Exploring the $\mathrm{KD45}_n$ Property of a Kripke Model After the Execution of an Action Sequence

Tran Cao Son and **Enrico Pontelli**
Computer Science Department
New Mexico State University

Chitta Baral and **Gregory Gelfond**
Department of Computer Science and Engineering
Arizona State University

Abstract

The paper proposes a condition for preserving the $\mathbf{KD45}_n$ property of a Kripke model when a sequence of update models is applied to it. The paper defines the notions of a *primitive update model* and a *semi-reflexive* $\mathbf{KD45}_n$ (or sr-$\mathbf{KD45}_n$) Kripke model. It proves that updating a sr-$\mathbf{KD45}_n$ Kripke model using a primitive update model results in a sr-$\mathbf{KD45}_n$ Kripke model, i.e., a primitive update model preserves the properties of a sr-$\mathbf{KD45}_n$ Kripke model. It shows that several update models for modeling well-known actions found in the literature are primitive. This result provides guarantees that can be useful in presence of multiple applications of actions in multi-agent system (e.g., multi-agent planning).

Introduction and Motivations

In a multi-agent action setting, agents need to not just reason about the properties of the world, but also about their *knowledge* and *beliefs* about the world and about other agents' own knowledge and beliefs. Among the various formalizations of multi-agent actions and their impact on a physical and mental world, the *action models* introduced in (Baltag and Moss 2004; Baltag, Moss, and Solecki 1998) and later extended to *update models* in (van Benthem, van Eijck, and Kooi 2006; van Ditmarsch, van der Hoek, and Kooi 2007) are the most widely accepted. For example, these action models and update models[1] have been employed in the study of epistemic planning problems in MAS (Bolander and Andersen 2011; Löwe, Pacuit, and Witzel 2011; van der Hoek and Wooldridge 2002; Baral et al. 2013).

However, there is something fundamental missing in these formulations. Let us consider a variant of the coin example from (Baltag and Moss 2004): there are two agents, 1 and 2, and a coin in a box. The coin lies heads up, but the two agents are unaware of this fact. Let us assume that agent 1 alone learns that the coin lies heads up (e.g., *by peeking into the box while 2 is looking away*). Intuitively, in the resulting state, we should be able to conclude that the following formulae are true ($\mathbf{K}_i\varphi$ and $\mathbf{B}_i\varphi$ represent that i knows that φ is true (or i knows φ, for short) and i believes that φ is true (i believes φ); h denotes the coin being heads up): $\mathbf{K}_1 h$ (1 *knows* h); $\neg\mathbf{K}_2 h \wedge \neg\mathbf{K}_2\neg h$ (2 does not know h); $\mathbf{K}_1(\neg\mathbf{K}_2 h \wedge \neg\mathbf{K}_2\neg h)$ (1 knows that 2 does not know h); $\mathbf{B}_2(\neg\mathbf{K}_1 h \wedge \neg\mathbf{K}_1\neg h)$ (2 believes that 1 does not know h); $\mathbf{K}_1(\mathbf{B}_2(\neg\mathbf{K}_1 h \wedge \neg\mathbf{K}_1\neg h))$ (1 knows that 2 believes that 1 does not know h); etc. Intuitively, and as well accepted, the distinction between knowledge and beliefs is that knowledge must be true in the actual world while beliefs could be false.

But when one looks at the formulations in (Baltag and Moss 2004; Baltag, Moss, and Solecki 1998) the Kripke models have only one accessibility relation and only one modality is defined; in (Baltag and Moss 2004) it is referred to as "knowledge". As one can notice, replacing $\mathbf{B}_2(\neg\mathbf{K}_1 h \wedge \neg\mathbf{K}_1\neg h)$ by $\mathbf{K}_2(\neg\mathbf{K}_1 h \wedge \neg\mathbf{K}_1\neg h)$ would not be intuitive. Also, the action and update model papers do not analyze the properties of the resulting Kripke models.

In this paper, we propose a method for drawing conclusions regarding both knowledge and beliefs of agents (e.g., the above mentioned formulae) after the execution of a sequence of actions, in the spirit of belief updates (and update models) rather than the belief revision based approaches[2] used in (Segerberg 1998; Baltag and Smets 2008; van Benthem 2007; van Ditmarsch 2005). We achieve this goal by noting that *knowledge* can be reduced to *true belief* in certain logics (Halpern, Samet, and Segev 2009). To see how this works, let us revisit the introductory example. The initial state of beliefs of agents 1 and 2 can be represented by the pointed Kripke model on the top-left of Fig. 1 (the actual world has a double circle). The logic of this model, together with the axiom $\mathbf{K}_i\varphi \leftrightarrow \mathbf{B}_i\varphi \wedge \varphi$ (indicating knowledge is true belief), has $\neg\mathbf{K}_i h \wedge \neg\mathbf{K}_i\neg h$ $(i = 1, 2)$ as a consequence. The update model representing *"1 peeks into the box while 2 is looking away"* is given in the bottom-left of Fig. 1 which encodes the view of the action occurrence of each agent. Intuitively, it says that 1 sees the action occurs (denoted by the event σ) while 2 does not (τ). The result of the update is the pointed Kripke model on the right. It is easy to check that the logic specified by the result of the update (with the axiom $\mathbf{K}_i\varphi \leftrightarrow \mathbf{B}_i\varphi \wedge \varphi$), entails the formulae in our initial example (e.g., 1 knows h, 1 knows that 2 believes that 1 does

[1] For the sake of simplicity, we will use the generic term *"update model"* to denote both "action models" and "update models." Furthermore, we will frequently use the terminology *"execution of an action"* to indicate the result of applying an update model.

[2] The belief revision based approach needs an additional ordering relation and is orthogonal to our approach.

not knows h, etc.).

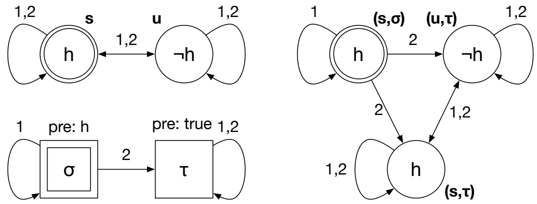

Figure 1: Knowledge as true belief

With our emphasis on distinguishing belief and knowledge, we also focus on what properties these modalities have and how they are preserved when a sequence of actions is executed. A limited attention to this has been paid in the literature—(Herzig, Lang, and Marquis 2005) being an exception. The focus has been mostly on properties of knowledge or beliefs of the agents after a single update—e.g., (Baltag and Moss 2004) shows that all agents have the correct knowledge about a formula φ after the execution of a public announcement. On the other hand, the more general question of whether the fundamental properties of beliefs (i.e., $\mathbf{KD45}_n$) and knowledge (i.e., $\mathbf{S4.4}_n$) are preserved by interesting classes of actions has been rarely touched.

In this paper, we provide a feasible approach for drawing conclusions about the knowledge and beliefs of agents after the execution of an arbitrary sequence of actions, i.e., after updating a Kripke model by a sequence of update models. This is particularly important in applications such as in security (e.g., exchanging a key between a group of agents without revealing it to outsiders) and multi-agent planning (e.g., executing an attack plan while making the enemy believe otherwise). To this end, we present a sufficient syntactic condition on update models and Kripke models under which the $\mathbf{KD45}_n$ property of the Kripke models is preserved after application of an update model. We observe that this result could be viewed as an extension of the result in (Halpern, Samet, and Segev 2009) to a dynamic setting. We demonstrate that update models found in the literature for representing several well-known actions satisfy the proposed sufficient condition. In the process, we identify some critical issues that need to be taken into consideration when update models are used.

Background

Logics of Knowledge and Belief. We consider the standard logics of knowledge and belief with a set of modalities $L_1, \ldots, L_k$ and use the notation from (Halpern, Samet, and Segev 2009). We use a collection $\mathcal{P}$ of *propositions* to describe the properties that characterize the world. $\mathcal{L}_{\mathcal{AG}}(L_1, \ldots, L_k)$ is the set of formulae defined as follows. Each $p \in \mathcal{P}$ is a formula. If φ and ψ are formulae then so are $\neg\varphi$, $\varphi \rightarrow \psi$, and $L_i\varphi$ ($i = 1, \ldots, k$). The connectives $\vee, \wedge, \leftrightarrow$ can be defined in terms of $\neg$ and $\rightarrow$. An *atomic formula* is a formula that does not contain a modal operator.

A logic Λ is a set of formulae in $\mathcal{L}_{\mathcal{AG}}(L_1, \ldots, L_k)$ that **(i)** contains all propositional tautologies; **(ii)** is closed under modus ponens; and **(iii)** is closed under substitution. A logic is *normal* if it contains the axioms $L_i(\varphi \rightarrow \psi) \rightarrow (L_i\varphi \rightarrow L_i\psi)$, referred to as $(\mathbf{K}_{L_i})$ for $i = 1, \ldots, k$, and is closed under *generalization*, i.e., if it contains φ, then it will also contain $L_i\varphi$. A logic generated by a set A of formulae (*axioms*) is the smallest normal logic containing A. For two sets of axioms Λ_1 and Λ_2, $\Lambda_1 + \Lambda_2$ is the smallest normal logic containing Λ_1 and Λ_2.

Let $\mathcal{AG} = \{1, 2, \ldots, n\}$ be a set of n agents. Each agent i is associated with a belief operator $\mathbf{B}_i$ and a knowledge operator $\mathbf{K}_i$. Our focus is the logic of belief, called $\mathbf{KD45}_n$, over the language $\mathcal{L}_{\mathcal{AG}}(\mathbf{B}_1, \ldots, \mathbf{B}_n)$ that is generated by the following axioms: **(D)** $\mathbf{B}_i\varphi \rightarrow \neg\mathbf{B}_i\neg\varphi$, **(4)** $\mathbf{B}_i\varphi \rightarrow \mathbf{B}_i\mathbf{B}_i\varphi$, and **(5)** $\neg\mathbf{B}_i\varphi \rightarrow \mathbf{B}_i\neg\mathbf{B}_i\varphi$ where $i \in \mathcal{AG}$ and $\varphi \in \mathcal{L}_{\mathcal{AG}}(\mathbf{B}_1, \ldots, \mathbf{B}_n)$.

It is shown in (Halpern, Samet, and Segev 2009) that knowledge can be reduced to true belief in $\mathbf{KD45}_n$. More specifically, the knowledge modality $\mathbf{K}_i$ is reducible to $\mathbf{B}_i$ in $\mathbf{KD45}_n$ by the axiom $\mathbf{K}_i\varphi \leftrightarrow (\mathbf{B}_i\varphi \wedge \varphi)$. This also means that we can remove the knowledge modal operator from the language yet still be able to derive conclusions about knowledge of agents.

A *Kripke frame* $\mathcal{F}$ is a tuple $\langle S, \mathcal{K}_1, \ldots, \mathcal{K}_n\rangle$ where S is a set of worlds (or *points*) and $\mathcal{K}_i \subseteq S \times S$ for $i \in \mathcal{AG}$, called the *accessibility relation* for i. A *Kripke model* $\mathcal{M}$ *based on the frame* $\mathcal{F}$ is a pair $(\mathcal{F}, \pi)$, where $\pi : \mathcal{F}[S] \rightarrow 2^{\mathcal{P}}$ is a function that associates an interpretation of $\mathcal{P}$ to each world in $\mathcal{F}$. For $\mathcal{M} = (\mathcal{F}, \pi)$, $\mathcal{M}[\pi]$ denotes π and $\mathcal{M}[S]$ and $\mathcal{M}[i]$ denote the set of worlds S and $\mathcal{K}_i$ of $\mathcal{F}$, respectively.

A *pointed Kripke model* (or *p-model*) is a pair $(\mathcal{M}, s)$, where $\mathcal{M}$ is a Kripke model and $s \in \mathcal{M}[S]$, called the *actual world*. We will often represent a Kripke model $\mathcal{M}$ by a directed graph whose nodes are the worlds in $\mathcal{M}[S]$ and the labeled edges are the members of $\mathcal{M}[i]$. The name of a world is drawn next to the node and the interpretation associated to it is given by the label of the node. Entailment of formulae in $\mathcal{L}_{\mathcal{AG}}(L_1, \ldots, L_n)$ in a p-model is defined next.

Definition 1 *Given a formula φ and a p-model $(\mathcal{M}, s)$:*
- $(\mathcal{M}, s) \models \varphi$ *if* $\mathcal{M}[\pi](s) \models \varphi$ *and* φ *is an atomic formula;*
- $(\mathcal{M}, s) \models L_i\varphi$ *if, for each* t *s.t.* $(s, t) \in \mathcal{K}_i$, $(\mathcal{M}, t) \models \varphi$;
- $(\mathcal{M}, s) \models \neg\varphi$ *if* $(\mathcal{M}, s) \not\models \varphi$;
- $(\mathcal{M}, s) \models \varphi_1 \rightarrow \varphi_2$ *if* $(\mathcal{M}, s) \models \neg\varphi_1$ *or* $(\mathcal{M}, s) \models \varphi_2$.

$\mathcal{M} \models \varphi$ denotes that $(\mathcal{M}, s) \models \varphi$ for each $s \in \mathcal{M}[S]$. φ is said to be valid in a frame $\mathcal{F}$ if $\mathcal{M} \models \varphi$ for every Kripke model $\mathcal{M}$ based on $\mathcal{F}$. The set of valid formulae in $\mathcal{F}$ is denoted with $Th(\mathcal{F})$. For a class of frames $\mathcal{S}$, $Th(\mathcal{S})$ is the set of formulae valid in each frame in $\mathcal{S}$. A logic Λ is sound for $\mathcal{S}$ if $\Lambda \subseteq Th(\mathcal{S})$; it is complete for $\mathcal{S}$ if $Th(\mathcal{S}) \subseteq \Lambda$. For a logic Λ, a frame $\mathcal{F}$ is said to be a Λ frame if $\Lambda \subseteq Th(\mathcal{F})$.

A relation $R \subseteq S \times S$ is *reflexive* iff $(u, u) \in R$ for every $u \in S$; *serial* iff for every $u \in S$ there exists some $v \in S$ such that $(u, v) \in R$; *transitive* iff $(u, v) \in R$ and $(v, z) \in R$ imply that $(u, z) \in R$; *Euclidean* iff $(u, v) \in R$ and $(u, z) \in R$ imply that $(v, z) \in R$.

Frames can be characterized by the properties of their accessibility relations. It is known that a frame $\mathcal{F} = (S, \mathcal{K}_1, \ldots, \mathcal{K}_n)$ is a $\mathbf{KD45}_n$ frame iff for every $i = 1, \ldots, n$, $\mathcal{K}_i$ is serial, transitive, and Euclidean. A Kripke model $\mathcal{M} = (\mathcal{F}, \pi)$ is said to be a $\mathbf{KD45}_n$ model if its frame $\mathcal{F}$ is a $\mathbf{KD45}_n$ frame.

Update Models. Update models describe transformations of (pointed) Kripke models according to a predetermined pattern. An update model uses structures similar to pointed Kripke models and they describe the effects of a transformation on p-models using an update operator (Baltag and Moss 2004; van Benthem, van Eijck, and Kooi 2006).

A set $\{p \rightarrow \varphi \mid p \in \mathcal{P}, \varphi \in \mathcal{L}_{\mathcal{AG}}(\mathbf{B}_1, \ldots, \mathbf{B}_n)\}$ is called an $\mathcal{L}_{\mathcal{AG}}(\mathbf{B}_1, \ldots, \mathbf{B}_n)$-*substitution* (or substitution, for short). For each substitution sub and each $p \in \mathcal{P}$, we assume that sub contains exactly one formula $p \rightarrow \varphi$. For simplicity of the presentation, we often omit $p \rightarrow p$ in a substitution. $SUB_{\mathcal{L}_{\mathcal{AG}}}$ denotes the set of all substitutions. A substitution is used to encode changes caused by an action occurrence. A formula $p \rightarrow \varphi$ in a substitution states the condition (φ) under which p will become true. For example, the action of flipping a coin can be represented by the substitution $\{h \rightarrow \neg h\}$ which says that h (the coin lies head up) is true iff $\neg h$ (the coin lies head down) were true.

Definition 2 (Update Model) *An* update model $\mathbf{\Sigma}$ *is a tuple* $\langle \Sigma, R_1, \ldots, R_n, pre, sub \rangle$ *where*

- Σ *is a set, whose elements are called* events*;*
- *each* R_i *is a binary relation on* Σ*;*
- $pre : \Sigma \rightarrow \mathcal{L}_{\mathcal{AG}}(\mathbf{B}_1, \ldots, \mathbf{B}_n)$ *is a function mapping each event* $a \in \Sigma$ *to a formula in* $\mathcal{L}_{\mathcal{AG}}(\mathbf{B}_1, \ldots, \mathbf{B}_n)$*; and*
- $sub : \Sigma \rightarrow SUB_{\mathcal{L}_{\mathcal{AG}}}$.

An update instance ω *is a pair* $(\mathbf{\Sigma}, e)$ *where* $\mathbf{\Sigma}$ *is an update model and* $e \in \Sigma$ *(or a* designated event*). An* update template *is a pair* $(\mathbf{\Sigma}, \Gamma)$ *where* $\mathbf{\Sigma}$ *is an update model with the set of events* Σ *and* $\Gamma \subseteq \Sigma$.

Intuitively, an update model represents different views of an action occurrence, associated to the observability of agents. Each view is represented by an event in Σ. The designated event in an update instance is the one that agents who are aware of the action occurrence will observe. Templates extend the notion of instance to capture non-deterministic actions and other non-simple actions. The relation R_i describes agent i's uncertainty about an action occurrence—i.e., if $(\sigma, \tau) \in R_i$ and event σ is performed, then agent i may believe that event τ is executed instead. pre defines the action precondition and sub specifies the changes of fluent values after the execution of an action. An update model is *serial* (resp., *reflexive*, *transitive*, *Euclidean*) if, for every $i \in \mathcal{AG}$, R_i is serial (resp., reflexive, transitive, Euclidean).

The update model in Fig. 1 is given by $\mathbf{\Sigma}_0 = \langle \{\sigma, \tau\}, \{(\sigma, \sigma), (\tau, \tau)\}, \{(\sigma, \tau), (\tau, \tau)\}, pre, sub \rangle$ where $pre(\sigma) = h$, $pref(\tau) = true$, and $sub(\sigma) = sub(\tau) = \emptyset$. It says that if the event σ occurs then 1 is certain that σ occurs while 2 thinks that τ occurs.

Definition 3 (Updates by an Update Model) *Let* $\mathbf{\Sigma} = \langle \Sigma, R_1, \ldots, R_n, pre, sub \rangle$ *be an update model and* $\mathcal{M} = (\mathcal{F}, \pi)$ *be a Kripke model. The* update operator *induced by* $\mathbf{\Sigma}$ *defines a Kripke model* $\mathcal{M}' = \mathcal{M} \otimes \mathbf{\Sigma}$*, where:*

- $\mathcal{M}'[S] = \{(s, \tau) \mid s \in \mathcal{M}[S], \tau \in \Sigma, (\mathcal{M}, s) \models pre(\tau)\}$*;*
- $((s, \tau), (s', \tau')) \in \mathcal{M}'[i]$ *iff* $(s, \tau), (s', \tau') \in \mathcal{M}'[S]$, $(s, s') \in \mathcal{M}[i]$ *and* $(\tau, \tau') \in R_i$*;*
- $\forall f \in \mathcal{P}. [\mathcal{M}'[\pi]((s, \tau)) \models f$ *iff* $f \rightarrow \varphi \in sub(\tau), (\mathcal{M}, s) \models \varphi]$.

The update of a p-model $(\mathcal{M}, s)$ given an update template $(\mathbf{\Sigma}, \Gamma)$ is a set of p-models, denoted by $(\mathcal{M}, s) \otimes (\mathbf{\Sigma}, \Gamma)$, where $(\mathcal{M}', s') \in (\mathcal{M}, s) \otimes (\mathbf{\Sigma}, \Gamma)$ iff it holds that $\mathcal{M}' = \mathcal{M} \otimes \mathbf{\Sigma}$ and $s' = (s, \tau)$ where $\tau \in \Gamma$ and $s' \in \mathcal{M}'[S]$.

Intuitively, the set $(\mathcal{M}, s) \otimes (\mathbf{\Sigma}, \Gamma)$ is the set of p-models encoding the result of the execution of the action, which is represented by the update template $(\mathbf{\Sigma}, \Gamma)$, in the p-model $(\mathcal{M}, s)$. It is easy to see that the p-model on the right of Fig. 1 is the unique element of $(\mathcal{M}_0, s) \otimes (\mathbf{\Sigma}, \{\sigma\})$ where $\mathcal{M}_0 = (\langle \{s, u\}, \mathcal{K}_1, \mathcal{K}_2 \rangle, \pi)$ with $\mathcal{K}_1 = \mathcal{K}_2 = \{(x, y) \mid x, y \in \{s, u\}\}$, $\pi(s)(h) = true$, and $\pi(u)(h) = false$.

We will often depict an update instance via a graph with rectangles representing events (double rectangles for designated events), and labeled edges representing each R_i.

Maintaining $\mathrm{KD45}_n$ via Update Models

In this section, we present a sufficient condition for the maintenance of the $\mathbf{KD45}_n$ property of a Kripke model $\mathcal{M}$ after the application of a sequence of update models $\mathbf{\Sigma}_1, \ldots, \mathbf{\Sigma}_k$. The next lemma shows that updating a p-model by an update model does preserve reflexivity, transitivity, and Euclidicity of a Kripke model if the update model also satisfies the corresponding property. Observe that, since we are exploring structural properties of the frames underlying updated models, we will not emphasize the designated events and the actual world whenever it is unnecessary.

Lemma 1 *Let* $\mathcal{M} = (\mathcal{F}, \pi)$ *be a Kripke model and* $\mathbf{\Sigma} = \langle \Sigma, R_1, \ldots, R_n, pre, sub \rangle$ *be an update model. Then, the following properties hold:*

- *if* $\mathcal{M}$ *and* $\mathbf{\Sigma}$ *are reflexive then* $\mathcal{M} \otimes \mathbf{\Sigma}$ *is reflexive;*
- *if* $\mathcal{M}$ *and* $\mathbf{\Sigma}$ *are transitive then* $\mathcal{M} \otimes \mathbf{\Sigma}$ *is transitive; and*
- *if* $\mathcal{M}$ *and* $\mathbf{\Sigma}$ *are Euclidean then* $\mathcal{M} \otimes \mathbf{\Sigma}$ *is Euclidean.*

The lemma shows that reflexivity, transitivity, and Euclidicity of a Kripke model can be easily maintained if the update model possesses the same property. As reflexivity implies seriality, the properties of $\mathbf{KD45}_n$ would be maintained if all update models used are reflexive, transitive, and Euclidean. This is, however, not always the case, as we will see later. Furthermore, seriality of $\mathcal{M} \otimes \mathbf{\Sigma}$ is not guaranteed even when $\mathcal{M}$ and $\mathbf{\Sigma}$ are serial, as shown next.

Example 1 *Consider the domain with two agents* 1 *and* 2 *and the set of propositions* $\{h\}$*. Let us consider the update model* $\mathbf{\Sigma}_1 = \langle \{\sigma\}, R_1, R_2, pre, sub \rangle$*, where* $R_1 = R_2 = \{(\sigma, \sigma)\}$*,* $pre(\sigma) = h$*, and* $sub(\sigma) = \emptyset$ *(Fig. 2, top left).*

Let $\mathcal{M}_1$ *be the Kripke model* $(\mathcal{F}_1, \pi_1)$*, where* $\mathcal{F}_1 = \langle \{s, u\}, \{(s, s), (u, u)\}, \{(s, u), (u, u)\} \rangle$*,* $\pi_1(s) = \{h\}$*, and* $\pi_1(u) = \emptyset$ *(Fig. 2, bottom left). As we will see later, the update model represents a public announcement that h is true.*

We can see that $\mathbf{\Sigma_1}$ *and* $\mathcal{M}_1$ *are serial, transitive, and Euclidean. However,* $\mathcal{M}_1 \otimes \mathbf{\Sigma}_1$ *(right, Fig. 2), is not serial, since there is no successor of* (s, σ) *with respect to the agent* 2.

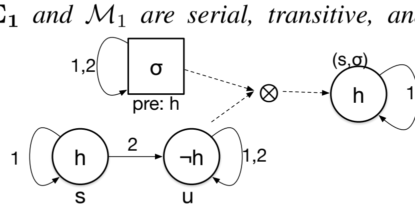

Figure 2: Lost of seriality

The above example also highlights that the update models of epistemic actions might not work properly for logics different from **S5**. We will return to this issue in Discussion Section at the end of the paper. The main reason for the loss of the seriality in $\mathcal{M}_1 \otimes \mathbf{\Sigma}_1$ lies in the fact that the precondition of σ in $\mathbf{\Sigma}_1$ is h and both R_1 and R_2 are reflexive. This indicates that for $\mathcal{M}_1 \otimes \mathbf{\Sigma}_1$ to maintain its seriality, in each world that satisfies the precondition of σ, the accessibility relation of any agent must contain a successor with the same interpretation. Intuitively, this means that the agent cannot have wrong beliefs. We will refer to this property as *semi-reflexivity* and formalize it as follows.

Let $\mathcal{M}$ be a Kripke model and let $u, v \in \mathcal{M}[S]$. Let us define $u \sim v$ iff $\mathcal{M}[\pi](u) \equiv \mathcal{M}[\pi](v)$. Intuitively, $u \sim v$ means that u and v are associated to the same interpretation over $\mathcal{P}$ by the function $\mathcal{M}[\pi]$.

Definition 4 *A Kripke model $\mathcal{M}$ is* semi-reflexive *if for every $i \in \mathcal{AG}$ and $u \in \mathcal{M}[S]$, there exists some v such that $(u, v) \in M[i]$ and $u \sim v$.*

It is easy to see that a Kripke model is semi-reflexive then it is serial. Furthermore, semi-reflexivity is a property of a Kripke model and differs seriality, transitivity, or reflexivity which are properties of the frame of the Kripke model. Although semi-reflexive Kripke models can avoid the problem mentioned in Example 1, updating a semi-reflexive Kripke model using an arbitrary update model can result in a non-semi-reflexive Kripke model as seen in the next example.

Example 2 *Consider again the domain in Example 1. Let $\mathbf{\Sigma}_2$ and $\mathcal{M}_2$ be the update model and Kripke model depicted in Fig. 3, top left and bottom left, respectively. The substitution function in $\mathbf{\Sigma}_2$ is given by $sub(\sigma) = \{h \rightarrow false\}$ and $sub(\tau) = \{h \rightarrow true\}$. It is easy to see that $\mathcal{M}_2 \otimes \mathbf{\Sigma}_2$ (Fig. 3, right) is not semi-reflexive ((s, σ) wrt. agent 2).*

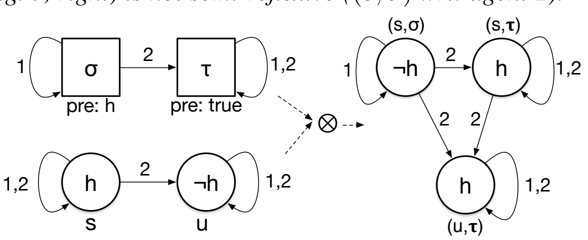

Figure 3: Loss of semi-reflexivity

In Example 2, $\mathcal{M}_2$ is reflexive. However, the accessibility relation for 2 in $\mathbf{\Sigma}_2$ is not. The key issue lies in that the substitution function in $\mathbf{\Sigma}_2$ assigns different interpretations for σ and τ. To maintain the semi-reflexivity of $\mathcal{M}_2$, there are two possibilities: the update model should be reflexive or there must be a link in the update model that allows the preservation of the semi-reflexivity of $\mathcal{M}_2$. This leads us to define a notion of a primitive update model:

Definition 5 *A serial, transitive, and Euclidean update model is* primitive *if the precondition $pre(\sigma)$ at every event σ is an atomic formula and for every $i \in \mathcal{AG}$ and $(\sigma, \tau) \in R_i$ such that $\sigma \neq \tau$ then either* (i) *$(\sigma, \sigma) \in R_i$; or* (ii) *$pre(\tau) = true$ and $sub(\tau) = sub(\sigma) = \emptyset$.*

A primitive update model $\mathbf{\Sigma}$ essentially guarantees that an agent cannot have false beliefs if she does not have false beliefs prior to the occurrence of the action represented by $\mathbf{\Sigma}$. This is proved in the next lemma.

Lemma 2 *If $\mathcal{M}$ is semi-reflexive and $\mathbf{\Sigma}$ is primitive then $\mathcal{M} \otimes \mathbf{\Sigma}$ is semi-reflexive.*

We are now ready to formalize a theorem that characterizes a sufficient condition for maintaining the belief structures of a Kripke model after the application of an update model.

Definition 6 *A Kripke model is a* sr-$\mathbf{KD45}_n$ *Kripke model if it is semi-reflexive, transitive, and Euclidean.*

Theorem 1 *For a sr-$\mathbf{KD45}_n$ Kripke model $\mathcal{M}$ and a primitive update model $\mathbf{\Sigma}$, $\mathcal{M} \otimes \mathbf{\Sigma}$ is a sr-$\mathbf{KD45}_n$ Kripke model.*

The proof of the theorem follows directly from Lemmas 1 and 2. This leads to the following consequence.

Corollary 1 *The result of applying a sequence of primitive update models to a sr-$\mathbf{KD45}_n$ Kripke model is a sr-$\mathbf{KD45}_n$ Kripke model.*

The corollary, together with the result in (Halpern, Samet, and Segev 2009), implies that we can assert the state-of-knowledge of the agents after the execution of a sequence of actions in terms of update models if the initial p-model is given by a sr-$\mathbf{KD45}_n$ Kripke model. We observe that the initial p-model in the majority of examples from the literature (e.g., the muddy children, the card game, the number game, etc.) are **S5** Kripke models (and, thus, sr-$\mathbf{KD45}_n$). This result is also significant to multi-agent planning approaches that employ update models (Bolander and Andersen 2011; Löwe, Pacuit, and Witzel 2011; van der Hoek and Wooldridge 2002) as the initial p-model in a multi-agent planning problem is often a **S5** Kripke model.

To support our claims, we need to show next that several classes of update models found in the literature (Baltag and Moss 2004; Baltag, Moss, and Solecki 1998; van Benthem, van Eijck, and Kooi 2006; van Ditmarsch, van der Hoek, and Kooi 2007) are indeed primitive.

Truthful Public Announcements. A truthful public announcement of a formula is an action that communicates to all agents that the formula is true. For example, in the muddy children example, the father informs his children that at least one of them is muddy. After the execution of a public announcement, all agents will know that the announced formula is true.

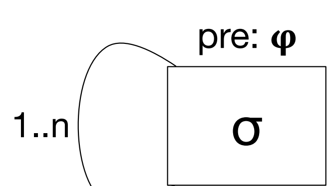

Figure 4: $\mathbf{\Sigma}^{Ann}(\mathcal{AG}; \emptyset; \varphi)$

A truthful public announcement of a formula φ to the agents in $\mathcal{AG}$ is often given in the literature by an update model $\mathbf{\Sigma}^{Ann}(\mathcal{AG}; \emptyset; \varphi) = \langle \{\sigma\}, R_1, \ldots, R_n, pre, sub \rangle$ where $R_1 = \ldots = R_n = \{(\sigma, \sigma)\}$, $pre(\sigma) = \varphi$, and $sub(\sigma) = \emptyset$ (Fig. 4). It is easy to see that if φ is an atomic formula then the update model $\mathbf{\Sigma}^{Ann}(\mathcal{AG}; \emptyset; \varphi)$ for truthful public announcement is a primitive update model.

Private Announcements. A private announcement of a formula φ to a group $A \subseteq \mathcal{AG}$ while other agents are unaware of its occurrence (e.g., the father whispering to one of his children that at least one of them is muddy) can be represented by an update model $\mathbf{\Sigma}^{Ann}(A;\emptyset;\varphi) = \langle\{\sigma,\tau\}, R_1,\ldots,R_n, pre, sub\rangle$ where $R_i = \{(\sigma,\sigma),(\tau,\tau)\}$ for $i \in A$, $R_i = \{(\sigma,\tau),(\tau,\tau)\}$ for $i \notin A$, $pre(\sigma) = \varphi$, $pre(\tau) = true$, and $sub(\sigma) = sub(\tau) = \emptyset$ (Fig. 5). We can easily check that if φ is an atomic formula then the update model $\mathbf{\Sigma}^{Ann}(A;\emptyset;\varphi)$ for private announcement is a primitive update model.

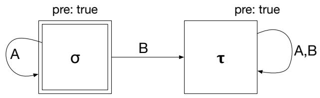

Figure 5: $\mathbf{\Sigma}^{Ann}(A;\emptyset;\varphi)$

Semi-Private Announcements. The occurrence of a private announcement action can be observed by some agents. In this situation, the agents—who observe the action occurrence—will be aware of the fact that the agents—to whom the announcement is made—have knowledge about the formula. However, they do not know the truth value of the formula. The action is referred to as a semi-private announcement. Let A be the set of agents to whom the announcement is made; let B the set of agents who observe the announcement, where $A \cap B = \emptyset$; let $C = \mathcal{AG} \setminus (A \cup B)$. A semi-private announcement to A with B observing can be formulated by the update model $\mathbf{\Sigma}^{Ann}(A;B;\varphi) = \langle\{\sigma,\tau,\epsilon\}, R_1,\ldots,R_n, pre, sub\rangle$ where $R_i = \{(\sigma,\sigma),(\tau,\tau),(\epsilon,\epsilon)\}$ for $i \in A$, $R_i = \{(\sigma,\sigma),(\tau,\tau),(\epsilon,\epsilon),(\sigma,\tau),(\tau,\sigma)\}$ for $i \in B$, $R_i = \{(\sigma,\epsilon),(\tau,\epsilon),(\epsilon,\epsilon)\}$ for $i \in C$, $pre(\sigma) = \varphi$, $pre(\tau) = \neg\varphi$, $pre(\epsilon) = true$, and $sub(\sigma) = sub(\tau) = sub(\epsilon) = \emptyset$ (Fig. 6). We can easily check that if φ is an atomic formula then the update model $\mathbf{\Sigma}^{Ann}(A;B;\varphi)$ for a semi-private announcement is a primitive update model.

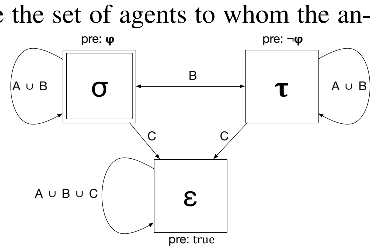

Figure 6: $\mathbf{\Sigma}^{Ann}(A;B;\varphi)$

Sensing Actions. A sensing action helps an agent to determine the truth value of a formula φ. In a multi-agent environment, the occurrence of a sensing action that determines the value of φ has a similar effect on the agents in the environment as a semi-private announcement. Indeed, the update model $\mathbf{\Sigma}^{Sensing}(A;B;\varphi)$ for a sensing action of a formula φ is identical to that of $\mathbf{\Sigma}^{Ann}(A;B;\varphi)$ where A and B are sets of agents that are fully aware and partially aware of the action occurrence (Fig. 7). The main distinction between a semi-private announcement and a sensing action lies in that the former will have only one designated event (σ) while the latter will have two (σ and τ). Again, we can check that $\mathbf{\Sigma}^{Sensing}(A;B;\varphi)$ is a primitive update model.

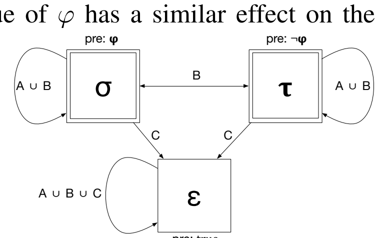

Figure 7: $\mathbf{\Sigma}^{Sensing}(A;B;\varphi)$

A subtle difference between the occurrence of an announcement action and a sensing action lies in the non-determinism of the state-of-knowledge of agents who are fully aware of the action occurrence.

Fully Observable Ontic Actions. An ontic action is different from announcement and sensing actions, in that it changes the "real" state of the world, and thereby changing the state of knowledge of the agents. For instance, the action of flipping a switch changes the position of the switch from *on* to $\neg$*on* (or *off*) and from $\neg$*on* to *on*. In general, the effects of an ontic action act can be given by the set C of formulae of the form $p \rightarrow \psi$ where ψ is an atomic formula characterizing the p-model of the world in which p would become true after the execution of act, i.e., by a substitution.

An occurrence of an ontic action that is fully observable by all agents can be modeled by an update model $\mathbf{\Sigma}^{On}(\mathcal{AG};\varphi;C) = (\{\sigma\}, R_1,\ldots,R_n, pre, sub)$ where $R_1 = \ldots = R_n = \{(\sigma,\sigma)\}$, $pre(\sigma) = \varphi$, and $sub(\sigma) = C$ where φ is the precondition of the action and C is the set of its effects. Again, we can see that if φ is an atomic formula then $\mathbf{\Sigma}^{On}(\mathcal{AG};\varphi;C)$ is a primitive update model.

Discussions

Our work in this paper is closely related to (Herzig, Lang, and Marquis 2005). Indeed, in our terminology, Proposition 1 in (Herzig, Lang, and Marquis 2005) states that the result of an update on a $\mathbf{KD45}_n$ Kripke model is a $\mathbf{KD45}_n$ Kripke model. As it turns out, the key difference between the two formalizations lies in the assumption made about the actions and the type of actions that are considered. We observe that the work in (Herzig, Lang, and Marquis 2005) considers only two types of actions, ontic actions and observation actions; furthermore, the latter type of actions is considered only in **S5** models. Our result, as shown in the previous section, is on the other hand applicable to all different types of actions. More importantly, it is assumed in (Herzig, Lang, and Marquis 2005) that the execution of an action in a given p-model of the world always results in some other p-model—this assumption means that actions can always be executed. This is, in our view, not a realistic assumption; for example, one cannot open a locked door without having a key for it. In our notation, this implies that for every update model $\mathbf{\Sigma} = \langle\Sigma, R_1,\ldots,R_n, pre, sub\rangle$, we have that $pre(\sigma) = true$ for every $\sigma \in \Sigma$. It is easy to see that, under this assumption, the update $\mathcal{M} \otimes \mathbf{\Sigma}$ is a $\mathbf{KD45}_n$ Kripke model if $\mathcal{M}$ and $\mathbf{\Sigma}$ are serial, transitive, and Euclidean.

Observe that the result of our work is also applicable for single agent domains. To the best of our knowledge, it is the first to distinguish between knowledge and beliefs of an agent in a single-agent domains. Previous works on reasoning about knowledge of an agent in the presence of incomplete information and sensing actions in single-agent domains (e.g., (Scherl and Levesque 2003; Son and Baral 2001)) do not investigate this distinction as they considered only the knowledge modal operator.

We conclude this section with an example that shows that the proposed sufficient condition does not cover the most general case of ontic actions and highlights a weakness of the approach using update models. We then discuss a possible solution for the issue raised by this example.

Example 3 *Consider a variation of Example 2 from (Baral et al. 2013): a domain with 3 agents $\mathcal{AG} = \{1, 2, 3\}$ and two propositions: o (*opened*) and h (*head*). The agents are in a room with a box that is not open ($\neg o$). The box contains a coin that lies head up (h), but none of the agent knows this. The initial p-model $(\mathcal{M}_0, s)$ encoding the knowledge of the agents and the actual world is given in Fig. 8 (left), where the double circle represents the actual world. The interpretation associated to each world is given by the label of the node. The names of the world (s and u) are given as text next to the node. On the right of Fig. 8 is an update template $(\mathbf{\Sigma}_{flip}, \{\sigma\})$, encoding the occurrence of the action* flip *that makes o true; agents 1 and 2 are aware of the action occurrence, while 3 is oblivious of it. The substitution in $\mathbf{\Sigma}_{flip}$ is given by $sub(\sigma) = \{o \rightarrow true\}$ and $sub(\tau) = \emptyset$. Observe that $\mathbf{\Sigma}_{flip}$ is not a primitive update model.*

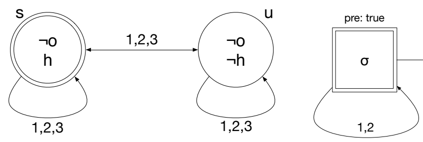

Figure 8: $(\mathcal{M}_0, s)$ (left) and $(\mathbf{\Sigma}_{flip}, \{\sigma\})$ (right)

After some inspection, we can see that updating $(\mathcal{M}_0, s)$ by $(\mathbf{\Sigma}_{flip}, \{\sigma\})$ results in the p-model $(\mathcal{M}_1, (s, \sigma))$ in Fig. 9. $\mathcal{M}_1$ is no longer a semi-reflexive Kripke model, but it is still a serial one.

Let us now update the p-model $(\mathcal{M}_1, (s, \sigma))$ with an update template, $(\mathbf{\Sigma}^{Ann}(\mathcal{AG}; \emptyset; o), \{\delta\})$, shown in Fig. 10 (left), representing the public announcement of o. The result of this update is shown on the right of Fig. 10. This resulting p-model is no longer serial. Even worse, the result renders that agent 3 becomes confused as his belief is inconsistent.

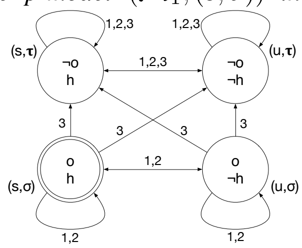

Figure 9: $(\mathcal{M}_1, (s, \sigma))$

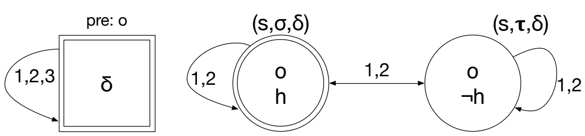

Figure 10: $(\mathbf{\Sigma}^{Ann}(\mathcal{AG}; \emptyset; o), \{\delta\})$ (left) and its result (right)

The above example shows that even when the initial p-model is **S5**, the execution of a sequence of non-primitive update models can result in the loss of the $\mathbf{KD45}_n$ property. This means that the proposed condition is one of the weakest sufficient conditions that guarantees the maintenance of the $\mathbf{KD45}_n$ property of the initial p-model. Observe that the issue displayed in this example can be attributed to the fact that agent 3 is oblivious of the occurrence of *flip*. If 3 was suspicious that the action might occur, then the update template would have been

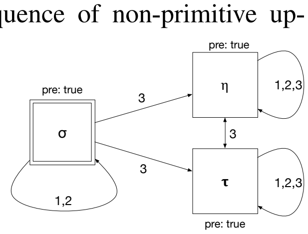

Figure 11: $(\mathbf{\Sigma}^2_{flip}, \{\sigma\})$

$(\mathbf{\Sigma}^2_{flip}, \{\sigma\})$ which is obtained from $(\mathbf{\Sigma}_{flip}, \{\sigma\})$ by adding an additional event η as in (Fig. 11) where $sub(\eta) = sub(\sigma)$. The result of updating $(\mathcal{M}_0, s)$ using $(\mathbf{\Sigma}^2_{flip}, \{\sigma\})$, $(\mathcal{M}_2, (s, \sigma))$ (Fig. 12), remains a sr-$\mathbf{KD45}_n$ p-model. We can verify that updating $(\mathcal{M}_2, (s, \sigma))$ with $(\mathbf{\Sigma}^{Ann}(\mathcal{AG}; \emptyset; o), \{\delta\})$ results in a sr-$\mathbf{KD45}_n$ p-model.

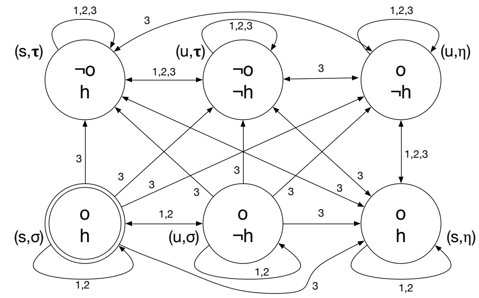

Figure 12: $(\mathcal{M}_2, (s, \sigma))$

The above example also shows that action models might yield counterintuitive results if applied to non-**S5** Kripke models. This is a legitimate concern, as the initial p-model $(\mathcal{M}_0, s)$ is a **S5** model. Thus, the question of whether previously developed action models are suitable for applications such as multi-agent planning (Bolander and Andersen 2011; Löwe, Pacuit, and Witzel 2011; van der Hoek and Wooldridge 2002) should be investigated. The above discussion also highlights the question of whether or not an agent should be considered as oblivious with respect to occurrences of ontic actions if he is not aware of the action occurrence. We believe that both of these issues are important for reasoning about effects of actions in multi-agent systems. They will be our main concerns in the near future.

Conclusions

Our main goal in this paper is to extend the result of (Halpern, Samet, and Segev 2009)—which states that knowledge is reducible to true beliefs in $\mathbf{KD45}_n$—to reasoning about knowledge and beliefs after the execution of an action sequence. To achieve this goal, we presented one of the weakest sufficient conditions on Kripke models and update models which preserves the $\mathbf{KD45}_n$ property of a Kripke model after it is updated by a sequence of update models. This provides a simple way for drawing conclusions regarding the knowledge of agents after the execution of an action sequence (via updates by update models) since knowledge is reducible to true belief in $\mathbf{KD45}_n$ models. On the one hand, the result opens the door for multi-agent planning systems proposed in the literature to attack planning problems which require the manipulation of not only knowledge but also beliefs of agents. On the other hand, this result can also be used as a guideline for the development of update models in practical applications. We related our work to others, elaborated an issue faced by our formalization, and identified problems that need to be investigated further. In addition to the mentioned issues, we intend to extend the result of this paper to other types of actions (e.g., lying, misleading, announcement of knowledge formula).

References

Baltag, A., and Moss, L. 2004. Logics for epistemic programs. *Synthese*.

Baltag, A., and Smets, S. 2008. A qualitative theory of dynamic interactive belief revision. In *Proc. of 7th LOFT*, Texts in Logic and Games 3, 13–60. Amsterdam University Press.

Baltag, A.; Moss, L.; and Solecki, S. 1998. The logic of public announcements, common knowledge, and private suspicions. In *7th TARK*, 43–56.

Baral, C.; Gelfond, G.; Pontelli, E.; and Son, T. C. 2013. Reasoning about the beliefs of agents in multi-agent domains in the presence of state constraints: The action language mal. In Leite, J.; Son, T. C.; Torroni, P.; van deer Torre, L.; and Woltran, S., eds., *Proceedings of the 14th International Workshop, Computational Logic in Multi-Agent Systems, CLIMA VIX, Coruna, Spain, September 16-18, 2013*, volume 8143 of *Lecture Notes in Computer Science*, 290–306. Springer.

Bolander, T., and Andersen, M. 2011. Epistemic Planning for Single and Multi-Agent Systems. *Journal of Applied Non-Classical Logics* 21(1).

Halpern, J.; Samet, D.; and Segev, E. 2009. Defining knowledge in terms of belief: The modal logic perspective. *Reviews of Symbolic Logic* 2(3):469–486.

Herzig, A.; Lang, J.; and Marquis, P. 2005. Action Progression and Revision in Multiagent Belief Structures. In *Sixth Workshop on Nonmonotonic Reasoning, Action, and Change (NRAC)*.

Löwe, B.; Pacuit, E.; and Witzel, A. 2011. Del planning and some tractable cases. In van Ditmarsch, H.; Lang, J.; and Ju, S., eds., *Logic, Rationality, and Interaction*, volume 6953 of *Lecture Notes in Computer Science*. Springer Berlin / Heidelberg. 179–192.

Scherl, R., and Levesque, H. 2003. Knowledge, action, and the frame problem. *Artificial Intelligence* 144(1-2).

Segerberg, K. 1998. Irrevocable belief revision in dynamic doxastic logic. *Notre Dame Journal of Formal Logic* 39(3):287–306.

Son, T. C., and Baral, C. 2001. Formalizing sensing actions - a transition function based approach. *Artificial Intelligence* 125(1-2):19–91.

van Benthem, J.; van Eijck, J.; and Kooi, B. P. 2006. Logics of communication and change. *Inf. Comput.* 204(11):1620–1662.

van Benthem, J. 2007. Dynamic logic of belief revision. *Journal of Applied Non-Classical Logics* 17(2):129–155.

van der Hoek, W., and Wooldridge, M. 2002. Tractable multiagent planning for epistemic goals. In *The First International Joint Conference on Autonomous Agents & Multiagent Systems, AAMAS 2002, July 15-19, 2002, Bologna, Italy, Proceedings*, 1167–1174. ACM.

van Ditmarsch, H.; van der Hoek, W.; and Kooi, B. 2007. *Dynamic Epistemic Logic*. Springer.

van Ditmarsch, H. 2005. Prolegomena to dynamic logic for belief revision. *Synthese (Knowledge, Rationality & Action)* 147:229–275.