ldots \langle !\varphi_n\rangle\psi$ at w, t, σ as above. □

It is important to observe that if some of $\varphi_1, \ldots, \varphi_n$ contain futuristic operators, then the equivalence in Proposition 6 is not guaranteed. To see this, note from clause 1 of Definition 2 above that for $\langle !\varphi_1, \ldots, \varphi_n\rangle\psi$ to be true, φ_1 must be true *relative to a hypothetical future in which all of $\varphi_1, \ldots, \varphi_n$ are publicly learned in that order, and then nothing else happens*. By contrast, for $\langle !\varphi_1\rangle \ldots \langle !\varphi_n\rangle\psi$ to be true, it is required that φ_1 be true *relative to a hypothetical future in which φ_1 is publicly learned, and then nothing else happens* and then $\langle !\varphi_2\rangle \ldots \langle !\varphi_n\rangle\psi$ is true at the next time step; but it is not required that φ_1 be true relative to a single hypothetical future in which all of $\varphi_1, \ldots, \varphi_n$ are learned. If φ_1 contains futuristic operators, then this can make a difference in truth values between $\langle !\varphi_1, \ldots, \varphi_n\rangle\psi$ and $\langle !\varphi_1\rangle \ldots \langle !\varphi_n\rangle\psi$. The moral is that if some of $\varphi_1, \ldots, \varphi_n$ contain futuristic operators, then the correct way to formalize the supposition of a sequence of updates of $\varphi_1, \ldots, \varphi_n$, of the kind made at the beginning of this section, is with $\langle !\varphi_1, \ldots, \varphi_n\rangle$ rather than $\langle !\varphi_1\rangle \ldots \langle !\varphi_n\rangle$.

Consider the fragment $\mathscr{L}_{SPAL}(!)$ of $\mathscr{L}_{SPAL}$ generated by the following grammar:

$$\varphi := p \mid \neg\varphi \mid (\varphi \wedge \varphi) \mid K_a\varphi \mid \langle !\varphi_1, \ldots, \varphi_n\rangle\varphi.$$

We can translate formulas of $\mathscr{L}_{SPAL}(!)$ to formulas of $\mathscr{L}_{PAL}$ in the obvious way, with $\tau(\langle !\varphi_1,\dots,\varphi_n\rangle\psi) = \langle !\tau(\varphi_1)\rangle\dots\langle !\tau(\varphi_n)\rangle\,\tau(\psi)$ in light of Proposition 6. For this fragment, we can capture SPAL semantics in PAL as follows.

Proposition 7 (From a Fragment of SPAL to PAL). *For any base model $M = \langle W,V\rangle$, sequential epistemic model $\mathscr{S} = \langle M,\sigma\rangle$, and $t \in \mathbb{N}$, define the epistemic model $\mathscr{S}_t = \langle W,\sigma(t),V\rangle$, where $\sigma(t)$ is the family of accessibility relations $\{R^a_t\}_{a\in \mathsf{Agt}}$ from σ at t. Then for any $t \in \mathbb{N}$, $w \in W$, and $\varphi \in \mathscr{L}_{SPAL}(!)$:*

$$M,w,t,\sigma \vDash_{SPAL} \varphi \quad \textit{iff} \quad \mathscr{S}_t, w \vDash_{PAL} \tau(\varphi).$$

Proof (Sketch). The proof is by induction on φ. The boolean cases are routine, and the atomic and K_a cases are guaranteed by the construction of $\mathscr{S}_t$. Suppose φ is of the form $\langle !\varphi_1,\dots,\varphi_n\rangle\psi$, so $\tau(\varphi) = \langle !\tau(\varphi_1)\rangle\dots\langle !\tau(\varphi_n)\rangle\,\tau(\psi)$. By Proposition 6, we have $M,w,t,\sigma \vDash_{SPAL} \varphi$ iff $M,w,t,\sigma \vDash_{SPAL} \langle !\varphi_1\rangle\dots\langle !\varphi_n\rangle\psi$, so it suffices to show that $M,w,t,\sigma \vDash_{SPAL} \langle !\varphi_1\rangle\dots\langle !\varphi_n\rangle\psi$ iff $\mathscr{S}_t, w \vDash_{PAL} \langle !\tau(\varphi_1)\rangle\dots\langle !\tau(\varphi_n)\rangle\,\tau(\psi)$. This equivalence follows easily from the inductive hypothesis and the definitions. $\square$

Finally, let us consider whether analogues of the PAL recursion axioms hold for $\langle !\varphi_1,\dots,\varphi_n\rangle$. For formulas from $\mathscr{L}_{SPAL}(!)$, they clearly do.

Proposition 8 (Recursion Axioms). *For every $\varphi_1,\dots,\varphi_n,\psi,\chi \in \mathscr{L}_{SPAL}(!)$ and $p \in$ At, the following are SPAL valid:*

1. $\langle !\varphi_1,\dots,\varphi_n\rangle p \leftrightarrow (\langle !\varphi_1,\dots,\varphi_n\rangle\top \wedge p)$;
2. $\langle !\varphi_1,\dots,\varphi_n\rangle\neg\psi \leftrightarrow (\langle !\varphi_1,\dots,\varphi_n\rangle\top \wedge \neg\langle !\varphi_1,\dots,\varphi_n\rangle\psi)$;
3. $\langle !\varphi_1,\dots,\varphi_n\rangle(\psi\wedge\chi) \leftrightarrow (\langle !\varphi_1,\dots,\varphi_n\rangle\psi \wedge \langle !\varphi_1,\dots,\varphi_n\rangle\chi)$;
4. $\langle !\varphi_1,\dots,\varphi_n\rangle K_a\psi \leftrightarrow (\langle !\varphi_1,\dots,\varphi_n\rangle\top \wedge K_a(\langle !\varphi_1,\dots,\varphi_n\rangle\top \to \langle !\varphi_1,\dots,\varphi_n\rangle\psi))$.

Proof. If one of the axioms is not SPAL-valid, then by Proposition 7, its translation into $\mathscr{L}_{PAL}$ is not PAL-valid. But it is straightforward to check that the translation of each axiom is PAL-valid (recall the axiomatization of PAL from §2). $\square$

What if we drop the restriction to $\mathscr{L}_{SPAL}(!)$ in Proposition 8? One can check that schema 1, the right-to-left direction of 2, and the left-to-right direction of 3 are valid for *all* $\varphi_1,\dots,\varphi_n,\psi,\chi \in \mathscr{L}_{SPAL}$. However, for the left-to-right direction of 2 and the right-to-left direction of 3, we can find falsifying instances if we consider formulas containing *futuristic* operators. The reason is that futuristic operators inside of $\langle !\varphi_1,\dots,\varphi_n\rangle$ can lead to *non-determinism*, as shown by the following example.

Example 8 (Non-determinism). Consider a model $\mathscr{S} = \langle W,V,\sigma\rangle$ with just four worlds, identifying each world with the set of atomic formulas true at that world: $\{p,q\}$, $\{p\}$, $\{q\}$, and $\emptyset$. Let R_0 be the universal relation on W. (The choice of R_i for $i > 0$ is irrelevant.) Then we claim that *all* of the following hold:

$$\mathscr{M},\{p,q\},0,R_0,R_1,\dots \vDash \langle !XKp \vee XKq\rangle(Kp\wedge\neg Kq) \tag{4}$$
$$\mathscr{M},\{p,q\},0,R_0,R_1,\dots \vDash \langle !XKp \vee XKq\rangle(Kq\wedge\neg Kp) \tag{5}$$
$$\mathscr{M},\{p,q\},0,R_0,R_1,\dots \vDash \langle !XKp \vee XKq\rangle(Kp\wedge Kq). \tag{6}$$

To see this, observe that the hypothetical relation S for time 1 in Figure 2 witnesses (4), while the hypothetical relation T for time 1 in Figure 3 witnesses (5). The witness for (6) is the hypothetical accessibility relation for time 1 where $\{p,q\}$ is related only to itself, while all the other states are related to each other.

This example shows that the choice of the S_i's in Definition 2 need not be unique. Thus, one could consider a variant of the semantics in which we require in part 3 of Definition 2 that ψ be true for *every* way of picking the S_i's that satisfies parts 1 and 2 of Definition 2. This would preclude (4)-(6) holding at once. Alternatively, we can accept the non-determinism by thinking of $\langle !\varphi_1,\ldots,\varphi_n\rangle\psi$ as saying something about what *could* happen, rather than what *necessarily would* happen.

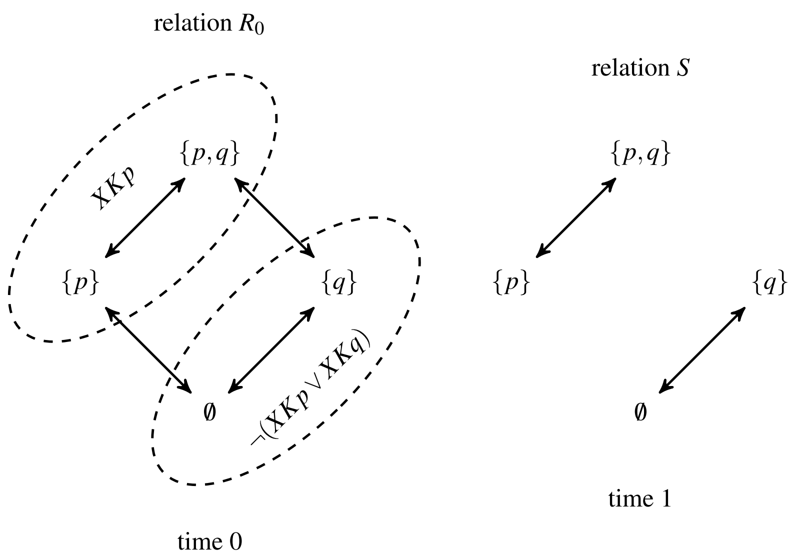

Fig. 2 For Example 8. Intended relations are the reflexive transitive closures of those displayed.

Finally, the semantics for the *arbitrary* hypothetical update operator $\langle !_n\rangle$ is as expected from the end of §4.1:

- $M,w,t,\sigma \vDash \langle !_n\rangle\psi$ iff $\exists\varphi_1,\ldots,\varphi_n \in \mathscr{L}^-_{SPAL}$: $\mathscr{M},w,t,\sigma \vDash \langle !\varphi_1,\ldots,\varphi_n\rangle\psi$.

Thus, $\langle !_n\rangle\psi$ says that there exists a sequence of n formulas and a hypothetical future of updates with those formulas that results in the truth of ψ. Recall that we take APAL formulas of the form $\langle !\rangle\psi$ as SPAL formulas of the form $\langle !_1\rangle\psi$, for which we will see an application in §5. Henceforth we drop the subscript for $n=1$.

One can check that analogues of Propositions 3, 4, and 7 hold for the language of APAL in place of PAL; thus, for the formulas that SPAL and APAL have in common (under translation), the SPAL semantics and the APAL semantics are equivalent. In this sense, SPAL is a conservative extension of APAL.

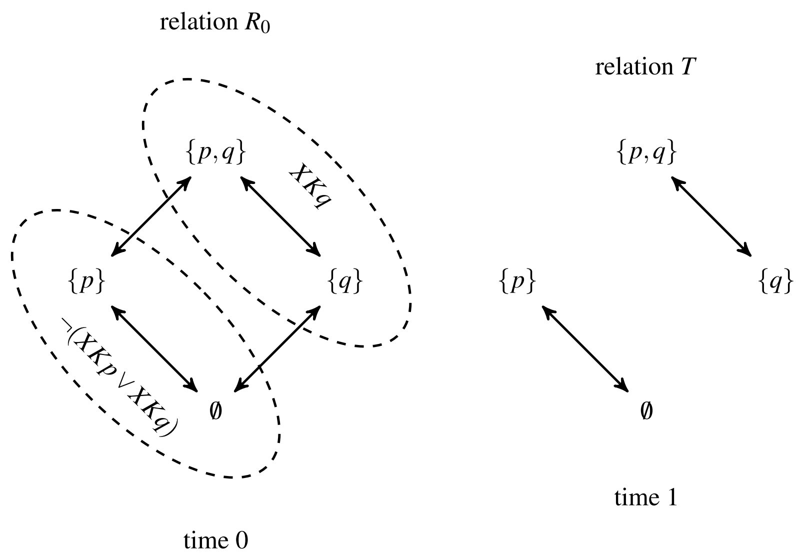

Fig. 3 For Example 8. Intended relations are the reflexive transitive closures of those displayed.

5 Applications

In this section, we sketch some sample applications of the framework of §4.

First, we will show that SPAL provides an enriched taxonomy of the important properties of formulas from a dynamic point of view. We begin with the following fundamental classifications for $\varphi \in \mathscr{L}_{SPAL}$:

- φ is *assimilable* iff $\langle\varphi\rangle\top$ is satisfiable; otherwise φ is *unassimilable*;
- φ is *always assimilable* iff $\varphi \to \langle !\varphi\rangle\top$ is valid.

For the first definition, note that if $\langle !\varphi\rangle\top$ is satisfiable, then so is $\langle\varphi\rangle\top$. However, the converse does not hold, because φ may describe a complicated future that we cannot suppose is brought about just by an update (or even a sequence of updates). For the second definition, note that the version $\varphi \to \langle\varphi\rangle\top$ makes little sense, since we cannot expect that whenever φ is true, the next epistemic state is in fact obtained by update with φ—though we may be able to hypothetically suppose it is.

Intuitively, an *assimilable* formula is a φ such that we can coherently conceive of an epistemic history in which φ is true at a time t and then between t and $t+1$ the agents update with φ. An *always assimilable* formula is a φ such that whenever it is true at a time t, we can coherently suppose that the agents update with φ from t to $t+1$. As we have seen in Example 4, not all formulas are *assimilable*: although we can conceive of a history in which $p \wedge X\neg Kp$ is true at a time t, we cannot conceive of one in which $p \wedge X\neg Kp$ is true at t and then between t and $t+1$ the agents update with $p \wedge X\neg Kp$, for that would result in knowledge of p at $t+1$, contradicting the requirement that $X\neg Kp$ be true at t. As an exercise, one can check that the following formula is *assimilable* but not *always assimilable*: $p \wedge X\neg Kq$.

In the standard taxonomy from the literature on PAL (see the references in §2), a formula φ is *successful* iff $\vDash_{PAL} \varphi \to \langle !\varphi \rangle \varphi$, which is equivalent to $\vDash_{PAL} \neg\langle !\varphi \rangle \neg\varphi$. Otherwise φ is *unsuccessful*. A formula φ is *self-refuting* iff $\vDash_{PAL} \varphi \to \langle !\varphi \rangle \neg\varphi$, which is equivalent to $\vDash_{PAL} \neg\langle !\varphi \rangle \varphi$. Neither of these equivalences hold for $\vDash_{SPAL}$. In SPAL, one has several flavors of successfulness. One could consider

(a) $\vDash_{SPAL} \neg\langle \varphi \rangle \neg\varphi$, or equivalently, $\vDash_{SPAL} \langle \varphi \rangle \top \to \langle \varphi \rangle \varphi$

with the descriptive operator, and analogues for $\langle !\varphi \rangle$. Another flavor is

(b) $\vDash_{SPAL} \varphi \to \langle !\varphi \rangle \varphi$

with the hypothetical operator. As above, $\varphi \to \langle \varphi \rangle \varphi$ makes little sense for the descriptive operator. Similarly, for flavors of self-refutation, one could consider

(c) $\vDash_{SPAL} \neg\langle \varphi \rangle \varphi$, or equivalently, $\vDash_{SPAL} \langle \varphi \rangle \top \to \langle \varphi \rangle \neg\varphi$

with the descriptive operator, and analogues for $\langle !\varphi \rangle$. Another flavor is

(d) $\vDash_{SPAL} \varphi \to \langle !\varphi \rangle \neg\varphi$

with the hypothetical operator. As above, $\varphi \to \langle \varphi \rangle \neg\varphi$ makes little sense.

Note that if φ is *unassimilable*, then φ automatically satisfies (a) and (c). Also note that if φ satisfies (b) or (d), then φ is *always assimilable*. One can check that $p \wedge \neg FKq$ is an example of an *assimilable* formula that satisfies (a) but not (b), because it is not *always assimilable*; and $p \wedge \neg Kp \wedge \neg FKq$ is an example of an *assimilable* formula that satisfies (c) but not (d), because it is not *always assimilable*.

Intuitively, a formula φ satisfies (a) iff whenever φ is assimilated by the agents, then after that update with φ, φ will (still) express a true proposition. By contrast, φ satisfies (b) iff whenever φ is true, we can coherently suppose that it is assimilated by the agents, and under this supposition, after the update with φ, φ will express a true proposition. Similar points apply to (c) and (d), but with *false* in place of *true*. We will not decide here how to apply the terms 'successful', 'self-refuting', and variants therefore to (a)-(d), since it is enough to make the distinctions.

The next important classifications are:

- φ is *ascertainable* iff $\langle\rangle KY\varphi$ is satisfiable;
- φ is *always ascertainable* iff $\varphi \to \langle ! \rangle KY\varphi$ is valid.

The notion of *always ascertainable* formulas is motivated by the problems with the notion of knowability from §3.1. There we pointed out that we cannot formalize the idea that *a true proposition can be known* with the principle $\varphi \to \langle ! \rangle K\varphi$ for $\varphi \in \mathscr{L}_{APAL}$. For with APAL semantics this principle says that if φ expresses a true proposition P_{initial} in some initial context, then there is a possible context-changing update that would result in the agent's knowing the proposition P_{new} that φ expresses in the new post-update context, with no guarantee that $P_{\text{new}} = P_{\text{initial}}$, so with no guarantee that the agent learns the original content of φ. By contrast, the principle $\varphi \to \langle ! \rangle KY\varphi$ for $\varphi \in \mathscr{L}_{APAL}$ with SPAL semantics says that if φ expresses a true proposition P_{initial} in some initial context, then there is a possible context-changing update that would result in the agent's knowing the proposition P_{initial}, which is expressed by $Y\varphi$ in the new post-update context. Thus, the agent does learn the original content of φ. This depends essentially on the assumption that $\varphi \in \mathscr{L}_{APAL}$, so φ does not contain futuristic operators. For if φ contains futuristic operators, which scan the *actual* future, and we consider a *hypothetical* update that departs from that future, then there is no guarantee that $Y\varphi$ expresses at $t+1$ in the hypothetical sequence the proposition that φ expresses at t in the actual sequence.[10]

Note that given the validity $\langle !\varphi\rangle\top \to \langle !\varphi\rangle KY\varphi$ observed in §4.2, if φ is *always assimilable*, then it is *always ascertainable*. However, the converse does not hold.

Example 9. Suppose there are just two worlds w and v, with a universal accessibility relation R_t at time t, and with p true only at w. Let $\varphi := XKp \vee XK\neg p$. We can make φ true at $\langle w,t\rangle$ with $R_{t+1} = \{\langle w,w,\rangle, \langle v,v\rangle\}$. But note that we cannot make $\langle !\varphi\rangle\top$ true at $\langle w,t\rangle$, since link-cutting cannot produce an epistemic relation R_{t+1} such that both φ is true at $\langle w,t\rangle$ *and* R_{t+1} comes from R_t by update with φ. Thus, φ is not always assimilable. Yet since $p \to \langle !p\rangle KYXKp$ and $\neg p \to \langle !\neg p\rangle KYXK\neg p$ are valid, it follows that $\varphi \to \langle !\rangle KY\varphi$ is valid, so φ is always ascertainable.

Returning now to Fitch's paradox, since in PAL and SPAL, $\neg Kp$ means that p is not known in the agent's *current* epistemic state, $p \wedge \neg Kp$ is not an example of an unascertainable formula. To the contrary, $p \wedge \neg Kp$ is *always ascertainable*:

$$(p \wedge \neg Kp) \to \langle !\rangle KY(p \wedge \neg Kp)$$

is valid.[11] Indeed, one can easily check the following.

Proposition 9 (Always Assimilable). *If $\varphi \in \mathscr{L}_{SPAL}$ does not contain futuristic operators, then φ is always assimilable and hence always ascertainable.*

But recall that Fitch's sentence was "*p* is true but no one knows, has known, or will know *p*." This is unascertainable. Indeed, just $p \wedge \neg FKp$ is unascertainable: $\langle\rangle KY(p \wedge \neg FKp)$ is unsatisfiable.[12] Thus, both assimilability and ascertainability

[10] Even if φ does not contain futuristic operators, the fact that $\langle !\rangle$ brings in a hypothetical future that may differ from the actual future means that we must be careful with the claim that $Y\varphi$ expresses at $t+1$ in the hypothetical sequence the "same proposition" that φ expresses at t in the actual sequence. This is correct if we mean that *the set of worlds Q at which $Y\varphi$ is true at $t+1$* in the hypothetical sequence is the same as *the set of worlds at which φ is true at t* in the actual sequence. But the "worlds" in Q may have different futures—i.e., with respect to what epistemic relations these worlds will stand in—in the hypothetical sequence vs. the actual sequence, so in a finer-grained sense, the set Q does not represent the same proposition relative to the hypothetical sequence and relative to the actual sequence. Still, the worlds in Q will have the same *past* up to t in both the hypothetical sequence and the actual sequence, so for a formula φ not containing futuristic operators, there is a reasonable sense in which the proposition expressed by $Y\varphi$ at $t+1$ in the hypothetical sequence is "the same" as the proposition expressed by φ at t in the actual sequence. This point deserves further discussion, but we do not have room for it here.

[11] Cf. Hintikka on $p \wedge \neg Kp$ in Example 1: "You may come to know that what I say *was* true."

[12] This uses the fact that we are treating p as an eternal sentence, so $KYp \to Kp$ is valid. If we were not treating p as eternal, then we would need to eternalize p with temporal operators: where $S\varphi := P\varphi \vee \varphi \vee F\varphi$ ("sometime, φ"), the formula $Sp \wedge \neg FKSp$ is unascertainable.

are non-trivial, and one may inquire into their necessary and sufficient syntactic conditions (cf. Holliday and Icard 2010). The fact that formulas not containing futuristic operators are always ascertainable suggests a question that we leave for the reader: for those philosophers who hold that statements about what will happen in the future lack a truth value, how can one prove that not every truth can be known?

Finally, we return to the designated student paradox. In §3.2, we criticized the formalization of the teacher's announcement as $S := (g_1 \wedge \neg K_1 g_1) \vee (g_2 \wedge \neg K_2 g_2) \vee (g_3 \wedge \neg K_3 g_3)$. The problem with S is that the teacher did not announce that the student with the gold star did not know, *before the announcement*, that he or she had it. Instead, the teacher essentially announced that the student with the gold star would not know that he or she had it, *even after the announcement*, since this would only be revealed upon the students breaking formation. The content of the teacher's announcement is better captured by the SPAL formula

$$X((g_1 \wedge \neg K_1 g_1) \vee (g_2 \wedge \neg K_2 g_2) \vee (g_3 \wedge \neg K_3 g_3)),$$

an announcement of which roughly amounts to "after this announcement, the student with the gold star won't know that he or she has it."

The formalizations with S vs. XS lead to very different predictions for what the students will know. The formula S is *always assimilable*, and as we saw in §3.2, if we start with the model $\mathscr{M}$ from Figure 1, reproduced in Figure 4 below, then after the agents update with S, they all know that student 3 does not have the gold star (for if she did have it, then she would have known she had it, contrary to S). By contrast, the formalization with XS does not support the elimination of 3. Unlike S, the formula XS is not always assimilable—and the very model $\mathscr{M}$ for the designated student paradox is a counterexample. Specifically, we have

$$\mathscr{M}, \{g_2\}, 0, R_0, \dots \nvDash \langle !XS \rangle \top.$$

As we encourage the reader to check, *there is no way of cutting links from the students' initial epistemic relations R_0^1, R_0^2, and R_0^3 such that with the new relations R_1^1, R_1^2, and R_1^3 for time* 1*, the formula XS expresses at time* 0 *a proposition P true at $\{g_2\}$ such that R_1^1, R_1^2, and R_1^3 are exactly the result of updating R_0^1, R_0^2, and R_0^3 by link-cutting with P*. What this means is that assuming the initial epistemic states of the students are as in $\mathscr{M}$ and assuming that the gold star is on the back of student 2, it is incoherent to suppose *both* that by announcing XS, the teacher expresses a *true* proposition P at t, *and* that from t to $t+1$ the agents update their epistemic states with that proposition P. Of course, the teacher may utter the words, but it cannot be that by doing so, the teacher expresses a true proposition with which the students update their knowledge upon hearing the announcement. Thus, it would be a mistake for student 2 to assume that the teacher does express such a true proposition and use this assumption to eliminate the possibility that student 3 has the gold star.[13]

[13] See Holliday 2016b for an analysis of the designated student paradox using static multi-agent epistemic logic. In that analysis, the assumptions about the initial epistemic states of the agents are not given by the model $\mathscr{M}$, but rather by a weaker set of syntactically specified assumptions.

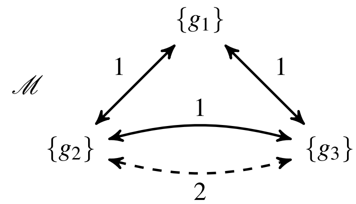

Fig. 4 Model for the designated student paradox.

6 Conclusion

Inspired by Hintikka's (1962) pioneering work on epistemic logic in epistemology, the purpose of this paper was to motivate (§3), introduce (§4), and apply (§5) the idea of *sequential epistemic logic*, building on dynamic epistemic logic. We implemented the general idea of sequential epistemic logic with a sequential epistemic analogue of APAL, dubbed SPAL. We defined the logic of SPAL semantically, leaving the problem of axiomatizing SPAL or fragments thereof to future research.

It should be stressed that the idea of sequential epistemic logic can be implemented in other ways, by looking at other ways of transforming epistemic accessibility relations. We looked only at one way: link-cutting update. But the program of dynamic epistemic logic provides a variety of ways of transforming epistemic relations, including a general approach using event models and product update (see the textbook van Ditmarsch et al 2008). Any way of transforming epistemic relations could in principle give rise to an associated sequential epistemic logic. In addition, the sequential epistemic models could be enriched with structure beyond accessibility relations, such as plausibility relations to be transformed for belief revision.

There is much more to say about the applications sketched in §5 and other potential applications. But hopefully our discussion here already shows how a sequential epistemic logic may illuminate the interplay of knowledge, time, and paradox.

Acknowledgements This paper is the written version of a talk with the same title that I gave at the LogiCIC mini-workshop on The Dynamics of Information States, organized by Ben Rodenhäuser in Amsterdam on June 19, 2014. I am grateful to Ben for the invitation to the workshop and to Alexandru Baltag and Sonja Smets for their encouraging feedback on the talk. I also benefited from the experience of co-teaching a course with Eric Pacuit on *Ten Puzzles and Paradoxes of Knowledge and Belief* at ESSLLI 2013 in Dusseldorf. It was during a sleepless jet-lagged night in Dusseldorf, before my lecture on dynamic epistemic logic, that ideas in this paper occurred to me. Finally, I appreciate having had the pleasure to meet Jaakko Hintikka at Stanford in March 2012, when he very kindly spent a lunch discussing epistemic logic with me.

For helpful comments on this paper, I thank Johan van Benthem and Hans van Ditmarsch.

References

Balbiani P, Baltag A, van Ditmarsch H, Herzig A, Hoshi T, de Lima T (2008) 'Knowable' as 'Known after an Announcement'. The Review of Symbolic Logic 1(3):305–334

Baltag A, Smets S (2010) Multi-Agent Belief Dynamics, URL http://www.vub.ac.be/CLWF/nasslli2010lecture3.pdf, Lecture Slides from NASSLLI 2010, Indiana University

van Benthem J (2004) What One May Come to Know. Analysis 64(2):95–105

van Benthem J (2011) Logical Dynamics of Information and Interaction. Cambridge University Press, New York

van Benthem J, Liu F (2007) Dynamic logic of preference upgrade. Journal of Applied Non-Classical Logics 17(2):157–182

van Benthem J, Gerbrandy J, Hoshi T, Pacuit E (2009) Merging Frameworks for Interaction. Journal of Philosophical Logic 38(5):491–526

Burgess J (2009) Can Truth Out? In: Salerno J (ed) New Essays on the Knowability Paradox, Oxford University Press, New York, pp 147–162

Cohen M (2015a) A Dynamic Epistemic Logic with a Knowability Principle. In: van der Hoek W, Holliday WH, Wang W (eds) Logic, Rationality, and Interaction: 5th International Workshop, LORI 2015, Taipei, Taiwan, October 28-30, 2015. Proceedings, Lecture Notes in Computer Science, vol 9394, Springer, Dordrecht

Cohen M (2015b) Dynamic Knowability: The Knowability Paradox in Dynamic Epistemic Logic. Master's thesis, The Munich Center for Mathematical Philosophy, Ludwig-Maximilians-Universität München

van Ditmarsch H, Kooi B (2006) The Secret of My Success. Synthese 151(2):201–232

van Ditmarsch H, van der Hoek W, Kooi B (2008) Dynamic Epistemic Logic. Springer, Dordrecht

van Ditmarsch H, van der Hoek W, Iliev P (2011) Everything is Knowable – How to Get to Know *Whether* a Proposition is True. Theoria 78(2):93–114

van Ditmarsch H, van der Hoek W, Ruan J (2013) Connecting dynamic epistemic and temporal epistemic logics. Logic Journal of the IGPL 21(3):380–403

van Ditmarsch H, Halpern JY, van der Hoek W, Kooi B (2015) Handbook of Epistemic Logic. College Publications, London

Egré P (2011) Epistemic Logic. In: Horsten L, Pettigrew R (eds) The Continuum Companion to Philosophical Logic, Continuum Companions, Continuum, New York, pp 503–542

Fitch FB (1963) A Logical Analysis of Some Value Concepts. The Journal of Symbolic Logic 28(2):135–142

Gerbrandy J (1999) Bisimulations on Planet Kripke. PhD thesis, University of Amsterdam, ILLC Dissertation Series DS-1999-01

Gerbrandy J (2007) The Surprise Examination in Dynamic Epistemic Logic. Synthese 155(1):21–33

Gerbrandy J, Groeneveld W (1997) Reasoning about Information Change. Journal of Logic, Language and Information 6(2):147–169

Gierasimczuk N (2010) Knowing One's Limits. Logical Analysis of Inductive Inference. PhD thesis, University of Amsterdam, ILLC Dissertation Series DS-2010-11

Herzig A, Lang J, Polacsek T (2000) A modal logic for epistemic tests. In: Horn W (ed) Proceedings of the 14th European Conference on Artificial Intelligence (ECAI 2000), IOS Press, Amsterdam, pp 553–557

Hintikka J (1962) Knowledge and Belief: An Introduction to the Logic of the Two Notions. Cornell University Press, Cornell

Holliday WH (2016a) On Being in an Undiscoverable Position. Thought: A Journal of Philosophy 5(1):33–40

Holliday WH (2016b) Simplifying the Surprise Exam, URL http://escholarship.org/uc/item/82w2d085, UC Berkeley Working Paper in Philosophy

Holliday WH (Forthcoming) Epistemic Logic and Epistemology. In: Hansson S, Hendricks V (eds) Handbook of Formal Philosophy, Springer, Dordrecht

Holliday WH, Icard TF (2010) Moorean Phenomena in Epistemic Logic. In: Beklemishev L, Goranko V, Shehtman V (eds) Advances in Modal Logic, vol 8, College Publications, London, pp 178–199

Holliday WH, Hoshi T, Icard TF (2011) Schematic Validity in Dynamic Epistemic Logic: Decidability. In: van Ditmarsch H, Lang J, Ju S (eds) Logic, Rationality and Interaction: Third International Workshop, LORI 2011, Guangzhou, China, October 10-13, 2011. Proceedings, Lecture Notes in Artificial Intelligence, vol 6953, Springer, Dordrecht, pp 87–96

Holliday WH, Hoshi T, Icard TF (2012) A Uniform Logic of Information Dynamics. In: Bolander T, Braüner T, Ghilardi S, Moss L (eds) Advances in Modal Logic, vol 9, College Publications, London, pp 348–367

Holliday WH, Hoshi T, Icard TF (2013) Information Dynamics and Uniform Substitution. Synthese 190(1):31–55

Hoshi T (2008) Epistemic Dynamics and Protocol Information. PhD thesis, Stanford University, ILLC Dissertation Series DS-2009-08

Hoshi T, Yap A (2009) Dynamic epistemic logic with branching temporal structures. Synthese 169(2):259–281

Kooi B (2007) Expressivity and completeness for public update logics via reduction axioms. Journal of Applied Non-Classical Logics 17(2):231–253

Lewis D (1969) Convention: A Philosophical Study. Harvary University Press, Cambridge, MA

Lewis D (1996) Elusive Knowledge. Australasian Journal of Philosophy 74(4):549–567

van Linder B, van der Hoek W, Meyer JJC (1994) Tests as Epistemic Updates: Pursuit of Knowledge. Tech. Rep. UU-CS-1994-08, Department of Computer Science, Utrecht University

Marcoci A (2010) The Surprise Examination Paradox in Dynamic Epistemic Logic. Master's thesis, ILLC Master of Logic Thesis Series MoL-2010-18

Pacuit E (2013) Dynamic Epistemic Logic I and II. Philosophy Compass 8(9):798–833

Plaza J (1989) Logics of public communications. In: Emrich M, Pfeifer M, Hadzikadic M, Ras Z (eds) Proceedings of the 4th International Symposium on Methodologies for Intelligent Systems: Poster Session Program, Oak Ridge National Laboratory, ORNL/DSRD-24, pp 201–216

Proietti C, Sandu G (2010) Fitch's paradox and *ceteris paribus* modalities. Synthese 173(1):75–87

Quine WVO (1960) Word and Object. MIT Press, Cambridge, MA

Renne B, Sack J, Yap A (2009) Dynamic Epistemic Temporal Logic. In: He X, Horty J, Pacuit E (eds) Logic, Rationality, and Interaction: Second International Workshop, LORI 2009, Chongqing, China, October 8-11, 2009. Proceedings, Lecture Notes in Artificial Intelligence, vol 5834, Springer, Dordrecht, pp 263–277

Renne B, Sack J, Yap A (2016) Logics of temporal-epistemic actions. Synthese 193(3):813–849

Sack J (2010) Logic for update products and steps into the past. Annals of Pure and Applied Logic 161(12):1431–1461

Sorensen R (1984) Recalcitrant Variations of the Prediction Paradox. Australasian Journal of Philosophy 69(4):355–362

Sorensen R (1988) Blindspots. Clarendon Press, Oxford

Stalnaker RC (1984) Inquiry. MIT Press, Cambridge, MA

Wansing H (2015) Knowability Remixed. In: Caret CR, Hjortland OT (eds) Foundations of Logical Consequence, Oxford University Press, New York, pp 322–328

Williamson T (2000) Knowledge and its Limits. Oxford University Press, New York

Yap A (2011) Dynamic Epistemic Logic and Temporal Modality. In: Girard P, Roy O, Marion M (eds) Dynamic Formal Epistemology, Synthese Library, vol 351, Springer, Dordrecht, pp 33–50