On the Semantics of Knowledge Update

Chitta Baral
Dept of Computer Science & Engineering
Arizona State University
Tempe, AZ 85287, USA
E-mail: chitta@asu.edu

Yan Zhang
School of Computing & IT
University of Western Sydney
Penrith South DC, NSW 1797, Australia
E-mail: yan@cit.uws.edu.au

Abstract

We consider the problem of how an agent's knowledge can be updated. We propose a formal method of knowledge update on the basis of the semantics of modal logic S5. In our method, an update is specified according to the minimal change on both the agent's actual world and knowledge. We then investigate the semantics of knowledge update and characterize several specific forms of knowledge update which have important applications in reasoning about change of agents' knowledge. We also discuss the persistence property of knowledge and ignorance associated with knowledge update.

1 Introduction and Motivation

The well-studied issues of belief updates and belief revision [Katsuno and Mendelzon, 1991] are concerned with the update and revision aspects of an agent's belief with respect to new beliefs. The notion of belief update has been used, and often serves as a guideline [Kartha and Lifschitz, 1994], in reasoning about the effect of (*world altering*) actions on the state of the world. Thus if ϕ represents the agent's belief about the world and the agent does an action that is supposed to make ψ true in the resulting world, then the agent's belief about the resulting world can be described by $\phi \diamond \psi$, where $\diamond$ is the update operator of choice.

Now let us consider reasoning about sensing actions [Scherl and Levesque, 1993; Son and Baral, 2000], which in their pure form, when executed, do not change the world, but change the agent's knowledge about the world. Let $sense_f$ be a sensing action whose effect is that after it is executed the agent knows whether f is true or not. This can be expressed as $Kf \vee K\neg f$, where K is the modal operator $Knows$. The current theory of belief updates does not tell us how to do updates with respect to such gain in knowledge due to a sensing action. (Note that we can not just have $\psi \equiv f \vee \neg f$ and use the the notion of belief update, as $f \vee \neg f$ is a tautology). One of our *goals* in this paper is to define a notion of *knowledge update*, analogous to belief update, where the original theory (ϕ) and the new theory (ψ) are in a language that can express knowledge. Such a notion would not only serve as a guideline to reason about pure and mixed sensing actions in presence of constraints, but also allow us to reason about actions corresponding to *forgetting*, and *ignorance*.

The structure of the rest of the paper is as follows. In Section 2 we start with describing the particular modal logic that we plan to use in expressing knowledge, and describe the notion of k-models analogous to 'models' in classical logic. We then define closeness between k-models and use it to define a particular notion of knowledge update. In Section 3 we present alternative characterizations of four particular knowledge update cases – *gaining knowledge, ignorance, sensing, and forgetting*, and show their equivalence to our original notion of knowledge update. Some of these alternative characterizations are based on the formulation of reasoning about sensing actions, and thus our equivalence results serve as justification of the intuitiveness of our definition of knowledge update. In Section 4 we explore sufficiency conditions that guarantee persistence of knowledge (or ignorance) during a knowledge update. In Section 5, we conclude and discuss future directions.

2 Closeness between k-models and Knowledge Update

In this section, we describe formal definitions for knowledge update. Our formalization will be based on the semantics of the propositional modal logic S5 with a single agent. In general, under Kripke semantics, a *Kripke structure* is a triple (W, R, π), where W is a set of possible worlds, R is an equivalence relation on W, and π is a truth assignment function that assigns a propositional valuation to each world in W. Given a Kripke structure $S = (W, R, \pi)$, a *Kripke interpretation* is a pair $M = (S, w)$, where $w \in W$ is referred to the *actual world* of M. The entailment relation $\models$ between Kripke interpretations and formulas is defined to provide semantics for formulas of S5 [Fagin and et al, 1995].

In the case of single agent, however, we may restrict ourselves to those S5 structures in which the relation R is universal, i.e. each world is accessible from every world, and worlds are identified with the set of atoms true at the worlds [Meyer and van der Hoek, 1995]. To simplify a comparison between two worlds (e.g. Definition 2), we may view an atom $p \in w$ iff $w \models p$. Therefore, in our context a Kripke structure (W, R, π) is uniquely characterized by W and we may simplify a Kripke interpretation as a pair (W, w) which we call a *k-model*, where w indicates the actual world of the agent and W presents all possible worlds that the agent may access. Note that w is in W for any k-model (W, w).

In our following description, we use $a, b, c, p, \cdots$ to denote primitive propositional atoms; $\phi, \psi, \upsilon, \cdots$ to denote propositional formulas without including modalities (we also call them *objective* formulas); and $\alpha, \beta, \mu, \cdots$ and T to denote formulas that may contain modal operator K. For convenience, we also use $T \equiv \alpha_1 \wedge \cdots \wedge \alpha_k$ to represent a finite set of formulas $\{\alpha_1, \cdots, \alpha_k\}$ and call T a *(knowledge) set*.

Definition 1 *(***S5 Semantics***) Let $\mathcal{P}$ be the set of all primitive propositions in the language. The entailment relation $\models$ under normal S5 semantics is defined as follows:*

1. *$(W, w) \models p$ iff p is primitive (i.e. $p \in \mathcal{P}$) and $w \models p$;*
2. *$(W, w) \models \alpha \wedge \beta$ iff $(W, w) \models \alpha$ and $(W, w) \models \beta$;*
3. *$(W, w) \models \neg\alpha$ iff it is not the case that $(W, w) \models \alpha$;*
4. *$(W, w) \models K\alpha$ iff $(W, w') \models \alpha$ for all $w' \in W$.*

Given a formula T, $M = (W, w)$ is called a k-model of T if $M \models T$. We use $KMod(T)$ to denote the set of all k-models of T. For objective formulas ϕ, $Mod(\phi)$ denotes the set of models of ϕ, and $w \models \phi$ denotes that w is a *model* of ϕ. For a formula α, we say that T *entails* α, denoted as $T \models \alpha$, if for every k-model M of T, $M \models \alpha$.

Now the basic problem of knowledge update that we would like to investigate is formally described as follows: given a *k-model* $M = (W, w)$, that is usually viewed as a *knowledge state* of an agent, and a formula μ - the agent's new knowledge that may contain modal operator K, how do we update M to another k-model $M' = (W', w')$ such that $M' \models \mu$ and M' is *minimally different* from M with respect to some criterion. To approach this problem, we first need to provide a definition of *closeness* between two k-models with respect to a given k-model.

Definition 2 *(***k-model Closeness***) Let $M = (W, w)$, $M_1 = (W_1, w_1)$ and $M_2 = (W_2, w_2)$ be three k-models. We say M_1 is* as close to M as *M_2, denoted as $M_1 \leq_M M_2$, if:*

1. *$(w_1 \setminus w \cup w \setminus w_1) \subseteq (w_2 \setminus w \cup w \setminus w_2)$; or*
2. *$w_1 = w_2$ and one of the following conditions holds:*
   - *(i) if $W \subseteq W_1$, then (a) there exist some ϕ and ψ such that $M \models K\phi$ and $M_2 \not\models K\phi$ and $M \not\models K\psi$ and $M_2 \models K\psi$, or (b) for any ϕ if $M \models K\phi$ and $M_1 \not\models K\phi$, then $M_2 \not\models K\phi$;*
   - *(ii) if $W_1 \subset W$, then condition (a) above is satisfied, or (c) for any ϕ if $M \not\models K\phi$ and $M_1 \models K\phi$, then $M_2 \models K\phi$;*
   - *(iii) if $W \not\subset W_1$ and $W_1 \not\subset W$, then conditions (b) and (c) above are satisfied.*

We denote $M_1 <_M M_2$ if $M_1 \leq_M M_2$ and $M_2 \not\leq_M M_1$.

In the above definition, condition 1 simply says that the symmetric differences between w and w_1 is not bigger than that between w and w_2, while in condition 2, (i), (ii) and (iii) express that different preferences are applied to compare knowledge between M_1 and M_2 with respect to M. For convenience, given a k-model M, if we denote $KM = \{\phi \mid \text{for all } w \in W, w \models \phi\}$, then (a) is equivalent to $KM \setminus KM_2 \neq \emptyset$ and $KM_2 \setminus KM \neq \emptyset$; (b) is equivalent to $KM \setminus KM_1 \subseteq KM \setminus KM_2$; and (c) is equivalent to $KM_1 \setminus KM \subseteq KM_2 \setminus KM$. Also (b) and (c) together present a difference on both knowledge decrease and increase between M_1 and M_2 in terms of M. It is also easy to see that $\leq_M$ is a partial ordering.

Note that during the comparison between two k-models, we give preference to the change of the actual world over the change of the knowledge about the world. (The closeness criterion between actual worlds that we use is the commonly used criterion [Winslett, 1988] based on symmetric difference). For instance, if the actual world of a k-model M_1 is closer to the actual world of M than the actual world of another k-model M_2, we will think M_1 is closer to M and the comparison of knowledge between M_1 and M_2 is ignored. Only when both M_1 and M_2 have the same actual world, we will compare the knowledge of M_1 and M_2 in terms of M. This seems to be intuitive to us. In fact, the comparison between actual worlds determines the *actual distance* between two k-models to the given k-model M. If M_1 and M_2 have the same actual distance to M, the *knowledge distance* is then taken into account.

Basically, condition (i) (or (ii) resp.) in Definition 2 defines a knowledge preference based on knowledge decrease (or increase resp). That is, if M_1 *only* looses knowledge from M (or *only* gains some knowledge to M, resp.), then M_1 is preferred over those k-models that have both knowledge decrease and increase from M, i.e. (a), and also preferred over those k-models that only loose more knowledge from M (or add more knowledge to M, resp.) than M_1 does, i.e. (b) or (c) respectively. Condition (iii), on the other hand, deals with the mixed situation that M_1 has both knowledge decrease and increase from M. In this case, a combined difference on knowledge decrease and increase is applied to determine the knowledge distance, i.e. (b) *and* (c). Conditions (i), (ii) and (iii) can be illustrated by the following figures respectively.

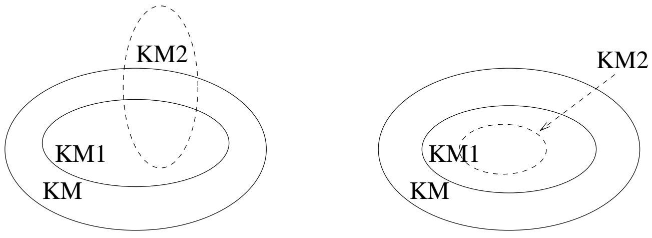

Figure 1: $M_1 \leq_M M_2$ under the condition $w_1 = w_2$ and $W \subseteq W_1$: (a) or (b) holds.

Definition 3 *(***k-model Update***) Let $M = (W, w)$ be a k-model and μ a formula. A k-model $M' = (W', w')$ is called a* possible resulting k-model *after updating M with μ if and only if the following conditions hold:*

1. *$M' \models \mu$;*
2. *there does not exist another k-model $M'' = (W'', w'')$ such that $M'' \models \mu$ and $M'' <_M M'$.*

We denote the set of all possible resulting k-models after updating M with μ as $Res(M, \mu)$.

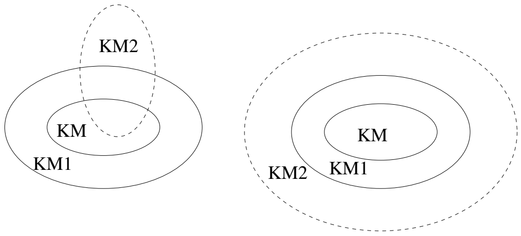

Figure 2: $M_1 \leq_M M_2$ under the condition $w_1 = w_2$ and $W_1 \subset W$: (a) or (c) holds.

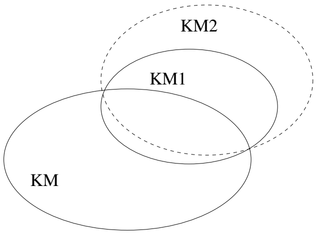

Figure 3: $M_1 \leq_M M_2$ under the condition $w_1 = w_2$, $W \not\subset W_1$ and $W_1 \not\subset W$: (b) and (c) hold.

Example 1 Let $T \equiv Kc \wedge \neg Ka \wedge \neg Kb \wedge K(a \vee b)$ and $\mu \equiv K\neg c$. We denote

$$w_0 = \{a, b, c\}, w_1 = \{a, c\}, w_2 = \{b, c\},$$
$$w_3 = \{c\}, w_4 = \{a, b\}, w_5 = \{a\},$$
$$w_6 = \{b\}, w_7 = \emptyset.$$

Clearly, $M_0 = (\{w_0, w_1, w_2\}, w_0)$ is a k-model of T. Consider the update of M_0 with μ. Let $M_1 = (\{w_4, w_5, w_6\}, w_4)$. Now we show that M_1 is a possible resulting k-model after updating M_0 with μ.

Since $(w_0 \setminus w_4 \cup w_4 \setminus w_0) = \{c\}$, we first consider any possible k-model $M' = (W', w')$ such that $(w_0 \setminus w' \cup w' \setminus w_0) \subset \{c\}$. Clearly, the only possible w' would be w_0 itself. Let $M' = (W', w_0)$, where W' is a subset of $\{w_0, \cdots, w_7\}$. However, since $c \in w_0$, there does not exist any W' such that $M' \models K\neg c$. Therefore, from Definition 2, only condition 2 can be used to find a possible M' such that $M' <_M M_1$. So we assume $M' = (W', w_4)$. On the other hand, from M_0 and M_1, it is easy to see that $KM_0 = \{c, a \vee b\}$ and $KM_1 = \{\neg c, a \vee b\}$[1]. Then we have $KM_0 \setminus KM_1 = \{c\}$ and $KM_1 \setminus KM_0 = \{\neg c\}$. Ignoring the detailed verifications, we can show that there does not exist such $M' = (W', w_4)$ satisfying $KM_0 \setminus KM' = KM' \setminus KM_0 = \emptyset$. ■

Based on the k-model update, updating a formula (knowledge set) T in terms of another formula μ is then achieved by updating every k-model of T with μ.

[1]For simplicity, here we only consider the *prime* formulas ϕ in KM in the sense that if $\phi \in KM$, then there is not another ψ such that $\models \psi \supset \phi$ and $\psi \in KM$.

Definition 4 *(**Knowledge Update**) Let T and μ be two formulas. The* update of T with μ*, denoted as $T \diamond \mu$, is defined by $KMod(T \diamond \mu) = \bigcup_{M \in KMod(T)} Res(M, \mu)$.*

3 Characterizing Knowledge Update

In this section, we first investigate basic properties of k-models, then provide alternative characterizations for different types of knowledge update and then show their equivalence with respect to our original characterization.

Proposition 1 *Let $M_1 = (W_1, w_1)$ and $M_2 = (W_2, w_2)$ be two k-models. Then the following properties hold:*

1. *$\phi \in KM_1$ iff $W_1 \subseteq Mod(\phi)$;*
2. *$W_1 \subseteq W_2$ iff $KM_2 \subseteq KM_1$;*
3. *$KM_1 = KM_2$ iff $W_1 = W_2$;*
4. *Let $M' = (W_1 \cup W_2, w')$, then $KM' = KM_1 \cap KM_2$;*
5. *Let $w' \in W_1 \cap W_2$ and $M' = (W_1 \cap W_2, w')$, then $KM_1 \cup KM_2 \subseteq KM'$.*

Given a set of k-models $\mathcal{S}$ and a k-model M, let $\leq_M$ be an ordering on $\mathcal{S}$ as we defined in Definition 2. By $Min(\mathcal{S}, \leq_M)$ we mean the set of all elements in $\mathcal{S}$ that are minimal with respect to ordering $\leq_M$. The following theorem shows that our knowledge update can be characterized by a minimal change criterion based on an ordering on k-models.

Theorem 1 *Let T and μ be two formulas. Then $KMod(T \diamond \mu) = \bigcup_{M \in KMod(T)} Min(KMod(\mu), \leq_M)$.*

Proof: To prove the result, we only need to show that for each k-model M of T, $Res(M, \mu) = Min(KMod(\mu), \leq_M)$. Let $M' \in Res(M, \mu)$. Since $M' \models \mu$, $M' \in KMod(\mu)$. On the other hand, according to Definition 3, for any $M'' \in KMod(\mu)$, we have $M' \leq_M M''$. That is, $M' \in Min(KMod(\mu), \leq_M)$. So $Res(M, \mu) \subseteq Min(KMod(\mu), \leq_M)$. Similarly, we can show $Min(KMod(\mu), \leq_M) \subseteq Res(M, \mu)$. ■

While the above theorem presents a general minimal change property of our knowledge update, we now give alternative characterizations of the update of T with μ when μ has some specific forms. These specific forms present some of the important features of knowledge update, and their alternative characterization is handy when the use of the notion of knowledge update becomes an overkill. For example, the alternative characterization of *sensing update* below is a much simpler characterization that is used in reasoning about sensing actions [Scherl and Levesque, 1993; Son and Baral, 2000]. We now introduce a notation that will be useful in our following discussions. Let W be a set of worlds and $w \in W$. By $W^{(w,\phi)}$, we denote the set $\{w' \mid w' \in W$ and $w' \models \phi$ iff $w \models \phi\}$.

Proposition 2 *Given T and $K\phi$ where ϕ is objective and $T \models \phi$. Then $M' = (W', w')$ is a k-model of $T \diamond K\phi$ if and only if there exists a k-model $M = (W, w)$ of T such that $w = w'$ and $W' = W^{(w,\phi)}$.*

The above proposition reveals an important property about knowledge update: to know some fact, the agent only needs to restrict the current possible worlds in each of her k-models, if this fact itself is already entailed by her current knowledge set. We call this kind of knowledge update *gaining knowledge update*.

Example 2 Let $T \equiv a \wedge \neg Ka \wedge Kb$. Suppose $w_0 = \{a, b\}$, $w_1 = \{a\}$, $w_2 = \{b\}$ and $w_3 = \emptyset$. Then T has one k-model $M = (\{w_0, w_2\}, w_0)$. Updating M with Ka, according to our k-model update definition, we have a unique resulting k-model $M' = (\{w_0\}, w_0)$. Indeed, this result is also obtained from Proposition 2. ■

As a contrary case to the gaining knowledge update, we now characterize an agent ignoring a fact from her knowledge set which we call *ignorance update*, i.e. updating T with $\neg K\phi$. From Definition 1, it is easy to see that $T \diamond \neg\phi \models \neg K\phi$. However, it should be noted that updating $T \diamond \neg\phi$ can *not* be used to achieve $T \diamond \neg K\phi$. Consider a k-model $M = (\{\{a, b\}, \{a\}\}, \{a, b\})$. Updating M with $\neg Ka$ we have a possible resulting k-model $M' = (\{\{a, b\}, \{a\}, \{b\}\}, \{a, b\})$, while updating M with $\neg a$ will lead to a possible result $M'' = (\{\{a, b\}, \{a\}, \{b\}\}, \{b\})$. Note that both M' and M'' entail $\neg Ka$, but $M' <_M M''$ according to Definition 2.

Proposition 3 *Given T and $\mu \equiv \neg K\phi$ where ϕ is objective. $M' = (W', w')$ is a k-model of $T \diamond \neg K\phi$ if and only if there exists a k-model $M = (W, w)$ of T such that*

(i) if $M \models K\phi$, then $w' = w$ and $W' = W \cup \{w^*\}$, where $w^* \models \neg\phi$;

(ii) otherwise, $w' = w$ and $W' = W$.

Example 3 Suppose $T \equiv \neg Ka \wedge \neg Kb \wedge K(a \vee b) \wedge Kc$ and the agent wants to ignore c. Let $w_0 = \{a, b, c\}$, $w_1 = \{a, c\}$, $w_2 = \{b, c\}$, $w_3 = \{c\}$, $w_4 = \{a, b\}$, $w_5 = \{a\}$, $w_6 = \{b\}$, $w_7 = \emptyset$. Clearly, T has three k-models: $M_0 = (\{w_0, w_1, w_2\}, w_0)$, $M_1 = (\{w_0, w_1, w_2\}, w_1)$, and $M_2 = (\{w_0, w_1, w_2\}, w_2)$. From Proposition 3, $T \diamond \neg Kc$ has the following twelve k-models: $(\{w_0, w_1, w_2, w_i\}, w_j)$, where $i = 4, 5, 6, 7$ and $j = 0, 1, 2$. ■

Now we consider the case when μ is of the form $K\phi \vee K\neg\phi$ where ϕ is objective. Updating T with this type of μ is particularly useful in reasoning about sensing actions [Scherl and Levesque, 1993; Son and Baral, 2000] where $K\phi \vee K\neg\phi$ represents the effect of a sensing action after whose execution, the agent will know either ϕ or its negation. We refer to such an update as a *sensing update*. The following proposition characterizes the update of T with a formula of the form $K\phi \vee K\neg\phi$. It is interesting to note that the sufficient and necessary condition for a k-model of $T \diamond (K\phi \vee K\neg\phi)$ is similar to the one presented in Proposition 2.

Proposition 4 *Given T and $\mu \equiv K\phi \vee K\neg\phi$ where ϕ is objective. $M' = (W', w')$ is a k-model of $T \diamond (K\phi \vee K\neg\phi)$ if and only if there exists a k-model $M = (W, w)$ of T such that $w = w'$ and $W' = W^{(w,\phi)}$ or $w = w'$ and $W' = W^{(w,\neg\phi)}$*

Proof: Let $\mu \equiv K\phi \vee K\neg\phi$ and $M' = (W^{(w,\phi)}, w)$. From the definition of $W^{(w,\phi)}$, it is easy to see that $M' \models \mu$. Consider any $M'' = (W'', w'')$ where $M'' \models \mu$ and $w'' \neq w$. According to the condition 1 of Definition 2, $M' <_M M''$. So M'' can not be a k-model of $T \diamond \mu$. In other words, each k-model of $T \diamond \mu$ must have a form of $M'' = (W'', w)$. Then from Theorem 1, to prove the result, it is sufficient to prove for k-model $M'' = (W'', w)$ where $W'' \neq W^{(w,\phi)}$ or $W'' \neq W^{(w,\neg\phi)}$, $M' \leq_M M''$. This can be shown in the same way as in the proof of Proposition 2. ■

Example 4 Suppose $T \equiv Kb \wedge \neg Ka \wedge \neg K\neg a$ represents the current knowledge of an agent. Note that T implies that the agent does not have any knowledge about a. Consider the update of T with $\mu \equiv Ka \vee K\neg a$ which can be thought of as the agent trying to reason – in the planning or plan verification stage – about a sensing action[2] that will give her the knowledge about a. Let $w_0 = \{a, b\}$, $w_1 = \{b\}$, $w_2 = \{a\}$ and $w_3 = \emptyset$. It is easy to see that $M_0 = (\{w_0, w_1\}, w_0)$ and $M_1 = (\{w_0, w_1\}, w_1)$ are two k-models of T. Then according to the above proposition, it is obtained that $M'_0 = (\{w_0\}, w_0)$ and $M'_1 = (\{w_1\}, w_1)$ are the two k-models of $T \diamond \mu$. ■

As another important type of knowledge update, we consider the update of T with $\mu \equiv \neg K\phi \wedge \neg K\neg\phi$. This update can be thought of as the result of an agent *forgetting* her knowledge about the fact ϕ. We will refer to such an update as a *forgetting update*. The following proposition shows that in order to forget ϕ from T, for each k-model of the current knowledge set, the agent only needs to expand the set of possible worlds of this model with exactly *one specific* world.

Proposition 5 *Given T and $\mu \equiv \neg K\phi \wedge \neg K\neg\phi$ where ϕ is objective. $M' = (W', w')$ is a k-model of $T \diamond \mu$ if and only if there exists a k-model $M = (W, w)$ of T such that*

(i) if $M \models K\phi$, then $w' = w$ and $W' = W \cup \{w^*\}$, where $w^* \models \neg\phi$;

(ii) if $M \models K\neg\phi$, then $w' = w$ and $W' = W \cup \{w^*\}$, where $w^* \models \phi$;

(iii) otherwise, $w' = w$ and $W' = W$.

Proof: Let $\mu \equiv \neg K\phi \wedge \neg K\neg\phi$ and $M = (W, w)$ be a k-model of T. Then it is easy to see that for two k-models $M' = (W', w')$ and $M'' = (W'', w'')$ such that $M' \models \mu$ and $M'' \models \mu$, and $w' = w$ and $w'' \neq w$, $M' <_M M''$. So M'' can not be a k-model of $T \diamond \mu$. In other words, a k-model of $T \diamond \mu$ must have a form $M' = (W', w)$.

From Theorem 1, to prove the result, we only need to show that for any k-model $M'' = (W'', w)$ such that $M'' \models \mu$ and $W'' \neq W \cup \{w^*\}$, $M' \leq_M M''$.

Let $M = (W \cup \{w^*\}, w)$, where $w^* \models \neg\phi$ if $M \models K\phi$ and $w^* \models \phi$ if $M \models K\neg\phi$. We first prove that for any k-model $M'' = (W'', w)$ such that $M'' \models \mu$ and W'' does not have a form of $W \cup \{w_i\}$, $M' \leq_M M''$.

Suppose $M \models K\phi$. Clearly $M' \models \mu$.

Case 1. Consider a k-model $M'' = (W'', w)$ where $W' =$

[2] Such reasoning is necessary in creating plans with sensing actions or verifying such plans. On the other hand after the execution of a sensing action the agent exactly knows either a or $\neg a$, and can simply use the notion of belief update.

$W \cup \{w^*\} \subseteq W''$. Note that $M'' \models \mu$ as well. However, from Proposition 1, we have $KM'' \subseteq KM' \subseteq KM$. So $M' \leq_M M''$ according to Definition 2 (i.e. condition (b)).

Case 2. Suppose $W'' \subset W'$ (proper set inclusion). Without loss of generality, we assume that $W'' = W \cup \{w^*\} \setminus \{w_j\}$ where $w_j \in W$. This follows that $W \not\subseteq W''$ and $W'' \not\subseteq W$. So it is the case that $KM \setminus KM'' \neq \emptyset$ and $KM'' \setminus KM \neq \emptyset$. On the other hand, we have $W \subseteq W'$, from Definition 2 (i.e. condition (a)), we have $M' \leq_M M''$.

Case 3. Now suppose $W'' \not\subseteq W'$ and $W' \not\subseteq W'$. Without loss of generality, we can assume that $W'' = W \cup \{w^*, w_i\} \setminus \{w_j\}$, where $w_j \in W$ and $w^*, w_j \notin W$. Again, this results to the situation that $W \not\subseteq W''$ and $W'' \not\subseteq W$. ¿From the above discussion, it implies that $M' \leq_M M''$.

Following the same way as above, we can prove that under the condition that $M \models K\neg\phi$ and $M' = (W \cup \{w^*\}, w)$ where $w^* \models \phi$, for any k-model $M'' = (W'', w)$ such that $M'' \models \mu$ and W'' does not have a form of $W \cup \{w_j\}$ $M' \leq_M M''$.

Now we show that for any k-model M'' that is of the form $M'' = (W \cup \{w_i\}, w)$ and w_i is any world such that $w_i \models \neg\phi$ if $M \models K\phi$ or $w_i \models \phi$ if $M \models K\neg\phi$, $M' \not\leq_M M''$ and $M'' \not\leq_M M'$. Suppose $M' \leq_M M''$. Since $W \subseteq W \cup \{w^*\}$, then according to Definition 2, condition (a) or (b) should be satisfied. As $W \subseteq W \cup \{w_i\}$, condition (a) can not be satisfied. So condition (b) must be satisfied. That is, for any ψ such that $M \models K\psi$ and $M' \not\models K\psi$, $M'' \models K\psi$. However, this implies that $KM \setminus KM' \subseteq KM \setminus KM''$, and also $KM \cap KM'' \subseteq KM \cap KM'$. From Proposition 1 (Results 2 and 4), it follows that $W \cup W' \subseteq W \cup W''$, that is, $W \cup \{w^*\} \subseteq W \cup \{w_i\}$. Obviously, this is not true. Similarly, we can show that $M'' \not\leq_M M'$. That means, both M' and M'' are in $Res(M, \mu)$. This completes our proof. ■

Example 5 Suppose $T \equiv Kb \wedge (Ka \vee K\neg a)$ represents the current knowledge of an agent. After executing a forgetting action the agent now would like to update her knowledge with $\mu \equiv \neg Ka \wedge \neg K\neg a$. Let $w_0 = \{a, b\}, w_1 = \{b\}, w_2 = \{a\}, w_3 = \emptyset$. It is easy to see that $M_0 = (\{w_0\}, w_0)$ and $M_1 = (\{w_1\}, w_1)$ are the two k-models of T. Then using Proposition 5, we conclude that $M_0' = (\{w_0, w_1\}, w_0)$, $M_1' = (\{w_0, w_3\}, w_0)$, $M_2' = (\{w_1, w_0\}, w_1)$, and $M_3' = (\{w_1, w_2\}, w_1)$ are the four k-models of $T \diamond \mu$. Note that $(\{w_0, w_2\}, w_0)$ cannot be a k-model of $T \diamond \mu$ according to Proposition 5. ■

4 Persistence of Knowledge and Ignorance

Like most systems that do dynamic modeling, the knowledge update discussed previously is non-monotonic in the sense that while adding new knowledge into a knowledge set, some previous knowledge in the set might be lost. However, it is important to investigate classes of formulas that are persistent with respect to an update, as this may partially simplify the underlying inference problem [Engelfreit, 1998; Zhang, 1999]. Given T and μ, a formula α is said to be *persistent* with respect to the update of T with μ, if $T \models \alpha$ implies $T \diamond \mu \models \alpha$. If α is of the form $K\phi$, we call this persistence as *knowledge persistence*, while if α is of the form $\neg K\phi$, we call it *ignorance persistence*. The question that we address now is that under what conditions, a formula α is persistent with respect to the update of T with μ.

As the update of T with μ is achieved based on the update of every k-model of T with μ, our task reduces to the study of persistence with respect to a k-model update. This is defined in the following definition.

Definition 5 *(***Persistence with respect to k-model Update***) Let μ and α be two formulas and M be a k-model. α is* persistent *with respect to the update of M with μ if for any $M' \in Res(M, \mu)$, $M \models \alpha$ implies $M' \models \alpha$.*

Clearly a formula α is persistent with respect to the update of T with μ if and only if for each k-model M of T, α is persistent with respect to the update of M with μ. To characterize the persistence property with respect to k-model updates, we first define a preference ordering on k-models in terms of a formula.

Definition 6 *(***Formula Closeness***) Let μ be a formula and M_1 and M_2 be two k-models. We say M_1 is* as close *to μ as M_2, denoted as $M_1 \leq_\mu M_2$, if one of the following conditions holds:*

1. *$M_1 \in KMod(\mu)$;*
2. *if $M_1, M_2 \notin KMod(\mu)$, then for any $M \in KMod(\mu)$, $M_1 \leq_M M_2$.*

We denote $M_1 <_\mu M_2$ if $M_1 \leq_\mu M_2$ and $M_2 \not\leq_\mu M_1$.

Intuitively, the above definition specifies a partial ordering to measure the closeness between two k-models to a formula. In particular, if M_1 is a k-model of μ, then M_1 is closer to μ than all other k-models (i.e. condition 1). If neither M_1 nor M_2 is a k-model of μ, then the comparison between M_1 and M_2 with respect to μ is defined based on the k-model preference ordering $\leq_M$ for each k-model M of μ (i.e. condition 2). Note that if both M_1 and M_2 are k-models of μ, we have $M_1 \leq_\mu M_2$ and $M_2 \leq_\mu M_1$, and both of them are equally close to μ.

Example 6 Let $\mu \equiv Ka \wedge Kb$, $w_0 = \{a, b\}$, $w_1 = \{b\}$, $w_2 = \{a\}$ and $w_3 = \emptyset$. Clearly, μ has one k-model $M = (\{w_0\}, w_0)$. Consider two k-models $M_1 = (\{w_0, w_1\}, w_0)$ and $M_2 = (\{w_1, w_2\}, w_1)$. Now let us compare which one of them is closer to μ. Since neither M_1 nor M_2 is a k-model of μ, we can use condition 2 in Definition 6 to compare M_1 and M_2. According to Definition 2, it is easy to see that $M_1 \leq_M M_2$ as $w_0 \setminus w_1 \cup w_1 \setminus w_0 = \{a\} \neq \emptyset$. Therefore, we conclude $M_1 \leq_\mu M_2$. Furthermore, we also have $M_1 <_\mu M_2$. ■

Proposition 6 *Let μ be a formula. For any two k-models M_1 and M_2, if $M_1 \leq_\mu M_2$, then $M_2 \models \mu$ implies $M_1 \models \mu$.*

Proof: Suppose $M_2 \models \mu$. Then $M_2 \in KMod(\mu)$. From Definition 6, we know that for any other k-model M', $M_2 \leq_\mu M'$. So $M_2 \leq_\mu M_1$. But we have $M_1 \leq_\mu M_2$. This implies that both M_1 and M_2 are equally close to μ. Hence, $M_1 \models \mu$. ■

Given a formula μ and a sequence of k-models $M_1, \cdots, M_k$, if the relation $M_1 \leq_\mu M_2 \leq_\mu \cdots \leq_\mu M_k$ holds, then it means that M_i is closer to μ than M_j, where

$i < j$. Now under this condition, if there is another formula α which satisfies the property that $M_j \models \alpha$ implies $M_i \models \alpha$ whenever $i < j$, we say that formula α is *persistent* with respect to formula μ. In other words, when k-models move closer to μ, α's truth value is preserved in these k-models. The following definition formalizes this idea.

Definition 7 *(* **$\leq_\mu$-persistence** *) Let α, μ be two formulas. We say that α is* $\leq_\mu$-persistent *if for any two k-models M_1 and M_2, $M_2 \models \alpha$ and $M_1 \leq_\mu M_2$ implies $M_1 \models \alpha$.*

Now we have the following important relationship between $\leq_\mu$-persistence and k-model update persistence.

Theorem 2 *Let α and μ be two formulas and M be a k-model. α is persistent with respect to the update of M with μ if α is $\leq_\mu$-persistent.*

Proof: Let M' be a k-model in $Res(M, \mu)$. Then we have $M' \in KMod(\mu)$. So for any k-model M'', we have $M' \leq_\mu M''$. So $M' \leq_\mu M$. Now suppose α is μ-persistent. It follows that $M \models \alpha$ implies $M' \models \alpha$. As M' is an arbitrary k-model in $Res(M, \mu)$, we can conclude that α is persistent with respect to the update of M with μ. ■

From Theorem 2, we have that $\leq_\mu$-persistence is a sufficient condition to guarantee a formula persistence with respect to a k-model update. As will be shown next, we can provide a unique characterization for μ-persistence. We first define the notion of ordering preservation as follows.

Definition 8 *(* **Ordering Preservation** *) Given two formulas α and β. We say that ordering $\leq_\alpha$* preserves *ordering $\leq_\beta$ if for any two k-models M_1 and M_2, $M_1 \leq_\alpha M_2$ implies $M_1 \leq_\beta M_2$.*

The intuition behind ordering preservation is clear. That is, if $\leq_\alpha$ preserves $\leq_\beta$, then for any two k-models M_1 and M_2, whenever M_1 is closer to α than M_2, M_1 will be closer to β than M_2 as well. Finally, we have the following important result to characterize μ-persistence.

Theorem 3 *Given two formulas α and μ, α is $\leq_\mu$-persistent if and only if $\leq_\mu$ preserves $\leq_\alpha$.*

Proof: ($\Rightarrow$) Suppose α is $\leq_\mu$-persistent. That is, for any two k-models M_1 and M_2, $M_1 \leq_\mu M_2$ and $M_2 \models \alpha$ implies $M_1 \models \alpha$. So under the constraint that α is μ-persistent, whenever $M_1 \leq_\mu M_2$, we have $M_1 \leq_\alpha M_2$. That means, $\leq_\mu$ preserves $\leq_\alpha$.

($\Leftarrow$) Suppose $\leq_\mu$ preserves $\leq_\alpha$. From Definition 8, we have that for any two k-models M_1 and M_2, $M_1 \leq_\mu M_2$ implies $M_1 \leq_\alpha M_2$. Now suppose $M_1 \leq_\mu M_2$. So we have $M_1 \leq_\alpha M_2$. From Proposition 6, we have that $M_2 \models \alpha$ implies $M_1 \models \alpha$. From this it follows that α is $\leq_\mu$-persistent. ■

5 Conclusion and Future Directions

In this paper we developed the notion of knowledge update – an analogous notion to belief update, that is useful in characterizing the knowledge change of an agent in presence of new knowledge. Our notion is particularly relevant in reasoning about actions and plan verifications when there are sensing actions. We presented simpler alternative characterization of knowledge update for particular cases, and showed its equivalence to the original characterization. Finally we discussed when particular knowledge (or ignorance) persists with respect to a knowledge update.

We believe our work here to be a starting point on knowledge update, and as evident from the research in belief update and revision in the past decade. A lot remains to be done in knowledge update. For example, issues such as iterative knowledge update, abductive knowledge update, minimal knowledge in knowledge update, etc. remain to be explored. Similarly, in regards to reasoning about actions, additional specific cases of knowledge update need to be identified and simpler alternative characterization for them will be needed to be developed. Finally, we are currently working on establishing complexity results for knowledge update.

Acknowledgement

The authors thank Norman Foo and Abhaya Nayak for useful discussions on this topic. The research of the firat author was supported by NSF under grants IRI 9501577 and 0070463, and part of this research was performed while he was visiting the University of Western Sydney. The research of the second author was supported in part by Australian Research Council under grant A49803542.

References

[Engelfreit, 1998] J. Engelfreit. Monotonicity and persistence in preferential logics. *Journal of Artificial Intelligence Research*, 8:1–12, 1998.

[Fagin and et al, 1995] R. Fagin and et al. *Reasoning about Knowledge*. MIT press, 1995.

[Kartha and Lifschitz, 1994] G. Kartha and V. Lifschitz. Actions with indirect effects: Preliminary report. In *Proceedings of KR-94*, pages 341–350, 1994.

[Katsuno and Mendelzon, 1991] H. Katsuno and A. Mendelzon. On the difference between updating a knowledge base and revising it. In *Proceedings of of KR-91*, pages 387–394, 1991.

[Meyer and van der Hoek, 1995] J-J. Meyer and W. van der Hoek. Epistemis logics for computer science and artificial intelligence. In *Cambridge Tracts in Theoretical Computer Science, 41*. Cambridge University Press, 1995.

[Scherl and Levesque, 1993] R. Scherl and H. Levesque. The frame problem and knowledge producing actions. In *Proceedings of AAAI-93*, pages 689–695, 1993.

[Son and Baral, 2000] T. Son and C. Baral. Formalizing sensing actions: a transition function based approach. *Artificial Intelligence*, 125:19–92, 2000.

[Winslett, 1988] M. Winslett. Reasoning about action using a possible models approach. In *Proceedings of AAAI-88*, pages 89–93, 1988.

[Zhang, 1999] Y. Zhang. Monotonicity in rule based update. In *Proceedings of the 1999 International Conference on Logic Programming (ICLP'99)*, pages 471–485. MIT Press, 1999.