Reasoning About Knowledge of Unawareness*

Joseph Y. Halpern
Computer Science Department
Cornell University, U.S.A.
e-mail: halpern@cs.cornell.edu

Leandro Chaves Rêgo
School of Electrical and Computer Engineering
Cornell University, U.S.A.
e-mail: lcr26@cornell.edu

Abstract

Awareness has been shown to be a useful addition to standard epistemic logic for many applications. However, standard propositional logics for knowledge and awareness cannot express the fact that an agent knows that there are facts of which he is unaware without there being an explicit fact that the agent knows he is unaware of. We propose a logic for reasoning about knowledge of unawareness, by extending Fagin and Halpern's *Logic of General Awareness*. The logic allows quantification over variables, so that there is a formula in the language that can express the fact that "an agent explicitly knows that there exists a fact of which he is unaware". Moreover, that formula can be true without the agent explicitly knowing that he is unaware of any particular formula. We provide a sound and complete axiomatization of the logic, using standard axioms from the literature to capture the quantification operator. Finally, we show that the validity problem for the logic is recursively enumerable, but not decidable.

1 Introduction

As is well known, standard models of epistemic logic suffer from the *logical omniscience* problem (first observed and named by Hintikka 1962): agents know all tautologies and all the logical consequences of their knowledge. This seems inappropriate for resource-bounded agents and agents who are unaware of various concepts (and thus do not know logical tautologies involving those concepts). Many approaches to avoiding this problem have been suggested. One of the best-known is due to Fagin and Halpern 1988 (FH from now on). It involves distinguishing *explicit knowledge* from *implicit knowledge*, using a syntactic awareness operator. Roughly speaking, implicit knowledge is the standard (S5) notion of knowledge; explicit knowledge amounts to implicit knowledge and *awareness*.

Since this approach was first introduced by FH, there has been a stream of papers on awareness in the economics literature (see, for example, (Dekel, Lipman, & Rustichini 1998; Halpern 2001; Heifetz, Meier, & Schipper 2003; Modica & Rustichini 1994; 1999)). The logics used in these papers cannot express the fact that an agent may (explicitly) know that he is unaware of some facts. Indeed, in the language of (Halpern 2001), all of these models are special cases of the original awareness model where awareness is *generated by primitive propositions*, that is, an agent is aware of a formula iff the agent is aware of all primitive propositions that appear in the formula. If awareness is generated by primitive propositions, then it is impossible for an agent to know that he is unaware of a specific fact.

Nevertheless, knowledge of unawareness comes up often in real-life situations. For example, when a primary physician sends a patient to an expert on oncology, he knows that an oncologist is aware of things that could help the patient's treatment of which he is not aware. Moreover, the physician is unlikely to know which specific thing he is unaware of that would improve the patient's condition (if he knew which one it was, he would not be unaware of it!). Similarly, when an investor decides to let his money be managed by an investment fund company, he knows the company is aware of more issues involving the financial market than he is (and is thus likely to get better results with his money), but again, the investor is unlikely to be aware of the specific relevant issues. Ghirardato 2001 pointed out the importance of dealing with unawareness and knowledge of unawareness in the context of decision-making, but did not give a formal model.

To model knowledge of unawareness, we extend the syntax of the logic of general awareness considered by FH to allow for quantification over variables. Thus, we allow formulas such as $X_i(\exists x \neg A_i x)$, which says that agent i (explicitly) knows that there exists a formula of which he is not aware. The idea of adding propositional quantification to modal logic is well known in the literature (see, for example, (Bull 1969; Engelhardt, van der Meyden, & Moses 1998; Fine 1970; Kaplan 1970; Kripke 1959)). However, as we explain in Section 3, because A_i is a syntactic operator, we are forced to give somewhat nonstandard semantics to the existential operator. Nevertheless, we are able to provide a sound and complete axiomatization of the resulting

*This work was supported in part by NSF under grants CTC-0208535, ITR-0325453, and IIS-0534064, by ONR under grants N00014-00-1-03-41 and N00014-01-10-511, and by the DoD Multidisciplinary University Research Initiative (MURI) program administered by the ONR under grant N00014-01-1-0795. The second author was also supported in part by a scholarship from the Brazilian Government through the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

logic, using standard axioms from the literature to capture the quantification operator. Using the logic, we can easily characterize the knowledge of the relevant agents in all the examples we consider.

The rest of the paper is organized as follows. In Section 2, we review the standard semantics for knowledge and awareness. In Section 3, we introduce our logic for reasoning about knowledge of unawareness. In Section 4 we axiomatize the logic, and in Section 5, we consider the complexity of the decision problem for the logic. We conclude in Section 6.

2 The FH model

We briefly review the FH Logic of General Awareness here, before generalizing it to allow quantification over propositional variables. The syntax of the logic is as follows: given a set $\{1, \dots, n\}$ of agents, formulas are formed by starting with a set $\Phi = \{p, q, \dots\}$ of primitive propositions, and then closing off under conjunction ($\wedge$), negation ($\neg$), and the modal operators K_i, A_i, X_i, $i = 1, \dots, n$. Call the resulting language $\mathcal{L}_n^{K,X,A}(\Phi)$. As usual, we define $\varphi \vee \psi$ and $\varphi \Rightarrow \psi$ as abbreviations of $\neg(\neg\varphi \wedge \neg\psi)$ and $\neg\varphi \vee \psi$, respectively. The intended interpretation of $A_i\varphi$ is "i is aware of φ". The power of this approach comes from the flexibility of the notion of awareness. For example, "agent i is aware of φ" may be interpreted as "agent i is familiar with all primitive propositions in φ" or as "agent i can compute the truth value of φ in time t".

Having awareness in the language allows us to distinguish two notions of knowledge: implicit knowledge and explicit knowledge. Implicit knowledge, denoted by K_i, is defined as truth in all states that the agent considers possible, as usual. Explicit knowledge, denoted by X_i, is defined as the conjunction of implicit knowledge and awareness.

We give semantics to formulas in $\mathcal{L}_n^{K,X,A}(\Phi)$ in awareness structures. A tuple $M = (S, \pi, \mathcal{K}_1, ..., \mathcal{K}_n, \mathcal{A}_1, \dots, \mathcal{A}_n)$ is an *awareness structure for n agents (over Φ)* if S is a set of states, $\pi : S \times \Phi \rightarrow \{\mathbf{true}, \mathbf{false}\}$ is an interpretation that determines which primitive propositions are true at each state, $\mathcal{K}_i$ is a binary relation on S for each agent $i = 1, \dots, n$, and $\mathcal{A}_i$ is a function associating a set of formulas with each state in S, for $i = 1, ..., n$. Intuitively, if $(s, t) \in \mathcal{K}_i$, then agent i considers state t possible at state s, while $\mathcal{A}_i(s)$ is the set of formulas that agent i is aware of at state s. The set of formulas the agent is aware of can be arbitrary. Depending on the interpretation of awareness one has in mind, certain restrictions on $\mathcal{A}_i$ may apply. (We discuss some interesting restrictions in the next section.)

Let $\mathcal{M}_n(\Phi)$ denote the class of all awareness structures for n agents over Φ, with no restrictions on the $\mathcal{K}_i$ relations and on the functions $\mathcal{A}_i$. We use the superscripts r, e, and t to indicate that the $\mathcal{K}_i$ relations are restricted to being reflexive, Euclidean,[1] and transitive, respectively. Thus, for example, $\mathcal{M}_n^{rt}(\Phi)$ is the class of all reflexive and transitive awareness structures for n agents.

[1] Recall that a binary relation $\mathcal{K}_i$ is Euclidean if (s, t), $(s, u) \in \mathcal{K}_i$ implies that $(t, u) \in \mathcal{K}_i$.

We write $(M, s) \models \varphi$ if φ is true at state s in the awareness structure M. The truth relation is defined inductively as follows:

$(M, s) \models p$, for $p \in \Phi$, if $\pi(s, p) = \mathbf{true}$
$(M, s) \models \neg\varphi$ if $(M, s) \not\models \varphi$
$(M, s) \models \varphi \wedge \psi$ if $(M, s) \models \varphi$ and $(M, s) \models \psi$
$(M, s) \models K_i\varphi$ if $(M, t) \models \varphi$
  for all t such that $(s, t) \in \mathcal{K}_i$
$(M, s) \models A_i\varphi$ if $\varphi \in \mathcal{A}_i(s)$
$(M, s) \models X_i\varphi$ if $(M, s) \models A_i\varphi$ and $(M, s) \models K_i\varphi$.

A formula φ is said to be *valid* in awareness structure M, written $M \models \varphi$, if $(M, s) \models \varphi$ for all $s \in S$. A formula is valid in a class $\mathcal{N}$ of awareness structures, written $\mathcal{N} \models \varphi$, if it is valid for all awareness structures in $\mathcal{N}$, that is, if $N \models \varphi$ for all $N \in \mathcal{N}$.

Consider the following set of well-known axioms and inference rules:

Prop. All substitution instances of valid formulas of propositional logic.

K. $(K_i\varphi \wedge K_i(\varphi \Rightarrow \psi)) \Rightarrow K_i\psi$.

T. $K_i\varphi \Rightarrow \varphi$.

4. $K_i\varphi \Rightarrow K_iK_i\varphi$.

5. $\neg K_i\varphi \Rightarrow K_i\neg K_i\varphi$.

A0. $X_i\varphi \Leftrightarrow K_i\varphi \wedge A_i\varphi$.

MP. From φ and $\varphi \Rightarrow \psi$ infer ψ (modus ponens).

Gen$_K$. From φ infer $K_i\varphi$.

It is well known that the axioms T, 4, and 5 correspond to the requirements that the $\mathcal{K}_i$ relations are reflexive, transitive, and Euclidean, respectively. Let $\mathbf{K}_n$ be the axiom system consisting of the axioms Prop, K and rules MP, and Gen$_K$. The following result is well known (see, for example, (Fagin *et al.* 1995) for proofs).

Theorem 2.1: *Let $\mathcal{C}$ be a (possibly empty) subset of $\{\mathrm{T}, 4, 5\}$ and let C be the corresponding subset of $\{r, t, e\}$. Then $\mathbf{K}_n \cup \{A0\} \cup \mathcal{C}$ is a sound and complete axiomatization of the language $\mathcal{L}_n^{K,X,A}(\Phi)$ with respect to $\mathcal{M}_n^C(\Phi)$.*

3 A logic for reasoning about knowledge of unawareness

To allow reasoning about knowledge of unawareness, we extend the language $\mathcal{L}_n^{K,X,A}(\Phi)$ by adding a countable set of propositional variables $\mathcal{X} = \{x, y, z, \dots\}$ and allowing universal quantification over these variables. Thus, if φ is a formula, then so is $\forall x\varphi$. As usual, we take $\exists x\varphi$ to be an abbreviation for $\neg\forall x\neg\varphi$. Let $\mathcal{L}_n^{\forall,K,X,A}(\Phi, \mathcal{X})$ denote this extended language.

We assume that $\mathcal{X}$ is countably infinite for essentially the same reason that the set of variables is always taken to be infinite in first-order logic. Without it, we seriously limit the expressive power of the language. For example, a formula such as $\exists x\exists y(\neg(x \Leftrightarrow y) \wedge A_1x \wedge A_1y)$ says that there are two distinct formulas that agent 1 is aware of. We can similarly define formulas saying that there are k distinct formulas that agent 1 is aware of. However, to do this we need k distinct primitive propositions.

Essentially as in first-order logic, we can define inductively what it means for a variable x to be *free* in a formula φ. If φ does not contain the universal operator $\forall$, then every occurrence of x in φ is free; an occurrence of x is free in $\neg\varphi$ (or $K_i\varphi$, $X_i\varphi$, $A_i\varphi$) iff the corresponding occurrence of x is free in φ; an occurrence of x in $\varphi_1 \wedge \varphi_2$ is free iff the corresponding occurrence of x in φ_1 or φ_2 is free; and an occurrence of x is free in $\forall y\varphi$ iff the corresponding occurrence of x is free in φ and x is different from y. Intuitively, an occurrence of a variable is free in a formula if it is not bound by a quantifier. A formula that contains no free variables is called a *sentence*.

Let $\mathcal{S}_n^{\forall,K,X,A}(\Phi,\mathcal{X})$ denote the set of sentences in $\mathcal{L}_n^{\forall,K,X,A}(\Phi,\mathcal{X})$. If ψ is a formula, let $\varphi[x/\psi]$ denote the formula that results by replacing all free occurrences of the variable x in φ by ψ. (If there is no free occurrence of x in φ, then $\varphi[x/\psi] = \varphi$.) We extend this notion of substitution to sequences of variables, writing $\varphi[x_1/\psi_1,\ldots,x_n/\psi_n]$. We say that ψ is *substitutable for x in φ* if, for all propositional variables y, if an occurrence of y is free in ψ, then the corresponding occurrence of y is free in $\varphi[x/\psi]$.

We want to give semantics to formulas in $\mathcal{L}_n^{\forall,K,X,A}(\Phi,\mathcal{X})$ in awareness structures (where now we allow $\mathcal{A}_i(s)$ to be an arbitrary subset of $\mathcal{S}_n^{\forall,K,X,A}(\Phi,\mathcal{X})$). The standard approach for giving semantics to propositional quantification ((Engelhardt, van der Meyden, & Moses 1998; Kripke 1959; Bull 1969; Kaplan 1970; Fine 1970)) uses *semantic valuations*, much like in first-order logic. A *semantic valuation* $\mathcal{V}$ associates with each propositional variable and state a truth value, just as an interpretation π associates with each primitive proposition and state a truth value. Then $(M,s,\mathcal{V}) \models x$ if $\mathcal{V}(x) = \mathit{true}$ and $(M,s,\mathcal{V}) \models \forall x\varphi$ if $(M,s,\mathcal{V}') \models \varphi$ for all valuations $\mathcal{V}'$ that agree with $\mathcal{V}$ on all propositional variables other than x.[2] We write $\mathcal{V} \sim_x \mathcal{V}'$ if $\mathcal{V}(y) = \mathcal{V}'(y)$ for all variables $y \neq x$.

Using semantic valuations does not work in the presence of awareness. If $\mathcal{A}_i(s)$ consists only of sentences, then a formula such as $\forall x A_i x$ is guaranteed to be false since, no matter what the valuation is, $x \notin \mathcal{A}_i(s)$. The valuation plays no role in determining the truth of a formulas of the form $A_i\psi$. On the other hand, if we allow $\mathcal{A}_i(s)$ to include any formula in the language, then $(M,s,\mathcal{V}) \models \forall x A_i(x)$ iff $x \in \mathcal{A}_i(s)$. But then it is easy to check that $(M,s,\mathcal{V}) \models \exists x A_i(x)$ iff $x \in \mathcal{A}_i(s)$, which certainly does not seem to capture our intuition.

We want to interpret $\forall x A_i(x)$ as saying "for all sentences $\varphi \in \mathcal{S}_n^{\forall,K,X,A}(\Phi,\mathcal{X})$, $A_i(\varphi)$ holds". For technical reasons (which we explain shortly), we instead interpret this it as "for all formulas $\varphi \in \mathcal{L}_n^{K,X,A}(\Phi)$, $A_i(\varphi)$ holds". That is, we consider only sentences with no quantification. To achieve this, we use *syntactic valuations*, rather than *semantic valuations*. A *syntactic valuation* is a function $\mathcal{V} : \mathcal{X} \to \mathcal{L}_n^{K,X,A}(\Phi)$, which assigns to each variable a sentence in $\mathcal{L}_n^{K,X,A}(\Phi)$.

We give semantics to formulas in $\mathcal{L}_n^{\forall,K,X,A}(\Phi,\mathcal{X})$ by by induction on the total number of free and bound variables, with a subinduction on the length of the formula. The definitions for the constructs that already appear in $\mathcal{L}_n^{K,X,A}(\Phi)$ are the same. To deal with universal quantification, we just consider all possible replacements of the quantified variable by a sentence in $\mathcal{L}_n^{K,X,A}(\Phi)$.

- If φ is a formula whose free variables are $x_1,\ldots,x_k$, then $(M,s,\mathcal{V}) \models \varphi$ if $(M,s,\mathcal{V}) \models \varphi[x_1/\mathcal{V}(x_1),\ldots,x_k/\mathcal{V}(x_k)]$
- $(M,s,\mathcal{V}) \models \forall x\varphi$ if $(M,s,\mathcal{V}') \models \varphi$ for all syntactic valuations $\mathcal{V}' \sim_x \mathcal{V}$.

Note that although $\varphi[x_1/\mathcal{V}(x_1),\ldots,\mathcal{V}(x_k)]$ may be a longer formula than φ, it involves fewer variables, since $\mathcal{V}(x_1),\ldots,\mathcal{V}(x_k)$ do not mention variables. This is why it is important that we quantify only over sentences in $\mathcal{L}_n^{K,X,A}(\Phi)$; if we were to quantify over all sentences in $\mathcal{S}_n^{\forall,K,X,A}(\Phi,\mathcal{X})$, then the semantics would not be well defined. For example, to determine the truth of $\forall xx$, we would have to determine the truth of $x[x/\forall xx] = \forall xx$. This circularity would make $\models$ undefined. In any case, given our restrictions, it is easy to show that $\models$ is well defined. Since the truth of a sentence is independent of a valuation, for a sentence φ, we write $(M,s) \models \varphi$ rather than $(M,s,V) \models \varphi$.

Under our semantics, the formula $K_i(\exists x(A_j x \wedge \neg A_i x))$ is consistent and that it can be true at state s even though there might be no formula ψ in $\mathcal{L}_n^{K,X,A}(\Phi)$ such that $K_i((A_j\psi \wedge \neg A_i\psi))$. This situation can happen if, at all states agent i considers possible, agent j is aware of something agent i is not, but there is no one formula ψ such that agent j is aware of ψ in all states agent i considers possible and agent i is not aware of ψ in all such states. By way of contrast, if $\exists x K_i(A_j x \wedge \neg A_i x)$ is true at state s, then there is a formula ψ such that $K_i(A_j\psi \wedge \neg A_i\psi)$ holds at s. The difference between $K_i\exists x(A_j x \wedge \neg A_i x)$ and $\exists x K_i(A_j x \wedge \neg A_i x)$ is essentially the same as the difference between $\exists x K_i\varphi$ and $K_i(\exists x\varphi)$ in first-order modal logic (see, for example, (Fagin *et al.* 1995) for a discussion).

The next example illustrates how the logic of knowledge of awareness can be used to capture some interesting situations.

Example 3.1: Consider an investor (agent 1) and an investment fund broker (agent 2). Suppose that we have two facts that are relevant for describing the situation: the NASDAQ index is more likely to increase than to decrease tomorrow (p), and Amazon will announce a huge increase in earnings tomorrow (q). Let $S = \{s\}$, $\pi(s,p) = \pi(s,q) = \mathbf{true}$, $\mathcal{K}_i = \{(s,s)\}$, $\mathcal{A}_1(s) = \{p, \exists x(A_2 x \wedge \neg A_1 x)\}$, and $\mathcal{A}_2(s) = \{p, q, A_2 q, \neg A_1 q, A_2 q \wedge \neg A_1 q\}$. Thus, both agents explicitly know that the NASDAQ index is more likely to increase than to decrease tomorrow. However, the broker also explicitly knows that Amazon will announce a huge increase in earnings tomorrow. Furthermore, the broker explicitly knows that he (broker) is aware of this fact and the

[2] We remark that the standard approach does not use separate propositional variables, but quantifies over primitive propositions. This makes it unnecessary to use valuations. It is easy to see that the definition we have given is equivalent to the standard definition. Using propositional variables is more convenient in our extension.

investor is not. On the other hand, the investor explicitly knows that there is something that the broker is aware of but he is not. That is,

$$(M,s,\mathcal{V}) \models X_1 p \wedge X_2 p \wedge X_2 q \wedge \neg X_1 q$$
$$\wedge X_2(A_2 q \wedge \neg A_1 q) \wedge X_1(\exists x(A_2 x \wedge \neg A_1 x)).$$

Since $X_2(A_2 q \wedge \neg A_1 q)$ implies $\exists x X_2(A_2 x \wedge \neg A_1 x)$, there is some formula x such that the broker knows that the investor is unaware of x although he is aware of x. However, since $(M,s,\mathcal{V}) \models \neg A_2(\exists x(A_2 x \wedge \neg A_1 x))$, it follows that $(M,s,\mathcal{V}) \models \neg X_2(\exists x(A_2 x \wedge \neg A_1 x))$. ∎

It may seem unreasonable that, in Example 3.1, the broker is aware of the formula $A_2 q \wedge \neg A_1 q$, without being aware of $\exists x(A_2 x \wedge \neg A_1 x)$. Of course, if the broker were aware of this formula, then $X_2((\exists x(A_2 x \wedge \neg A_1 x))$ would hold at state s. This example suggests that we may want to require various properties of awareness. Here are some that are relevant in this context:

- Awareness is *closed under existential quantification* if $\varphi \in \mathcal{A}_i(s)$, $\varphi = \varphi'[x/\psi]$ and $\psi \in \mathcal{L}_n^{K,X,A}(\Phi)$, then $(\exists x \varphi') \in \mathcal{A}_i(s)$.
- Awareness is *generated by primitive propositions* if, for all agents i, $\varphi \in \mathcal{A}_i(s)$ iff all the primitive propositions that appear in φ are in $\mathcal{A}_i(s) \cap \Phi$. That is, an agent is aware of φ iff she is aware of all the primitive propositions that appear in φ.
- *Agents know what they are aware of* if, for all agents i and all states s,t such that $(s,t) \in \mathcal{K}_i$ we have that $\mathcal{A}_i(s) = \mathcal{A}_i(t)$.

Closure under existential quantification does *not* hold in Example 3.1. It is easy to see that it is a consequence of awareness being generated by primitive propositions. As shown by Halpern 2001 and Halpern and Rêgo 2005, a number of standard models of awareness in the economics literature (e.g., (Heifetz, Meier, & Schipper 2003; Modica & Rustichini 1999)) can be viewed as instances of the FH model where awareness is taken to be generated by primitive propositions and agents know what they are aware of. While assuming that awareness is generated by primitive propositions seems like quite a reasonable assumption if there is no existential quantification in the language, it does not seem quite so reasonable in the presence of quantification. For example, if awareness is generated by primitive propositions, then the formula $A_i(\exists x \neg A_i x)$ is valid, which does not seem to be reasonable in many applications. For some applications it may be more reasonable to instead assume only that awareness is *weakly generated by primitive propositions*. This is the case if, for all states s and agents i,

- $\neg\varphi \in \mathcal{A}_i(s)$ iff $\varphi \in \mathcal{A}_i(s)$;
- $\varphi \wedge \psi \in \mathcal{A}_i(s)$ iff $\varphi, \psi \in \mathcal{A}_i(s)$;
- $K_i\varphi \in \mathcal{A}_i(s)$ iff $\varphi \in \mathcal{A}_i(s)$;
- $A_i\varphi \in \mathcal{A}_i(s)$ iff $\varphi \in \mathcal{A}_i(s)$;
- $X_i\varphi \in \mathcal{A}_i(s)$ iff $\varphi \in \mathcal{A}_i(s)$;
- if $\forall x \varphi \in \mathcal{A}_i(s)$, then $p \in \mathcal{A}_i(s)$ for each primitive proposition p that appears in $\forall x \varphi$;
- if $\varphi[x/\psi] \in \mathcal{A}_i(s)$ for some formula $\psi \in \mathcal{L}_n^{K,X,A}(\Phi)$, then $\exists x \varphi \in \mathcal{A}_i(s)$.

If the language does not have quantification, then awareness is weakly generated by primitive propositions iff it is generated by primitive propositions. However, with quantification in the language, while it is still true that if awareness is generated by primitive propositions then it is weakly generated by primitive propositions, the converse does not necessarily hold. For example, if $\mathcal{A}_1(s) = \emptyset$ for all s, then awareness is weakly generated by primitive propositions. Intuitively, not being aware of any formulas is consistent with awareness being weakly generated by primitive propositions. However, if agent 1's awareness is generated by primitive propositions, then, for example, $\exists x A_j x$ must be in $\mathcal{A}_1(s)$ for all s and all agents j.

4 Axiomatization

In this section, we provide a sound and complete axiomatization of the logic described in the previous section. We show that, despite the fact that we have a different language and used a different semantics for quantification, essentially the same axioms characterize our definition of quantification as those that have been shown to characterize the more traditional definition. Indeed, our axiomatization is very similar to the multi-agent version of an axiomatization given by Fine 1970 for a variant of his logic where the range of quantification is restricted.

Consider the following axioms for quantification:

$1_\forall$. $\forall x \varphi \Rightarrow \varphi[x/\psi]$ if ψ is a quantifier-free formula substitutable for x in φ.

$\mathrm{K}_\forall$. $\forall x(\varphi \Rightarrow \psi) \Rightarrow (\forall x \varphi \Rightarrow \forall x \psi)$.

$\mathrm{N}_\forall$. $\varphi \Rightarrow \forall x \varphi$ if x is not free in φ.

Barcan. $\forall x K_i \varphi \Rightarrow K_i \forall x \varphi$.

$\mathrm{Gen}_\forall$. From φ infer $\forall x \varphi$.

These axioms are almost identical to the ones considered by Fine 1970, except that we restrict $1_\forall$ to quantifier-free formulas; Fine allows arbitrary formulas to be substituted (provided that they are substitutable for x). $\mathrm{K}_\forall$ and $\mathrm{Gen}_\forall$ are analogues to the axiom K and rule of inference Gen_K in $\mathbf{K}_n$. The Barcan axiom, which is well-known in first-order modal logic, captures the relationship between quantification and K_i.

Let $\mathbf{K}_n^\forall$ be the axiom system consisting of the axioms in $\mathbf{K}_n$ together with $\{$A0, $1_\forall$, $\mathrm{K}_\forall$, $\mathrm{N}_\forall$, Barcan, $\mathrm{Gen}_\forall\}$.

Theorem 4.1 : *Let $\mathcal{C}$ be a (possibly empty) subset of $\{\mathrm{T}, 4, 5\}$ and let C be the corresponding subset of $\{r, t, e\}$. If Φ is countably infinite, then $\mathbf{K}_n^\forall \cup \mathcal{C}$ is a sound and complete axiomatization of the language $\mathcal{L}_n^{\forall,K,X,A}(\Phi, \mathcal{X})$ with respect to $\mathcal{M}_n^{\forall,C}(\Phi, \mathcal{X})$.*[3]

Proof: Showing that a provable formula φ is valid can be done by a straightforward induction on the length of the

[3]We remark that Prior 1956 showed that, in the context of first-order modal logic, the Barcan axiom is not needed in the presence of the axioms of **S5** (that is T, 4, and 5). The same argument works here.

proof of φ, using the fact that all axioms are valid in the appropriate set of models and all inference rules preserve validity.

In the standard completeness proof for modal logic, a canonical model M^c is constructed where the states are maximal consistent sets of formulas. It is then shown that if s_V is the state corresponding to the maximal consistent set V, then $(M^c, s_V) \models \varphi$ iff $\varphi \in V$. This will not quite work in our logic. We would need to define a canonical valuation function to give semantics for formulas containing free variables. We deal with this problem by considering states in the canonical model to consist of maximal consistent sets of sentences. There is another problem in the presence of quantification since there may be a maximal consistent set V of sentences such that $\neg\forall x\varphi \in V$, but $\varphi[x/\psi]$ for all $\psi \in \mathcal{L}_n^{K,X,A}(\Phi)$. That is, there is no witness to the falsity of $\forall x\varphi$ in V. We deal with this problem by restricting to maximal consistent sets V that are *acceptable* in the sense that if $\neg\forall x\varphi \in V$, then $\neg\varphi[x/q] \in V$ for some primitive proposition $q \in \Phi$. This argument requires Φ to be infinite. The details can be found in the appendix. ∎

To understand why, in general, we need to assume Φ is countably infinite, consider the case where there is only one agent and $\Phi = \{p\}$. Let φ be the formula that essentially forces the S5 axioms to hold:

$$\forall x(Kx \Rightarrow KKx) \wedge (\neg Kx \Rightarrow K\neg Kx).$$

As is well-known, the S5 axioms force every formula in $\mathcal{L}_1^K$ to be equivalent to a depth-one formula (i.e., one without nested K's). Thus, it is not hard to show that there exists a finite set F of formulas in $\mathcal{L}_1^{K,X,A}(\{p\})$ such that for all formulas ψ with one free variable y and no quantification, we have for $C \subseteq \{r, e, t\}$

$$\mathcal{M}_1^{\forall,C}(\Phi,\mathcal{X}) \models (\varphi \wedge \forall x Ax) \Rightarrow (\forall y\psi \Leftrightarrow \wedge_{\sigma\in F}\psi[y/\psi]).$$

Thus, if Φ has only one primitive proposition, then there are circumstances under which universal quantification is equivalent to a finite conjunction. We can construct similar examples if Φ is an arbitrary finite set of propositions even if there is more than one agent. (In the latter case, we add a formula φ' to the antecedent that says that agent 1 knows that all agents have the same knowledge that he does: $\forall x K_1 \wedge_{i=2}^n (K_i x \Leftrightarrow K_1 x)$.) Thus, if Φ is finite, we would need extra axioms to capture the fact universal quantification can sometimes be equivalent to a finite conjunction. We remark that this phenomenon of needing additional axioms if Φ is finite has been observed before in the literature (cf. (Fagin, Halpern, & Vardi 1992; Halpern & Lakemeyer 2001)).

If we make further assumptions about the awareness operator, these can also be captured axiomatically. For example, as shown by FH, the assumption that agents know what they are aware of corresponds to the axioms

$$A_i\varphi \Rightarrow K_iA_i\varphi \text{ and}$$
$$\neg A_i\varphi \Rightarrow K_i\neg A_i\varphi.$$

It is not hard to check that awareness being generated by primitive propositions can be captured by the following axiom:

$$A_i\varphi \Leftrightarrow \wedge_{\{p\in\Phi:\ p \text{ occurs in } \varphi\}}A_ip.$$

In this axiom, the empty conjunction is taken to be vacuously true, so that $A_i\varphi$ is vacuously true if no primitive propositions occur in φ.

We can axiomatize the fact that awareness is weakly generated by primitive propositions using the following axioms:

A1. $A_i(\varphi \wedge \psi) \Leftrightarrow A_i\varphi \wedge A_i\psi$.

A2. $A_i\neg\varphi \Leftrightarrow A_i\varphi$.

A3. $A_iX_j\varphi \Leftrightarrow A_i\varphi$.

A4. $A_iA_j\varphi \Leftrightarrow A_i\varphi$.

A5. $A_iK_j\varphi \Leftrightarrow A_i\varphi$.

A6. $A_i\varphi \Rightarrow A_ip$ if $p \in \Phi$ occurs in φ.

A7. $A_i\varphi[x/\psi] \Rightarrow A_i\exists x\varphi$, where $\psi \in \mathcal{L}_n^{K,X,A}(\Phi)$.

As noted in (Fagin *et al.* 1995), the first five axioms capture awareness generated by primitive propositions in the language $\mathcal{L}_n^{K,X,A}(\Phi)$; we need A6 and A7 to deal with quantification. A7 captures the fact that awareness is closed under existential quantification.

5 Complexity

Since the logic is axiomatizable, the validity problem is at worst recursively enumerable. As the next theorem shows, the validity problem is no better than r.e.

Theorem 5.1: *The problem of deciding if a formula in the language $\mathcal{L}_n^{\forall,K,X,A}(\Phi,\mathcal{X})$ is valid in $\mathcal{M}_n^C(\Phi,\mathcal{X})$ is r.e.-complete, for all $C \subseteq \{r,t,e\}$ and $n \geq 1$.*

Proof: The fact that deciding validity is r.e. follows immediately from Theorem 2.1. For the hardness result, we show that, for every formula φ in first-order logic over a language with a single binary predicate can be translated to a formula $\varphi^t \in \mathcal{L}_1^{\forall,K,A}(\Phi,\mathcal{X})$ such that φ is valid over relational models iff φ^t is valid in $\mathcal{M}_n^{\emptyset}(\Phi,\mathcal{X})$ (and hence in $\mathcal{M}_n^C(\Phi,\mathcal{X})$, for all $C \subseteq \{r,t,e\}$. The result follows from the well-known fact that the validity problem for first-order logic with one binary predicate is r.e. The details of the proof of this theorem and the remaining results in this section can be found in the full paper (Halpern & Rêgo 2006c). ∎

Theorem 5.1 is somewhat surprising, since Fine 1970 shows that his logic (which is based on S5) is decidable. It turns out that each of the following suffices to get undecidability: (a) the presence of the awareness operator, (b) the presence of more than one agent, or (c) not having $e \in C$ (i.e., not assuming that the $\mathcal{K}$ relation satisfies the Euclidean property). The fact that awareness gives undecidability is the content of Theorem 5.1; Theorem 5.2 shows that having $n \geq 2$ or $e \notin C$ suffices for undecidability as well. On the other hand, Theorem 5.3 shows that if $n = 1$ and $e \in C$, then the problem is decidable. Although, as we have observed, our semantics is slightly differently from that of Fine, we believe that corresponding results hold in his setting. Thus, he gets decidability because he does not have awareness, restricts to a single agent, and considers S5 (as opposed to say, S4).

Let $\mathcal{L}_n^{\forall,K}(\Phi,\mathcal{X})$ consist of all formulas in $\mathcal{L}_n^{\forall,K,X,A}(\Phi,\mathcal{X})$ that do not mention the A_i or X_i operators.

Theorem 5.2: *The problem of deciding if a formula in the language $\mathcal{L}_n^{\forall,K}(\Phi)$ is valid in $\mathcal{M}_n^C(\Phi)$ is r.e.-complete if $n \geq 2$ or if $e \notin C$.*

Theorem 5.3 : *The validity problem for the language $\mathcal{L}_1^{\forall,K}(\Phi, \mathcal{X})$ with respect to the structures in $\mathcal{M}_1^C(\Phi, \mathcal{X})$ for $C \supseteq \{e\}$ is decidable.*

Interestingly, the role of the Euclidean property in these complexity results mirrors its role in complexity for $\mathcal{L}_n^K$, basic epistemic logic without awareness or quantification. As we have shown (Halpern & Rêgo 2006a), the problem of deciding if a formula in the language $\mathcal{L}_n^K(\Phi)$ is valid in $\mathcal{M}_n^C(\Phi)$ is PSPACE complete if $n \geq 2$ or $n \geq 1$ and $e \notin C$; if $n = 1$ and $e \in C$, it is co-NP-complete.

6 Conclusion

We have proposed a logic to model agents who are able to reason about their lack of awareness. We have shown that such reasoning arises in a number of situations. We have provided a complete axiomatization for the logic, and examined the complexity of the validity problem.

Our original motivation for considering knowledge of unawareness came from game theory. Notions like Nash equilibrium do not make sense in the presence of lack of awareness. Intuitively, a set of strategies is a Nash equilibrium if each agent would continue playing the same strategy despite knowing what strategies the other agents are using. But if an agent is not aware of the moves available to other agents, then he cannot even contemplate the actions of other players. In a companion paper (Halpern & Rêgo 2006b), we show how to generalize the notion of Nash equilibrium so that it applies in the presence of (knowledge of) unawareness. Feinberg 2004 has already shown that awareness can play a significant role in analyzing games. (In particular, he shows that a small probability of an agent not being aware of the possibility of defecting in finitely repeated prisoners dilemma can lead to cooperation.) It is not hard to show that knowledge of unawareness can have a similarly significant impact. Consider, for example, a chess game. If we interpret "lack of awareness" as "unable to compute" (cf. (Fagin & Halpern 1988)), then although all players understand in principle all the moves that can be made, they are certainly not aware of all consequences of all moves. Such lack of awareness has strategic implications. For example, in cases where the opponent is under time pressure, experts will make deliberately make moves that lead to positions that are hard to analyze. (In our language, these are positions where there is a great deal of unawareness.) The logic we have presented here provides an initial step to modeling the reasoning that goes on in games with (knowledge of) unawareness. To fully model what is going on, we need to capture probability as well as awareness and knowledge. We do not think that there will be any intrinsic difficulty in extending our logic to handle probability along the lines of the work in (Fagin & Halpern 1994; Fagin, Halpern, & Megiddo 1990), although we have not checked the details.

A Proof of Theorem 4.1

Theorem 4.1: *Let $\mathcal{C}$ be a (possibly empty) subset of $\{\mathrm{T}, 4, 5\}$ and let C be the corresponding subset of $\{r, t, e\}$. Then $\mathbf{K}_n^\forall \cup \mathcal{C}$ is a sound and complete axiomatization of the language $\mathcal{L}_n^{\forall,K,X,A}(\Phi, \mathcal{X})$ with respect to $\mathcal{M}_n^{\forall,C}(\Phi, \mathcal{X})$.*

Proof: We give the proof only for the case $\mathcal{C} = \emptyset$; the other cases follow using standard techniques (see, for example, (Fagin *et al.* 1995; Hughes & Cresswell 1996)). Showing that a provable formula φ is valid can be done by a straightforward induction on the length of the proof of φ, using the fact that all axioms in $\mathbf{K}_n^\forall$ are valid in $\mathcal{M}_n^{\forall,\emptyset}(\Phi, \mathcal{X})$ and all inference rules preserve validity in $\mathcal{M}_n^{\forall,\emptyset}(\Phi, \mathcal{X})$.

As we said in the main text, we prove completeness by modifying the standard canonical model construction, restricting to acceptable maximal consistent sets of sentences. Thus, the first step in the proof is to guarantee that every consistent sentence is included in an acceptable maximal consistent set of sentences.

If q is a primitive proposition, we define $\varphi[q/x]$ and the notion of x being substitutable for q just as we did for the case that q is a propositional variable.

Lemma A.1: *If $\mathbf{K}_n^\forall \cup \mathcal{C} \vdash \varphi$ and x is substitutable for q in φ, then $\mathbf{K}_n^\forall \cup \mathcal{C} \vdash \forall x \varphi[q/x]$.*

Proof: We first show by induction on the length of the proof of φ that if z is a variable that does not appear in any formula in the proof of φ, then $\mathbf{K}_n^\forall \cup \mathcal{C} \vdash \varphi[q/z]$. If there is a proof of φ of length one, then φ is an instance of an axiom. It is easy to see that $\varphi[q/z]$ is an instance of the same axiom. (We remark that it is important in the case of axioms $\mathrm{N}_\forall$ and $1_\forall$ that z does not occur in φ.) Suppose that the lemma holds for all φ' that have a proof of length no greater than k, and suppose that φ has a proof of length $k+1$ where z does not occur in any formula of the proof. If the last step of the proof of φ is an axiom, then φ is an instance of an axiom, and we have already dealt with this case. Otherwise, the last step in the proof of φ is an application of either MP, Gen_K, or $\mathrm{Gen}_\forall$. We consider these in turn.

If MP is applied at the last step, then there exists some φ', such that φ' and $\varphi' \Rightarrow \varphi$ were previously proved and, by assumption, z does not occur in any formula of their proof. By the induction hypothesis, both $\varphi'[q/z]$ and $(\varphi' \Rightarrow \varphi)[q/z] = \varphi'[q/z] \Rightarrow \varphi[q/z]$ are provable. The result now follows by an application of MP.

The argument for Gen_K and $\mathrm{Gen}_\forall$ is essentially identical, so we consider them together. Suppose that Gen_K (resp., $\mathrm{Gen}_\forall$) is applied at the last step. Then φ has the form $K_i\varphi'$ (resp., $\forall y \varphi'$) and there is a proof of length at most k for φ' where z does not occur in any formula in the proof. Thus, by the induction hypothesis, $\varphi'[q/z]$ is provable. By applying Gen_K (resp., $\mathrm{Gen}_\forall$), it immediately follows that $\varphi[q/z]$ is provable.

This completes the proof that $\varphi[q/z]$ is provable. By applying $\mathrm{Gen}_\forall$, it follows that $\forall z \varphi[q/z]$ is provable. Since x is substitutable for q in φ, x must be substitutable for z in $\varphi[q/z]$. Thus, by applying the axiom $1_\forall$ and MP, we can prove $\varphi[q/x]$. The fact that $\forall x \varphi[q/x]$ is provable now follows from $\mathrm{Gen}_\forall$. ∎

Lemma A.2: *Every $\mathbf{K}_n^{\forall}$-consistent sentence φ is contained in some acceptable maximal $\mathbf{K}_n^{\forall}$-consistent set of sentences.*

Proof: We first show that if Δ is a finite $\mathbf{K}_n^{\forall}$-consistent set of sentences, $\neg\forall x\psi \in \Delta$, and q is a primitive proposition that does not appear in any sentence in Δ, then $\Delta \cup \neg\varphi[x/q]$ is $\mathbf{K}_n^{\forall}$-consistent. Suppose not. Then there exist sentences $\beta_1, \ldots, \beta_k \in \Delta$ such that

$$\mathbf{K}_n^{\forall} \vdash (\beta_1 \wedge \ldots \wedge \beta_k) \Rightarrow \psi[x/q].$$

By Lemma A.1, we have

$$\mathbf{K}_n^{\forall} \vdash \forall x((\beta_1 \wedge \ldots \wedge \beta_k) \Rightarrow \psi). \quad (1)$$

Now applying $\mathrm{K}_{\forall}$ and $\mathrm{N}_{\forall}$, and using the fact that x is not free in $\beta_1 \wedge \ldots \wedge \beta_k$ (since $\beta_1, \ldots, \beta_k$ are sentences), it easily follows that

$$\mathbf{K}_n^{\forall} \vdash (\beta_1 \wedge \ldots \wedge \beta_k) \Rightarrow \forall x\psi.$$

Since $\beta_1, \ldots, \beta_k, \neg\forall x\psi \in \Delta$, this contradicts the assumption that Δ is $\mathbf{K}_n^{\forall}$-consistent.

We now use standard techniques to construct an acceptable maximal $\mathbf{K}_n^{\forall}$-consistent set of sentences containing φ. Consider an enumeration $\{\psi_1, \psi_2, \ldots\}$ of the sentences in $\mathcal{S}_n^{\forall,K,X,A}$. We construct a sequence of $\mathbf{K}_n^{\forall}$-consistent sets of sentences $\Delta_0, \Delta_1, \ldots$. Let $\Delta_0 = \{\varphi\}$. For all $k \geq 1$, if $\Delta_{k-1} \cup \{\psi_k\}$ is not $\mathbf{K}_n^{\forall}$-consistent, $\Delta_k = \Delta_{k-1}$. If $\Delta_{k-1} \cup \{\psi_k\}$ is $\mathbf{K}_n^{\forall}$-consistent, if ψ_k is not of the form $\neg\forall y\psi'$, then $\Delta_k = \Delta_{k-1} \cup \{\psi_k\}$, while if ψ_k is of the form $\neg\forall y\psi'$, then $\Delta_k = \Delta_{k-1} \cup \{\psi_k, \neg\psi'[y/q]\}$ for some $q \in \Phi$ not occurring in $\Delta_{k-1} \cup \{\psi_k\}$. By our earlier argument, it easily follows that each set Δ_k is $\mathbf{K}_n^{\forall}$-consistent. Let Δ be the union of the Δ_n's. It is easy to see that Δ is an acceptable maximal $\mathbf{K}_n^{\forall}$-consistent set of sentences that contains φ. ∎

For a set Γ of formulas, define $\Gamma/K_i = \{\psi : K_i\psi \in \Gamma\}$.

Lemma A.3: *If Γ is a $\mathbf{K}_n^{\forall}$-consistent set of sentences containing $\neg K_i\varphi$, then $\Gamma/K_i \cup \{\neg\varphi\}$ is $\mathbf{K}_n^{\forall}$-consistent.*

Proof: This is a standard modal logic argument; see, for example, (Hughes & Cresswell 1996, Lemma 6.4). We omit details here. ∎

Lemma A.4: *If Γ is an acceptable maximal $\mathbf{K}_n^{\forall}$-consistent set of sentences and $\neg K_i\varphi \in \Gamma$, then there exists an acceptable maximal $\mathbf{K}_n^{\forall}$-consistent set of sentences Δ such that $(\Gamma/K_i \cup \{\neg\varphi\}) \subseteq \Delta$.*

Proof: We first show that if $\Gamma/K \cup \{\gamma_1, \ldots, \gamma_n, \neg\forall x\psi\}$ is a $\mathbf{K}_n^{\forall}$-consistent set of sentences, then there exists some $q \in \Phi$ such that $\Gamma/K_i \cup \{\gamma_1, \ldots, \gamma_n, \neg\forall x\psi, \neg\psi[x/q]\}$ is $\mathbf{K}_n^{\forall}$-consistent. For suppose not. Then, for all q, there must exist $\beta_1, \ldots, \beta_k \in \Gamma/K_i$ such that

$$\mathbf{K}_n^{\forall} \vdash (\beta_1 \wedge \ldots \wedge \beta_k) \Rightarrow (\gamma \Rightarrow \psi[x/q]),$$

where $\gamma = \gamma_1 \wedge \ldots \wedge \gamma_n \wedge \neg\forall x\psi$. By Gen_K, we have that

$$\mathbf{K}_n^{\forall} \vdash K_i((\beta_1 \wedge \ldots \wedge \beta_k) \Rightarrow (\gamma \Rightarrow \psi[x/q])).$$

Applying axiom K, and using the fact that $K(\beta_1 \wedge \ldots \wedge \beta_k) \in \Gamma$ and Γ is maximal, we have that $K_i(\gamma \Rightarrow \psi[x/q]) \in \Gamma$ for all q. Since Γ is acceptable, it must be the case that $\forall x K_i(\gamma \Rightarrow \psi) \in \Gamma$. By the Barcan axiom, it follows that $K_i\forall x(\gamma \Rightarrow \psi) \in \Gamma$. Since γ is a sentence, applying $\mathrm{K}_{\forall}$, $\mathrm{N}_{\forall}$, K, and Gen_K, it follows that $K_i\forall x(\gamma \Rightarrow \psi) \Rightarrow K_i(\gamma \Rightarrow \forall x\psi)$ is provable in $\mathbf{K}_n^{\forall}$. Hence, $K_i(\gamma \Rightarrow \forall x\psi) \in \Gamma$. Thus, $\gamma \Rightarrow \forall x\psi \in \Gamma/K_i$. But this contradicts the consistency of $\Gamma/K \cup \{\gamma_1, \ldots, \gamma_n, \neg\forall x\psi\}$.

We now proceed much as in the proof of Lemma A.2. Given an enumeration $\psi_1, \psi_2, \ldots$ of the sentences in $\mathcal{S}_n^{\forall,K,X,A}$, we construct a sequence of $\mathbf{K}_n^{\forall}$-consistent sets of sentences $\Delta_0, \Delta_1, \ldots$. Let $\Delta_0 = \Gamma/K_i \cup \{\neg\varphi\}$. (Note that Lemma A.3 implies that Δ_0 is $\mathbf{K}_n^{\forall}$-consistent.) For all $k \geq 1$, if $\Delta_{k-1} \cup \{\psi_k\}$ is not $\mathbf{K}_n^{\forall}$-consistent, $\Delta_k = \Delta_{k-1}$. If $\Delta_{k-1} \cup \{\psi_k\}$ is $\mathbf{K}_n^{\forall}$-consistent, if ψ_k is not of the form $\neg\forall y\psi'$, then $\Delta_k = \Delta_{k-1} \cup \{\psi_k\}$, while if ψ_k is of the form $\neg\forall y\psi'$, then $\Delta_k = \Delta_{k-1} \cup \{\psi_k, \neg\psi'[y/q]\}$ for some $q \in \Phi$ such that $\Delta_{k-1} \cup \{\psi_k, \neg\psi'[y/q]\}$ is consistent. (Such a q exists by our earlier argument.) It is easy to see that $\Delta = \cup_n\Delta_n$ is the desired acceptable maximal $\mathbf{K}_n^{\forall}$-consistent set of sentences. ∎

We are now able to prove the following key lemma.

Lemma A.5: *If φ is a $\mathbf{K}_n^{\forall}$-consistent sentence, then φ is satisfiable in $\mathcal{M}_n^{\forall,\emptyset}(\Phi, \mathcal{X})$.*

Proof: Let $M^c = (S, \mathcal{K}_1, ..., \mathcal{K}_n, \mathcal{A}_1, \ldots, \mathcal{A}_n, \pi)$ be a canonical awareness structure constructed as follows

- $S = \{s_V : V$ is an acceptable maximal $\mathbf{K}_n^{\forall}$-consistent set of sentences$\}$;
- $\pi(s_V, p) = \begin{cases} 1 & \text{if } p \in V, \\ 0 & \text{if } p \notin V; \end{cases}$
- $\mathcal{A}_i(s_V) = \{\varphi : A_i\varphi \in V\}$;
- $\mathcal{K}_i(s_V) = \{s_W : V/K_i \subseteq W\}$.

We show as usual that if ψ is a sentence, then

$$(M^c, s_V) \models \psi \text{ iff } \psi \in V. \quad (2)$$

Note that this claim suffices to prove Lemma A.5 since, if φ is a $\mathbf{K}_n^{\forall}$-consistent sentence, by Lemma A.2, it is contained in an acceptable maximal $\mathbf{K}_n^{\forall}$-consistent set of sentences.

We prove (2) by induction of the depth of nesting of $\forall$, with a subinduction on the length of the sentence.

The base case is if ψ is a primitive proposition, in which case (2) follows immediately from the definition of π. For the inductive step, given ψ, suppose that (2) holds for all formulas ψ' such that either the depth of nesting for $\forall$ in ψ' is less than that in ψ, or the depth of nesting is the same, and ψ' is shorter than ψ. We proceed by cases on the form of ψ.

- If ψ has the form $\neg\psi'$ or $\psi_1 \wedge \psi_2$, then the result follows easily from the inductive hypothesis.
- If ψ has the form $A_i\psi'$, then note that φ' is a sentence and $(M^c, s_V) \models A_i\psi'$ iff $\psi' \in \mathcal{A}_i(s_V)$ iff $A_i\psi' \in V$.
- If ψ has the form $K_i\psi'$, then if $\psi \in V$, then $\psi' \in W$ for every W such that $s_W \in \mathcal{K}_i(s_V)$. By the induction hypothesis, $(M^c, s_W) \models \psi'$ for every $s_W \in \mathcal{K}_i(s_V)$, so $(M^c, s_V) \models K_i\psi'$. If $\psi \notin V$, then $\neg\psi \in V$ since V is a maximal $\mathbf{K}_n^{\forall}$-consistent set. By Lemma A.4, there exists

an acceptable maximal $\mathbf{K}_n^{\forall}$-consistent set of sentences W such that $(V/K_i \cup \{\neg\psi'\}) \subseteq W$. By the induction hypothesis, $(M^c, s_W) \not\models \psi'$. Thus, $(M^c, s_V) \not\models K_i\psi'$.

- If ψ has the form $X_i\psi'$, the argument is immediate from the preceding two cases and the observation that $(M, s_V) \models X_i\psi'$ iff both $(M, s_V) \models K_i\psi'$ and $(M, s_V) \models A'_\psi$, while $X_i\psi' \in V$ iff both $K_i\psi' \in V$ and $A_i\psi' \in V$.
- Finally, suppose that $\psi = \forall x\psi'$. If $\psi \in V$ then, by axiom $1_\forall$, $\psi'[x/\varphi] \in V$ for all $\varphi \in \mathcal{L}_n^{K,X,A}(\Phi)$. The depth of nesting of $\psi'[x/\varphi]$ is less than that of $\forall x\psi'$, so by the induction hypothesis $(M, s_V) \models \psi'[x/\varphi]$ for all $\varphi \in \mathcal{L}_n^{K,X,A}(\Phi)$. By definition, $(M, s_V) \models \psi$, as desired. If $\psi \notin V$ then, since V is acceptable, there exists a primitive proposition $q \in \Phi$ such that $\psi'[x/q] \notin V$. By the induction hypothesis, $(M^c, s_V) \not\models \psi'[x/q]$. Thus, $(M^c, s_V) \not\models \psi$, as desired.

∎

To finish the completeness proof, suppose that φ is valid in $\mathcal{M}_n^{\forall,\emptyset}(\Phi, \mathcal{X})$. Then, consider two cases: (1) φ is a sentence, and (2) φ is not a sentence. If (1), then $\neg\varphi$ is a sentence and is not satisfiable in $\mathcal{M}_n^{\forall,\emptyset}(\Phi, \mathcal{X})$. So, by Lemma A.5, $\neg\varphi$ is not $\mathbf{K}_n^{\forall}$-consistent. Thus, φ is provable in $\mathbf{K}_n^{\forall}$. If (2) and $\{x_1, \ldots, x_k\}$ is the set of free variables in φ, then $\forall x_1 \ldots \forall x_k\varphi$ is a valid sentence. Thus, by case (1), $\forall x_1 \ldots \forall x_k\varphi$ is provable in $\mathbf{K}_n^{\forall}$. Applying $1_\forall$ repeatedly it follows that φ is provable in $\mathbf{K}_n^{\forall}$, as desired. ∎

References

Bull, R. 1969. On modal logic with propositional quantifiers. *Journal of Symbolic Logic* 34:257–263.

Dekel, E.; Lipman, B.; and Rustichini, A. 1998. Standard state-space models preclude unawareness. *Econometrica* 66:159–173.

Engelhardt, K.; van der Meyden, R.; and Moses, Y. 1998. Knowledge and the logic of local propositions. In *Theoretical Aspects of Rationality and Knowledge: Proc. Seventh Conference (TARK 1998)*. 29–41.

Fagin, R., and Halpern, J. Y. 1988. Belief, awareness, and limited reasoning. *Artificial Intelligence* 34:39–76.

Fagin, R., and Halpern, J. Y. 1994. Reasoning about knowledge and probability. *Journal of the ACM* 41(2):340–367.

Fagin, R.; Halpern, J. Y.; Moses, Y.; and Vardi, M. Y. 1995. *Reasoning about Knowledge*. Cambridge, Mass.: MIT Press. A slightly revised paperback version was published in 2003.

Fagin, R.; Halpern, J. Y.; and Megiddo, N. 1990. A logic for reasoning about probabilities. *Information and Computation* 87(1/2):78–128.

Fagin, R.; Halpern, J. Y.; and Vardi, M. Y. 1992. What can machines know? On the properties of knowledge in distributed systems. *Journal of the ACM* 39(2):328–376.

Feinberg, Y. 2004. Subjective reasoning—games with unawareness. Technical Report Resarch Paper Series #1875, Stanford Graduate School of Business.

Fine, K. 1970. Propositional quantifiers in modal logic. *Theoria* 36:336–346.

Ghirardato, P. 2001. Coping with ignorance: unforeseen contingencies and non-additive uncertainty. *Economic Theory* 17:247–276.

Halpern, J. Y., and Lakemeyer, G. 2001. Multi-agent only knowing. *Journal of Logic and Computation* 11(1):41–70.

Halpern, J. Y., and Rêgo, L. C. 2005. Interactive unawareness revisited. In *Theoretical Aspects of Rationality and Knowledge: Proc. Tenth Conference (TARK 2005)*, 78–91.

Halpern, J. Y., and Rêgo, L. C. 2006a. Characterizing the NP-PSPACE gap in the satisfiability problem for modal logic. Available at arxiv.org/cs.LO/0603019.

Halpern, J. Y., and Rêgo, L. C. 2006b. Extensive games with possibly unaware players. In *Proc. Fifth International Joint Conference on Autonomous Agents and Multiagent Systems*.

Halpern, J. Y., and Rêgo, L. C. 2006c. Reasoning about knowledge of unawareness. Available at arxiv.org/cs.LO/0603020.

Halpern, J. Y. 2001. Alternative semantics for unawareness. *Games and Economic Behavior* 37:321–339.

Heifetz, A.; Meier, M.; and Schipper, B. 2003. Multi-person unawareness. In *Theoretical Aspects of Rationality and Knowledge: Proc. Ninth Conference (TARK 2003)*, 148–158. An extended version with the title "Interactive unawareness" will appear in the *Journal of Economic Theory*.

Hintikka, J. 1962. *Knowledge and Belief*. Ithaca, N.Y.: Cornell University Press.

Hughes, G., and Cresswell, M. 1996. *A New Introduction to Modal Logic*. London: Routledge.

Kaplan, D. 1970. S5 with quantifiable propositional variables. *Journal of Symbolic Logic* 35:355.

Kripke, S. 1959. A completeness theorem in modal logic. *Journal of Symbolic Logic* 24:1–14.

Modica, S., and Rustichini, A. 1994. Awareness and partitional information structures. *Theory and Decision* 37:107–124.

Modica, S., and Rustichini, A. 1999. Unawareness and partitional information structures. *Games and Economic Behavior* 27(2):265–298.

Prior, A. N. 1956. Modality and quantification in S5. *Journal of Symbolic Logic* 21:60–62.