Carnegie Mellon University

DIETRICH COLLEGE OF HUMANITIES AND SOCIAL SCIENCES

Dissertation

Submitted in Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

Title: Endogenizing Epistemic Actions in Dynamic Epistemic Logic

Presented by: William Nalls

Accepted by: Department of Philosophy

Name: Adam Bjorndahl
Adam Bjorndahl (Sep 9, 2021 08:46 EDT) Date: Sep 9, 2021

Name: Teddy Seidenfeld
Teddy Seidenfeld (Sep 9, 2021 10:16 EDT) Date: Sep 9, 2021

Name: Eric Pacuit (Oct 19, 2021 15:05 EDT) Date: Oct 19, 2021

Approved by the Committee on Graduate Degrees

(Dean) Date: 10/25/21

Endogenizing Epistemic Actions in Dynamic Epistemic Logic

Will Nalls

August 27, 2021

Contents

Acknowledgements

This dissertation is the culmination of work which was encouraged and supported by a network of people; I will thank them here.

First and foremost, I express my gratitude to my advisor, Adam Bjorndahl, who oversaw this project. Conversations with Adam shaped and substantiated this dissertation, and inspired progress. I have learned much from Adam during my time with the department – not only about logic, but also about teaching and academic culture. I am very thankful to have had such a capable researcher and patient instructor as my mentor.

I thank Eric Pacuit, of my committee, for thoughtful feedback and discussion along the way. I also thank Teddy Seidenfeld, of my committee, for the friendly support and stimulating conversations over the years, concerning my dissertation and otherwise.

Much of my intellectual growth over the course of this dissertation is happily attributed to the CMU philosophy department, whose generous support and guidance over these years I gratefully acknowledge. In particular, I thank Wilfried, my master's advisor, for teaching me how to sharpen my analytic thinking; I thank Mara and David, for making me feel welcome in the department; and I thank Joel, Christina, and Mandy for friendly conversations and collaborations.

Often unsung heroes in the production of research exemplified by this dissertation are the administrative staff to the department. I thank Jan, Ro, and Jackie for their efforts in maintaining the department and their assistance to me throughout the program; my time in the department was happier for their presence. I grateful to Mary Grace in particular for her valuable help time and again; I appreciate the care she puts into fostering a friendly environment in the department, and have enjoyed working with her on various projects over the years.

Outside of the department, I thank Larry Moss for generously inviting me to talk about my work in Bloomington; conversations with Larry inspired progress and confidence in my work. I also thank Hannes Leitgeb, Olivier Roy, and Conor Mayo-Wilson, for helping me to enter the world of analytic philosophy, and encouraging me to pursue this PhD.

My time in Pittsburgh has witnessed as much personal growth as intellectual growth; this dissertation is a product of both, and so also indebted to the community outside of my academic life. I thank my parents, Leslie and Floyd, for helping me get to the place where I could pursue a PhD, and for their support during it. I thank Ben, Joe, and Sarah for just doing their respective sibling things. I thank the web of friends I have found in Pittsburgh for all of the attention, care, and delights they have given me over my time here. In particular, I thank Aubrey, Mike, Lillian, Patrick, and Rainy for the laughs and love. Chelsea and Eric get a shoutout, too.

Finally, and crucially, I thank Hillary, my partner, who played an essential part in my finishing this work. The world is a brighter place when I am at your side.

1 Introduction

This dissertation studies the representation of higher-order uncertainty in epistemic logic. We consider a series of examples which feature this concept, and formulate a criticism of the *action logic* framework. This criticism motivates a change to the update mechanism of action logic; we call the resulting logic *action-epistemic logic*, and demonstrate its answer to the criticism stated. We explore the properties of action-epistemic logic, prove completeness, and apply the logic in an analysis of the 'No Miracles' principle in epistemic logic.

We begin the dissertation by providing some background in section 1.1 for the reader lacking context; this is by no means a thorough history of the field or a technically complete exposition, but should help to situate this dissertation within the field of formal epistemology. Those already familiar with epistemic logic may jump to section 1.2.

1.1 Background

Formal epistemology is a subfield of epistemology which treats philosophical questions about knowledge and belief using technical apparatus. *Epistemic logic* is an umbrella term which includes those investigations in formal epistemology which make use of logical frameworks. Most, if not all, of the work in epistemic logic may be traced back to the ideas contained in Hintikka's *Knowledge and Belief: An Introduction to the Logic of the Two Notions* [3]. In this seminal book, Hintikka proposes the syntactic representation of statements involving belief and knowledge using a logical language. To take a simple example, suppose that we allow the letter p to symbolize some proposition we are interested in – for now, let p symbolize the claim 'Pluto is not a planet' – then we may build more complex statements in the following way:

$$(1) K_a p \quad (2) \neg B_a p$$

(1) reads 'a knows that Pluto is not a planet' while (2) reads 'a does not believe that Pluto is not a planet'. The symbols K_a and B_a are called *epistemic operators* while the lowercase letter p is an atom of the language, intended to represent basic propositions. More complex statements in the language may be built from expressions of this sort using standard logical connectives like $\wedge, \vee, \neg, \rightarrow$, and so forth.

As Hintikka demonstrates, the syntactic representation of epistemic concepts in this way opens the door to a precise analysis of the features of these concepts and of their relationships with one another. For example, a central concept in Hintikka's study is the notion of *consistency*. Given some collection of statements which may include epistemic and doxastic claims, how should we go about deciding whether the collection is consistent? Returning to our example above, it seems obvious that it would be inconsistent to maintain both (1) and (2) simultaneously. However, this has nothing to with the *content* of p – rather, it has to do with the relationship between the forms of (1) and (2). In particular, in any case where a *knows* some claim p, it seems reasonable to think that a ought also *believe* that p, in order to be consistent. A strength of this approach is that insights such as these may be stated precisely as 'axioms'; the present insight might be encoded with the axiom $K_a\varphi \rightarrow B_a\varphi$, where we use φ to mean *any* statement in our language to which an epistemic operator may apply [3]. A collection of formulas is consistent with respect to some set of axioms if it satisfies each of the axioms.[1] The syntactic representation of epistemic concepts in this way also facilitates the representation of *higher-order* epistemic statements, such as 'b knows that a knows that Pluto is not a planet':

$$K_b K_a p$$

Thus, principles concerning higher-order epistemic concepts may also be straightforwardly formulated in our language; for instance, the principle 'if you believe some fact φ, then you believe that you believe φ' might be presented as the rule 'if a collection of epistemic claims contains $B_a\varphi$, then it must also contain $B_a B_a \varphi$ (in order to be admissible)'.

In this thesis, we will focus on a particular way of semantically evaluating these formulas called *possible worlds semantics*. In possible worlds semantics, we consider a set of 'possible worlds', where each possible world is a full description of how the world might be, in the respects which are under study; in the present work, we opt for the terminology 'state descriptions', 'states of affairs', or simply 'states', over 'possible worlds'. Under these semantics, to know some fact p is to only consider possible those states in which p is true. For each epistemic agent j, we introduce a binary *accessibility* relation R_j over the set of states under consideration. To say that the world w' is epistemically *accessible* to an agent from w, notated $(w, w') \in R_j$, just means that when the possible world w is the 'actual' one, the agent j maintains that the possible world w' may be the actual one. Taken together, the collection of states with their descriptions and the accessibility relations comprise an *epistemic model*, notated M, and every formula built up in the way above is either true or false at each state in this model.[2] *Epistemic logic* (EL) studies the properties of these models in an abstract setting, treating questions such as: what formulas will be generally true (valid) in all epistemic models, or in certain classes of epistemic models? How do the properties of the accessibility relations correspond to these collections of formulas? Is there a systematic way of generating the formulas which are true in these classes of epistemic models?

Knowledge is a dynamic concept; people learn things, and what they know changes. A natural extension to epistemic logic is to examine *changes* in epistemic states, or how the transmission of information may bear on what an agent knows or believes. In terms of possible worlds semantics, the transmission of information corresponds to a transformation on epistemic models: to represent an agent j learning p, we move from a model where j does not know p, to a model where they do. There are many ways of understanding and implementing this transformation, dependent on the kind of information transmission one wishes to represent; *dynamic epistemic logic* is a subfield of epistemic logic which studies information transmission using the framework of epistemic models.

One of the first systematic attempts to add such dynamics to this framework considers how one might represent 'public announcements' in epistemic logic ([6, 21]).[3] In *Public Announcement Logic* (PAL), to represent the public announcement

[1]There are also 'rules of inference' to consider, but we omit this detail for the present discussion.

[2]These models are also called 'Kripke models', in virtue of the work by Saul Kripke [4]; however, some incipient traces of possible worlds semantics may be found in Carnap's *Modalities and Quantification* [1]. See [12] for a thorough history of possible worlds semantics.

[3]A logic of public announcements was independently developed by Gerbrandy and Groeneveld

of some fact p in a scenario represented by epistemic model M, we simply prune away any states in which p is not true.[4] In the resulting model, all agents know that p, and know that the other agents know that p (and so on). In PAL, the standard epistemic language is enriched with announcement 'operators'; for instance, the claim 'after a public announcement of p, φ will be true' is represented with the formula $[p]\varphi$. This formula is evaluated by seeing whether φ is true once all non-p states have been pruned from the model.

Of course, there are many ways to transmit and receive information besides public announcements; in this thesis, we consider a more general framework which admits transmissions of information that are not necessarily common knowledge to all agents. *Action Logic* (AL) is a framework for reasoning about how knowledge and belief change on the basis of incoming information [9, 13, 22].[5] In this logic, uncertainty about epistemic actions is explicitly encoded in a structure called an *action model*; this allows for the representation of scenarios in which agents can be uncertain about which action has taken place. This logic subsumes PAL in the sense that every public announcement corresponds to a trivial action model in which it is common knowledge that every agent can distinguish the epistemic action performed. In addition to public announcements, AL can encode asymmetrical distributions of information where one agent learns something that another does not.

Like in PAL, AL enriches the basic epistemic language with dynamic operators; if A is an action model which captures the agents' uncertainty involving an epistemic action σ, then the statement 'after σ is performed (given context A), φ will be true' is represented using the formula $[A, \sigma]\varphi$. This formula is also evaluated on an epistemic model using a dynamic procedure: AL includes an 'update mechanism' which, given an epistemic model M and an action model A, produces an epistemic model M^A that encodes the epistemic consequences of A, given the starting point M.

In this thesis, we will criticize the AL framework and propose a change to the update semantics. We now turn to a summary of the contributions made by the present work.

1.2 Expository remarks

By design, AL is well-equipped to capture uncertainty about actions themselves; however, we argue in this thesis that the AL framework is ill-suited to the representation of *higher-order* uncertainty about actions. Roughly speaking, this is because the action model that captures uncertainty about actions is itself effectively common knowledge amongst the agents, making it awkward to encode, for example, one agent's uncertainty about another agent's uncertainty about actions, or indeed, even one agent's uncertainty about their own ability to distinguish different epistemic actions.

in [8].

[4]It should be noted that in the original presentation, Plaza introduces a binary propositional connective, rather than a model transformation; in later presentations of PAL, however, such as in [22], formulas involving public announcements are evaluated using transformations on epistemic models.

[5]Terminology varies; some authors instead use *Dynamic Epistemic Logic* to refer to this framework [26, §2.2.2]. We follow van Ditmarsch et al. [22] in using it instead as an umbrella term for a collection of thematically related logics of information change, including Action Logic.

We expose this difficulty through a series of motivating examples which have the following flavor:

Example 1. Anne and Bob work in an appliance repair shop. Anne is instructing Bob on how to repair a machine with which he is unfamiliar, before sending him to the assignment. Importantly, Anne relays that the appropriate action to take depends on whether the machine is in state κ (p) or in state γ ($\neg p$); she describes both of the appropriate actions in detail.

Anne is thorough in her instruction, but forgets one crucial detail before sending Bob off to repair the machine: by convention, she and the other experienced technicians call the states of the machine p and $\neg p$, but on the machine itself, the state is indicated by a pair of vertically oriented lights. Both lights will be on (σ_p) if the machine is in state p, and only the bottom lights will be on ($\sigma_{\neg p}$) if the machine is in state $\neg p$. Of course, it may be that Bob has learned this convention already from another technician – in this case (q), Bob will learn the state of the machine (p or $\neg p$) from seeing the state of the light; but if not ($\neg q$), he will not learn the state of the machine from seeing the light.

The crucial feature of this example is the uncertainty surrounding Bob's ability to parse the incoming information: whether he can distinguish between the signal for p (σ_p) and the signal for $\neg p$ ($\sigma_{\neg p}$) depends on the proposition q. In the AL framework, however, epistemic actions are either always distinguishable ($\sigma_p \sim_b \sigma_{\neg p}$) or never distinguishable ($\sigma_p \not\sim_b \sigma_{\neg p}$); there is no way for the indistinguishability relation to be dependent on a proposition internal to the model.

A related feature which AL struggles to capture is *Anne's* uncertainty about Bob's epistemic capabilities – since $\sigma_p \sim_b \sigma_{\neg p}$ is either true everywhere or true nowhere in the action model, there is no room for Anne to have doubts about Bob's abilities.

We demonstrate that although AL *can* capture this sort of higher-order uncertainty about actions, this can only be done by expanding the action model in such a way as to essentially "pre-encode" the desired uncertainty; this makes choosing an appropriate action model for any given application problematically post hoc (example 5). Furthermore, we show that in such cases small variations in the background epistemic conditions require corresponding alterations to the action model in order to ensure that the "pre-encoded" uncertainty maintains the right form (examples 7 and 6). These observations seriously undermine the practical applicability of AL as a tool for reasoning about information updates.

In response to these challenges, we reformulate the semantics by "endogenizing" the action model; that is, we allow each agent's uncertainty about actions to be state-dependent, and therefore itself subject to uncertainty (section 3.1). Revisiting our examples, we show that these revised semantics completely circumvent the earlier difficulties; our semantics capture *formally* the informal process underlying the aforementioned post hoc expansion of the action model (section 3.2). These revisions also prompt a change to the update mechanism which underlies the AL framework. The altered update which we present captures, intuitively, the idea that when an epistemic action is performed, an agent may learn about their ability to distinguish that action from others (section 3.4). This type of learning cannot be directly encoded in AL, where each agent's epistemic abilities is common knowledge in virtue of its encoding in the action model.

We call the resulting framework *Action-epistemic logic* (AEL) and show that this logic generalizes AL. Interestingly, AEL is also no more expressive than AL in the following sense: every model produced by the update mechanism of AEL may be simulated using the underlying epistemic model paired with a suitable action model (sections 3.3 and 3.4). However, when higher-order uncertainty about actions is involved in an example, the corresponding action model will generally be much larger than the AEL model.

In section 4, we present some completeness results for the logic AEL. We first introduce constants $\xi_{j,\sigma,\sigma'}$ into the basic epistemic language which reflect the state-dependent indistinguishability of actions; the resulting language is $\mathcal{L}_{K\xi}$. We axiomatize $\mathcal{L}_{K\xi}$ in section 4.1.1. We extend this language with an update operator $[\sigma]$ to create the dynamic language $\mathcal{L}^{\Sigma}_{K[\sigma]\xi}$. We then give a completeness result for $\mathcal{L}^{\Sigma}_{K[\sigma]\xi}$ via a set of reduction axioms which show that every formula in $\mathcal{L}^{\Sigma}_{K[\sigma]\xi}$ is equivalent to a formula in $\mathcal{L}_{K\xi}$ (section 4.1.2). We also consider the language $\mathcal{L}^{\Sigma}_{K[\sigma]C\xi}$, which is the dynamic language $\mathcal{L}^{\Sigma}_{K[\sigma]\xi}$ expanded with a common knowledge operator, C (section 4.2). Standard reduction techniques are not available in the presence of common knowledge, and so we instead strengthen our language to that of *propositional dynamic logic*, and establish a completeness result in that setting.

In section 5, we consider an application of AEL towards the analysis of the 'No Miracles' principle in epistemic logic. This principle appears primarily in the context of the more general *epistemic temporal logic* framework, and captures, in part, what sort of learning is available under the AL framework. As we will show, this principle is violated in AEL. We begin the section with a brief history of the principle in epistemic logic (section 5.1). We then show, using a technique involving a non-reductive completeness proof from [31], that AEL is characterized by a weakened version of the No Miracles principle (section 5.2).

The idea of representing higher-order uncertainty about epistemic actions by encoding extra information about the agents (and their perceptions of such actions) into the state space is a natural one; it also occurs in [34], in which Bolander et al. study a more general class of announcements that may not be entirely public in that (loosely speaking) some agents may not be "listening". In essence, their semantics work by encoding into each possible world whether or not each agent is "paying attention" at that world. As the authors show, such "attention-based" announcements can also be described using action models. The present work is more general in that we begin with the full action model framework rather than PAL, and we encode into each world the full spectrum of each agent's uncertainty regarding epistemic actions, not just whether or not they are attentive.

2 Action Logic

2.1 Introduction to Action Logic

We begin by reviewing the foundational definitions and motivations of AL, largely following van Ditmarsch et al. [22, Chapter 6]. This logic is an extension of standard epistemic logic, so we begin there.

In standard epistemic logic, we fix a countable set of primitive propositions PROP and a finite set of agents G. Let $\mathcal{L}_K$ denote the language recursively defined as follows:

$$\varphi := p \mid \neg\varphi \mid \varphi \wedge \psi \mid K_j\varphi,$$

where $p \in$ PROP and $j \in G$. We read $K_j\varphi$ as "agent j knows that φ". Thus, $\mathcal{L}_K$ is a language for reasoning about the knowledge of the agents in G. The other Boolean connectives can be defined in the usual way; we write $\hat{K}_j$ to abbreviate $\neg K_j \neg$, and read $\hat{K}_j\varphi$ as "agent j considers it possible that φ".

An **epistemic model** is a structure of the form $M = \langle W, \{\sim_j : j \in G\}, V\rangle$, where:

- W is a (nonempty) set of *states*,
- each $\sim_j$ is an equivalence relation on W,
- V : PROP $\to 2^W$ is a *valuation function*.

Intuitively, V specifies for each primitive proposition those states where it is true, while the relations $\sim_j$ capture indistinguishability from the perspective of agent j. These intuitions are formalized in the following semantic clauses for the evaluation of formulas on pointed models:[6]

$$\begin{array}{lll}
(M,w) \models p & \text{iff} & w \in V(p) \\
(M,w) \models \neg\varphi & \text{iff} & (M,w) \not\models \varphi \\
(M,w) \models \varphi \wedge \psi & \text{iff} & (M,w) \models \varphi \text{ and } (M,w) \models \psi \\
(M,w) \models K_j\varphi & \text{iff} & (\forall w' \in [w]^j)((M,w) \models \varphi),
\end{array}$$

where $[w]^j$ denotes the equivalence class of w under $\sim_j$. Thus, the Boolean connectives are interpreted as usual, and $K_j\varphi$ is true at w precisely when φ is true at all states that agent j cannot distinguish from w. Insisting that the indistinguishability relations be equivalence relations results in a logic of knowledge that is *factive* and *fully introspective*.[7] For a more thorough development of differing logics of knowledge we direct the reader to [7].

The framework of epistemic logic is appropriate for reasoning about scenarios involving uncertain agents. Consider the following simple example.

Example 2. Colleagues Anne and Bob are discussing whether a particular company policy passed in this morning's board meeting: both are ignorant of whether or not the policy passed. Taking p to be the proposition "the policy passed", this scenario may be represented with the simple epistemic model M_0 presented in Figure 1.[8]

[6] A pointed model is a model M together with some state from the model, w. We sometimes write $w \models \varphi$, rather than $(M,w) \models \varphi$, when M is clear from context.

[7] That is, the logic witnesses the following validities: $\models K_j\varphi \to \varphi$, $\models K_j\varphi \to K_jK_j\varphi$, and $\neg K_j\varphi \to K_j\neg K_j\varphi$.

[8] Unless stated otherwise, all relations are equivalence relations, so reflexive loops and edges implied by transitivity are assumed to be present, even when suppressed in the diagrams.

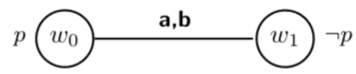

Figure 1: M_0 – Anne and Bob's initial uncertainty

In this diagram and others like it, the circles represents the states of the model (in this case, w_0 and w_1), the formulas listed beside a state are true at that state (in this case, p and $\neg p$ are true at w_0 and w_1, respectively), and edges between states are labelled with those agents that cannot distinguish them (in this case, the edge labelled with a and b indicates that neither Anne nor Bob can distinguish between these two possible states). Indeed, we have $M_0 \vDash (\neg K_a p \wedge \neg K_a \neg p) \wedge (\neg K_b p \wedge \neg K_b \neg p)$, so this model properly captures the ignorance of Anne and Bob regarding p.

On top of the basic epistemic framework, AL adds a layer of structure that aims to capture the epistemic dynamics of information update. An **action model** is a structure of the form:

$$A = \langle \Sigma, \{\approx_j : j \in G\}, Pre \rangle$$

- Σ is a (nonempty) set of *epistemic actions*,
- $\approx_j$ is an equivalence relation on Σ,
- $Pre : \Sigma \to \mathcal{L}_K$ is a *precondition function.*

Intuitively, the relation $\approx_j$ captures indistinguishability of actions from the perspective of agent j, while the function Pre captures the background conditions $Pre(\sigma)$ that must hold for a given action σ to be successfully performed. In short, an action model specifies a set of epistemic actions that can be executed together with their preconditions and the extent to which they can be individuated by the agents.

To formalize these intuitions, we must define the process by which an epistemic model M is updated based on the performance of an epistemic action from A. This is captured in the **update model**:

$$M^A = \langle W_\Sigma, \{\sim'_j : j \in G\}, V' \rangle$$

- $W_\Sigma = \{(w, \sigma) : w \vDash Pre(\sigma)\}$,
- $(w_0, \sigma_0) \sim'_j (w_1, \sigma_1)$ iff $w_0 \sim_j w_1$ and $\sigma_0 \approx_j \sigma_1$,
- $(w, \sigma) \in V'(p)$ iff $w \in V(p)$.

Thus, the states of the updated model consist of those state-action pairs (w, σ) such that the precondition of the action σ is satisfied by the state w in M; intuitively, (w, σ) represents the state of affairs w after σ has been performed. The definition of V' ensures that such "updated states" satisfy the same primitive propositions as they did before (corresponding to the intuition that epistemic actions can change the *information* agents have access to, but cannot change basic facts about the world). Finally, the definition of $\sim'_j$ specifies that updated state-action pairs are indistinguishable for agent j precisely when the constituent states and actions were indistinguishable for j in M and A, respectively. It is easy to see that M^A is an epistemic model.

The language for AL, $\mathcal{L}_{K[A,\sigma]}$, extends the basic epistemic language $\mathcal{L}_K$ with an update operator:

$$\varphi := p \mid \neg\varphi \mid \varphi \wedge \psi \mid K_j\varphi \mid [A,\sigma]\varphi$$

$[A,\sigma]\varphi$ is read "if action σ (in the context A) can be performed, then afterwards, φ is true".[9] This language therefore lets us reason about agents' knowledge and how it can change as a result of epistemic actions. Formulas of $\mathcal{L}_{K[A,\sigma]}$ can be interpreted in epistemic models as before, with the additional semantic clause for $[A,\sigma]$ given by:

$$(M,w) \vDash [A,\sigma]\varphi \quad \text{iff} \quad (M,w) \vDash Pre(\sigma) \text{ implies } (M^A,(w,\sigma)) \vDash \varphi.$$

To make these definitions clear, we expand on our previous example.

Example 3. Anne and Bob are discussing whether a particular company policy passed (p). Their friend, Carl, comes along with news of p; however, he announces that for security purposes, he can only tell Bob whether or not p, and cannot tell Anne. He then takes one of two actions: he either tells Bob that p is true (σ_p), or he tells Bob that p is false ($\sigma_{\neg p}$). Anne watches, knowing that Carl is telling Bob whether or not p is true, but too far away to hear. Since we know Carl to be an honest fellow, the precondition of σ_p is p, and the precondition of $\sigma_{\neg p}$ is $\neg p$.

This is all captured in the action model A_0 depicted in Figure 2. In diagrams of action models, the circles represent actions (in this case, σ_p and $\sigma_{\neg p}$), the formulas listed beside the actions are the corresponding preconditions (in this case, p and $\neg p$ for σ_p and $\sigma_{\neg p}$, respectively), and the edges represent action indistinguishability (in this case, Bob can distinguish the two actions, but Anne cannot).

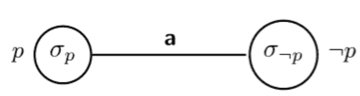

Figure 2: A_0 – Carl's communication to Bob

If we perform the update procedure on M_0 and A_0, then we arrive at the updated model $M_0^{A_0}$ depicted in Figure 3.

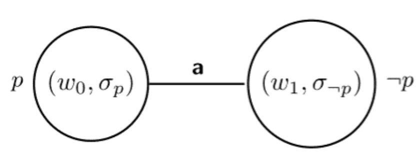

Figure 3: $M_0^{A_0}$ – After Carl's announcement

[9]The "if-then" construction here is interpreted as a standard material conditional, i.e., $[A,\sigma]\varphi$ is vacuously true when σ cannot be performed.

Thus we see that when we update the epistemic model with the action model, the result is a model where Bob knows whether or not p, while Anne does not: $M_0^{A_0} \models K_b p \vee K_b \neg p$, and $M_0^{A_0} \models \neg(K_a p \vee K_a \neg p)$, as desired. $\square$

Note that Carl's announcement is not *public* since Anne cannot tell whether he announced p or $\neg p$ while Bob can.[10]. This demonstrates that AL can capture epistemic dynamics that PAL cannot.[11] Nonetheless, as we have claimed, AL suffers from limitations of its own; we turn now to a discussion of these limitations.

2.2 Limitations of Action Logic

The limitations we have in mind can be summarized quite simply: the AL framework treats the action model as common knowledge. Indeed, the updated model M^A imports much of the structure of the action model: in it, all agents come to know what actions the *other* agents can distinguish. The result is that AL has difficulty capturing scenarios involving *higher-order uncertainty*—e.g., uncertainty about what other agents know. Although this limitation can be overcome, in a sense, by expanding the action model, we will see that in general this is not an appealing solution. Moreover, the examples we consider demonstrate that the AL framework is not as modular as it might appear to be: adjustments to the epistemic model will often require corresponding adjustments to the action model to preserve the intended semantic interpretations. To illustrate these points, we return to the scenario presented in Example 3.

Example 4. Suppose again that Carl comes along and announces that he has news of whether p; however, instead of speaking so that only Bob can hear him, Carl speaks plainly for all to hear, but he delivers the message in French. As it happens, Bob speaks French and Anne does not.

As before, the apparent actions that Carl might take are telling Bob that p and telling Bob that $\neg p$, which presumably have preconditions p and $\neg p$, respectively. Bob can distinguish these actions, as he speaks French, while Anne cannot. This reasoning produces the same action model A_0 depicted in Figure 2, which therefore produces the same updated model $M_0^{A_0}$ shown in Figure 3.

As expected, then, just as in Example 3, Bob ends up knowing whether or not p, and Anne does not. An unexpected result, however, is that Anne *knows this former fact*, and Bob knows the latter. That is, we have:

$$M_0^{A_0} \models K_a(K_b p \vee K_b \neg p)$$

and

$$M_0^{A_0} \models K_b(\neg(K_a p \vee K_a \neg p)).$$

The former says that Anne knows that Bob knows whether p, while the latter says that Bob knows that Anne is uncertain about p. But there was no assumption

[10]This particular sort of announcement, where not all parties are privy to the information exchanged, but all parties are *aware* of the information exchange, is sometimes referred to as "semi-private" in the literature ([13], [30]).

[11]In fact, AL subsumes PAL: to capture a public announcement of φ, consider the action model A_φ consisting of a single node σ with the precondition φ. Then, given an epistemic model M and a state w therein, one can show that the resulting models $(M^{A_\varphi}, (w, \sigma))$ and $(M|_\varphi, w)$ are bisimilar: updating with A_φ effectively deletes the states in M where φ is false.

that Bob knows that Anne cannot understand Carl's message, nor that Anne knows that Bob can. That is, we did not explicitly stipulate whether or not either knew about the other's (in)ability to speak French. To capture this, the model must be refined. □

Loosely speaking, Bob's ability to speak French and Anne's inability to speak French are represented in the structure of the action model A_0 in Figure 2, and this is why they effectively become common knowledge in the updated model. But, of course, we may want to capture a scenario where one or both are *uncertain* about whether or not the other speaks French; indeed, uncertainties of this sort play an important role in everyday reasoning. For the sake of simplicity, let us begin by aiming only to remove the consequence in the updated model that Anne knows that Bob knows whether p.

Example 5. A relevant proposition here is that Bob speaks French; call this q. If we wish to account for Anne's uncertainty about q, we ought to expand the initial epistemic model M_0 to include the possible values this proposition might have. The result is the model M_1 depicted in Figure 4.

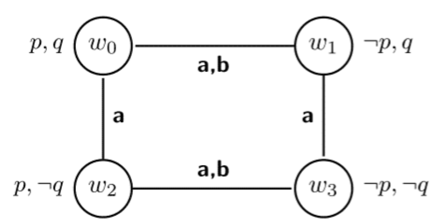

Figure 4: M_1 – An expanded model of Anne and Bob's initial uncertainty

Anne does not know whether or not p is true, and also does not know whether or not q is true; thus, there are a-edges between all 4 nodes. We assume that Bob does know whether or not he speaks French, but, as before, does not know whether or not p; thus there are horizontal b-edges, but no vertical b-edges.

It is easy to check that updating M_1 with A_0 produces an epistemic model $M_1^{A_0}$ in which Bob knows (at every state) the true value of p, and thus Anne knows that Bob knows this. So it seems we must modify the action model as well; for example, we can add a third action, σ, corresponding intuitively to the "unsuccessful" announcement in which Carl speaks his piece but no one (including Bob) understands him. The precondition for σ should therefore be $\neg q$: that Bob does not speak French. Furthermore, the preconditions for the actions σ_p and $\sigma_{\neg p}$ ought to be strengthened to include q, since these actions now represent Carl telling Bob p and $\neg p$, respectively, with the stipulation that *Bob understands what was said*. Bob will be able to distinguish any of these three actions, since he knows whether or not he speaks French, and, given that he speaks French, he knows which announcement Carl is making. Anne, on the hand, will not be able to distinguish any of the three actions—it all sounds the same to her. All this is captured by the action model A_1 given in Figure 5.

The updated model $M_1^{A_1}$ is shown in Figure 6.

As expected, we now have that Bob knows whether p only in those states where σ_p or $\sigma_{\neg p}$ was performed—that is, those states where q ("Bob speaks French") was

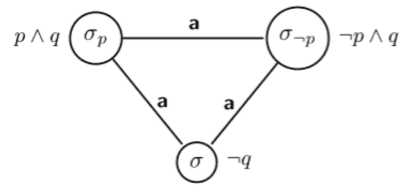

Figure 5: A_1 – An expanded action model

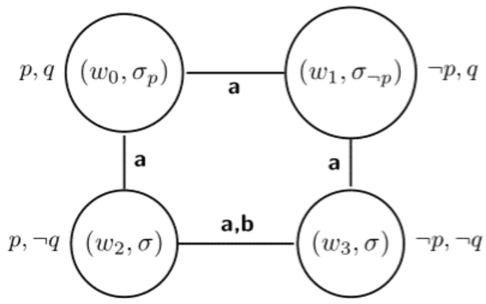

Figure 6: $M_1^{A_1}$ – After Carl's announcement

true—and Anne does not know whether p at all. Furthermore, we have:

$$(w_0, \sigma_p) \vDash \hat{K}_a(K_b p \vee K_b \neg p) \wedge \hat{K}_a(\neg K_b p \wedge \neg K_b \neg p)$$

This reads: Anne considers it possible that Bob knows whether or not p, and also considers it possible that he does not. Thus, by expanding the initial epistemic model to include Anne's uncertainty about Bob's ability to speak French, as well as adding a third node to the action model corresponding (roughly speaking) to an "unsuccessful" announcement from Carl, we are able to capture the second-order uncertainty that we set out to capture. □

One might reasonably feel some discomfort regarding the introduction of σ into the action model. On at least one intuition for what constitutes an "action", Carl's announcement (in French) of p ought to count as the *same* action regardless of who hears it or what languages they might understand. In other words, one might object to distinguishing σ_p from σ on the grounds that it builds into the ontology of actions properties that really have nothing to do with actions, but rather with agents.[12]

This philosophical objection could perhaps be swept aside if the underlying formalism actually did the job we wanted it to: it is hard to make the case that vague ontological concerns ought to receive priority over mathematical efficacy. What we now aim to demonstrate, however, is that this technique of expanding the action model is *not* an effective tool for capturing higher-order uncertainty.

Example 6. We alter the example 5 in only one respect: we assume now that Anne *does* speak French. This requires no change to the initial epistemic model M_1 (since whether or not Anne speaks French is not represented explicitly in this model), but

[12]A similar objection to AL is articulated in [25] section 3.2.1; here, the authors discuss the property of AL which guarantees that any *Epistemic Temporal Logic* (ETL) model generated by AL (in the sense of the correspondence established in [23]) will satisfy the *No Miracles* principle. The alternative semantics we propose below widens the class of ETL models which can be generated to include models which do *not* satisfy the No Miracles principle.

it does, intuitively, require us to re-work the action model A_1. In particular, the Anne-edge connecting σ_p and $\sigma_{\neg p}$ no longer seems appropriate, since now Anne *can* understand what Carl announces.

Now we are faced with a somewhat awkward question—should there be an Anne-edge between σ_p and σ? Intuitively, there should be, since σ is supposed to encode the fact that Bob does not understand Carl's announcement, and Anne is not supposed to be able to tell whether he does or not. Similar reasoning leads us to leave the Anne-edge between $\sigma_{\neg p}$ and σ in place, so the resulting relation fails to be transitive.

Perhaps we could relax the requirements placed on the relations $\approx_i$ in action models to accommodate this type of problem, but in fact there is a deeper issue here that suggests an alternative resolution: it is easy to see that *any* reflexive relation $\approx'_a$ for Anne on the set $\{\sigma_p, \sigma_{\neg p}, \sigma\}$ produces an action model A'_1 such that the updated model $M_1^{A'_1}$ satisfies $(w_2, \sigma) \sim_a (w_3, \sigma)$. But this misrepresents the situation: in state w_2, Carl's announcement must have been that the policy passed, p; as such, after his announcement, Anne should no longer be uncertain about p.

The problem here lies with σ: it was introduced originally to represent the possibility of an "unsuccessful" announcement by Carl. But in the present context, Carl's announcement is always at least partially successful, in that it always informs Anne of the truth value of p. The natural fix to this problem is another adjustment to the action model: we "split" the action σ into two actions, $\sigma_{\neg p \neg q}$ and $\sigma_{p \neg q}$ (and relabel the other actions for clarity). The new action model A_2 is shown in Figure 7.

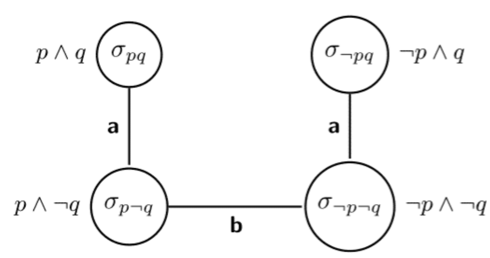

Figure 7: A_2 – The adjusted action model for when Anne speaks French.

These new actions $\sigma_{p\neg q}$ and $\sigma_{\neg p \neg q}$ might be thought of as corresponding to situations where Carl announces p and Bob does not understand, and where Carl announces $\neg p$ and Bob does not understand, respectively. Anne can distinguish announcements based on their content, but not based on whether Bob understands them. Bob can distinguish announcements based on whether he understands them and, provided he understands them, based on their content as well. The updated model $M_1^{A_2}$ is given in Figure 8.

As expected, Anne has learned whether or not p in the updated model (since she heard Carl), but she continues to be ignorant as to whether or not Bob speaks French (and, in turn, whether or not Bob has learned p). $\square$

Note the duplication of effort in the construction of A_2. The initial epistemic model M_1 already encodes the possibilities regarding Bob's ability to speak French and Anne's uncertainty about this. Yet our action model recapitulates this structure with actions that incorporate not just what Carl says, but also whether Bob understands it or not. Moreover, once our background assumptions are fixed (such

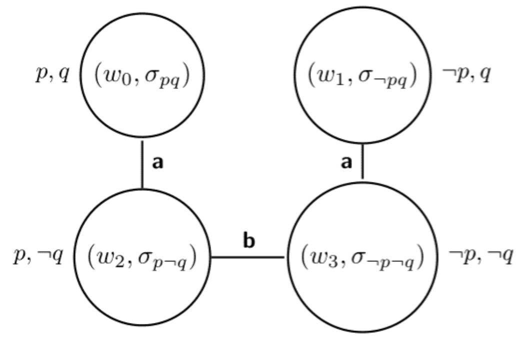

Figure 8: $M_1^{A_2}$ – The updated model when Anne speaks French.

as whether Anne understands French or not), edges (and nodes!) in A_2 are determined, essentially, by examining M_1 and reading off what the uncertainties ought to be. Thus, while AL gives the impression of a clean, modular division between epistemic states and actions, in practice the two seem to be quite tangled, with unavoidable redundancies in their representations. To drive this point home, we sketch one further example.

Example 7. Consider an expanded epistemic model M_2 in which we take not just Bob's but also *Anne's* knowledge of French as endogenous: that is, suppose we also wish to represent Bob as being uncertain of whether or not Anne speaks French. A simple model of such a scenario might consist in eight states representing the possible combinations of truth values for primitive propostions p, q, and r, where p and q are interpreted as before and r stands for the proposition "Anne speaks French".

The action model A_2 of Example 6 is again inadequate. To see why, consider whether there ought to be an Anne-edge connecting σ_{pq} and $\sigma_{\neg pq}$. Intuitively, whether Anne can distinguish Carl announcing p from Carl announcing $\neg p$ depends on the state (i.e., it depends on whether or not Anne can speak French); it is not a fixed and unchanging truth that can be hard-coded into the model. As such, in order to capture this with a fixed action model we require, yet again, a proliferation of actions: e.g., actions of the form $\sigma_{pq\neg r}$, corresponding to something like Carl announcing p and Bob but not Anne understanding it. □

These examples make it clear that the AL formalism is not well-suited to the practical task of building models to represent scenarios in which second-order knowledge is relevant: in addition to specifying the initial epistemic model, one must construct alongside it an elaborate space of actions fine-tuned to the specifics of the epistemic setting. These actions, rather than corresponding in a natural way with concrete events in the world (like Carl making an announcement), are individuated by details about agents' perceptions of them that seem less like part of the actions themselves and more like part of the background epistemic situation. We turn now to our proposed solution to this problem, which simplifies the action space considerably and imports the representation of higher-order uncertainty into the epistemic model, which is designed to handle it.

3 Action-Epistemic Logic

3.1 New Semantics

We propose a new semantics for modeling information update that subsumes AL and is able to capture higher-order uncertainty without the proliferation of actions illustrated in the previous section. In essence, we endogenize the action model, making the distinguishability of the actions state-dependent. It is therefore natural in our revised framework to drop the notion of a separate action model altogether. We call the resulting framework *action-epistemic logic* (AEL).

In addition to a countable collection of primitive propositions PROP and a finite set of agents G, fix a set Σ of *epistemic actions* together with a *precondition* function $Pre : \Sigma \to \mathcal{L}_K$ specifying the precondition for each action as before. A **action-epistemic model** (over Σ and Pre) is a tuple

$$M = \langle W, \{\sim_j : j \in G\}, \{f_j : j \in G\}, V\rangle$$

where $\langle W, \{\sim_j : j \in G\}, V\rangle$ is an epistemic model, and for each agent j, $f_j : W \to 2^{\Sigma\times\Sigma}$ is a function from states to relations on actions. Intuitively, $(\sigma, \sigma') \in f_j(w)$ means that at state w, if action σ is performed, then agent j cannot rule out σ' being the action performed. The f_j functions constitute the crucial novel component of our framework which allows action indistinguishability to vary from state to state.

We will use action-epistemic models to interpret the language $\mathcal{L}^{\Sigma}_{K[\sigma]}$ recursively defined by

$$\varphi := p \mid \neg\varphi \mid \varphi \wedge \psi \mid K_j\varphi \mid [\sigma]\varphi$$

where $\sigma \in \Sigma$ and $[\sigma]\varphi$ is read, as before, "if action σ can be performed, then afterwards, φ is true".[13]

In order to interpret the update modalities, we must define the correct notion of an update model:

$$M^+ = \langle W^+, \{\sim^+_j : j \in G\}, \{f^+_j : j \in G\}, V^+\rangle,^{14}$$

where:

$$\begin{array}{rcl} W^+ & = & \{(w,\sigma) \ : \ w \vDash Pre(\sigma)\} \\ (w,\sigma) \sim^+_j (w',\sigma') & \Leftrightarrow & w \sim_j w' \ \text{ and } \ (\sigma,\sigma') \in f_j(w) \\ f^+_j((w,\sigma)) & = & f_j(w) \\ (w,\sigma) \in V^+(p) & \Leftrightarrow & w \in V(p). \end{array}$$

Note that the clause for $\sim^+_j$ very closely resembles the clause for $\sim'_j$ above, in Action Logic; here, we require that $(\sigma, \sigma') \in f_j(w)$, which serves as the analogue to the requirement $\sigma \approx_j \sigma'$. Thus, although we have relativized the indistinguishability of actions to states, we have not really altered the mechanism which produces the

[13] We no longer need to "tag" the update modality with an action model since our framework does not employ action models; this turns out to have important implications in our proof of completeness.

[14] This notation does not specify *which* action σ the update is being performed with respect to, since (as in AL) our update procedure effectively performs all available updates simultaneously.

update model (though we return to this point in Section 3.4). It is easy to see that M^+ is itself a action-epistemic model over Σ and Pre provided the relations $\sim_j^+$ are equivalence relations. In general, this need not be the case, but we can formulate natural conditions which suffice to ensure that each $\sim_j^+$ is an equivalence relation.

An **action-epistemic frame** is a tuple $F = \langle W, \{\sim_j : j \in G\}, \{f_j : j \in G\}\rangle$, i.e., it is an action-epistemic model without a valuation function. We say that action-epistemic model M is based on frame F just when $M = \langle F, V\rangle$ for some V. Consider the following frame properties:

(Act-Equiv) Each $f_j(w)$ is an equivalence relation on Σ.

(Act-Intro) If $w \sim_j w'$, then $f_j(w) = f_j(w')$.

Proposition 8. *If M is based on a frame which satisfies (Act-Equiv) and (Act-Intro), then M^+ is an action-epistemic model.*

Proof. As noted, it suffices to show that $\sim_j^+$ is an equivalence relation. Reflexivity follows immediately from reflexivity of $\sim_j$ and $f_j(w)$. For symmetry, suppose that $(w, \sigma) \sim_j^+ (w', \sigma')$. Then $w \sim_j w'$ so also $w' \sim_j w$; moreover, by (Act-Intro) we have $(\sigma, \sigma') \in f_j(w) = f_j(w')$, so (Act-Equiv) implies $(\sigma', \sigma) \in f_j(w')$, whence $(w', \sigma') \sim_j^+ (w, \sigma)$. For transitivity, suppose that $(w, \sigma) \sim_j^+ (w', \sigma')$ and $(w', \sigma') \sim_j^+ (w'', \sigma'')$. Clearly $w \sim_j w''$. Moreover, we have $(\sigma, \sigma') \in f_j(w)$ and $(\sigma', \sigma'') \in f_j(w')$; by (Act-Intro) $f_j(w) = f_j(w')$, and by (Act-Equiv) this relation is transitive, so we deduce that $(\sigma, \sigma'') \in f_j(w)$, which shows that $(w, \sigma) \sim_j^+ (w'', \sigma'')$, as desired.[15] $\square$

We consider (Act-Equiv) a natural assumption, given a straightforward reading of 'indistinguishability'.[16] (Act-Intro) states that an agent is introspective with respect to their ability to distinguish actions; that is, if they are actually able/unable to distinguish actions σ and σ', then at all states which they consider possible, they are able/unable to distinguish actions σ and σ'. One does not need to look far for examples where this does not hold; we revisit this principle in section 3.4, where we consider how to update action-epistemic models which do not satisfy it.

We presently only consider action-epistemic models M which satisfy both (Act-Equiv) and (Act-Intro); thus M^+ will again be an action-epistemic model, and we may define

$$(M, w) \vDash [\sigma]\varphi \quad \text{iff} \quad (M, w) \vDash Pre(\sigma) \text{ implies } (M^+, (w, \sigma)) \vDash \varphi.$$

This completes our specification of the new semantics for epistemic actions. It is not hard to see that update by an action model is essentially a special case of update in this framework: given an action model A, one simply defines each f_j to be the constant function such that $f_j(w) = \approx_j$. This directly realizes the intuition that the action model is common knowledge in AL.

[15] In fact, (Act-Intro) is stronger than it needs to be to establish Proposition 8: since $\sim_j$ is assumed to be symmetric; weakening (Act-Intro) to:

(Act-Intro$^-$) If $w \sim_j w'$, then $f_j(w) \subseteq f_j(w')$

has no effect. Of course, this argument would not work for weaker epistemic logics where agents are not assumed to be negatively introspective and $\sim_j$ is not assumed to be symmetric.

[16] This same reading motivates the assumption that $\sim_j$ is an equivalence relation. Of course, there are counterexamples to this principle—for instance, scenarios where the appropriate indistinguishability relation ought to violate transitivity—but for present purposes we ignore such cases.

3.2 Revisiting the examples

We now revisit the problematic examples of section 2.2 and show that the new framework we have developed actually addresses the deficiencies we demonstrated. Let $\Sigma = \{\sigma_p, \sigma_{\neg p}\}$ and set $Pre(\sigma_p) = p$ and $Pre(\sigma_{\neg p}) = \neg p$, corresponding to the two intuitive actions of Carl announcing p or announcing $\neg p$, respectively. Recall that in Example 5, Carl delivers his message in French, we assume that Anne does not know French, and we let q represent the proposition that Bob knows French. We therefore define the action-epistemic model $\tilde{M}_1$ that we will use to reason about this scenario by extending the epistemic model M_1 depicted in Figure 4. In particular, we set

- $f_a(w_0) = f_a(w_1) = f_a(w_2) = f_a(w_3) = \Sigma \times \Sigma$,
- $f_b(w_0) = f_b(w_1) = id_\Sigma$, and
- $f_b(w_2) = f_b(w_3) = \Sigma \times \Sigma$,

where id_X denotes the identity relation on X. Thus, this action-epistemic model encodes the fact that Anne can never distinguish the two actions, whereas Bob can distinguish them just in case he speaks French.

It is easy to see that the epistemic part of $\tilde{M}_1^+$ looks exactly like the model $M_1^{A_1}$ depicted in Figure 6, except with the nodes (w_2, σ) and (w_3, σ) relabeled (w_2, σ_p) and $(w_3, \sigma_{\neg p})$, respectively. In other words, our update produces the "right" epistemic results, and it does so using a simple and natural set of actions and without requiring any fine-tuning of the model beyond the basic association between states where Bob speaks French and states where he can distinguish Carl's two possible announcements.

Next consider the scenario of Example 6, which is just like the previous one except it is assumed that Anne *does* speak French. To capture this, we need only change one line of the previous specifications for $\tilde{M}_1$:

- $f_a(w_0) = f_a(w_1) = f_a(w_2) = f_a(w_3) = id_\Sigma$.

This corresponds directly to the assumption that Anne knows French (i.e., this is valid in the model), so she can always distinguish the two actions in question. Call this action-epistemic model $\tilde{M}_1'$. Now as before, it is straightforward to check that the epistemic part of $\tilde{M}_1'^+$ looks exactly like the model $M_1^{A_2}$ given in Figure 8, provided we replace every instance of σ_{pq} and $\sigma_{p\neg q}$ with σ_p, and every instance of $\sigma_{\neg pq}$ and $\sigma_{\neg p\neg q}$ with $\sigma_{\neg p}$. So again, without the confusion of defining new, abstract actions, our framework reproduces the intended epistemic consequences of Carl's announcement.

Finally, it is not hard to figure out how to define a action-epistemic model $\tilde{M}_2$ extending the epistemic model M_2 of Example 7:

- $f_a(w) = \begin{cases} id_\Sigma & \text{if } w \models r \\ \Sigma \times \Sigma & \text{if } w \models \neg r, \end{cases}$
- $f_b(w) = \begin{cases} id_\Sigma & \text{if } w \models q \\ \Sigma \times \Sigma & \text{if } w \models \neg q. \end{cases}$

3.3 Comparing updates

Interestingly, any action-epistemic update from a model which satisfies (Act-Intro) and (Act-Equiv) may be simulated by an action model:

Theorem 9. *Given any action-epistemic model M over Σ which satisfies (Act-Intro) and (Act-Equiv), there is an action model A^M such that for any $(w, \sigma) \in M^+$, there is a bisimilar state in the updated model M^{A^M}.*[17]

Proof. Let M be an action-epistemic model M over Σ which satisfies (Act-Intro) and (Act-Equiv); we define A^M and construct a bisimulation between M^+ and M^{A^M}. Define $A^M = \langle \Sigma^M, \{\approx_j^M : j \in G\}, Pre^M \rangle$ as follows:[18]

$$\begin{aligned} \Sigma^M &= \{\sigma^w : (w, \sigma) \in M^+\} \\ \sigma^w \approx_j^M \sigma'^{w'} &\Leftrightarrow (\sigma, \sigma') \in f_j(w) \\ Pre^M(\sigma^w) &= th(w) \end{aligned}$$

We now show that the relation B which relates each $(w, \sigma) \in M^+$ to the state $(w, \sigma^w) \in M^{A^M}$ is a bisimulation. It is clear, from the definition of A^M, that $(w, \sigma^w) \in M^{A^M}$; it is also clear that the propositional requirement for bisimulations is satisfied by this relation. ($\Rightarrow$) suppose that $(w, \sigma) \sim_j^+ (w', \sigma')$. Then $(\sigma, \sigma') \in f_j(w)$, and so $\sigma^w \approx_j^M \sigma'^{w'}$ in A^M. Since it must also be true that $w \sim_j w'$ in M, we have that $(w, \sigma^w) \sim_j' (w', \sigma'^{w'})$ in M^{A^M}. ($\Leftarrow$) suppose that $(w, \sigma^w) \sim_j' (w', \sigma'^{w'})$ in M^{A^M}. Then we must have that $w \sim_j w'$ in M and $\sigma^w \approx_j^M \sigma'^{w'}$ in A^M. By the definition of $\approx_j^M$, this implies that $(\sigma, \sigma') \in f_j(w)$ and so it follows that $(w, \sigma) \sim_j^+ (w', \sigma')$ in M^+.

$\square$

Note that, in general, Σ^M will be significantly larger than Σ. The existence of an appropriate A^M should not come as a surprise; action logic captures a very general notion of update. To illustrate this, we demonstrate the conditions under which, given epistemic models M_1, M_2, there is an action model A such that M_1^A represents M_2. We present three theorems, ordered by the strength of the assumptions placed on M_1 and M_2; in theorem 11 we assume that $|M_1|_{\mathcal{L}_K}$ is finite and M_1 is $\mathcal{L}_K$-coarse and in theorem 12 we assume that M_1 is finite and non-redundant. In both cases, these conditions are relatively weak, and show that action logic is capable of simulating a wide class of transformations on epistemic models.[19]

Before stating these theorems, we introduce some terminology: given an epistemic model M, let $|M|_{\mathcal{L}_K}$ denote the number of equivalence classes in M under the relation $\{(w, v) | th(w) = th(v)\}$, where $th(w)$ is the theory of w (the set of formulas in $\mathcal{L}_K$ which are true at M, w). We call an epistemic model M *non-redundant* just in case any two states in the model disagree on some formula of the basic epistemic language; if M is non-redundant, then it follows that $|M| = |M|_{\mathcal{L}_K}$. We say that

[17] We write M^{A^M} to mean the epistemic component of M updated according to the procedure from action logic on action model A^M.

[18] We write $th(w)$ to mean the theory of w in M in the basic epistemic language.

[19] In the theorem, we use epistemic models with only one agent, but this result may by easily generalized to a multi-agent setting.

M is $\mathcal{L}_K$-*coarse* if it satisfies the condition: if $th(w) = th(v)$, then $w \sim v$. In our S5 setting, every non-redundant model is trivially $\mathcal{L}_K$-coarse. We now state and prove these theorems, where $M_1 = \langle W_1, \sim_1, V_1\rangle$ and $M_2 = \langle W_2, \sim_2, V_2\rangle$ are epistemic models:

Theorem 10. *The following conditions are equivalent: (1) there is an action model A such that M_2 is isomorphic to some submodel $X \subseteq M_1^A$ (2) there is a function $g : W_2 \to W_1$ such that (i) $w \in V_2(p)$ iff $g(w) \in V_1(p)$ and (ii) if $w \sim_2 w'$, then $g(w) \sim_1 g(w')$.*

Theorem 11. *If $|M_1|_{\mathcal{L}_K}$ is finite and M_1 is $\mathcal{L}_K$-coarse, then the following conditions are equivalent: (1) there is an action model A such that M_2 is bisimilar to M_1^A (2) there is a function $g : W_2 \to W_1$ such that (i) $w \in V_2(p)$ iff $g(w) \in V_1(p)$ and (ii) if $w \sim_2 w'$, then $g(w) \sim_1 g(w')$.*

Theorem 12. *If M_1 and M_2 are finite and non-redundant, then the following conditions are equivalent: (1) there is an action model A such that M_1^A is isomorphic to M_2 (2) there is a function $g : W_2 \to W_1$ such that (i) $w \in V_2(p)$ iff $g(w) \in V_1(p)$ and (ii) if $w \sim_2 w'$, then $g(w) \sim_1 g(w')$.*

Proof. Theorem 10.
(1) $\Rightarrow$ (2). Suppose that (1) is true of M_1, M_2; let Φ be an isomorphism from M_2 to $X \subseteq M_1^A$. Define $g : W_2 \to W_1$ as follows:[20]

$$g(w) := pr_1(\Phi(w))$$

We show that g satisfies (i) and (ii).

$$w \in V_2(p) \text{ iff } \Phi(w) \in V_{M^A}(p) \text{ iff } pr_1(\Phi(w)) \in V_1(p)$$

$$w \sim_2 w' \text{ iff } \Phi(w) \sim_1^A \Phi(w') \;\Rightarrow\; pr_1(\Phi(w)) \sim_1 pr_1(\Phi(w'))$$

(2) $\Rightarrow$ (1). Let g witness that (2) is true of M_1, M_2. We define an action model $A = \langle \Sigma, \{\approx_j : j \in G\}, Pre\rangle$ as follows:

$$\Sigma = W_2$$

$$w \approx w' \text{ iff } w \sim_2 w'$$

$$Pre(w) = \top$$

We identify the set $X = \{(g(w), w) | w \in W_2\} \subseteq M_1^A$ and define the isomorphism $\Phi(w) = (g(w), w)$.

□

In order to prove theorem 11, we first show a lemma:

Lemma 13. *If B is a bisimulation between (countable) epistemic models M and N, then there is a function $f_B : W_M \to W_N$ such that if $w \sim_M w'$, then $f_B(w) \sim_N f_B(w')$.*

[20] Where pr_1 is the first projection: $pr_1((x, y)) = x$.

Proof. Lemma 13.
For any isolated state w in M (states which are only indistinguishable from themselves), we may define $f_B(w)$ as any element of $B(w)$. Consider any cluster of size $k > 2$ in M, $\{c_i | 1 \leq i \leq k\}$. Let $f_B(c_1)$ be any arbitrary element $v_1 \in B(c_1)$. Since $(c_1, v_1) \in B$ and $c_1 \sim_M c_2$, there must exist $v_2 \in B(c_2)$ with $v_1 \sim v_2$. Let $f_B(c_2) = v_2$. Iterate this process to define f_B. It is immediate that the desired property holds.[21]

□

Proof. Theorem 11.
(1) ⇒ (2). Suppose that (1) is true of M_1, M_2; let $B \subseteq M_2 \times M_1^A$. We set $g = f_B$ the function implied by lemma 13.
Both properties are immediate.
(2) ⇒ (1). Let g witness that (2) is true of M_1, M_2. We define an action model $A = \langle \Sigma, \{\approx_j : j \in G\}, Pre \rangle$ as follows:

$$\Sigma = W_2$$

$$w \approx w' \text{ iff } w \sim_2 w'$$

$$Pre(w) = \bigwedge_{w' \in W_1, th(w') \neq th(g(w))} \varphi_{M_1, g(w), w'}$$

where $\varphi_{M,x,x'}$ is a formula which (M, x) satisfies but (M, x') does not. This conjunction will be a formula since $|M_1|_{\mathcal{L}_K}$ is finite. Note that in the resulting update model, M_1^A, states are of the form (w, v) for $w \in W_1, v \in W_2$. We define the bisimulation B as $\{(v, (w, v)) | (w, v) \in W_1^A\}$. Suppose that $v \in V_2(p)$; then $g(v) \in V_1(p)$. Since $M_1, w \vDash Pre(v)$, $th(w) = th(g(v))$ and so $w \in V_1(p)$; thus $M_1^A, (w, v) \vDash p$.
Consider $(v, (w, v)) \in B$. For the forward direction, suppose that $v \sim v'$; then $g(v) \sim g(v')$. Since $th(w) = th(g(v))$ and M_1 is $\mathcal{L}_K$-coarse, $w \sim g(v)$ and so $w \sim g(v')$. Thus $(w, v) \sim (g(v'), v')$ (and $(g(v'), v') \in B(v')$).
For the backwards direction, suppose that $(w, v) \sim (w', v')$. Then $v \sim_2 v'$ and $(w', v') \in B(v')$ and we are done.

□

Proof. Theorem 12.
(1) ⇒ (2). Suppose that (1) is true of M_1, M_2; let A be the involved action model and let B be a bisimulation from M_2 to M_1^A. For each $w \in W_2$, let $B_w = \{(v, \sigma) \in M_1^A | (w, (v, \sigma)) \in B\}$; from each set B_w, select an arbitrary member b_w. Define $g : W_2 \to W_1$ as follows:

$$g(w) := pr_1(b_w)$$

In words, $g(w)$ is the W_1 component of an arbitrary member of B_w, or the set of states in M_1^A which it is bisimilar to.
We show that g satisfies (i) and (ii).

$$w \in V_2(p) \text{ iff } B(w) \subseteq V_{M_1^A}(p) \text{ iff } pr_1(B(w)) \subseteq V_1(p) \text{ iff }^* b_w \in V_1(p)$$

For the backwards direction of $*$, consider that all of the states in $pr_1(B(w))$ must satisfy all of the same propositions. For (ii), suppose that $w \sim_2 w'$. It follows that

[21] Recall that all relations here are equivalence relations.

there exists $(v, \sigma) \in M_1^A$ such that $(v, \sigma) \in B_{w'}$ and $b_w \sim_1^A (v, \sigma)$; then by the construction of the product update, it follows that $pr_1(b_w) \sim_1 v$.
(2) $\Rightarrow$ (1). Let g witness that (2) is true of M_1, M_2. We define an action model $A = \langle \Sigma, \{\approx_j : j \in G\}, Pre \rangle$ as follows:

$$\Sigma = W_2$$

$$w \approx w' \text{ iff } w \sim_2 w'$$

$$Pre(w) = \bigwedge_{w' \in W_1, w' \neq g(w)} \varphi_{M_1, g(w), w'}$$

where $\varphi_{M,x,x'}$ is a formula which (M, x) satisfies but (M, x') does not (this is guaranteed to exist when M is non-redundant). Note that in the resulting update model, M_1^A, states are of the form $(g(w), w)$ for $w \in W_2$. Now the desired isomorphism Φ is simply the second projection pr_2:

$$pr_2((g(w), w)) = w$$

We verify that pr_2 is an isomorphism. It is easy to see that it is a bijection; suppose that $(g(w), w) \in V_1^A(p)$. This is equivalent to $g(w) \in V_1(p)$ which in turn is equivalent to $w \in V_2(p)$. Suppose that $w \sim_2 w'$. By the properties of g, $g(w) \sim_1 g(w')$, and by the definition of A, $w \approx w'$. Thus, $(g(w), w) \sim_1^A (g(w'), w')$. The other direction goes similarly. □

Note that when proving (2) $\Rightarrow$ (1) above, we chose an appropriate A by essentially 'copying' M_2. This is reminiscent of the reasoning invoked in modeling example 6 above: we examined what the resulting epistemic relations should be, and encoded them in the action model.

3.4 Generalized Semantics

Note that in all of the examples we have been considering so far, (Act-Equiv) and (Act-Intro) have been satisfied, and so the update model produced from our representation of the scenario has been again an action-epistemic model. As noted above, however, counterexamples to (Act-Intro) are easy to come by, such as the following scenario:

Example 14. Anne and Bob are communicating about the viability of a company which is going public tomorrow. For security, they are encrypting their messages to one another. Bob is waiting for an email from Anne – the email will tell him that he should purchase stock (p) or that he should not purchase ($\neg p$). Unfortunately, Bob is not great at remembering or recording passwords – he has an idea for what the password to his private key might be, but he is not sure, and will not know until he tries it. Thus, when Anne's message arrives, Bob is unsure that he will be able to decrypt the message; if he recalls the password correctly (q), then he will decrypt it, and otherwise ($\neg q$) he will not.

We represent the scenario with the action-epistemic model M:
where $\Sigma = \{\sigma_p, \sigma_{\neg p}\}$, $Pre(\sigma_p) = p$, $Pre(\sigma_{\neg p}) = \neg p$, and:

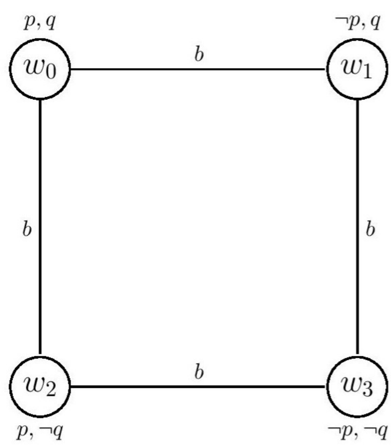

Figure 9: M, the scenario before Bob receives the message

$$f_a(w_0) = f_a(w_1) = f_a(w_2) = f_a(w_3) = \Sigma \times \Sigma$$
$$f_b(w_0) = f_b(w_1) = \Sigma \times \Sigma$$
$$f_b(w_2) = f_b(w_3) = id_\Sigma$$

Note that this model violates (Act-Equiv) because $w_0 \sim_b w_2$, but $(\sigma_p, \sigma_{\neg p}) \in f_b(w_0)$ and $(\sigma_p, \sigma_{\neg p}) \notin f_b(w_2)$. Now, consider the epistemic component of the ensuing update model M^+:

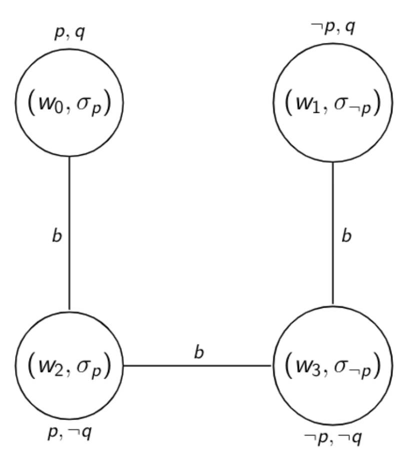

Figure 10: The epistemic part of M^+, the scenario after Bob receives the message

As we would expect, in the subset of states where q is false (that is, where Bob does not recall the correct password), Bob cannot distinguish the state where p is true from the state where p is false. On the other hand, Bob *can* distinguish w_0 and w_1, because q is true there. However, the update model also contains an edge between w_0 and w_2 – this is to say that Bob cannot distinguish the state where he forgot his password, and received (but did not understand) the message that p, from the state where he did *not* forget his password, and received (and did understand) the message that p.

This seems wrong, however – surely Bob knows, after receiving the message that p, whether he remembered his password or not. This should be apparent to him as soon as he decrypts (or fails to decrypt) the message. Perhaps more obviously: if

Bob rules out the possibility that $\sigma_{\neg p}$ was performed (this possibility will be ruled out at, for example, (w_0, σ_p)), then Bob should not consider possible any state where he thinks $\sigma_{\neg p}$ might have been performed (this is a live possibility for Bob at, for example, (w_2, σ_p)). We generalize this intuition as follows: after the update, two states (w, σ) and (w', σ') should be indistinguishable to Bob only if the epistemic actions Bob might have mistaken for σ at w are the same actions Bob might have mistaken for σ' at w'. We codify this in a generalized semantics for the updated relation:[22]

$$(w, \sigma) \sim_j^{\#} (w', \sigma') \Leftrightarrow w \sim_j w' \text{ and } f_j(w)[\sigma] = f_j(w')[\sigma']$$

We first note that this is indeed a generalization of the update semantics for AEL (and so also AL): $\sim_j^{\#}$ specializes to $\sim_j^{+}$ whenever both (Act-Equiv) and (Act-Intro) are satisfied.

Proposition 15. *If M satisfies (Act-Equiv) and (Act-Intro), then $\sim_j^{\#} = \sim_j^{+}$.*

Proof. Suppose that M satisfies (Act-Equiv) and (Act-Intro). We show that under the assumption that $w \sim_j w'$, the condition (#) $f_j(w)[\sigma] = f_j(w')[\sigma']$ is equivalent to the condition (+) $(\sigma, \sigma') \in f_j(w)$. The forwards direction is straightforward; since $f_j(w')$ is reflexive, $\sigma' \in f_j(w')[\sigma'] = f_j(w)[\sigma]$ and so we conclude (+). For the backwards direction, we assume (+) and show (#). Note that the reflexivity of $f_j(w)$ and $f_j(w')$, taken together with (+), implies that $\sigma' \in f_j(w)[\sigma]$ and $\sigma \in f_j(w')[\sigma']$. ($\subseteq$) suppose that $\sigma'' \in f_j(w)[\sigma]$. This implies that σ, σ', and σ'' are all in the same cell of the partition $f_j(w)$. By (Act-Intro), $f_j(w) = f_j(w')$, and so $\sigma'' \in f_j(w')[\sigma']$. The ($\supseteq$) direction is similar. □

For this reason, we drop the notation $\sim_j^{\#}$ and simply redefine $\sim_j^{+}$ (from the definition of the updated model):

$$(w, \sigma) \sim_j^{+} (w', \sigma') \ \Leftrightarrow \ w \sim_j w' \text{ and } f_j(w)[\sigma] = f_j(w')[\sigma']$$

The generalized semantics also obviate the need for assuming (Act-Intro):

Proposition 16. *If M satisfies (Act-Equiv), then M^+ is an action-epistemic model.*

Thus AEL is capable of representing scenarios where action-introspection fails – that is, where an agent is not aware of their abilities to distinguish epistemic actions. We now revisit example 14; the update model for M in figure 9 is now M^+ in figure 11.

The edge between (w_0, σ_p) and (w_2, σ_p) has been removed in virtue of the fact that $\sigma_{\neg p} \in f_b(w_0)[\sigma_p]$ but $\sigma_{\neg p} \notin f_b(w_2)[\sigma_p]$. This represents the fact that when Bob receives message σ_p, he either learns that he can distinguish σ_p, or he learns that he cannot.

We might express the content of this generalization as follows: when an agent j successfully perceives some action σ, they learn not only the content of σ, but also the fact that σ was perceivable to them. Indeed, we may express this fact by expanding our language; we introduce the constants $\xi_{j,\sigma,\sigma'}$ for every $\sigma, \sigma' \in \Sigma$ and $j \in G$ and give the semantics for these constants as follows:

[22] We write $R[x]$ to indicate the image of x in relation R: $R[x] = \{y | (x, y) \in R\}$.

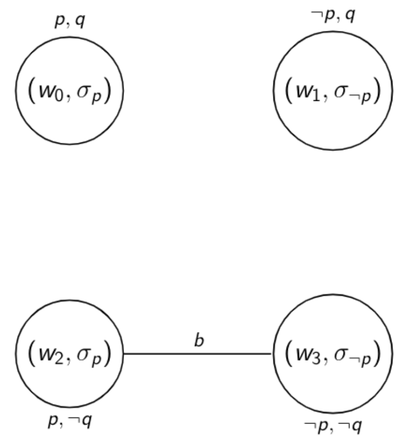

Figure 11: M^+, the scenario after Bob receives the message

$$(M, w) \vDash \xi_{j,\sigma,\sigma'} \text{ iff } (\sigma, \sigma') \in f_j(w)$$

Plainly, $\xi_{j,\sigma,\sigma'}$ is true at state w just when σ and σ' are indistinguishable to j at w. We extend the languages $\mathcal{L}_K$ and $\mathcal{L}^{\Sigma}_{K[\sigma]}$ with these constants to produce the languags $\mathcal{L}_{K\xi}$ and $\mathcal{L}^{\Sigma}_{K[\sigma]\xi}$, respectively. With these constants in hand, we may express the above fact as the following proposition:

Proposition 17. *For any $\sigma, \sigma' \in \Sigma$ and $j \in G$, the following formula is valid: $[\sigma](K_j\xi_{j,\sigma,\sigma'} \vee K_j\neg\xi_{j,\sigma,\sigma'})$.*

Proof. Suppose that $(M^+, (w, \sigma)) \vDash \xi_{j,\sigma,\sigma''}$, and consider any (w', σ') such that $(w, \sigma) \sim^+_j (w', \sigma')$. By the definition of $\sim^+_j$, we know that $f_j(w)[\sigma] = f_j(w')[\sigma']$. Thus since $(\sigma, \sigma'') \in f_j(w)$, we have that $\sigma'' \in f_j(w')[\sigma']$. By reflexivity, $\sigma \in f_j(w)[\sigma] = f_j(w')[\sigma']$, and since $f_j(w')$ is an equivalence relation, we have that $(\sigma, \sigma'') \in f_j(w')$. The case for $\neg\xi_{j,\sigma,\sigma''}$ is similar.

□

Interestingly, even under the generalized semantics, any action-epistemic update may be simulated by an action model:

Theorem 18. *Given any action-epistemic model M over Σ which satisfies (Act-Equiv), there is an action model A^M such that for any $(w, \sigma) \in M^+$, there is a bisimilar state in the updated model M^{A^M}.*[23]

Given an action-epistemic model M over Σ, we define A^M as before, with one modification:

$$\sigma^w \approx^M_j \sigma'^{w'} \Leftrightarrow (\sigma, \sigma') \in f_j(w) \text{ and } f_j(w)[\sigma] = f_j(w')[\sigma']$$

There is a sense in which AEL offers *only* an improvement in efficiency, but does not capture any new dynamic epistemic scenarios. It was mentioned above that any epistemic model paired with an action model may be viewed as a action-epistemic model with constant f_j functions: an agent's ability to distinguish actions is constant over states. This means that an update of an epistemic model M by an action model

[23]We write M^{A^M} to mean the epistemic component of M updated according to the procedure from action logic on action model A^M.

A may be readily captured by the update of a corresponding action-epistemic model $\tilde{M}$; the updated models M^A and $\tilde{M}^+$ will be bisimilar.[24]

There is a correspondence of the same nature in the other direction, too: given any dynamic model $\tilde{M}$, there is an action model A such that $\underset{\sim}{M}^A$ is bisimilar to $\tilde{M}^+$.

The result of this fact is that no *new* epistemic scenarios are captured by AEL – each may be translated back into action logic. However, in many cases this will involve the creation of action models with many more actions than the action-epistemic model had: whenever an action-epistemic model contains indistinguishability functions which are non-constant, new actions must be introduced in the creation of the corresponding action model.

[24] *Bisimilarity* is a notion of equivalence in modal logic; here, it guarantees that the models agree on all formulas of $\mathcal{L}_K$.

4 Completeness

In this section, we prove several soundness and completeness results for AEL. We axiomatize AEL by using a standard reduction technique in section 4.1; this involves axiomatizing the language *without* the update operator, and then designing a meaning-preserving translation from the full language to this fragment. This translation requires that we extend both languages with the ξ constants introduced above; the translation is from $\mathcal{L}^{\Sigma}_{K[\sigma]\xi}$ to $\mathcal{L}_{K\xi}$.

In section 4.2, we axiomatize the language with a common knowledge operator, C_B, where we read $C_B\varphi$ as 'it is common knowledge among the agents in B that φ'. This involves an encoding of action-epistemic logic in propositional dynamic logic (PDL), and an appeal to the completeness of PDL.

4.1 Completeness for AEL

It is well known that both public announcement logic and action logic reduce to epistemic logic; for any formula of $\mathcal{L}_{K[\varphi]}$ or $\mathcal{L}_{K[A,\sigma]}$, there is a formula of $\mathcal{L}_K$ which is equivalent in all epistemic models. In this sense, PAL and AL are no more expressive than EL: they merely offer (very convenient!) shorthand for formulas corresponding to epistemic dynamics.

Unlike PAL and EL, the language of action-epistemic logic does not offer a reduction to the basic epistemic language—it is easy to produce pairs of action-epistemic models whose epistemic parts are bisimilar but which satisfy different formulas of $\mathcal{L}^{\Sigma}_{K[\sigma]}$ at bisimilar states. This rules out the tactic—used to establish the completeness of PAL and AL—of reducing the language to that of EL, and appealing to EL's completeness. Instead, we adapt this tactic by reducing to a stronger language, and showing completeness for that language.

$\mathcal{L}^{\Sigma}_{K[\sigma]\xi}$ and $\mathcal{L}_{K\xi}$ denote the languages $\mathcal{L}^{\Sigma}_{K[\sigma]}$ and $\mathcal{L}_K$, respectively, augmented with the additional constants $\xi_{j,\sigma,\sigma'}$ whose semantics we gave at the end of the previous section. Intuitively, $\xi_{j,\sigma,\sigma'}$ says that if action σ is performed, agent j cannot rule out that σ' was the action performed. Thus, $\mathcal{L}_{K\xi}$ can talk about both knowledge of the agents and action indistinguishability. Somewhat surprisingly, $\mathcal{L}^{\Sigma}_{K[\sigma]\xi}$ is reducible to $\mathcal{L}_{K\xi}$: every formula of $\mathcal{L}^{\Sigma}_{K[\sigma]\xi}$ is equivalent to a formula in $\mathcal{L}_{K\xi}$, and this equivalence can be captured by reduction schemes that allow us to provide a sound and complete axiomatization of $\mathcal{L}^{\Sigma}_{K[\sigma]\xi}$ with respect to the class of all action-epistemic models that satisfy (Act-Equiv). It can be shown, however, that $\mathcal{L}^{\Sigma}_{K[\sigma]\xi}$ is strictly more expressive than $\mathcal{L}^{\Sigma}_{K[\sigma]}$. We turn to the axiomatization of $\mathcal{L}^{\Sigma}_{K[\sigma]\xi}$ now.

4.1.1 Soundness and completeness of $\mathcal{L}_{K\xi}$

Fix a finite set of actions Σ together with a precondition function $Pre : \Sigma \to \mathcal{L}_K$. We begin by axiomatizing $\mathcal{L}_{K\xi}$ and then turn to $\mathcal{L}^{\Sigma}_{K[\sigma]\xi}$. Consider the following axioms:

1. All instantiations of propositional tautologies
2. $K_j(\varphi \to \psi) \to (K_j\varphi \to K_j\psi)$
3. $K_j\varphi \to \varphi$

4. $K_j\varphi \to K_jK_j\varphi$

5. $\neg K_j\varphi \to K_j\neg K_j\varphi$

6. ξ axioms:

   (a) $\xi_{j,\sigma,\sigma}$

   (b) $\xi_{j,\sigma,\sigma'} \to \xi_{j,\sigma',\sigma}$

   (c) $\xi_{j,\sigma,\sigma'} \to (\xi_{j,\sigma',\sigma''} \to \xi_{j,\sigma,\sigma''})$

7. From φ and $\varphi \to \psi$, infer ψ.

8. From φ, infer $K_j\varphi$.

Soundness is easy, while completeness can be established by a fairly standard canonical model construction.

Let $M^c = \langle W^c, \{\sim_j^c : j \in G\}, \{f_j^c : j \in G\}, V^c\rangle$ be defined as follows:

- $M^c = \{\Gamma \subseteq \mathcal{L}_{K\xi} : \Gamma \text{ is maximally consistent}\}$
- $\Gamma \sim_j^c \Delta$ iff $\{K_j\varphi : K_j\varphi \in \Gamma\} = \{K_j\varphi : K_j\varphi \in \Delta\}$
- $(\sigma, \sigma') \in f_j^c(\Gamma)$ iff $\xi_{j,\sigma,\sigma'} \in \Gamma$
- $\Gamma \in V^c(p)$ iff $p \in \Gamma$.

We note that M^c is, as desired, an action-epistemic model over (Σ, Pre) which satisfies (Act-Equiv). That $\sim_j^c$ is an equivalence relation follows from standard proofs of the completeness of $\mathsf{S5}$. Axioms 6a-c guarantee that M^c satisfies condition (Act-Equiv).

The equivalence:

$$(M^c, \Gamma) \vDash \varphi \text{ iff } \varphi \in \Gamma$$

(i.e., the Truth Lemma) is proved in the standard way by structural induction.

4.1.2 Soundness and completeness of $\mathcal{L}^{\Sigma}_{K[\sigma]\xi}$

Let pre_σ abbreviate the epistemic formula $Pre(\sigma)$, and consider the following additional axioms and inference rule:

9. Action axioms:

   (a) $[\sigma](\varphi \to \psi) \to ([\sigma]\varphi \to [\sigma]\psi)$

   (b) $[\sigma]p \leftrightarrow (pre_\sigma \to p)$

   (c) $[\sigma]\neg\varphi \leftrightarrow (pre_\sigma \to \neg[\sigma]\varphi)$

   (d) $[\sigma](\varphi \wedge \psi) \leftrightarrow ([\sigma]\varphi \wedge [\sigma]\psi)$

   (e) $[\sigma]K_j\varphi \leftrightarrow \Big(pre_\sigma \to \bigwedge_{\sigma' \in \Sigma} \big(\xi_{j,\sigma,\sigma'} \to [j, \sigma, [\sigma']\varphi]\big)\Big)$

10. From φ, infer $[\sigma]\varphi$.

Axiom 9e uses the following abbreviations:

$$\xi_{j,\sigma,\Sigma'} := \bigwedge_{\sigma'\in\Sigma'} \xi_{j,\sigma,\sigma'} \wedge \bigwedge_{\sigma''\notin\Sigma'} \neg\xi_{j,\sigma,\sigma''}$$

$$[j,\sigma,\varphi] := \bigwedge_{\Sigma'\subseteq\Sigma} \big(\xi_{j,\sigma,\Sigma'} \rightarrow K_j(\xi_{j,\sigma,\Sigma'} \rightarrow \varphi)\big)$$

$\xi_{j,\sigma,\Sigma'}$ is true at w when Σ' is exactly the set of actions which j cannot distinguish from σ at w. We will write $\Sigma_{j,w,\sigma}$ to indicate the unique subset of Σ such that $w \vDash \xi_{j,\sigma,\Sigma_{j,w,\sigma}}$.[25] $[j,\sigma,\varphi]$ is true at w if, for any state w' which (1) is j-reachable from w and (2) satisfies $\xi_{j,\sigma,\Sigma_{j,w,\sigma}}$, w' satisfies φ.

The soundness of most of the above axioms is immediate; we show the soundness of (9e).

($\Rightarrow$) Suppose that $(M,w) \vDash [\sigma]K_j\varphi$ and assume that $(M,w) \vDash pre_\sigma$ (otherwise the equivalence is trivial). Let $\sigma' \in \Sigma$ be such that $w \vDash \xi_{j,\sigma,\sigma'}$. Consider any state w' such that $w \sim_j w'$ and $w' \vDash \xi_{j,\sigma,\Sigma_{j,w,\sigma}}$. We wish to show that $w' \vDash [\sigma']\varphi$. By supposition, $(M^+,(w,\sigma)) \vDash K_j\varphi$. Since $w' \vDash \xi_{j,\sigma,\Sigma_{w,\sigma}}$, and $\sigma' \in \Sigma_{j,w,\sigma}$, we have that $f_j(w)[\sigma] = f_j(w')[\sigma']$. Thus, $(w,\sigma) \sim^+_j (w',\sigma')$; it follows that $(w',\sigma') \vDash \varphi$, which means that $w' \vDash [\sigma']\varphi$.

($\Leftarrow$) Suppose that $w \vDash \Big(pre_\sigma \rightarrow \bigwedge_{\sigma'\in\Sigma} \big(\xi_{j,\sigma,\sigma'} \rightarrow [j,\sigma,[\sigma']\varphi]\big)\Big)$, and that $w \vDash pre_\sigma$. We wish to show that $(w,\sigma) \vDash K_j\varphi$. Consider any $(w',\sigma') \sim^+_j (w,\sigma)$; we will show that $(w',\sigma') \vDash \varphi$. Since $(\sigma,\sigma') \in f_j(w)$, we know that $w \vDash \xi_{j,\sigma,\sigma'}$, so by supposition, $w \vDash [j,\sigma,[\sigma']\varphi]$. By the semantics of $\sim^+_j$, we know that $f_j(w)[\sigma] = f_j(w')[\sigma']$; this implies that $w' \vDash \xi j,\sigma,\Sigma_{j,w,\sigma}$, and so $w' \vDash [\sigma']\varphi$. This then implies that $(w',\sigma') \vDash \varphi$, as desired.

For completeness, consider the following translation:

$$\begin{array}{rcl}
t(p) &=& p \\
t(\xi_{j,\sigma,\sigma'}) &=& \xi_{j,\sigma,\sigma'} \\
t(\neg\varphi) &=& \neg t(\varphi) \\
t(\varphi\wedge\psi) &=& t(\varphi)\wedge t(\psi) \\
t(K_a\varphi) &=& K_a t(\varphi) \\
t([\sigma]p) &=& pre_\sigma \rightarrow p \\
t([\sigma]\neg\varphi) &=& pre_\sigma \rightarrow \neg t([\sigma]\varphi) \\
t([\sigma](\varphi\wedge\psi)) &=& t([\sigma]\varphi)\wedge t([\sigma]\psi) \\
t([\sigma]K_j\varphi) &=& pre_\sigma \rightarrow \bigwedge_{\sigma'\in\Sigma} \big(\xi_{j,\sigma,\sigma'} \rightarrow [j,\sigma,t([\sigma']\varphi)]\big) \\
t([\sigma][\sigma']\varphi) &=& t([\sigma]t([\sigma']\varphi))
\end{array}$$

Proposition 19. *For all formulas $\varphi \in \mathcal{L}^{\Sigma}_{K[\sigma]\xi}$, $t(\varphi)$ is provably equivalent to φ and $t(\varphi) \in \mathcal{L}_{K\xi}$.*

Proof. We proceed by induction on the *action nesting depth* of φ, defined in the

[25] Note that this is trivially identical to the set $f_j(w)[\sigma]$.

obvious way:

$$
\begin{aligned}
d(p) &= 0 \\
d(\xi_{j,\sigma,\sigma'}) &= 0 \\
d(\neg\varphi) &= d(\varphi) \\
d(\varphi \wedge \psi) &= \max(d(\varphi), d(\psi)) \\
d(K_j\varphi) &= d(\varphi) \\
d([\sigma]\varphi) &= d(\varphi) + 1.
\end{aligned}
$$

The case $d(\varphi) = 0$ is immediate. So suppose the result holds for all formulas with nesting depth less than n, and let $\varphi \in \mathcal{L}^{\Sigma}_{K[\sigma]\xi}$ be such that $d(\varphi) \leq n$.

We now proceed via a subinduction on the *weight* of φ, defined as follows:

$$
\begin{aligned}
w(p) &= 1 \\
w(\xi_{j,\sigma,\sigma'}) &= 1 \\
w(\neg\varphi) &= w(\varphi) + 1 \\
w(\varphi \wedge \psi) &= max(w(\varphi), w(\psi)) + 1 \\
w(K_j\varphi) &= w(\varphi) + 1 \\
w([\sigma]\varphi) &= w(\varphi) + 1
\end{aligned}
$$

The base case where $w(\varphi) = 1$ is again immediate. So suppose inductively the result holds for formulas of weight less than $w(\varphi)$. The proof now breaks into cases depending on the structure of φ, since this determines which recursive clause of the definition of t is relevant. The inductive steps corresponding to the Boolean connectives and the K_a modalities are straightforward, so we move to the case where $\varphi = [\sigma]\psi$; this in turn naturally breaks into several subcases depending on the structure of ψ:

- If $\psi = p$, then $t(\varphi) = t([\sigma]p) = pre_\sigma \rightarrow p$ and we are done by axiom (9b).

- If $\psi = \neg\chi$, then $t(\varphi) = t([\sigma]\neg\chi) = pre_\sigma \rightarrow \neg t([\sigma]\chi)$. Clearly $w([\sigma]\chi) < w(\varphi)$, so by the inductive hypothesis we know that $t([\sigma]\chi) \in \mathcal{L}_{K\xi}$ and is provably equivalent to $[\sigma]\chi$. It follows immediately that $t(\varphi) \in \mathcal{L}_{K\xi}$ and, by axiom (9c), that $t(\varphi)$ is provably equivalent to φ.

- If $\psi = \chi_1 \wedge \chi_2$, then $t(\varphi) = t([\sigma]\chi_1) \wedge t([\sigma]\chi_2)$. Clearly $w([\sigma]\chi_1) < w(\varphi)$ and $w([\sigma]\chi_2) < w(\varphi)$, so by the inductive hypothesis we know that $t([\sigma]\chi_1) \in \mathcal{L}_{K\xi}$ and $t([\sigma]\chi_2) \in \mathcal{L}_{K\xi}$, and they are provably equivalent to $[\sigma]\chi_1$ and $[\sigma]\chi_2$, respectively. It follows immediately that $t(\varphi) \in \mathcal{L}_{K\xi}$ and, by axiom (9d), that $t(\varphi)$ is provably equivalent to φ.

- If $\psi = K_j\chi$, then $t(\varphi) = t([\sigma]K_j\chi) = pre_\sigma \rightarrow \bigwedge_{\sigma' \in \Sigma} \big(\xi_{j,\sigma,\sigma'} \rightarrow [j, \sigma, t([\sigma']\chi)]\big)$. Clearly, for each $\sigma' \in \Sigma$, $w([\sigma']\chi) < w(\varphi)$, so by the inductive hypothesis we know that $t([\sigma']\chi) \in \mathcal{L}_{K\xi}$ and is provably equivalent to $[\sigma']\chi$. It follows immediately that $t(\varphi) \in \mathcal{L}_{K\xi}$ and, by axiom (9e), that $t(\varphi)$ is provably equivalent to φ, as desired.

- Finally, if $\psi = [\sigma']\chi$, then $t(\varphi) = t([\sigma][\sigma']\chi) = t([\sigma]t([\sigma']\chi))$. Let $\tilde{\psi} = t([\sigma']\chi)$; clearly $w([\sigma']\chi) < w(\varphi)$, so by the inductive hypothesis we know that $\tilde{\psi} =$

$t([\sigma']\chi) \in \mathcal{L}_{K\xi}$ and $\tilde{\psi}$ is provably equivalent to $[\sigma']\chi$. So we have $t(\varphi) = t([\sigma]t([\sigma']\chi)) = t([\sigma]\tilde{\psi})$. Now the weight of $\tilde{\psi}$ may be very large, so we can't apply our inner inductive hypothesis again here. However, since $\tilde{\psi} \in \mathcal{L}_{K\xi}$, it is easy to see that $d([\sigma]\tilde{\psi}) = 1 < d(\varphi)$, so we can appeal to our *outer* inductive hypothesis to conclude that $t(\varphi) = t([\sigma]\tilde{\psi}) \in \mathcal{L}_{K\xi}$ and is provably equivalent to $[\sigma]\tilde{\psi}$. Moreover, since $\tilde{\psi}$ is provably equivalent to $[\sigma']\chi$, using axiom (9a) it is easy to show that $[\sigma]\tilde{\psi}$ is provably equivalent to $[\sigma][\sigma']\chi$, whence $t(\varphi)$ is provably equivalent to φ, as desired. □

The argument for completeness now runs as follows: consider some valid formula $\varphi \in \mathcal{L}^{\Sigma}_{K[\sigma]\xi}$. By 19, $t(\varphi)$ is provably equivalent to φ and so must also be valid (by soundness). Since $t(\varphi) \in \mathcal{L}_{K\xi}$, the completeness of $\mathcal{L}_{K\xi}$ ensures there is some proof of $t(\varphi)$ using axioms 1-9; we may combine this proof with the proof of the equivalence of φ and $t(\varphi)$ to get a proof of φ using the axioms for $\mathcal{L}^{\Sigma}_{K[\sigma]\xi}$, which concludes our argument.

4.1.3 Completeness with Act-Intro

If we restrict our attention to the class of models which satisfy (Act-Intro), then we may modify our axiom system in the following respect to get a sound and complete axiomatization.

We add the following axiom to 6:[26]

$$6(d) \quad \xi_{j,\sigma,\sigma'} \rightarrow K_j \xi_{j,\sigma,\sigma'}$$

The soundness of 6(d) in the presence of (Act-Intro) is easy to see: suppose that $\xi_{j,\sigma,\sigma'}$ is true at M, w, and that $w \sim_j w'$. This implies that $(\sigma, \sigma') \in f_j(w)$, and by Act-Intro, that $f_j(w) = f_j(w')$; thus, $M, w' \models \xi_{j,\sigma,\sigma'}$, as desired.

We also confirm that the resulting canonical model satisfies (Act-Intro): suppose that $\Gamma \sim^c_j \Gamma'$ and that $(\sigma, \sigma') \in f^c_j(\Gamma)$. By definition, this implies that $\xi_{j,\sigma,\sigma'} \in \Gamma$ and by axiom 6(d), $K_j\xi_{j,\sigma,\sigma'} \in \Gamma$. Then by the definition of $\sim^c_j$, $K_j\xi_{j,\sigma,\sigma'} \in \Gamma'$, and by axiom 3, $\xi_{j,\sigma,\sigma'} \in \Gamma'$. Thus, $(\sigma, \sigma') \in f^c_j(\Gamma')$, as desired; the other direction is analogous.

In the presence of 6(d), 9(e) becomes equivalent to a much simpler form:

$$9(e*) \quad [\sigma]K_j\varphi \leftrightarrow \Big(pre_\sigma \rightarrow \bigwedge_{\sigma' \in \Sigma} \left(\xi_{j,\sigma,\sigma'} \rightarrow K_j[\sigma']\varphi\right)\Big)$$

4.2 Completeness for AEL with common knowledge

In this section, we discuss the consequences of introducing *common knowledge* to our language and semantics. We first discuss how the introduction of common knowledge impacts related languages and their axiomatizations, and then show that in our case, a completeness proof is achievable.

Common knowledge is an important concept in the study of knowledge and strategic reasoning. Standard game-theoretic analyses assume common knowledge of rationality between players—determining a player's best strategy is made tractable by the

[26]We might think of this as 'positive introspection' of the ξ formulas. Negative introspection, given by the schema $\neg\xi_{j,\sigma,\sigma'} \rightarrow K_j\neg\xi_{j,\sigma,\sigma'}$, is derivable from this axiom in the presence of the S5 axioms for K_j.

assumption that all other players are rational, and that this fact is common knowledge. Some research has systematically weakened this assumption to determine its effect on strategy analysis. Outside of these fields of research, common knowledge is ubiquitous in everyday reasoning. The efficacy of conversations, for instance, is contingent on some method of communication which is common knowledge to all participating parties.

As such, common knowledge is a valuable concept to represent in an epistemic logic. We turn to a formal characterization of the concept now. Intuitively, 'it is common knowledge to everyone in group B that φ' will be true just when all the conjuncts of 'everyone in B knows φ, and everyone in B knows that everyone in B knows φ, and ...' are simultaneously true. We can make this precise by introducing shorthand for 'everybody in group B knows that φ' and its iterations:

$$E_B\varphi = \bigwedge_{j \in B} K_j\varphi$$
$$E_B^n\varphi = \underbrace{E_B E_B \dots E_B}_{n \text{ times}}\varphi$$

Using this, we can provide semantics for the common knowledge operator, C_B:

$$(M, w) \vDash C_B\varphi \text{ iff } (\forall n \in \mathbb{N})\ (M, w) \vDash E_B^n\varphi$$

It is helpful to think of the semantics of this operator in terms of 'edge traversal': $C_B\varphi$ is true at a state w when every state reachable via only B-edges is a state which satisfies φ.

Several prominent epistemic languages have been extended with a common knowledge operator—for instance, $\mathcal{L}_K$, $\mathcal{L}_{K[\varphi]}$, and $\mathcal{L}_{K[A,\sigma]}$ are extended to $\mathcal{L}_{KC}$, $\mathcal{L}_{K[\varphi]C}$, and $\mathcal{L}_{K[A,\sigma]C}$, respectively, by adding the clause $C_B\varphi$, where $B \subseteq G$. We define the languages $\mathcal{L}_{KC\xi}$ and $\mathcal{L}^{\Sigma}_{K[\sigma]C\xi}$ by extending $\mathcal{L}_{K\xi}$ and $\mathcal{L}^{\Sigma}_{K[\sigma]\xi}$, respectively, in the same way.

The logics associated with $\mathcal{L}_{KC}$, $\mathcal{L}_{K[\varphi]C}$, and $\mathcal{L}_{K[A,\sigma]C}$ have been given sound and complete axiomatizations in [22, Chapter 7]. The inclusion of a common knowledge operator, however, necessitates a dramatically different tactic when establishing completeness for logics which include dynamic operators. To provide a sound and complete axiom system for PAL *without* common knowledge, a straightforward reduction suffices—one shows that every formula in $\mathcal{L}_{K[\varphi]}$ is equivalent to a formula in $\mathcal{L}_K$ via a set of reduction axioms, and invokes completeness of the axiom system $S5$.[27] The resulting axiom system for PAL consists in the axioms of $S5$ plus these reduction axioms. One might think that a similar approach can be taken for $\mathcal{L}_{K[\varphi]C}$: we show that every formula in $\mathcal{L}_{K[\varphi]C}$ is equivalent to a formula in $\mathcal{L}_{KC}$, and appeal to a sound and complete axiomatization of the $\mathcal{L}_{KC}$. This proves impossible, however: interestingly, the introduction of C_B to both languages does not preserve the reducibility of $\mathcal{L}_{K[\varphi]}$ to $\mathcal{L}_K$. In particular, $[\psi]C_B\varphi$ cannot be transformed into an equivalent formula of $\mathcal{L}_{KC}$. Intuitively, the reason is that this formula will be true at w just when all states which are reachable by $\psi - B$ paths—paths consisting of only B-edges and ψ-states—satisfy φ after the update. However, there is no mechanism in $\mathcal{L}_{KC}$ for quantifying over this set. Note that as a consequence of this

[27] A method of proving completeness of PAL without reduction axioms is presented in [32].

irreducibility, $\mathcal{L}_{K[A,\sigma]C}$ must also be irreducible to $\mathcal{L}_{KC}$, since public announcement logic can be expressed within action logic.
Since completeness by reduction is not an option for PAL with common knowledge (PALC), the standard method—the construction of a canonical model—is used to show completeness. This method is used successfully for both PALC and ALC (AL with common knowledge), but it requires substantial adaptation to accommodate common knowledge. This is because, roughly, the infinitary nature of $C_B\varphi$ makes both logics non-compact, and so the standard canonical model does not suffice. This problem is overcome by catering the canonical model to the formula in question; for details, see [22, Chapter 7].
Turning to action-epistemic logic (AELC) with common knowledge, we see that completeness by reduction is also out of reach, due to the simple fact that ALC can be captured in AELC, and $\mathcal{L}_{K[A,\sigma]C}$ cannot be reduced to $\mathcal{L}_{KC}$.[28] Instead, we might try to construct canonical models in the same way done for PALC and ALC.
Unfortunately, this, too, proves problematic. As mentioned above, in this construction, the canonical model we construct is particular to the unprovable formula for which a counterexample is desired. The problem arises in the proof of the *Truth Lemma*, which asserts, roughly, that our canonical model for φ has the property necessary to serve as a counterexample to φ.[29] The difficulty arises with the treatment of formulas of the form $[\alpha][\beta]\varphi$—these cases require that we look to the *update* canonical model, which lies outside of the scope of the inductive hypothesis. Several details are relevant to this problem that we do not mention here, but as of now, we see no way forward in this proof.
Another option, however, is afforded by an alternative proof in [14]. In this short article, Kooi and van Benthem discuss the difficulties with proving completeness when common knowledge is introduced to dynamic logics. They introduce a new operator, called *relativized common knowledge*, expressed by the formula $C_B(\psi, \varphi)$. $C_B(\psi, \varphi)$ is true at a state w just when every ψ-B-path (all paths consisting of only B-edges, and encountering only ψ-states) from w ends in a state which satisfies φ.[30] Although the operator resists a completely natural interpretation, it offers a solution to the reducibility problem discussed above: public announcement logic with relativized common knowledge (PALRC) is reducible to epistemic logic with relativized common knowledge (S5RC). This is because the relativized common knowledge operator captures precisely the conditions for when $[\psi]C_B\varphi$ holds, which was the problematic case in the attempt at reduction above. Axiomatizing S5RC is a straightforward adaptation of the axiomatization for S5C; with this, completeness has been established for PALRC. Somewhat counterintuitively, it is by *strengthening* PALC that an appropriate reduction proof is found, and completeness established. This is similar to the approach taken to establish completeness for AEL.
While relativized common knowledge works well in the case for PAL, it will not work for AL or AEL. This is because $[\alpha]C_B\varphi$ evades reduction, even when relativized common knowledge is in play. The conditions for the satisfaction of $[\alpha]C_B\varphi$ involve

[28] This establishes only that $\mathcal{L}^{\Sigma}_{K[\sigma]C}$ cannot be reduced to $\mathcal{L}_{KC}$; however, it follows immediately that $\mathcal{L}^{\Sigma}_{K[\sigma]C\xi}$ cannot be reduced to $\mathcal{L}_{K[A,\sigma]C}$.

[29] In more detail: the states of the canonical model are maximal consistent sets of formulas from the language in question. The Truth Lemma shows that if a formula ψ is a member of some state Γ, then in the canonical model, $\Gamma \vDash \psi$.

[30] Note that the standard common knowledge operator may be recovered by simply setting $\psi = \top$.

paths in both the epistemic model *and* the action model—or in the case of action-epistemic logic, paths in W and properties of $\{f_j | j \in G\}$—and the language without an update operator is not powerful enough to capture these conditions, even with relativized common knowledge.

In [15], van Eijck demonstrates another means of showing completeness for ALC. This method is via *Propositional Dynamic Logic* (PDL), a logic designed for the description of programs which is prominent in computer science. In the language of PDL, $\mathcal{L}_{PDL}$, modalities correspond to programs; formulas are evaluated in Kripke structures, and the execution of a program is indicated by a relation over the possible states of affairs. The set of programs which may be invoked by formulas is closed under certain operations; these closure conditions ensure that if each agent's relation is included in the set as a program, then there exists a program in the set which may be interpreted as the relation for common knowledge for some subset of agents. Van Eijck shows that $\mathcal{L}_{K[A,\sigma]C}$ may be expressed using $\mathcal{L}_{PDL}$ extended with *program transformations*, which are functions on the space of programs. A sound and complete axiomatization is already known for PDL with program transformations; van Eijck combines this axiomatization with his reducibility result to provide an complete axiomatization for ALC.[31]

We take a similar tactic here. We extend PDL to *action-epistemic PDL* by introducing an update operator to the logic and interpreting formulas on action-epistemic models which satisfy (Act-Equiv). The language of PDL supplemented with update operators will be notated by $\mathcal{L}^{\Sigma}_{PDL[\sigma]}$; the language which additionally includes ξ formulas will be $\mathcal{L}^{\Sigma}_{PDL[\sigma]\xi}$. This language is rich enough to capture $\mathcal{L}^{\Sigma}_{K[\sigma]C\xi}$. We then show, using an adaptation of van Eijck's proof, that the dynamic modalities included in $\mathcal{L}^{\Sigma}_{PDL[\sigma]\xi}$ do not increase expressive power:

Theorem 20. *There is a truth-preserving translation from $\mathcal{L}^{\Sigma}_{PDL[\sigma]\xi}$ to $\mathcal{L}^{\Sigma}_{PDL\xi}$.*

Lastly, we append the reduction axioms used in this proof onto a complete axiomatization for PDL (with ξ operators) for a complete axiomatization of $\mathcal{L}^{\Sigma}_{K[\sigma]C\xi}$.

4.2.1 Defining $\mathcal{L}^{\Sigma}_{PDL[\sigma]\xi}$

We begin by defining the language of PDL, and then extend the language with an update operator and the ξ formulas. As before, this language is relative to a fixed finite set of agents G, and a fixed finite set of actions Σ with preconditions Pre. We enumerate $\Sigma : \Sigma = \{\sigma_0, \ldots, \sigma_{n-1}\}$. i and j will be used to denote the indices of actions, and will range over $[0, n)$. $\mathcal{L}_{PDL}$ is generated by the following:

$$p \mid \neg\varphi \mid \varphi \wedge \psi \mid [\pi]\varphi$$

where $\pi \in \Pi$; Π is the set of programs generated inductively with the following clauses:

$$a \in G \mid \pi; \pi' \mid \pi \cup \pi' \mid \pi^* \mid ?\varphi$$

where φ is a formula in $\mathcal{L}_{PDL}$. Here $\pi; \pi'$ denotes sequential composition, $\pi \cup \pi'$ denotes non-deterministic choice, π^* denotes arbitrary iteration, and $?\varphi$ denotes a

[31] An abbreviated form of this proof is also presented in [16]. This proof was inspired by one similar in spirit presented in [14] – this proof uses *automata* instead of program transformations.

test. For any $B \subseteq G$, We will abbreviate $\bigcup_{a \in B} a$ with B. $\mathcal{L}^{\Sigma}_{PDL\xi}$ includes all previous formula clauses as well as the clause $\xi_{a,\sigma_i,\sigma_j}$ for $a \in G$, $\sigma_i, \sigma_j \in \Sigma$. $\mathcal{L}^{\Sigma}_{PDL[\sigma]\xi}$ includes this additional clause and the clause $[\sigma]\varphi$ for $\sigma \in \Sigma$ and $\varphi \in \mathcal{L}^{\Sigma}_{PDL[\sigma]\xi}$. Lastly, a program transformation is a function $r : \Pi \to \Pi$.

We interpret the formulas of $\mathcal{L}^{\Sigma}_{PDL[\sigma]\xi}$ in an action-epistemic model, $M = \langle W, \{\sim_j: j \in G\}, \{f_j : j \in G\}, V\rangle$, over Σ and Pre. Semantics for boolean operators, the update operator, and $\xi_{a,\sigma_i,\sigma_j}$ are as before, and semantics for the remaining formulas are as follows:[32]

- $(M,w) \vDash [\pi]\varphi$ iff $(\forall v \in R^M_\pi(w))((M,v) \vDash \varphi)$

where R^M_π is defined inductively, using the relations $\sim_a$ in M, as follows:

- $R^M_a = \sim_a$
- $R^M_{\pi;\pi'} = \{(w,w'')|(\exists w')((w,w') \in \pi \ \wedge \ (w',w'') \in \pi')\}$
- $R^M_{\pi\cup\pi'} = \{(w,w')|(w,w') \in \pi \ \vee \ (w,w') \in \pi'\}$
- $R^M_{\pi^*} = \{(w,w')|(\exists m \geq 0)(\exists w_0, \ldots, w_m)(w = w_0 \wedge w' = w_m \wedge (\forall i \in [1,m])(w_i, w_{i+1}) \in \pi)\}$
- $R^M_{?\varphi} = \{(w,w)|M, w \vDash \varphi\}$

Thus, $R^{M^+}_\pi$ will also be defined inductively using the relations $\sim^+_a$.

This language is sufficient to capture the language $\mathcal{L}^{\Sigma}_{K[\sigma]C\xi}$ via the following interpretation (this interpretation is used in [15] and [16]): we interpret all formulas as themselves except in the following cases:

$$K_a\varphi := [a]\varphi$$

$$C_B\varphi := [B^*]\varphi$$

4.2.2 Reducing to $\mathcal{L}^{\Sigma}_{PDL\xi}$

Now we show that $\mathcal{L}^{\Sigma}_{PDL[\sigma]\xi}$ may be reduced to $\mathcal{L}^{\Sigma}_{PDL\xi}$. We do this by providing a truth-preserving translation t from the former to the latter, defined as follows on the formulas and programs of $\mathcal{L}^{\Sigma}_{PDL[\sigma]\xi}$:

[32] The overlap in notation between a program operator $[\pi]$ and an epistemic action $[\sigma]$ is unfortunate; however, we will make it clear which we mean, with every use.

$$
\begin{aligned}
t(p) &= p \\
t(\xi_{j,\sigma,\sigma'}) &= \xi_{j,\sigma,\sigma'} \\
t(\neg\varphi) &= \neg t(\varphi) \\
t(\varphi \wedge \psi) &= t(\varphi) \wedge t(\psi) \\
t([\pi]\varphi) &= [t(\pi)]t(\varphi) \\
t([\sigma_i]p) &= pre_{\sigma_i} \to p \\
t([\sigma_i]\neg\varphi) &= pre_{\sigma_i} \to \neg t([\sigma_i]\varphi) \\
t([\sigma_i](\varphi \wedge \psi)) &= t([\sigma]\varphi) \wedge t([\sigma_i]\psi) \\
t([\sigma_i][\pi]\varphi) &= \bigwedge_{j=0}^{n-1} T_{\sigma_i,\sigma_j}(\pi)[\sigma_j]t(\varphi) \\
t([\sigma_i][\sigma_j]\varphi) &= t([\sigma_i]t([\sigma_j]\varphi))
\end{aligned}
$$

$$
\begin{aligned}
t(a) &= a \\
t(\pi;\pi') &= t(\pi);t(\pi') \\
t(\pi \cup \pi') &= t(\pi) \cup t(\pi') \\
t(\pi^*) &= t(\pi)^* \\
t(?\varphi) &= ?t(\varphi)
\end{aligned}
$$

Almost all formula cases are direct analogues of the translation of $\mathcal{L}^{\Sigma}_{K[\sigma]\xi}$ to $\mathcal{L}_{K\xi}$. The new case, for $[\sigma_i][\pi]\varphi$, uses a class of *program transformations*, which we motivate now. Intuitively, this formula is true at (M, w) when (assuming that $w \vDash Pre_{\sigma_i}$) for any pair (w', σ_j) in M^+ such that $((w, \sigma_i), (w', \sigma_j)) \in R^{M^+}_{\pi}$, we have that $(M^+, (w', \sigma_j)) \vDash \varphi$. We need to characterize this condition while remaining within $\mathcal{L}^{\Sigma}_{PDL\xi}$; to do this, we need a way of quantifying over worlds which are possibly the result of executing π at (w, σ_i) in the update model. We take these possibilities by cases, according to which action σ_j was performed at the result (w', σ_j). Thus, we create a transformation T_{σ_i,σ_j} on programs parametrized by the starting action σ_i and ending action σ_j. We want every member (w, w') of $R^M_{T_{\sigma_i,\sigma_j}(\pi)}$ to correspond to a tuple $((w, \sigma_i), (w', \sigma_j)) \in W^+ \times W^+$ which occurs in the interpretation of π in the update model. That is, we want a program $T_{\sigma_i,\sigma_j}(\pi)$ such that $(w, w') \in R^M_{T_{\sigma_i,\sigma_j}(\pi)}$ if and only if $((w, \sigma_i), (w', \sigma_j)) \in R^{M^+}_{\pi}$.
We define T_{σ_i,σ_j} on programs inductively.

$$
\begin{aligned}
T_{\sigma_i,\sigma_j}(a) &= \bigcup_{\sigma_j \in \Sigma' \subseteq \Sigma} ?(pre_{\sigma_i} \wedge \xi_{a,\sigma_i,\Sigma'}); a; ?(\xi_{a,\sigma_j,\Sigma'}) \\
T_{\sigma_i,\sigma_j}(\pi \cup \pi') &= T_{\sigma_i,\sigma_j}(\pi) \cup T_{\sigma_i,\sigma_j}(\pi') \\
T_{\sigma_i,\sigma_j}(\pi;\pi') &= \bigcup_{\sigma_k \in \Sigma} (T_{\sigma_i,\sigma_k}(\pi); T_{\sigma_k,\sigma_j}(\pi')) \\
T_{\sigma_i,\sigma_j}(\pi^*) &= K_{\sigma_i,\sigma_j,n}(\pi) \\
T_{\sigma_i,\sigma_j}(?\varphi) &= \begin{cases} ?(\varphi \wedge [\sigma_i]\varphi) & \text{if } i = j \\ ?\bot & \text{otherwise} \end{cases}
\end{aligned}
$$

In the last clause, we use another transformation, $K_{\sigma_i,\sigma_j,k}$, which is defined recursively over k. The intuition behind this transformation is that $K_{\sigma_i,\sigma_j,n}(\pi)$ should contain $(w,w') \in R^M_\pi$ if and only if $((w,\sigma_i),(w',\sigma_j)) \in R^{M+}_{\pi^*}$.[33] We interpret $K_{\sigma_i,\sigma_j,k}(\pi)$ as a program which contains (w,w') if there is a sequence of members in $R^{M+}_{\pi^*}$ which lead from (w,σ_i) to (w',σ_j) without crossing any tuples containing actions which have index k or greater.[34]

$$K_{\sigma_i,\sigma_j,0}(\pi) = \begin{cases} ?\top \cup T_{\sigma_i,\sigma_j}(\pi) \text{ if } i = j \\ T_{\sigma_i,\sigma_j}(\pi) \text{ if } i \neq j \end{cases}$$

$$K_{\sigma_i,\sigma_j,k+1}(\pi) = \begin{cases} (K_{\sigma_k,\sigma_k,k}(\pi))^* \text{ if } i = j = k \\ (K_{\sigma_k,\sigma_k,k}(\pi))^*; K_{\sigma_k,\sigma_j,k}(\pi) \text{ if } i = k \neq j \\ (K_{\sigma_i,\sigma_k,k}(\pi); (K_{\sigma_k,\sigma_k,k}(\pi))^*) \text{ if } i \neq j = k \\ (K_{\sigma_i,\sigma_j,k}(\pi) \cup (K_{\sigma_i,\sigma_k,k}(\pi); (K_{\sigma_k,\sigma_k,k}(\pi))^*; K_{\sigma_k,\sigma_j,j}(\pi))) \text{ if } i \neq j \neq k \end{cases}$$

We show that the clauses given define an operator with the desired behavior, summarized in the following lemma:

Lemma 21. *Suppose that $(w,w') \in R^M_{T_{\sigma_i,\sigma_j}(\pi)}$ iff $((w,\sigma_i),(w',\sigma_j)) \in R^{M+}_\pi$. Then $(w,w') \in R^M_{K_{i,j,k}(\pi)}$ just when there is a (possibly empty) sequence of π-steps[35] in M^+ from (w,σ_i) to (w',σ_j) which does not have any intermediate states with an action σ_l where $l \geq k$.*[36]

$k = 0$:: Suppose that $i = j$. Then $K_{\sigma_i,\sigma_j,0}(\pi) = ?\top \cup T_{\sigma_i,\sigma_j}(\pi)$. Consider $(w,w) \in R^M_{?\top}$ – this corresponds to the empty sequence of π steps from (w,σ_i) to itself. Suppose $(w,w') \in R^M_{T_{\sigma_i,\sigma_j}(\pi)}$; this is equivalent, by our premise, to $((w,\sigma_i),(w',\sigma_j)) \in R^{M+}_\pi$, which is a single π-step and so trivially satisfies the requirement on intermediate states.
Suppose that $i \neq j$. Then $K_{\sigma_i,\sigma_j,0}(\pi) = T_{\sigma_i,\sigma_j}(\pi)$, and the second of the two cases considered above applies.

For the induction step, we assume that the result holds up to k; we show it for $k+1$. We consider the same cases used to define $K_{\sigma_i,\sigma_j,k}$ above.
Suppose that $i = j = k$. Any π-sequence from (w,σ_k) to (w',σ_k) in M^+ which does not pass through any intermediate states of the form (u,σ_l) for $l \geq k+1$ may be viewed as a chain of π-sequences with the following characteristics:

- Each π-sequence begins with a state of the form (u,σ_k), and ends with a state of the form (v,σ_k).
- The first sequence begins at (w,σ_k), and the last sequence ends at (w',σ_k).
- No sequence encounters an intermediate state of the form (u,σ_l) with $l \geq k$.

[33] Recall that n is the size of Σ.

[34] For the first clause (when $k = 0$), the original presentation splits into two cases – $i = j$ and $i \neq j$. I removed the case because I think it's irrelevant here, but we may want to talk about this.

[35] In the following, we will use the term 'π-sequence' to denote a chain of π steps. More formally, a π-sequence in an action-epistemic model M' is a sequence of states in M', $(s_0,\ldots,s_b)$, such that for $i \in [0,b)$, $(s_i,s_{i+1}) \in R^{M'}_\pi$.

[36] The case of zero π steps corresponds to the trivial π^* path.

By the induction hypothesis, each π-sequence of the described sort from (u, σ_k) to (v, σ_k) exists iff $(u, v) \in R^M_{K_{\sigma_k,\sigma_k,k}(\pi)}$. Thus, the full chain π-sequences from (w, σ_k) to (w', σ_k) in M^+ exists iff $(w, w') \in R^M_{(K_{\sigma_k,\sigma_k,k}(\pi))^*} = R^M_{K_{\sigma_i,\sigma_j,k+1}(\pi)}$.
Suppose that $i = k \neq j$. Any π from (w, σ_i) to $(w', \sigma_j) \in M^+$ which does not pass through any intermediate states of the form (u, σ_l) for $l \geq k+1$ may be viewed as a chain of π-sequences with the following characteristics:

- Every π-sequence in the chain, *except for the last*, begins with a state of the form (u, σ_k), and ends with a state of the form (v, σ_k). The last π-sequence in the chain begins with a state of the form (u, σ_k), and ends with a state of the form (v, σ_j).
- The first sequence of the chain begins at (w, σ_k), and the last sequence ends at (w', σ_j).
- No sequence encounters an intermediate state of the form (u, σ_l) with $l \geq k$.

By the induction hypothesis, a π-sequence of the first sort (i.e., not including the last sequence of the chain) from (u, σ_k) to (v, σ_k) exists iff $(u, v) \in R^M_{(K_{\sigma_i,\sigma_j,k}(\pi))}$, and a π-sequence of the second sort (i.e., the last sequence in the chain) from (u, σ_k) to (v, σ_j) exists iff $(u, v) \in R^M_{K_{\sigma_k,\sigma_j,k}(\pi)}$. Thus, the full chain from (w, σ_i) to (w', σ_j) exists if and only if $(w, w') \in R^{M^+}_{(K_{\sigma_i,\sigma_j,k}(\pi))^*;K_{\sigma_k,\sigma_j,k}(\pi)} = R^{M^+}_{K_{\sigma_i,\sigma_j,k+1}(\pi)}$.
The other cases work similarly. $\square$
An easy consequence of this lemma is the following:

Lemma 22. *Suppose that* $(w, w') \in R^M_{T_{\sigma_i,\sigma_j}(\pi)}$ *iff* $((w, \sigma_i), (w', \sigma_j)) \in R^{M^+}_\pi$. *Then* $(w, w') \in R^M_{K_{\sigma_i,\sigma_j,n}(\pi)}$ *iff* $((w, \sigma_i), (w', \sigma_j)) \in R^{M^+}_{\pi^*}$.

We use this lemma to show that our translation is truth-preserving in the case for $[\sigma_i][\pi]\varphi$, which is stated in the following theorem:

Theorem 23. $M, w \vDash [\sigma_i][\pi]\varphi$ *if and only if* $M, w \vDash \bigwedge_{\sigma_j \in \Sigma_0} [T_{\sigma_i,\sigma_j}(\pi)][\sigma_j]\varphi$.

We show this by induction over the complexity of π.

$\pi = a \in G$:: suppose the left hand side, and take some $\sigma_j \in \Sigma$. We show that $M, w \vDash [\bigcup_{\sigma_j \in \Sigma' \subseteq \Sigma} ?(pre_{\sigma_i} \wedge \xi_{a,\sigma_i,\Sigma'}); a; \xi_{a,\sigma_j,\Sigma'}][\sigma_j]\varphi$. Suppose that $M, w \vDash pre_{\sigma_i} \wedge \xi_{a,\sigma_i,\Sigma'}$, and consider any w' such that $w \sim_a w'$ and $M, w \vDash \xi_{a,\sigma_j,\Sigma'}$; suppose that $w' \vDash pre_{\sigma_j}$. We know that $(w, \sigma_i) \in M^+$ and $(w, \sigma_i) \sim^+_a (w', \sigma_j)$, and so by hypothesis, $(w', \sigma_j) \vDash \varphi$. Thus $w' \vDash [\sigma_j]\varphi$ as desired.
Suppose the right hand side, and that $w \vDash pre_{\sigma_i}$. Consider some (w', σ_j) such that $(w, \sigma_i) \sim^+_a (w', \sigma_j)$. This implies that for some $\Sigma' \subseteq \Sigma$ with $\sigma_j \in \Sigma'$, we have that $w \vDash \xi_{a,\sigma_i,\Sigma'}$ and $w' \vDash \xi_{a,\sigma_j,\Sigma'}$. Thus, $(w, w') \in R^M_{T_{\sigma_i,\sigma_j}(a)}$, and so by our hypothesis, $w' \vDash [\sigma_j]\varphi$, and $(w', \sigma_j) \vDash \varphi$.

$\pi = ?\psi$:: suppose the left hand side. Since $T_{i,j}(?\psi)$ is non-trivial only when $i = j$, we only need to show that $w \vDash [?(pre_{\sigma_i} \wedge [\sigma_i]\psi)][\sigma_i]\varphi$. Suppose $w \vDash pre_{\sigma_i} \wedge [\sigma_i]\psi$; then $(w, \sigma_i) \vDash \psi$ and so by the hypothesis, $(w, \sigma_i) \vDash \varphi$, as desired.
Suppose the right hand side, and that $w \vDash pre_{\sigma_i}$, and that $(w, \sigma_i) \vDash \psi$. By hypothesis, $w \vDash [?(pre_{\sigma_i} \wedge [\sigma_i]\psi)][\sigma_i]\varphi$. w satisfies the test formula, and so $(w, \sigma_i) \vDash \varphi$ and we are done.

$\pi = \pi_1; \pi_2$::
We show by a sequence of equivalent formulas, starting with the right hand side:

$$\bigwedge_{\sigma_j \in \Sigma_0} [\bigcup_{\sigma_k \in \Sigma_0} (T_{\sigma_i,\sigma_k}(\pi_1); T_{\sigma_k,\sigma_j}(\pi_2))][\sigma_j]\varphi$$

$$\bigwedge_{\sigma_j \in \Sigma_0} \bigwedge_{\sigma_k \in \Sigma_0} [(T_{\sigma_i,\sigma_k}(\pi_1); T_{\sigma_k,\sigma_j}(\pi_2))][\sigma_j]\varphi$$

$$\bigwedge_{\sigma_k \in \Sigma_0} [T_{\sigma_i,\sigma_k}(\pi_1)] \bigwedge_{\sigma_j \in \Sigma_0} [T_{\sigma_k,\sigma_j}(\pi_2)][\sigma_j]\varphi$$

$$\bigwedge_{\sigma_k \in \Sigma_0} [T_{\sigma_i,\sigma_k}(\pi_1)][\sigma_k][\pi_2]\varphi$$

$$[\sigma_k][\pi_1][\pi_2]\varphi$$

$$[\sigma_k][\pi_1; \pi_2]\varphi$$

Where lines three and four are justified by the induction hypothesis. The case where $\pi = \pi_1 \cup \pi_2$ is similar.

$\pi = \pi^*$. The induction hypothesis establishes the premise of Lemma 22, by which we conclude that $(w, w') \in R^M_{K_{i,j,n}}$ if and only if $((w, \sigma_i), (w', \sigma_j)) \in R^{M^+}_{\pi^*}$. From this the desired conclusion follows straightforwardly. $\square$

Now we give the reduction scheme from $\mathcal{L}^{\Sigma}_{PDL[\sigma]\xi}$ to $\mathcal{L}^{\Sigma}_{PDL\xi}$.

$$t(p) = p$$

$$t(\xi_{a,\sigma,\sigma'}) = \xi_{a,\sigma,\sigma'}$$

$$t(\neg\varphi) = \neg t(\varphi)$$

$$t(\varphi \wedge \psi) = t(\varphi) \wedge t(\psi)$$

$$t([\pi]\varphi) = [r(\pi)]t(\varphi)$$

$$t([\pi]\varphi) = [r(\pi)]t(\varphi)$$

$$t([\sigma]p) = pre_\sigma \rightarrow p$$

$$t([\sigma]\neg\varphi) = pre_\sigma \rightarrow \neg[\sigma]t(\varphi)$$

$$t([\sigma](\varphi \wedge \psi)) = t([\sigma]\varphi) \wedge t([\sigma]\psi)$$

$$t([\sigma_i][\pi]\varphi) = \bigwedge_{\sigma_j \in \Sigma} T_{i,j}(r(\pi))t([\sigma_j]\varphi)$$

$$t([\sigma][\sigma']\varphi = t([\sigma]t([\sigma']\varphi))$$

$$r(a) = a$$

$$r(?\varphi) = ?t(\varphi)$$

$$r(\pi; \pi') = r(\pi); r(\pi')$$

$$r(\pi \cup \pi') = r(\pi) \cup r(\pi')$$

$$r(\pi^*) = r(\pi)^*$$

The correctness of this translation may be proved as before, by defining an appropriate weight and depth on formulas. We then induct over the depth and weight of the formula; all cases are identical to those before except for $[\pi]\varphi$ and $[\sigma_i][\pi][\varphi]$. For these cases, we may induct over the structure of π. We leave this proof to the reader. With this, we have shown theorem 20.

4.2.3 Completeness for $\mathcal{L}^{\Sigma}_{PDL[\sigma]\xi}$

Our axiom system is all of the axioms and rules of PDL, axioms 6a-c, 7 from above (where 7 has been suitably adapted to the present language: $\xi_{j,\sigma,\sigma'} \to [\pi]\xi_{j,\sigma,\sigma'}$), plus the following reduction axioms:

$$[\sigma]p \leftrightarrow pre_\sigma \to p$$

$$[\sigma]\neg\varphi \leftrightarrow pre_\sigma \to \neg[\sigma]\varphi$$

$$[\sigma](\varphi \wedge \psi) \leftrightarrow [\sigma]\varphi \wedge [\sigma]\psi$$

$$[\sigma][\pi]\varphi \leftrightarrow \bigwedge_{\sigma_j \in \Sigma} T_{\sigma_i,\sigma_j}(\pi)[\sigma_j]\varphi$$

The soundness of the last axiom was establish by Theorem 23. With this, we have established completeness for $\mathcal{L}^{\Sigma}_{PDL[\sigma]\xi}$ with respect to the above collection of axioms: take any non-theorem φ of $\mathcal{L}^{\Sigma}_{PDL[\sigma]\xi}$. We know that φ is provably equivalent to $t(\varphi)$, which is in the language $\mathcal{L}^{\Sigma}_{PDL\xi}$. Since the axiomatization for $\mathcal{L}^{\Sigma}_{PDL\xi}$ is complete (we include axioms 6a-c and 7 in its adjusted form), there is some action-epistemic model (M, w) such that $M, w \vDash \neg t(\varphi)$, and we are done.

4.2.4 Discussion

It is worthwhile to note that this proof, too, falls in the paradigm we have already witnessed: to axiomatize some language $\mathcal{L}$, we reduce it to a stronger language, $\mathcal{L}'$, and appeal to reducibility/completeness results for $\mathcal{L}'$. For AEL without common knowledge, this took the form of supplementing our language with the ξ formulas – once these were added, it was possible to reduce to a language which did not include the update operator. It was clear what we were gaining by adding ξ formulas to the language: it allowed us to express, in the object language, statements about the indistinguishability of actions. Once action indistinguishability could be described in the language, we were able to translate formulas which include update operators into ones which do not.

It is not as clear, however, what strengthening to $\mathcal{L}_{PDL}$ gives us: how much stronger is $\mathcal{L}^{\Sigma}_{PDL\xi}$ than $\mathcal{L}_{KC\xi}$, $\mathcal{L}_{K[A,\sigma]C\xi}$, or $\mathcal{L}^{\Sigma}_{K[\sigma]C\xi}$? The language of PDL allows us to perform operations on programs – in this context, the programs act as epistemic operators. For instance, in PDL, we may take the programs a and b – whose corresponding modalities translate to "Anne knows" and "Bob knows", respectively – and create a new program, $a \cup b$, whose modality translates to "Anne and Bob know". The formula $[a \cup b]\varphi$ has an equivalent in the basic epistemic language with $K_a\varphi \wedge K_b\varphi$. The inclusion of an iteration operator, $*$, lets us express common knowledge as a transformation on a set of programs. However, the program transformations permitted in PDL also produce formulas which have no equivalent in $\mathcal{L}_{K[A,\sigma]C}$. One such example is the formula $[(a;b)^*]\varphi$. Formulas of this sort have no natural reading in an epistemic interpretation, and so it is unclear *how* PDL is increasing our expressive power, besides affording us a means of expressing common knowledge.

Similarly, the inclusion of relativized common knowledge, in the completeness proof for $\mathcal{L}_{K[\varphi]C}$, also marks a real increase in expressive power. The relativized common knowledge operator is suited specifically to address the reduction clause for formulas of the form $[\psi]C_B\varphi$, but it does not have a straightforward epistemic reading. In fact, one can show that $\mathcal{L}_{KRC}$ is *strictly* more expressive than $\mathcal{L}_{K[\varphi]C}$, prompting the question: are completeness proofs for these languages possible *without* these appeals to strictly stronger languages?[37] And if they are not, how much of an increase in expressive power is required?

It is worth noting that in the proof of this result, the models which demonstrate the difference in expressive power are not S5 (and cannot be, for this version of the proof); it may be that if we restrict our attention to S5 models, then the relationships between these languages changes – this is currently under investigation.

[37]For a proof of this result, see theorem 8.66 in [22].

5 The ‘No Miracles’ Principle

In this section, we use the proposed semantics to analyze a well-known principle in epistemic logic. The ‘No Miracles’ principle arises in the context of the more general *Epistemic Temporal Logic* framework, and delimits the type of learning which is possible under the AL update mechanism [10, 17]. Roughly put, the principle states that information is required in order for learning to take place, or that there can be no ‘miracles’. More precisely, the principle is a condition on how the indistinguishability of epistemic actions is to be understood in the AL framework: indistinguishable epistemic actions cannot be leveraged to learn anything new.

It is easily seen that the No Miracles principle is not validated in the AEL framework. Consider the treatment of example 14 in section 3.4. Although σ_p cannot be distinguished from itself at state w_0 (i.e., $(\sigma_0, \sigma_0) \in f_b(w_0)$), σ_p may be leveraged to distinguish between w_0 and w_1 (in the update model M^+ given in figure 11, $(w_0, \sigma_p) \not\sim_b (w_1, \sigma_p)$).

Of course, the intuition underlying the No Miracles principle seems to be honored in the AEL framework – learning does not come about ‘miraculously’, even if this technical articulation of the principle is not sound. This prompts the question: is there a suitable weakening of the principle which *is* validated by the AEL framework? In this section, we provide such a weakening, and show that this formulation of the No Miracles principle characterizes the AEL update mechanism.

To do this, we make use of an alternative axiomatization technique developed in [32] and [31]. This technique is used in [31] to give a reduction-free axiomatization of DEL – that is, an axiomatization of DEL which does *not* work by reducing the language to a static one – and in the process, establishes the No Miracles principle above as characteristic of the DEL product update. Using this same technique, we will develop an alternative (reduction-free) axiomatization of AEL; in presenting this axiomatization, we will formulate a weakened version of the No Miracles principle which is validated by the AEL framework.

We begin by giving a brief history of the principle in section 5.1; to our knowledge, this is the first such documentation of the origins of the principle. In this brief history, we analyze the No Miracles principle into independent components, and demonstrate how AEL fails to validate it. In section 5.2, we apply the technique used in [31] in order to formulate an appropriate weakening of the No Miracles principle, and show that it characterizes the AEL update mechanism. We conclude with a discussion of this result, and how it may bear on the analysis of security protocols.

5.1 A brief history of the ‘No Miracles’ principle

The *No Miracles* principle finds its origins in computer science literature as the technical ‘No Learning’ property, in the context of *distributed systems* as in [5].[38] A distributed system consists of (1) a set of *processors*, each of which may be in a different state; a ‘global state’ of the system corresponds to a complete assignment of processors to states. (2) sequences of these global states are called *runs*; each run is interpreted as a possible evolution of the state of the system (that is, the global state). We interpret the nth member of a run as the state of the system at

[38] These later come to be known as ‘interpreted systems’ in [7].

the nth step of that run. At any step n in a run, a process i can eliminate from possibility all those runs which disagree with the present run on i's state (at step n). A process 'learns' if it reduces the number of runs it considers possible. Thus, the 'No Learning' principle simply states that the number of runs a process considers possible never decreases. Equivalently: if runs v and v' cannot be distinguished by process i at time n, then v and v' cannot be distinguished by process i at any future time $n' > n$. This principle has been shown to play a non-trivial role in the complexity of the validity problem for the logic(s) associated with distributed systems.[39]

We see the No Miracles principle first emerge in [10], wherein Johan van Benthem demonstrates how a dynamic logical framework may be applied in the study of game trees. Given a set G of agents, a set A of actions that the agents can take, and a countable set PROP of propositions, a *game* is a relational model of the form:

$$M = \langle W, \{R_a | a \in A\}, \{\sim_i | i \in G\}, V \rangle$$

where:

- W is the set of (global) states the game can reach.
- R_a is an asymmetric relation on W where we interpret $(w, w') \in R_a$ (also written as $w \to_a w'$) as 'executing action a in state w leads to state w''.[40]
- $\sim_i$ is an equivalence relation on W where we interpret $(w, w') \in \sim_i$ to mean 'agent i cannot distinguish between states w and w'.
- $V : \text{PROP} \to 2^W$ is a valuation function.

We define the *stage* k of the game M to be the restriction of M to the set $\{w | len(w) = k\}$. These models encode the agents' uncertainty about how the game might progress: if $w \to_a w'$, for instance, then the epistemic effects of action a on agent i can be assessed by looking at the worlds which i cannot distinguish from w'. We can restrict our attention to games of a particular sort by placing restrictions on R_a and $\sim_i$.

This constitutes a sort of 'bird's eye view' of the game: all game states, actions, and uncertainty relations are simultaneously presented. One might reasonably ask, however, whether the epistemic effects of actions may be *generated* from a sufficiently detailed description of the possible actions in the game.

Van Benthem shows that one can do precisely this by using the *product update* from Action Logic (AL) [22, 9]. In AL, the description of the actions takes the form of a relation over the set of actions, for each agent: $\approx_i \subseteq A \times A$ (for now, we assume that $\approx_i$ is an equivalence relation).[41] We read $(a, a') \in \approx_i$ as 'agent i cannot distinguish between actions a and a''. The product update in AL is a formal procedure which, given a model of the 'initial' state of affairs and description of the actions, produces a model of the actions' effects. The central insight of the product update is simple to state: if an agent cannot distinguish between initial states w

[39] The logic presented in [5] is the *propositional model logic of knowledge and time*; this language includes standard epistemic and temporal operators.

[40] We assume that the union of these relations makes M a tree; thus, every state w has a unique length from the root $len(w)$ (this is the number of 'moves' made so far).

[41] The reader familiar with Action Logic will note that we are missing a *precondition function* in this description; to simplify matters, we omit this from the present discussion, but introduce it below.

and w', and also cannot distinguish between the actions a and a', then agent cannot distinguish between the states which result from performing a and a' in w and w', respectively.

Adapting this procedure to the setting of games, van Benthem uses the product update to generate the uncertainty relation down the game tree as follows: for any w, w' in stage k with $w \to_a v$ and $w' \to_{a'} v'$, we have that:

$$v \sim_i v' \text{ iff } w \sim_i w' \text{ and } a \approx_i a'$$

In other words, the uncertainty between two 'future' states is factored into (1) uncertainty between the 'parents' (predecessors) of those states and (2) uncertainty between the actions which respectively bring about these future states.

Naturally, not all games can be generated using the product update. As van Benthem shows, the following three properties characterize the set of games which can be generated in this way:

- *Perfect Recall*: if $w \to_a v$, $w' \to_{a'} v'$, and $v \sim_i v'$, then $w \sim_i w'$.
- *Uncertainty Propagation*: if $w \to_a v$, $w' \to_a v'$, and $w \sim_i w'$, then $v \sim_i v'$.
- *Uniformity*: if $w \to_a v$, $w' \to_{a'} v'$, and $v \sim_i v'$, then for any w'', w''', v'', v''', if $w'' \sim_i w'''$, $w'' \to_a v''$, and $w''' \to_{a'} v'''$, then $v'' \sim_i v'''$.

Perfect Recall requires that agents do not forget distinctions that they have already made – no actions can lead from distinguishable states to indistinguishable ones. *Uncertainty propagation* is simply a restriction of No Learning to the performance of an individual action: if the same action a is performed in two indistinguishable states, then those states will remain indistinguishable after a is executed in them, respectively. Van Benthem summarizes this principle with the remark that "there are no 'miracles'". *Uniformity* states that actions are either always distinguishable, or never are; in essence, this restricts the scope of inquiry to only those actions whose effect is context-independent.

A later paper establishes a closely related characterization of AL in the context of *Epistemic Temporal Logic* [23]. In this paper, *Uncertainty propagation* and *Uniformity* are combined into the 'No Miracles' principle, which, together with Perfect Recall (and a property called 'Synchronicity': each agent knows the number of actions which has led to the present state), characterizes the AL product update. A succinct explanation of the name comes in a different paper: "unless a 'miracle' happens, uncertainty of agents cannot be erased by the same event" [17, p. 4].

We now present the No Miracles principle using notation consistent with that of AEL which is presented in this thesis: we will use σ (rather than a) as the variable for an action, and we write (w, σ) to signify the resulting state after σ has been performed at w.[42]

(**No Miracles**) for any states w_0, w_1, w_3, w_4 and actions σ, σ', P1-3 jointly imply C:

$$\left.\begin{array}{ll} (P1) & w_0 \sim_i w_1 \\ (P2) & (w_0, \sigma), (w_1, \sigma') \text{ exist} \\ (P3) & (w_3, \sigma) \sim_i (w_4, \sigma') \end{array}\right| \Rightarrow \left| \begin{array}{ll} (C) & (w_0, \sigma) \sim_i (w_1, \sigma') \end{array}\right.$$

[42]In the game notation above: (w, σ) is the unique state w' such that $w \to_\sigma w'$.

We analyze (**No Miracles**) into the following conditions, which are jointly equivalent to the principle:[43]

(*reflexivity*) if $w \sim_i w'$, then for any σ, $(w, \sigma) \sim_i (w', \sigma)$.

(*context-independence*) two actions σ, σ' are either distinguishable at all states, or at no states

(*uncertainty propagation*) if σ and σ' cannot be distinguished at w and w', and $w \sim_i w'$, then $(w, \sigma) \sim_i (w', \sigma')$

It is easy to see that these conditions are satisfied by the AL product update: (*reflexivity*) the relation $\approx_i$ over $\Sigma \times \Sigma$ is an equivalence relation, thus reflexive; (*context-independence*) $\approx_i$ does not depend on the choice of world; (*uncertainty propagation*) this condition is essentially a reformulation of how the relation is generated in the update model, where we interpret "σ and σ' cannot be distinguished" as $\sigma \approx_i \sigma'$. Taken together, these three conditions precisely delimit the kinds of dynamic scenarios which AL studies.

It is also easy to see that there are natural epistemic scenarios which (**No Miracles**) excludes from analysis, and therefore resist a natural interpretation in the AL framework. For ease of exposition, we include another example (similar to example 4) here:

Example 24. Anne needs to send Bob classified information. Their encryption protocol works such that each time Anne sends an encrypted message, Bob enters a password which they have agreed on previously in order to decrypt the message; each message permits only one attempt. Today, Anne sends Bob an encrypted message saying whether proposition p is true. Unfortunately, Bob smudged the password which they last agreed on, and can't make out whether the password ends in 'l' or '1'. Bob guesses 'l'; if he's correct (q), then he will be able to decipher Anne's message and say whether p or not p; if he's not ($\neg q$), then the information in the message will be lost.

In a natural representation of this example, we might consider two actions, σ_p and $\sigma_{\neg p}$, to represent the epistemic actions available to Anne here – she can send the message p, or send the message $\neg p$. Now we must ask: are σ_p and $\sigma_{\neg p}$ indistinguishable to Bob? This would seem to depend on whether Bob is correct (q) or not ($\neg q$); but according to (*context-independence*), no such dependence is permitted between propositions internal to the model and the indistinguishability of epistemic actions. This restriction corresponds to the fact that in AL, the indistinguishability relation over Σ is not state-dependent. To see that (*reflexivity*) is also violated, suppose that Anne actually sends the message σ_p. Bob will be able to distinguish the state where he does learn that p, from the state where he does not; so this action leads from indistinguishable states, for Bob, to distinguishable ones.

[43] It is worthwhile to note that there is some variation in the definition of the No Miracles principle; for example, in [17], (*context-independence*) is omitted in the statement of the principle. In some statements of the principle, (*reflexivity*) is omitted but follows as a consequence of the principle and the background assumption that the indistinguishability relation $\sim_i$ is reflexive [23, 27] while other statements include (*reflexivity*) explicitly [17]. However, all statements of the principle in a characterization of the AL product update include all three parts.

5.2 An alternative (reduction-free) axiomatization of AEL

As we mentioned in section 4.1, it is typical to show that a system completely axiomatizes a semantics by means of *reduction* to a static language, and appeal to a completeness result on that static language. In PAL, for example, one devises a set of 'reduction axioms' which produce, for any PAL-formula, an equivalent formula in the (static) language of epistemic logic. This method is useful for proving completeness, but also sheds light on the logic itself: these reduction axioms characterize how the dynamic operator interacts with the knowledge modality.

However, one might rightly ask whether these particular dynamic logics *can* be axiomatized without reducing to a static language. This question has been answered in [32] and [31], where PAL and AL are axiomatized, respectively, without a reduction to a static language. This is achieved by recognizing that a public announcement or epistemic action may be simulated *inside* of an epistemic model itself; one simply imports the 'update states' into the epistemic model along with an update relation for every announcement or action under consideration.

We employ this same technique here to axiomatize AEL without reducing to a static language. The resulting axiomatization will feature a weaker version of the No Miracles principle, which we will explore. We begin by summarizing the result from [31] with respect to AL.

We now turn to establishing a corresponding result for AEL. We begin by internalizing the update relation into the action-epistemic model:
An **extended action-epistemic model** over Σ and Pre is any tuple $E = \langle W, \{\sim_j : j \in G\}, \{f_j : j \in G\}, \{\rightarrow_\sigma : \sigma \in \Sigma\}, V\rangle$ where:

- $\overline{E} = \langle W, \{\sim_j : j \in G\}, \{f_j : j \in G\}, V\rangle$ is an action-epistemic model.
- $\rightarrow_\sigma \subseteq W \times W$ is a partial function which satisfies the following conditions:
  1. If $w \rightarrow_\sigma w'$, then for any proposition p, $w \in V(p)$ iff $w' \in V(p)$.
  2. If $w \rightarrow_\sigma w'$, then for any agent j, $f_j(w) = f_j(w')$.
  3. $dom(\rightarrow_\sigma) = [\![Pre(\sigma)]\!]_{\overline{E}}$.[44]

The $\rightarrow_\sigma$ relation is intended to represent the execution of an action σ: $w \rightarrow_\sigma w'$ indicates that performing action σ at state w results in state w'. We require that $\rightarrow_\sigma$ is a partial function because actions in AEL (and AL) are deterministic; conditions 1 and 2 correspond to features of the action-epistemic update.[45] On this interpretation, we can check the consequences of an action without leaving the model: if we want to check whether φ is a consequence of executing σ at w, for instance, we need only check the $\rightarrow_\sigma$-successors of w. This suggests that in an extended action-epistemic model, we interpret $[\sigma]\varphi$ as a box-like operator, instead. Thus we define another semantics for $\mathcal{L}^{\Sigma}_{K[\sigma]\xi}$; we will notate this semantics with $\Vdash$. We interpret $\Vdash$ identically to $\vDash$ on all formulas except for the dynamic ones:

$$(E, w) \Vdash [\sigma]\varphi \quad \text{iff} \quad \text{for all } w' \text{ such that } w \rightarrow_\sigma w',\ w' \vDash \varphi$$

With this in hand, we now turn to proving completeness for the semantics $\vDash$ for action-epistemic logic. Let $\mathcal{E}$ denote the class of all extended action-epistemic models over Σ and Pre. We take the following steps, mimicking the strategy taken in [31]:

[44] Recall that $Pre(\sigma) \in \mathcal{L}_K$.

[45] In the action-epistemic update, (1) facts about the world (propositions) do not change, (2) facts about whether two actions may be distinguished by an agent do not change, and (3) an action is performed (exactly) when its precondition is satisfied.

find a class $\mathcal{N}$ of extended action-epistemic models on which $\vDash$ and $\Vdash$ are equivalent (theorem 28); give a complete axiomatization for $\Vdash$ on the class $\mathcal{N}$ (theorem 32); and conclude completeness for $\vDash$ (theorem 34).

We say that an extended action-epistemic model E is *normal* if it satisfies the following properties (for any $j \in G$; $\varphi, \psi \in \mathcal{L}^{\Sigma}_{K[\sigma]\xi}$; $p \in \text{PROP}$; and $\sigma, \sigma' \in \Sigma$):[46]

No Miracles+: for any states w, w', v, v' and action σ, σ', (P1)-(P3) jointly imply (C):

$$\begin{array}{rl} (P1) & w \sim_j w' \\ (P2) & w \to_\sigma v \text{ and } w' \to_{\sigma'} v' \\ (P3) & f_j(w)[\sigma] = f_j(w')[\sigma'] \\ (C) & v \sim_j v' \end{array}$$

Perfect Recall+: for any states w, v, v' and action σ, (P1) and (P2) jointly imply (C):

$$\begin{array}{rl} (P1) & w \to_\sigma v \\ (P2) & v \sim v' \\ (C) & \text{There exist } w', \sigma' \text{ such that:} \\ & (1)\ w \sim_j w',\ (2)\ w' \to_{\sigma'} v', \text{ and } (3)\ f_j(w)[\sigma] = f_j(w')[\sigma'] \end{array}$$

We now set out to show that $\Vdash$ and $\vDash$ agree on all formulas of $\mathcal{L}^{\Sigma}_{K[\sigma]\xi}$ over the class $\mathcal{N}$ of normal extended action-epistemic models. To do this, we first formulate an appropriate notion of bisimulation, and state a series of lemmata.

Let (E, x) and (F, y) be two pointed action-epistemic models over Σ and Pre. A **Σ-bisimulation** between (E, x) and (F, y) is a bisimulation $B \subseteq W_E \times W_F$ containing (x, y) such that for any $(w, v) \in B$: for any $\sigma, \sigma' \in \Sigma$, $(\sigma, \sigma') \in f^E_j(w)$ iff $(\sigma, \sigma') \in f^F_j(v)$.[47][48]

Lemma 25. *If B is a Σ-bisimulation between action-epistemic models E and F, then the following is a Σ-bisimulation between E^+ and F^+:*

$$B^+ := \{((x, \sigma), (y, \sigma)) | (x, y) \in B \textit{ and } E, x \vDash Pre(\sigma)\}$$

Proof. Suppose that B is such a bisimulation, and define B^+ as above. 1 and 2 are easily seen to be satisfied. For 3, suppose that $(x, \sigma) \sim_j (x', \sigma')$. Then in E, $x \sim^E_j x'$ and $f^E_j(x)[\sigma] = f^E_j(x')[\sigma']$. Since B is a Σ-bisimulation, there must be a y' such that $(x', y') \in B$ and $y \sim^F_j y'$. We must also have that $E, x' \vDash Pre(\sigma')$; then since $(x', y') \in B$, it follows that $F, y' \vDash Pre(\sigma)$. Thus, $((y, \sigma), (y', \sigma')) \in B^+$. Furthermore, since $(x', y') \in B$, $f^E_j(x') = f^F_j(y')$. Since $f^E_j(x) = f^F_j(y)$, it follows

[46] We will refer to these two principles using the shorthand (NM+) and (PR+), respectively.

[47] Recall that B is a bisimulation just when for any $(w, v) \in B$: (1) w and v agree on all propositions; (2a) if $w \sim_j w'$, then there exists v' such that $v \sim_j v'$ and $(w', v') \in B$; (2b) vice-versa.

[48] For our present purposes, we only need a notion of bisimulation between two action-epistemic models over the *same* (Σ, Pre). However, we may also formulate a more general notion of bisimulation, between action-epistemic models (E, x) over (Σ, Pre) and (F, y) over (Σ', Pre'). In this case, we extend the relation B to include entries of the form (σ, σ') where σ and σ' (1) have equivalent preconditions, and (2) cohere with respect to the action-indistinguishability functions f_j. This is similar to the *action emulation* found in [28].

that $f_j^F(y)[\sigma] = f_j^F(y')[\sigma']$, and so $(y,\sigma) \sim_j^+ (y',\sigma')$, as desired. The other direction is similar.

□

Lemma 26. *Suppose that B is a Σ-bisimulation between E, x and F, y. Then for any formula $\varphi \in \mathcal{L}^{\Sigma}_{K[\sigma]\xi}$, $E, x \vDash \varphi$ iff $F, y \vDash \varphi$.*

Proof. We prove by induction on the structure of the formula; since the present notion of bisimulation is stronger than the standard one, the boolean cases and the K_j case follow immediately. The case for $\xi_{\sigma,\sigma',j}$ is also trivial, by condition 2. We now treat the case $\varphi := [\sigma]\psi$. Suppose that $E, x \vDash [\sigma]\psi$. Suppose additionally that $E, x \vDash \neg Pre(\sigma)$; since $Pre(\sigma)$ is a formula in the language $\mathcal{L}_K$, it follows from our induction hypothesis that $F, y \vDash \neg Pre(\sigma)$; thus $F, y \vDash [\sigma]\psi$. Suppose instead that $E, x \vDash Pre(\sigma)$; then it follows that $E^+, (x,\sigma) \vDash \psi$. By the same reasoning above, we also know that $F, y \vDash Pre(\sigma)$. Then by lemma 25, we have that $((x,\sigma),(y,\sigma) \in B^+$ where B^+ is a bisimulation between E^+ and F^+. We appeal to the induction hypothesis to conclude that $F^+, (y,\sigma) \vDash \varphi$. Thus $F, y \vDash [\sigma]\psi$.

□

Lemma 27. *If E is a normal extended action-epistemic model, then for any formula $\varphi \in \mathcal{L}^{\Sigma}_{K[\sigma]\xi}$ and $w \to_\sigma v$ in E:*

$$\overline{E}^+, (w,\sigma) \vDash \varphi \textit{ if and only if } \overline{E}, v \vDash \varphi$$

Proof. We appeal to lemma 26, showing that the following is a Σ-bisimulation between $\overline{E}^+, (w,\sigma)$ and $\overline{E}, v$:

$$B = \{((w,\sigma), v) | w \to_\sigma v\}$$

Conditions 1 and 2 are satisfied trivially by the definition of an extended action-epistemic model. For condition 3, suppose that $((w,\sigma), v) \in B$ and $(w,\sigma) \sim_j^+ (w',\sigma')$. Then $\overline{E}, w' \vDash Pre(\sigma')$, and so by the properties of $\to'_\sigma$, $w' \to_{\sigma'} v'$ for some v'. Furthermore, we know that in $\overline{E}$, $w \sim_j w'$ and $f_j(w)[\sigma] = f_j(w')[\sigma']$. By (NM+), then, $v \sim_j v'$ as desired.

For condition 4, suppose that $((w,\sigma), v) \in B$ and $v \sim_j v'$. Since $w \to_\sigma v$ and $v \sim_j v'$, by (PR+) there must exist a w', σ' such that $w' \to_{\sigma'} v'$ and $w \sim_j w'$ and $f_j(w)[\sigma] = f_j(w')[\sigma]$. Clearly, $((w',\sigma'), v') \in B$, and we are done.

□

Theorem 28. *Given a normal extended action-epistemic model E and state $w \in E$, for any $\varphi \in \mathcal{L}^{\Sigma}_{K[\sigma]\xi}$, $(E, w) \Vdash \varphi$ iff $(\overline{E}, w) \vDash \varphi$.*

Proof. Let E be a normal extended action-epistemic model over Σ, Pre. We show the result by induction over $\mathcal{L}^{\Sigma}_{K[\sigma]\xi}$; the only non-trivial case to show is that of $[\sigma]\varphi$:

$$(E, w) \Vdash [\sigma]\varphi \;\Leftrightarrow\; (\overline{E}, w) \vDash [\sigma]\varphi$$

Suppose that $E, w \Vdash \neg Pre(\sigma)$; since $Pre(\varphi) \in \mathcal{L}_K$, by the IH we have that $\overline{E}, w \vDash \neg Pre(\sigma)$ and so the two agree on the formula $[\sigma]\varphi$.

Suppose now that $E, w \Vdash Pre(\sigma)$, and thus $\overline{E}, w \vDash Pre(\sigma)$. It follows that there is some $v \in W$ such that $w \to_\sigma v$. To show the desired equivalence, it suffices to show that:

$$E, v \Vdash \varphi \;\Leftrightarrow\; \overline{E}^{+}, (w, \sigma) \vDash \varphi$$

By the induction hypothesis and lemma 27 (respectively), we have this immediately:

$$E, v \Vdash \varphi \;\Leftrightarrow\; \overline{E}, v \vDash \varphi \;\Leftrightarrow\; \overline{E}^{+}, (w, \sigma) \vDash \varphi$$

□

We now present the axiomatization AEL (relative to Σ and Pre) for the entailment operator $\Vdash$. The rules of inference for this axiom system will be modus ponens, and necessitation for each of the modalities. Before presenting AEL, we define the following abbreviations (these will be used in axioms 7b and 7c):

$$\xi_{j,\sigma,\Sigma'} := \bigwedge_{\sigma' \in \Sigma'} \xi_{j,\sigma,\sigma'} \wedge \bigwedge_{\sigma'' \notin \Sigma'} \neg \xi_{j,\sigma,\sigma''}$$

$$\langle j, \sigma, \varphi \rangle := \bigvee_{\Sigma' \subseteq \Sigma} \big(\xi_{j,\sigma,\Sigma'} \wedge \hat{K}_j (\xi_{j,\sigma,\Sigma'} \wedge \varphi) \big)$$

$\xi_{j,\sigma,\Sigma'}$ is true at w when Σ' is exactly the set of actions which j cannot distinguish from σ at w. We will write $\Sigma_{j,w,\sigma}$ to indicate the unique subset of Σ such that $w \vDash \xi_{j,\sigma,\Sigma_{j,w,\sigma}}$.[49] $\langle j, \sigma, \varphi \rangle$ is true at w if there is some j-reachable state w' which satisfies $\xi_{j,\sigma,\Sigma_{j,w,\sigma}}$ and φ. In other words, $\langle j, \sigma, \varphi \rangle$ is true at w if φ is true at some world w' which cannot be distinguished from w after the performance of σ. With these in hand, we turn to the axiom system AEL:[50]

1. All propositional tautologies.
2. $K(\varphi \to \psi) \to (K_j \varphi \to K_j \psi)$
3. $[\sigma](\varphi \to \chi) \to ([\sigma]\varphi \to [\sigma]\chi)$
4. $(p \to [\sigma]p) \wedge (\neg p \to [\sigma]\neg p)$
5. $Pre(\sigma) \leftrightarrow \langle \sigma \rangle \top$
6. (a) $\xi_{j,\sigma,\sigma}$
   (b) $\xi_{j,\sigma,\sigma'} \to \xi_{j,\sigma',\sigma}$
   (c) $\xi_{j,\sigma,\sigma'} \to (\xi_{j,\sigma',\sigma''} \to \xi_{j,\sigma,\sigma''})$
7. (a) $\langle \sigma \rangle \varphi \leftrightarrow (Pre(\sigma) \wedge [\sigma]\varphi)$
   (b) $\big(\xi_{j,\sigma,\sigma'} \wedge \langle j, \sigma', \langle \sigma' \rangle \varphi \rangle\big) \to [\sigma]\hat{K}_j \varphi$
   (c) $\langle \sigma \rangle \hat{K}_j \varphi \to \bigvee_{\sigma' \in \Sigma'} \big(\xi_{j,\sigma,\sigma'} \wedge \langle j, \sigma', \langle \sigma' \rangle \varphi \rangle\big)$

It is easy to see that AEL is sound:

Theorem 29. AEL *is sound with respect to the semantics for* $\Vdash$ *on the class of normal extended action-epistemic models.*

We now show completeness. We define the extended canonical model $E^C = \langle W^C, \{\sim_j^C : j \in G\}, \{f_j^C : j \in G\}, \{\to_\sigma^C : \sigma \in \Sigma\}, V^C \rangle$ as follows:

[49] Note that this is trivially identical to the set $f_j(w)[\sigma]$.

[50] These are axiom schemata, where $j \in G$; $\varphi, \psi \in \mathcal{L}^{\Sigma}_{K[\sigma]\xi}$; $p \in$ PROP; and $\sigma, \sigma' \in \Sigma$.

$$
\begin{array}{lll}
\Gamma \in W^C & \text{iff} & \Gamma \text{ is a maximally consistent set (with respect to AEL) of } \mathcal{L}^{\Sigma}_{K[\sigma]\xi} \\
\Gamma \sim^C_j \Gamma' & \text{iff} & K_j\varphi \in \Gamma \text{ implies } \varphi \in \Gamma' \\
(\sigma, \sigma') \in f^C_j(\Gamma) & \text{iff} & \xi_{j,\sigma,\sigma'} \in \Gamma \\
\Gamma \rightarrow^C_\sigma \Gamma' & \text{iff} & [\sigma]\varphi \in \Gamma \text{ implies } \varphi \in \Gamma' \\
\Gamma \in V^C(p) & \text{iff} & p \in \Gamma
\end{array}
$$

Before proceeding, we note that, since states are maximally consistent sets, lines 2 and 4 have the equivalent formulations:

$$
\begin{array}{lll}
\Gamma \sim^C_j \Gamma' & \text{iff} & \varphi \in \Gamma' \text{ implies } \hat{K}_j\varphi \in \Gamma \\
\Gamma \rightarrow^C_\sigma \Gamma' & \text{iff} & \varphi \in \Gamma' \text{ implies } \langle\sigma\rangle\varphi \in \Gamma
\end{array}
$$

E^C is a normal extended action-epistemic model:

Lemma 30. *E^C satisfies (NM+) and (PR+).*

Proof. **(NM+)** suppose that the four premises hold:

1. $\Gamma \rightarrow^C_\sigma \Delta$
2. $\Gamma \sim^C_j \Gamma'$
3. $f^C_j(\Gamma)[\sigma] = f^C_j(\Gamma')[\sigma']$
4. $\Gamma' \rightarrow^C_{\sigma'} \Delta'$

We show the consequent: $\Gamma' \sim^C_j \Delta'$. Suppose that $\varphi \in \Delta'$; we show that $\hat{K}_j\varphi \in \Gamma'$. We show that the antecedent of axiom 7b is satisfied at Γ. The first conjunct follows immediate from premise (3) and axioms 6a-c. Since $\varphi \in \Delta'$, we have that $\langle\sigma'\rangle\varphi \in \Delta$. Since $\Gamma \sim^C_j \Delta$, it follows that $\hat{K}_j\langle\sigma'\rangle\varphi \in \Gamma$, and thus $\langle j, \sigma', \langle\sigma'\rangle\varphi\rangle \in \Gamma$, which is the second conjunct of the antecedent of 7b. Thus $[\sigma]\hat{K}_j\varphi \in \Gamma$ and so $\hat{K}_j\varphi \in \Gamma'$.

(PR+) suppose that $\Gamma \rightarrow^C_\sigma \Delta \sim^C_j \Delta'$; we find Γ' and σ' such that (a) $\Gamma \sim^C_j \Gamma' \rightarrow^C_{\sigma'} \Delta'$, and (b) $f_j(\Gamma)[\sigma] = f_j(\Gamma')[\sigma']$. Let φ be any member of Δ'; then it follows from our assumption that $\langle\sigma\rangle\hat{K}_j\varphi \in \Gamma$. By axiom 7c, there must be some $\sigma' \in \Sigma$ such that:

$$\xi_{j,\sigma,\sigma'} \wedge \langle j, \sigma', \langle\sigma'\rangle\varphi\rangle \in \Gamma$$

Decomposing the second part of the formula, there must exist a $\Sigma' \subseteq \Sigma$ such that:

$$\xi_{j,\sigma',\Sigma'} \wedge \hat{K}_j(\xi_{j,\sigma',\Sigma'} \wedge \langle\sigma'\rangle\varphi) \in \Gamma$$

Consider the following set Φ:

$$\Phi := \{\langle\sigma'\rangle\varphi : \varphi \in \Delta'\} \cup \{\psi : K_j\psi \in \Gamma\} \cup \{\xi_{j,\sigma',\Sigma'}\}$$

Clearly, any MCS $\Sigma' \supseteq \Phi$ would satisfy (a) and (b) above. Thus we show that Φ is consistent; by the Lindenbaum Lemma, it can be extended to an MCS Γ'.

Suppose for contradiction that Φ were not consistent. Then there must be $\varphi_0, \ldots, \varphi_n \in \Delta'$ and $K_j\psi_0, \ldots, K_j\psi_m \in \Gamma$ such that:[51]

$$\vdash_{AEL} (\psi_0 \wedge \ldots \wedge \psi_m) \rightarrow ([\sigma']\neg\varphi_0 \vee \ldots \vee [\sigma']\neg\varphi_n)$$

and by necessitation and axiom 2:

$$\vdash_{AEL} (K_j\psi_0 \wedge \ldots \wedge K_j\psi_m) \rightarrow K_j([\sigma']\neg\varphi_0 \vee \ldots \vee [\sigma']\neg\varphi_n)$$

Since the antecedent occurs in Γ, we have that $K_j([\sigma']\neg\varphi_0 \vee \ldots \vee [\sigma']\neg\varphi_n) \in \Gamma$.

By axioms 2 and 3 and necessitation (for both modalities), it follows that $K_j[\sigma'](\neg\varphi_0 \vee \ldots \vee \neg\varphi_n) \in \Gamma$. Since $\Gamma \rightarrow^C_\sigma \Delta \sim^C_j \Delta'$, it follows that $\neg\varphi_0 \vee \ldots \vee \neg\varphi_n \in \Delta'$; this contradicts the assumption that $\varphi_0, \ldots, \varphi_n \in \Delta'$. Thus, we conclude that Φ is consistent and can be extended to the MCS Γ' which has desired properties (a) and (b).

□

The truth lemma is straightforward to show, as both K_j and $\langle\sigma\rangle$ are normal modal operators, and so we state it here without proof:[52]

Lemma 31. *For any $\varphi \in \mathcal{L}^\Sigma_{K[\sigma]\xi}$ and $\Gamma \in W^C$: $\varphi \in \Gamma$ iff $M^C, \Gamma \Vdash \varphi$.*

Completeness for $\Vdash$ then follows immediately:

Theorem 32. *For any $\varphi \in \mathcal{L}^\Sigma_{K[\sigma]\xi}$: If $\nvdash_{AEL} \varphi$, then $\nVdash \varphi$.*

Proof. Consider some $\varphi \in \mathcal{L}^\Sigma_{K[\sigma]\xi}$ such that $\nvdash_{AEL} \varphi$. Thus $\{\neg\varphi\}$ is consistent and can be extended to a MCS $\Gamma_{\neg\varphi}$ such that $(M^C, \Gamma_{\neg\varphi}) \Vdash \neg\varphi$ and so $\nVdash \varphi$.

□

Since E^C is a normal extended model, we may use theorem 28 to extend completeness to $\vDash$:

Lemma 33. *For any $\varphi \in \mathcal{L}^\Sigma_{K[\sigma]\xi}$ and $\Gamma \in W^C$: $\varphi \in \Gamma$ iff $(\overline{M^C}, \Gamma) \vDash \varphi$*

Theorem 34. *For any $\varphi \in \mathcal{L}^\Sigma_{K[\sigma]\xi}$: If $\nvdash_{AEL} \varphi$, then $\nvDash \varphi$.*

We have shown that the axiom system AEL completely axiomatizes the entailment operator $\Vdash$ on the class of normal extended action-epistemic models $\mathcal{N}$ and, via theorem 28, also axiomatizes the entailment operator $\vDash$ on the class of all action-epistemic models.

An immediate result of theorem 28 is the following:

Theorem 35. *For any $\varphi \in \mathcal{L}^\Sigma_{K[\sigma]\xi}$, if $\vDash \varphi$, then $\Vdash_{\mathcal{N}} \varphi$.*

In other words, all validities of action-epistemic logic are also validities of $\Vdash$ over the class of normal extended action-epistemic models. However, we make the stronger claim that the semantics $\Vdash$ over the class of normal extended action-epistemic models capture *exactly* the validities of AEL. To show the required direction ($\Vdash_{\mathcal{N}} \varphi \Rightarrow \vDash \varphi$), we use the following lemma:

[51] We can ignore $\xi_{j,\sigma',\Sigma'}$ because if $K_j\neg\xi_{j,\sigma',\Sigma'} \in \Gamma$ were true, then $\neg\xi_{j,\sigma',\Sigma'} \in \Gamma$ would contradict $\xi_{j,\sigma,\sigma'}, \xi_{j,\sigma,\Sigma'} \in \Gamma$ via axioms 6a-c.

[52] For a standard proof of the truth lemma, see [22] or [11].

Lemma 36. *For any pointed action-epistemic model (M, x), there exists a pointed normal extended action-epistemic model (E^∞, x) such that for any formula $\varphi \in \mathcal{L}^{\Sigma}_{K[\sigma]\xi}$, $(M, x) \vDash \varphi$ iff $(E^\infty, x) \Vdash \varphi$.*

Proof. Let (M, x) be a pointed action-epistemic model; we first construct an appropriate action-epistemic model (M^∞, x). We introduce the notation $M^n = \langle W^n, \{\sim^n_j : j \in G\}, \{f^n_j : j \in G\}, V^n\rangle$, where $n \in \mathbb{N}$, defined recursively as follows:

$$M^0 = M \qquad M^{n+1} = (M^n)^+$$
$$M^\infty = \langle W^\infty, \{\sim^\infty_j : j \in G\}, \{f^\infty_j : j \in G\}, V^\infty\rangle$$

$W^* = \{(w, \overline{\sigma}) \in W \times \Sigma^* : (w, \overline{\sigma}) \in W^{len(\overline{\sigma})}\}$
$(w, \overline{\sigma}) \sim^\infty_j (w', \overline{\sigma'})$ iff $len(\overline{\sigma}) = len(\overline{\sigma'}) = k$ and $(w, \overline{\sigma}) \sim^k_j (w', \overline{\sigma'})$
$f^\infty_j = f_j$
$(w, \overline{\sigma}) \in V^\infty(p)$ iff $w \in V(p)$

It is immediate that (M, x) and (M^∞, x) agree on all formulas of $\mathcal{L}^{\Sigma}_{K[\sigma]\xi}$ via the identity bisimulation on (M, x). We now extend M^∞ to an extended action epistemic model:

$$E^\infty = \langle W^\infty, \{\sim^\infty_j : j \in G\}, \{f^\infty_j : j \in G\}, \{\rightarrow_\sigma : \sigma \in \Sigma\}, V^\infty\rangle$$

where $(w, \overline{\sigma}) \rightarrow_{\sigma*} (w, \overline{\sigma'})$ iff $\overline{\sigma'} = (\overline{\sigma}, \sigma*)$. Clearly, $\overline{E^\infty} = M^\infty$, and so by theorem 28, for any $\varphi \in \mathcal{L}^{\Sigma}_{K[\sigma]\xi}$:

$$(M, x) \vDash \varphi \text{ iff } (M^\infty, x) \vDash \varphi \text{ iff } (E^\infty, x) \Vdash \varphi$$

□

Theorem 37. *For any $\varphi \in \mathcal{L}^{\Sigma}_{K[\sigma]\xi}$, $\vDash \varphi$ iff $\Vdash_{\mathcal{N}} \varphi$.*

Proof. To show the remaining direction, suppose that $\nVdash_{\mathcal{N}} \varphi$; let (M, x) be the action-epistemic model which witnesses this. By lemma 36, there is a normal extended action-epistemic model (E^∞, x) such that $(E^\infty, x) \vDash \neg\varphi$, and so $\nvDash \varphi$. □

Thus, the semantics $\Vdash$ over the class of normal extended action-epistemic models capture exactly the validites of AEL. This suggests that the normality conditions on the relations $\{\rightarrow_\sigma : \sigma \in \Sigma\}$ characterize the AEL product update. In particular, (PR+) and (NM+) capture the mechanism producing $\sim^+_j$ in the product update.

5.3 Discussion

The move from (PR) to (PR+) involves only a direct adaptation of the principle to the AEL context. In the setting of AL, (PR) states that 'ignorance cannot emerge from nowhere'[53]: if two states (w, σ) and (w', σ') in the update model are related for j, then their respective predecessors w and w' in the original model must also be j-related, and the actions σ and σ' must also be j-related. Moving to the AEL framework, we need to adapt the second conjunct of the consequent to account for the world-dependence of epistemic action distinguishability. In the spirit of the principle, we need to require *only* that the actions are indistinguishable for j at the present worlds; they do not need to be indistinguishable for j at every world. Thus,

[53]Another way of putting this is that 'agents do not forget' (hence the name 'Perfect Recall').

we change the second conjunct of the consequent to state that $(\sigma, \sigma') \in f_j(w)$. However, we also need to account for the means of learning introduced by our adaptation to the semantics: the agent might also distinguish on the basis of what they learn about σ and σ' at w and w', respectively. In particular, in order for $(w, \sigma) \sim_j^+ (w', \sigma')$, it must also be the case that $f_j(w)[\sigma] = f_j(w')[\sigma']$, which is the final piece of the consequent to (PR+). In spirit, however, (PR+) remains true to (PR): (PR+) states, in the context of AEL, that agents do not forget.

The move from (NM) to (NM+) involves a more substantial change. (NM+) is a natural weakening of the No Miracles principle which permits the analysis of scenarios like example 24. Let us consider the intuitions corresponding to the formulations of No Learning and No Miracles which we saw in section 5.1:

1. No epistemic actions can eliminate uncertainty.
2. Indistinguishable epistemic actions cannot be leveraged to eliminate uncertainty.
    (a) If two actions take (anywhere in the model) indistinguishable worlds to distinguishable ones, then the states which they respectively produce can always be distinguished.
    (b) Performing the same action in two indistinguishable states does not result in those states being distinguishable.

Intuition 1 underlies the 'No Learning' principle, and is clearly too strong for both the AL and the AEL framework, which seek to represent, among other things, learning. Intuition 2 underlies the formulation of (NM) which characterizes the update procedure in AL; in particular, (a) corresponds to the (*reflexivity*) component of the statement of (**No Miracles**) in section 5.1, and (b) corresponds to the (*context-independence*) component. As we saw above, these components of (NM) prevent AL from encoding examples like 24. To see in what form intuition 2 might be preserved by our weakened principle, we restate (NM+) in the form of (*uncertainty propagation*), the third component of our decomposition of (**No Miracles**):

> (*uncertainty propagation+*) if σ and σ' cannot be distinguished at w and w' *and* σ and σ' are indistinguishable from the same set of actions at w and w' respectively, and $w \sim_i w'$, then $(w, \sigma) \sim_i (w', \sigma')$

While (*context-independence*) and (*reflexivity*) are 'directly' disavowed by the change to (NM+), (*uncertainty propagation*) is preserved in a moderated form. In particular, the additional premise reflects the feature of the AEL framework that an agent may learn about their ability to distinguish some epistemic action σ by the very performance of that action. We might summarize the change to (NM+) with the following adaption to the original simple sentiment: '*epistemically equivalent* actions cannot be leveraged to learn anything new', where two epistemic actions are epistemically equivalent at a state if they are each indistinguishable from the same set of epistemic actions at that state. It is this insight which corresponds to the altered update semantics underlying AEL – that this richer notion of epistemic action indistinguishability is relevant in the assessment of some forms learning, particularly in the case when one learns about one's own epistemic abilities.

5.3.1 Verification of security protocols

We conclude this discussion with a brief note on how the above bears on the application of the AEL framework to the verification of security protocols. Both ETL and AL have been proposed as frameworks for the verification of security protocols. In particular, both have served as the basis for model checkers in order to test security protocols.[54] The idea behind the use of these frameworks in this context is straightforward, and involves two steps:[55]

1. The articulation of desired safety properties in the relevant epistemic language. Desirable properties for a protocol might involve notions such as *anonymity* and *secrecy*, which lend themselves to articulation in a (sufficiently rich) language with epistemic and temporal (or dynamic) operators. Basic facts – such as the identity of a participant in the protocol – may be encoded as propositions, and more complex concepts are built using these operators.

2. Given a formal description of a protocol, design an appropriate model which captures the relevant features of the protocol, and check whether the articulated properties hold.

However, it has been pointed out that AL is limited in the kinds of protocols which it can model ([25, pp.65-66]). In particular, AL can only model protocols which satisfy the No Miracles principle studied above; an example of a protocol which does not satisfy (NM) may be extracted from example 14, wherein Anne sends Bob an encrypted message. The protocol here is simple: Anne and Bob have an agreed-upon password which encrypts strings; Anne sends Bob a message and Bob decrypts the message using the password. The difficulty in representing this situation arises when we introduce Bob's uncertainty as to whether he has the correct password: since Bob's ability to distinguish Anne's actions (messaging that p or messaging that $\neg p$) depends on the password being correct, a model faithful to this situation will violate No Miracles.

As we have seen, AEL is capable of representing a wider class of protocols, including the example 14; any protocol which satisfies the weaker (NM+) principle may be studied in the AEL framework.

We would be remiss not to mention that model checking in AL is (somewhat notoriously) computationally expensive.[56] This costliness results, in part, from execution of the update mechanism and creation of the update model. Since AEL still centrally involves an update mechanism (with more 'moving parts', not less), it seems to us that AEL will suffer from a similar level of costliness. However, there has been promising recent work on improvements in the efficiency of model checking in AL, both generally ([36]) and in the particular setting of security protocols ([37]). It is a subject of ongoing investigation whether these improvements in efficiency may translate to the AEL setting.

[54] For an example of a model checker which automatically constructs ETL models from a formal specification of a protocol, see [24]. Early explorations of using AL to verify security protocols may be found in [19] and [20]; a more recent study of the topic is undertaken in [33].

[55] We give only a broad outline of the method here; for a more detailed discussion of how epistemic logics may be used to model check security protocols, see [25].

[56] For a survey of these results and a study of the costliness in restricted settings, see [36] (forthcoming in the Journal of Applied Logics).

6 Concluding Remarks

We conclude with a summary of the thesis and possible extensions to the work.

In a series of examples, we outlined some difficulties AL faces with the representation of higher-order uncertainty involving epistemic actions. The type of uncertainty which we considered is typified in the sentence 'Anne is unsure whether Bob will be able to decipher her message'. We found that in order to represent higher-order uncertainty in dynamic scenarios, we need to introduce artificial actions to the action model; these actions import information already contained in the epistemic model. An undesired consequence is that small changes to the epistemic part of scenario required corresponding changes to the epistemic actions in the action model.

To address this difficulty we targeted the rigidity of the action model; in particular, we relativized the indistinguishability of actions to the worlds in which they are performed. This was done via the introduction of an indistinguishability function f_j for each agent j; this enrichment of the framework also called for an adaptation to the update semantics to accommodate the possibility of agents learning about their own epistemic abilities. The resulting logic, AEL, encodes examples of the above sort handily; we demonstrated that while AL can simulate any action-epistemic model, AEL may encode such examples significantly more succinctly and effectively.

We gave several completeness results for AEL and related languages. We introduced constants $\xi_{j,\sigma,\sigma'}$ into the basic epistemic language which reflect the state-dependent indistinguishability of actions to make language $\mathcal{L}_{K\xi}$; we then extended this language with an update operator $[\sigma]$ to create the dynamic language $\mathcal{L}^{\Sigma}_{K[\sigma]\xi}$. We then gave a complete axiom system of $\mathcal{L}_{K\xi}$, and showed that via a system of reduction axioms, $\mathcal{L}^{\Sigma}_{K[\sigma]\xi}$ is also completely axiomatized via this axiom system. We also consider the language $\mathcal{L}^{\Sigma}_{K[\sigma]C\xi}$, which is the dynamic language $\mathcal{L}^{\Sigma}_{K[\sigma]\xi}$ expanded with a common knowledge operator; we then showed completeness by strengthening our language to that of *propositional dynamic logic*, to produce the language $\mathcal{L}^{\Sigma}_{PDL[\sigma]\xi}$ and establishing a completeness result in that setting.

Lastly, we considered a theoretical application of AEL in the study of the No Miracles principle in epistemic logic. The No Miracles principle has been shown to characterize the AL product update by delimiting what sort of indistinguishability between epistemic actions can be represented in the framework. An analysis of the principle revealed that some of its underlying intuitions are directly violated in the AEL framework; however, the core intuition of the principle (there is no 'spontaneous', or 'miraculous', learning) is preserved in the framework. To show this, we provided a reduction-free axiomatization of AEL which features a weakened version of No Miracles; we show that this version actually characterizes the product update in AEL. We concluded this application with a note on how this result may bear on the verification of security protocols using model checkers in dynamic epistemic logic.

6.1 Possible extensions

A natural question is whether this adaptation to the semantics of AL may apply to the study of higher-order uncertainty in settings involving agents making decisions. For example, *epistemic game theory* is a field in which higher-order uncertainty plays a primary role in the analysis of games: how an agent will play a game depends on how they think their opponent will play, how the agent thinks their opponent thinks the agent will play, and so on.[57] In such settings, it is typical to represent an agent's values with a real-valued *utility function*, and to represent an agent's beliefs using a *probability* distribution. In particular, the relational representation of an agent's uncertainties in the epistemic model is replaced with a probability distribution over the states; an agent's uncertainty about epistemic actions is also captured by a probability distribution over those actions. What, then, should we make of higher-order uncertainty involving epistemic actions, like the examples in this thesis? In the single-agent case (involving higher-order uncertainty about one's *own* epistemic abilities), there is a relatively straightforward answer provided in Savage's classic *The Foundations of Statistics* [2, p.58].[58] Savage notes that, in the case of personal probabilities, 'higher-order' probability distributions for a single agent – that is, probability distributions over probability distributions over events of interest – collapse to the first order, simply by taking the composite of these distributions.[59] Thus, in a single-agent setting which represents uncertainty using probability distributions, the notion of higher-order uncertainty (involving epistemic actions or otherwise) reduces to first-order uncertainty. However, higher-order uncertainty in a *multi-agent* setting does not collapse in this same way; in this case, agent j's uncertainty about agent k's epistemic abilities is represented as a probability distribution over possible probability distributions which may represent k's epistemic state.

The central insight of this dissertation is that, in order to represent this higher-order uncertainty in a relational setting, we must endogenize the agents' ability to discern epistemic actions.[60] We make the fact 'whether j can distinguish actions σ and σ'' something which is true or false at states within the model; this permits the representation of situations where agents learn about others' (or their own) epistemic abilities. This accounts for the dynamics involved in agents' knowledge of the agents' epistemic abilities; however, this does not account for dynamics which may occur at the level of the agents' epistemic abilities *themselves*. In particular, we might wonder about situations where agent j *initially* cannot discern σ and σ', but then – through information she receives from an external source, or through introspection – she comes to acquire this ability. For example, consider Bob in example 14 where he is unsure whether he has correctly remembered the password; suppose that Bob has some time before needing to enter the password, and uses some recollection techniques which help him remember the password. In this case, his ability to distinguish σ_p and $\sigma_{\neg p}$ shouuld *change* over time. By assumption, our action-distinguishability functions are fixed in the update mechanism – $f_j^+((w,\sigma)) = f_j(w)$ whenever $f_j^+((w,\sigma))$ is defined. Of course, there is a straightforward way of relaxing this assumption: we may simply redefine f_j to be a function with domain

[57] For a helpful and thorough introduction to epistemic game theory and its scope of study, see [35].

[58] My thanks go to Teddy Seidenfeld for this observation.

[59] We give a trivial example here: suppose that we are interested in the likelihood of an event B, but are unsure whether to follow distribution p_1, which assigns B likelihood 1/8, or p_2, which assigns B likelihood 3/8. Our uncertainty between these distributions is captured by a 'higher-order' probability distribution, q, which tells us that p_1 and p_2 have equal chances of being correct. Taking the composite of these functions tells us to assign likelihood $1/2 * p_1(B) + 1/2 * p_2(B) = 1/4$ to B. This method generalizes to any (finite) level of 'higher-order' distribution.

[60] My thanks also go to Teddy for raising these discussion points.

$W \times \Sigma^*$. This way, for each world w and sequence of actions σ^*, f_j specifies j's ability to distinguish epistemic actions at w after σ^* has been performed.

This adaptation is promising, but not a fully satisfactory response – while this permits an agent's action-indistinguishability function to 'change' over time, it provides no insight into *how* this function might change. Consider another minor extension to the situation in 14: suppose that, in addition to modeling Bob's ability to recall the password after some effort, we also wish to model *Anne*'s uncertainty about whether Bob may do this successfully. To model this uncertainty, we would need to represent the different ways in which Bob's abilities might evolve over time – in our framework, this translates most naturally to uncertainty regarding the values of f_b (defined on $W \times \Sigma^*$). However, there is only *one* action-distinguishability function for each agent, and so each function is common knowledge in the model; Anne's uncertainty in this case must be built into the states themselves, rather than the functions. Extending the framework to overcome this limitation is the subject of current investigation.

The axiomatizations provided in 4 raise some questions about AEL which have not yet been answered. In particular, the introduction of the ξ constants plays an essential role in establishing completeness for AEL; it is clear that $\mathcal{L}^{\Sigma}_{K[\sigma]}$ cannot be reduced to $\mathcal{L}_K$ without this extension. However, this prompts the question: is there a weaker extension of the dynamic language which affords a complete axiomatization? The completeness proof for AEL with common knowledge raises a similar question. It was pointed out that extending the language to PDL involved a large jump in expressive power: is there a weaker extension whereby we can axiomatize AEL with common knowledge?

Lastly, this thesis does not treat questions concerning the complexity of the logic. It is well-known that the model-checking problem for public announcement logic is in P ([14]), and the satisfiability problem is PSPACE-complete ([18]); the model-checking problem for AL is PSPACE-complete and the satisfiability problem of AL is NEXPTIME-complete ([29]). Preliminary investigations indicate that the complexity of AEL coincides with that of AL, but this not yet known.

A related but distinct question involves the succinctness of AEL – this topic is also treated in [18] with respect to PAL.[61] In [18], Lutz shows that even though $\mathcal{L}_{K[\varphi]}$ (the language of public announcement logic) is reducible to $\mathcal{L}_K$, it can formulate certain properties exponentially more succinctly than the latter. However, this *kind* of succinctness formulated in [18] is not relevant in the case of AEL; AEL does not present a more succinct *language* than that of AL. Rather, the improvement in efficiency occurs in the size of the model itself. AEL may, in many cases, require fewer epistemic actions (a smaller Σ) than AL to represent a dynamic scenario (see examples 6 and 7). Thus, the question in the present case may be articulated using the result from theorem 9, which states that the AEL update may always be simulated by AL: how much larger will the action model in this simulation need to be, generally? Although this improvement in succinctness is intuitively clear in examples which violate the No Miracles principle, formulating a precise bound for this phenomenon has proved elusive, and is part of ongoing work.

[61] My thanks go to Eric Pacuit for raising this discussion point.

7 References

[1] R. Carnap. "Modalities and quantification". In: *The Journal of Symbolic Logic* 11.2 (1946), pp. 33–64.

[2] L.J. Savage. *The Foundations of Statistics*. New York: John Wiley & Sons, 1954.

[3] J. Hintikka. *Knowledge and Belief: An Introduction to the Logic of the Two Notions*. Cornell University Press, 1962.

[4] S. Kripke. "Semantical analysis of modal logic". In: *Zeitschrift fur Mathematische Logik und Grundlagen der Mathematik* 9 (1963), pp. 67–96.

[5] Joseph Y Halpern and Moshe Y Vardi. "The complexity of reasoning about knowledge and time. I. Lower bounds". In: *Journal of Computer and System Sciences* 38.1 (1989), pp. 195–237.

[6] J. Plaza. "Logics of public announcements". In: *Proceedings 4th International Symposium on Methodologies for Intelligent Systems*. 1989.

[7] R. Fagin et al. *Reasoning About Knowledge*. A slightly revised paperback version was published in 2003. Cambridge, Mass.: MIT Press, 1995.

[8] J. Gerbrandy and W. Groeneveld. "Reasoning about information change". In: *Journal of logic, language and information* 6.2 (1997), pp. 147–169.

[9] A. Baltag, L. S. Moss, and S. Solecki. "The logic of public announcements, common knowledge, and private suspicions". In: *Proceedings of the conference on Theoretical Aspects of Rationality and Knowledge (TARK 98)*. Ed. by I. Gilboa. 1998, pp. 43–56.

[10] J. van Benthem. "Games in Dynamic-Epistemic Logic". In: *Bulletin of Economic Research* 53.4 (2001), pp. 219–248.

[11] P. Blackburn, M. de Rijke, and Y. Venema. *Modal Logic*. Cambridge Tracts in Theoretical Computer Science, No. 53. Cambridge, U.K.: Cambridge University Press, 2001.

[12] J. Copeland. "The genesis of possible worlds semantics". In: *Journal of Philosophical logic* 31.2 (2002), pp. 99–137.

[13] A. Baltag and L. S. Moss. "Logics for Epistemic Programs". In: *Synthese* 139.2 (2004), pp. 165–224. DOI: `10.1023/B:SYNT.0000024912.56773.5e`.

[14] B. Kooi and J. van Benthem. "Reduction axioms for epistemic actions". In: *AiML-2004: Advances in Modal Logic* (2004), pp. 197–211.

[15] J. van Eijck. "Reducing Dynamic Epistemic Logic to PDL by Program Transformation". In: *Software Engineering* (2004).

[16] J. van Eijck J. van Benthem and B. Kooi. "Common Knowledge in Update Logics". In: *Proceedings of the 10th Conference on Theoretical Aspects of Rationality and Knowledge*. National University of Singapore. 2005, pp. 253–261.

[17] J. van Benthem and E. Pacuit. "The Tree of Knowledge in Action: Towards a Common Perspective". In: *Advances in Modal Logic* 6 (2006), pp. 87–106.

[18] C. Lutz. "Complexity and succinctness of public announcement logic". In: *Proceedings of the fifth international joint conference on Autonomous agents and multiagent systems*. 2006, pp. 137–143.

[19] J. van Eijck and S. Orzan. "Epistemic verification of anonymity". In: *Electronic Notes in Theoretical Computer Science* 168 (2007), pp. 159–174.

[20] F. Dechesne and Y. Wang. "Dynamic epistemic verification of security protocols: framework and case study". In: *A Meeting of the Minds: Proceedings of the Workshop on Logic, Rationality, and Interaction, Beijing*. Citeseer. 2007.

[21] J. Plaza. "Logics of Public Communications". In: *Synthese* 158 (2007), pp. 165–179. DOI: 10.1007/s11229-007-9168-7.

[22] H. van Ditmarsch, W. van der Hoek, and B. Kooi. *Dynamic Epistemic Logic*. Springer, 2008.

[23] J. van Benthem et al. "Merging frameworks for interaction". In: *Journal of Philosophical logic* 38.5 (2009), pp. 491–526.

[24] I. Boureanu, M. Cohen, and A. Lomuscio. "Automatic verification of temporal-epistemic properties of cryptographic protocols". In: *Journal of Applied Non-Classical Logics* 19.4 (2009), pp. 463–487.

[25] Francien Dechesne and Yanjing Wang. "To know or not to know: epistemic approaches to security protocol verification". In: *Synthese* 177.1 (2010), pp. 51–76.

[26] B. Kooi and E. Pacuit. "Logics of Rational Interaction". In: *Dynamic Formal Epistemology*. Ed. by P. Girard, O. Roy, and M. Marion. Springer, 2011, pp. 5–32.

[27] C. Dégremont, B. Löwe, and A. Witzel. "The synchronicity of dynamic epistemic logic". In: *Proceedings of the 13th Conference on Theoretical Aspects of Rationality and Knowledge*. 2011, pp. 145–152.

[28] J. Ruan J. van Eijck and T. Sadzik. "Action emulation". In: *Synthese* 185.1 (2012), pp. 131–151.

[29] G. Aucher and F. Schwarzentruber. "On the Complexity of Dynamic Epistemic Logic". In: *Proceedings of the conference on Theoretical Aspects of Rationality and Knowledge (TARK 2013)*. hal-00856468. Chennai, India, 2013.

[30] J. Seligman, F. Liu, and P. Girard. "Facebook and the epistemic logic of friendship". In: *arXiv preprint arXiv:1310.6440* (2013).

[31] Y. Wang and G. Aucher. "An Alternative Axiomatization of DEL and its Applications". In: *Twenty-Third International Joint Conference on Artificial Intelligence*. 2013.

[32] Y. Wang and Q. Cao. "On axiomatizations of public announcement logic". In: *Synthese* 190.1 (2013), pp. 103–134.

[33] M. Gattinger and J. van Eijck. "Towards model checking cryptographic protocols with dynamic epistemic logic". In: *Proc. LAMAS*. Citeseer. 2015, pp. 1–14.

[34] T. Bolander et al. "Announcements to Attentive Agents". In: *Journal of Logic, Language, and Information* 25.1 (2016), pp. 1–25. DOI: 10.1007/s10849-015-9234-3.

[35] E. Pacuit and O. Roy. “Epistemic Foundations of Game Theory”. In: *The Stanford Encyclopedia of Philosophy*. Ed. by Edward N. Zalta. Summer 2017. Metaphysics Research Lab, Stanford University, 2017.

[36] R. de Haan and I. van de Pol. “On the Computational Complexity of Model Checking for Dynamic Epistemic Logic with S5 Models”. In: *CoRR* abs/1805.09880 (2018). arXiv: `1805.09880`. URL: `http://arxiv.org/abs/1805.09880`.

[37] X. Chen and H. Deng. “Efficient Verification of Cryptographic Protocols with Dynamic Epistemic Logic”. In: *Applied Sciences* 10.18 (2020), p. 6577.