Alternative Semantics for Unawareness

Joseph Y. Halpern*

Cornell University
Computer Science Department
Ithaca, NY 14853
E-mail: halpern@cs.cornell.edu http://www.cs.cornell.edu/home/halpern

Modica and Rustichini [1994] provided a logic for reasoning about knowledge where agents may be unaware of certain propositions. However, their original approach had the unpleasant property that nontrivial unawareness was incompatible with partitional information structures. More recently, Modica and Rustichini [1999] have provided an approach that allows for nontrivial unawareness in partitional information structures. Here it is shown that their approach can be viewed as a special case of a general approach to unawareness considered by Fagin and Halpern [1988]. *Journal of Economic Literature* Classification Numbers: D80, D83.

1. INTRODUCTION

The standard approach to reasoning about knowledge [Fagin, Halpern, Moses, and Vardi 1995] implicitly assumes that agents are (commonly known to be) aware of all the relevant propositions. However, in decision theory under uncertainty, we must often deal with unforeseen contingencies. (See [Dekel, Lipman, and Rusticchini 1998] for a discussion of issues related to unforeseen contingencies and further references.) Thus, we would like to be able to model agents who are not necessarily aware of all the relevant propositions.

In the economics literature, an initial attempt to do this was made by Modica and Rustichini [1994] (MR from now on). They provided a logic where awareness was defined in terms of knowledge. An agent was aware of p if he either knew p or knew he did not know p. Unfortunately, the original MR approach had the unpleasant property that nontrivial unawareness was incompatible with partitional information structures. More recently [1999], MR provided an approach that allows for nontrivial unawareness in partitional information structures.

*Supported in part by NSF under grant IRI-96-25901.

As MR observe, in the Artificial Intelligence/Computer Science literature, Fagin and Halpern [1988] (FH from now on) had earlier introduced a model of knowledge and unawareness. However, in this model, unawareness is introduced as a separate modal operator, independent of knowledge, rather than being defined in terms of knowledge. Nevertheless, as I show here, by making some assumptions on the connection between knowledge and awareness (assumptions that are actually discussed explicitly by FH), the MR definition becomes a special case of the FH definition.

The rest of this paper is organized as follows. In the next section, I briefly review the FH definitions and in Section 3, I briefly review the MR definitions. The fact that the MR definitions can be viewed as a special case of the FH definitions is proved in Section 4. In Section 5, complete axiomatizations are provided for the FH notions of knowledge and (un)awareness. I do not give much in the way of motivation here for the FH and MR definitions here; I would encourage the reader to consult the original papers for details.

2. THE FH DEFINITIONS

The intuition behind the FH approach is that there is *implicit* knowledge—the notion traditionally considered in the AI and economics literature—where agents are aware of all propositions and can do perfect reasoning with respect to these propositions, and *explicit knowledge* which, roughly speaking, is intended to capture those conclusions of which agents are explicitly aware. Explicit knowledge implies implicit knowledge, but the converse does not hold in general.

To capture this intuition, FH start with a set $\Phi = \{p, q, \ldots\}$ of primitive propositions, and build more complicated formulas by closing under the propositional connectives $\wedge$ and $\neg$ and the modal operators K (for implicit knowledge), X (for explicit knowledge) and A (for awareness).[1] Let $\mathcal{L}^{KXA}(\Phi)$ denote the resulting language; let $\mathcal{L}^{XA}(\Phi)$, (resp., $\mathcal{L}^{KA}(\Phi)$; $\mathcal{L}^{K}(\Phi)$; $\mathcal{L}^{X}(\Phi)$) denote the sublanguages of $\mathcal{L}^{KXA}(\Phi)$ where the only modal operators are X and A (resp., K and A; K; X). I typically omit the parenthesized Φ if it does not play an important role. As usual, $\varphi \vee \psi$ is taken to be an abbreviation for $\neg(\neg\varphi \wedge \neg\psi)$ and $\varphi \Rightarrow \psi$ taken to be an abbreviation for $\neg\varphi \vee \psi$.

An *awareness structure* is a tuple $M = (S, \rho, \mathcal{K}, \mathcal{A})$, where S is a set of states or possible worlds, ρ associates with each primitive proposition $p \in \Phi$

[1] In [Fagin and Halpern 1988], L and B are used for implicit and explicit knowledge, respectively. I am following the notation of [Fagin, Halpern, Moses, and Vardi 1995] here. For consistency with MR, I consider only the single-agent case here, although in fact, FH consider the multiagent case.

an event $\rho(p) \subseteq S$ (intuitively, $\rho(p)$ is the set of states where p is true),[2] $\mathcal{K}$ is a binary relation on S, and $\mathcal{A}$ associates with each state $s \in S$ a set $\mathcal{A}(s)$ of formulas. Intuitively, the formulas in $\mathcal{A}(s)$ are the formulas that the agent is aware of at state s. The binary relation $\mathcal{K}$ describes the agent's possibility relation; if $(s,t) \in \mathcal{K}$ then in state s the agent considers state t possible. If $\mathcal{K}$ is an equivalence relation, then $(S, \mathcal{K})$ gives us the standard partitional model, familiar in the economics literature. By defining $\mathcal{K}(s) = \{t : (s,t) \in \mathcal{K}\}$, we can view $\mathcal{K}$ as defining what his been called a *possibility correspondence* in the economics literature [Geanakoplos 1989], i.e., a function from states to sets of states. An awareness structure is *reflexive* (resp., *reflexive and transitive*; *partitional*) if $\mathcal{K}$ is reflexive (resp., reflexive and transitive; an equivalence relation). Let $\mathcal{M}$ (resp., $\mathcal{M}^r$; $\mathcal{M}^{rt}$; $\mathcal{M}^{rst}$) denote the class of all (resp., reflexive; reflexive and transitive; partitional) awareness structures.

The awareness function $\mathcal{A}$ gives awareness structures a syntactic component. $\mathcal{A}(s)$ can be arbitrary (although I shall shortly consider some restrictions on it). In general, it is possible for both φ and $\neg\varphi$ to be in $\mathcal{A}(s)$, for only one of φ and $\neg\varphi$ to be in $\mathcal{A}(s)$, or for neither φ nor $\neg\varphi$ to be in $\mathcal{A}(s)$. It is also possible, for example, that $\varphi \wedge \psi$ is in $\mathcal{A}(s)$ but neither ψ nor φ is in $\mathcal{A}(s)$.

I write $(M,s) \models_{FH} \varphi$ if φ is true at state s in awareness structure M. (The subscript FH is used to distinguish this definition from the one given by MR, described in the next section.) The truth relation $\models_{FH}$ is defined inductively. The definition in the case of primitive propositions, conjunctions, negations, and for formulas $K\varphi$ is standard. The definition for $A\varphi$ just says that $A\varphi$ is true at s exactly if φ is one of the formulas in $\mathcal{A}(s)$ (i.e., the formulas that the agent is aware of at state s). Finally, $X\varphi$ is defined as $K\varphi \wedge A\varphi$: the agent explicitly knows φ exactly if he is aware of φ and implicitly knows φ. You cannot have explicit knowledge about formulas of which you are not aware.

$(M,s) \models_{FH} p$ (for $p \in \Phi$) iff $s \in \rho(p)$
$(M,s) \models_{FH} \varphi \wedge \psi$ iff both $(M,s) \models_{FH} \varphi$ and $(M,s) \models_{FH} \psi$
$(M,s) \models_{FH} \neg\varphi$ iff $(M,s) \not\models_{FH} \varphi$
$(M,s) \models_{FH} K\varphi$ iff $(M,t) \models_{FH} \varphi$ for all $t \in \mathcal{K}(s)$
$(M,s) \models_{FH} A\varphi$ iff $\varphi \in \mathcal{A}(s)$
$(M,s) \models_{FH} X\varphi$ iff $(M,s) \models_{FH} A\varphi$ and $(M,s) \models_{FH} K\varphi$.

It is easy to see that $\mathcal{L}^{KA}$ is equivalent in expressive power to $\mathcal{L}^{KXA}$. Since X is definable in terms of K and A, it does not add anything in the presence

[2] I follow MR in using ρ for this component rather than π, as used in [Fagin and Halpern 1988; Fagin, Halpern, Moses, and Vardi 1995]. In general, the notation used here is a compromise between that of FH and MR.

of these two other operators. However, I focus here on the sublanguages $\mathcal{L}^X$ and $\mathcal{L}^{XA}$, where the X operator clearly plays a nontrivial role.

What exactly does awareness mean and what should the relationship between knowledge and awareness be? MR have one particular interpretation in mind, which leads them to define $A\varphi$ as $X\varphi \vee X\neg X\varphi$.[3] Given their assumptions about the X operator a number of properties of awareness follow, such as the fact that an agent is aware of φ iff she is aware of $\neg\varphi$. In FH, awareness is taken as a primitive notion, open to a number of interpretations. The appropriate application depends on the application. One possible interpretation is certainly that suggested by MR, but it might also be reasonable to consider a "computational" notion of awareness, where awareness of φ corresponds to "being able to compute the truth of φ (perhaps within some time or space bounds)". Such an interpretation leads to a resource-bounded notion of knowledge. Under this interpretation, an agent may not be aware of tautologies, for example, and the agent may be aware of φ without being aware of $\neg\varphi$ (adding a negation may be just enough to make the formula too long for the agent to deal with).

Of course, once we have a concrete interpretation in mind, it will suggest some restrictions on $\mathcal{A}$ (or, equivalently, some axioms that awareness should satisfy). Some typical restrictions on $\mathcal{A}$ considered by FH include the following.

- Awareness is *closed under subformulas* if $\varphi \in \mathcal{A}(s)$ and ψ is a subformula of φ, then $\psi \in \mathcal{A}(s)$. Note that this makes sense if we are reasoning about a computer program that will never compute the truth of a formula unless it has computed the truth of all its subformulas. But it is also easy to imagine a program that knows that $\varphi \vee \neg\varphi$ is true without needing to compute the truth of φ. Note that if an agent's awareness is closed under subformulas, then the agent's knowledge is closed under implication. More precisely, it is easy to check that $(X\varphi \wedge X(\varphi \Rightarrow \psi)) \Rightarrow X\psi$ is valid if the agent's awareness is closed under implication. This suggests that taking awareness to be closed under implication is inappropriate to the extent that one is thinking in terms of a computational notion of awareness.

- Awareness is *generated by primitive propositions* if, at each state s, there is some subset $\Phi(s) \subseteq \Phi$ of primitive propositions such that $\mathcal{A}(s)$ consists of exactly those formulas that mention only primitive propositions that appear in $\Phi(s)$. This seems much in the spirit of MR's intuition for awareness. Note that if awareness is generated by primitive propositions, then it is closed under subformulas.

[3]MR actually write $K\varphi \vee K\neg K\varphi$, but their K acts more like the explicit knowledge operator X of this paper.

• An agent is *self-reflective* if $\varphi \in \mathcal{A}(s)$ implies $A\varphi \in \mathcal{A}(s)$. This corresponds to the axiom $A\varphi \Rightarrow AA\varphi$. Again, note that if awareness is generated by primitive propositions, then an agent will be self-reflective in this sense.

• An agent *knows which formulas he is aware of* if $(s,t) \in \mathcal{K}$, then $\mathcal{A}(s) = \mathcal{A}(t)$. This corresponds to the axioms $A\varphi \Rightarrow KA\varphi$ and $\neg A\varphi \Rightarrow K\neg A\varphi$: that is, if the agent is aware of φ, then she knows it and if she is not aware of φ, she knows that too.

If awareness is generated by primitive propositions, then it is characterized by MR's definition; that is, $A\varphi$ can be viewed as an abbreviation of $X\varphi \vee (\neg X\varphi \wedge X\neg X\varphi)$.

Lemma 2.1. *If M is a partitional awareness structure where awareness is generated by primitive propositions, then*

$$M \models_{FH} A\varphi \Leftrightarrow (X\varphi \vee (\neg X\varphi \wedge X\neg X\varphi)).$$

Proof. Suppose $(M,s) \models_{FH} A\varphi$. Clearly we have $(M,s) \models_{FH} K\varphi$ or $(M,s) \models_{FH} \neg K\varphi$. If $(M,s) \models_{FH} K\varphi$, then it is immediate from the definition that $(M,s) \models_{FH} X\varphi$. If $(M,s) \models_{FH} \neg K\varphi$, then it is immediate from the definition that $(M,s) \models_{FH} \neg X\varphi$. Moreover, since we are restricting to partitional models, by negative introspection we have that $(M,s) \models_{FH} K\neg K\varphi$. Thus, for all $t \in \mathcal{K}(s)$, we have $(M,t) \models_{FH} \neg K\varphi$, and hence $(M,t) \models_{FH} \neg X\varphi$ It follows that $(M,s) \models_{FH} K\neg X\varphi$. Since awareness is generated by primitive propositions in M and the same primitive propositions appear in both φ and $\neg X\varphi$, from $(M,s) \models_{FH} A\varphi$ it follows that $(M,s) \models_{FH} A\neg X\varphi$. Thus, if $(M,s) \models_{FH} \neg K\varphi$, we must have $(M,s) \models_{FH} X\neg X\varphi$. We have just shown that if $(M,s) \models_{FH} A\varphi$, then $(M,s) \models_{FH} X\varphi \vee (\neg X\varphi \wedge X\neg X\varphi)$.

The converse is even easier. Clearly if $(M,s) \models_{FH} X\varphi$, by definition, $(M,s) \models_{FH} A\varphi$. Since awareness is generated by primitive propositions, from $(M,s) \models_{FH} X\neg X\varphi$ it also follows that $(M,s) \models_{FH} A\varphi$. This completes the proof. ∎

Thus, if we restrict to partitional awareness structures where awareness is generated by primitive propositions, we can define A in terms of X as MR do. That is, $\mathcal{L}^{XA}$ is equivalent in expressive power to $\mathcal{L}^{X}$. An even stronger connection between the definitions of FH and MR is given in Section 4. To state it, I need to review the definitions of MR's semantic structures.

3. THE MR DEFINITIONS

MR also start with a set Φ of primitive propositions, and then close off under the modal operator X.[4] They give semantics to formulas in $\mathcal{L}^X$ using *generalized standard models* (GSMs). A GSM is a tuple $M = (S, S', \rho, \mathcal{P}, \pi)$ whose components satisfy the following conditions:

- S, as before, is a set of states.
- S' is also a set of states; moreover $S' = \cup_{\Psi \subseteq \Phi} S'_\Psi$, where the sets S'_Ψ are disjoint. Intuitively, the agent is aware of only the formulas with propositions in Ψ at the states in S'_Ψ.
- $\rho(p)$, as before, is a subset of S, for each $p \in \Phi$.
- $\mathcal{P}$ is a generalized possibility correspondence; $\mathcal{P}(s)$ is a subset of S' for each $s \in S$.
- π is a projection from S to S' (i.e., an onto map from S to S') such that (1) if $\pi(s) = \pi(t) \in S'_\Psi$ then (a) s and t agree on the truth values of all primitive propositions in Ψ, that is, $s \in \rho(p)$ iff $t \in \rho(p)$ for all $p \in \Psi$ and (b) $\mathcal{P}(s) = \mathcal{P}(t)$ and (2) if $\pi(s) \in S'_\Psi$ then $\mathcal{P}(s) \subseteq S'_\Psi$.

We can extend $\mathcal{P}$ to a map $\mathcal{P}^*$ defined on $S \cup S'$ as follows: if $s' \in S'$ and $\pi(s) = s'$, define $\mathcal{P}^*(s') = \mathcal{P}(s)$. Condition 1(b) on π guarantees that this extension is well defined. A GSM is *reflexive* (resp., *reflexive and transitive*; *partitional*) if $\mathcal{P}^*$ restricted to S' is reflexive (resp., reflexive and transitive; partitional). Similarly, we can extend ρ to a map ρ^* defined on $S \cup S'$ by defining $\rho^*(p) = \rho(p) \cup \cup_{\Psi \subseteq \Phi} \{s' \in S'_\Psi : p \in \Psi, \pi^{-1}(s') \subseteq \rho(p)\}$. That is, if $s' \in S_\Psi$, then $s' \in \rho(p)$ iff $p \in \Psi$ and $s \in \rho(p)$ for some (and hence, by condition 1(a), all) $s \in S$ such that $\pi(s) = s'$.

We can now define a satisfaction relation $\models_{MR}$ on GSMs. States in both S and S' are allowed on the left-hand side of $\models_{MR}$, but they are treated differently in the case of negation:[5]

$(M, s) \models_{MR} p$ (for $p \in \Phi$) iff $s \in \rho^*(p)$
$(M, s) \models_{MR} \varphi \wedge \psi$ iff both $(M, s) \models_{MR} \varphi$ and $(M, s) \models_{MR} \psi$
$(M, s) \models_{MR} \neg\varphi$ iff $(M, s) \not\models_{MR} \varphi$ and either $s \in S$ or $s \in S'_\Psi$ and $\varphi \in \mathcal{L}^X(\Psi)$
$(M, s) \models_{MR} X\varphi$ iff $(M, t) \models_{MR} \varphi$ for all $t \in \mathcal{P}^*(s)$.

MR extend the definition of $\models$ to $\mathcal{L}^{XA}$ by taking $A\varphi$ to be an abbreviation for $X\varphi \vee X\neg X\neg\varphi$.

The following simple lemma, which will be of use in the proof of the main result, casts some light on the connection between the states s and $\pi(s)$.

---

[4]MR use lower case k and a for knowledge and unawareness. I am using upper case for consistency with the notation of FH and explicit knowledge (X) rather than implicit knowledge (K), since the notion of knowledge in [Modica and Rustichini 1999] is closer to FH's explicit knowledge. MR also have an operator $\Diamond$, but since $\Diamond\varphi$ is defined as an abbreviation for $\neg X\neg\varphi \wedge A\varphi$, I ignore $\Diamond$ here, with no loss of generality.

[5]My presentation is slightly different from that of MR, but is easily seen to be equivalent.

LEMMA 3.1. *Given a GSM* $M = (S, S', \rho, \mathcal{P}, \pi)$ *and* $s \in S$ *such that* $\pi(s) \in S'_\Psi$*, then (a) if* $\varphi \in \mathcal{L}^X(\Psi)$*, then* $(M, s) \models_{MR} \varphi$ *iff* $(M, \pi(s)) \models_{MR} \varphi$ *and (b) if* $\varphi \notin \mathcal{L}^X(\Psi)$*, then* $(M, \pi(s)) \not\models_{MR} \varphi$*.*

Proof. A straightforward induction on the structure of φ; left to the reader. ▌

4. RELATING GSMS AND AWARENESS STRUCTURES

Awareness (resp., an awareness structure M) is said to be *propositionally determined* if it (resp., the awareness function in M) is generated by primitive propositions and the agent knows what formulas he is aware of. In this section, I prove that partitional GSMs and propositionally determined awareness structures are equivalent in a strong sense. Assume that we are working over a fixed set Φ of primitive propositions.

THEOREM 4.1.

(a)Given a propositionally determined awareness structure $M = (S, \rho, \mathcal{K}, \mathcal{A})$ *there is a GSM* $M' = (S, S', \rho, \mathcal{P}, \pi)$ *such that for all formulas* $\varphi \in \mathcal{L}^X(\Phi)$*, we have*

$$(M, s) \models_{FH} \varphi \text{ iff } (M', s) \models_{MR} \varphi.$$

Moreover, if M *is reflexive (resp., reflexive and transitive, partitional), then so is* M'*.*

(b)Given a GSM $M = (S, S', \rho, \mathcal{P}, \pi)$*, there is a propositionally determined awareness structure* $M' = (S, \rho, \mathcal{K}, \mathcal{A})$ *such that for all formulas* $\varphi \in \mathcal{L}^X(\Phi)$*, we have*

$$(M, s) \models_{MR} \varphi \text{ iff } (M', s) \models_{FH} \varphi.$$

Moreover, if M *is reflexive (resp., reflexive and transitive, partitional), then so is* M'*.*

Proof. For part (a), given a propositional awareness structure $M = (S, \rho, \mathcal{K}, \mathcal{A})$, for each state $s \in S$, define $\Psi(s) = \mathcal{A}(s) \cap \Phi$. Let $S' = S$ be a disjoint copy of S (where $s' \in S'$ is the copy of $s \in S$), let $S'_\Psi = \{s' \in S' : \Psi(s) = \Psi\}$, let $\mathcal{P}(s) = \{t' : (s, t) \in \mathcal{K}\}$, and let $\pi(s) = s'$. Let $M' = (S, S', \rho, \mathcal{P}, \pi)$. Since M is propositionally determined, it follows that if $\pi(s) \in S'_\Psi$, then $\mathcal{P}(s) \subseteq S'_\Psi$. Thus, M' is a GSM. Moreover, it is immediate that M' is reflexive (resp., reflexive and transitive, partitional) if M is.

I now show by induction on the structure of formulas that for all states $s \in S$ and formulas $\varphi \in \mathcal{L}^X(\Phi)$, we have

$$(M,s) \models_{FH} \varphi \text{ iff } (M',s) \models_{MR} \varphi.$$

If φ is the primitive proposition p, this is immediate from the definitions. If φ is of the form $\varphi_1 \wedge \varphi_2$ or $\neg\varphi'$, then it follows immediately from the induction hypothesis. Now suppose that φ is of the form $X\varphi'$. Let Φ_φ consist of all the primitive propositions that appear in φ. Then we have $(M,s) \models_{FH} X\varphi'$ iff $(M,t) \models_{FH} \varphi'$ for all $t \in \mathcal{K}(s)$ and (since awareness in M is propositionally determined) $\Phi_\varphi \subseteq \mathcal{A}(s)$. By the induction hypothesis, the latter condition is true iff $(M,t) \models_{MR} \varphi'$ for all $t \in \mathcal{K}(s)$ and $\pi(s) \in S'_\Psi$ for some $\Psi \supseteq \Phi_\varphi$. By Lemma 3.1, it follows that this is true iff $(M,\pi(t)) \models_{MR} \varphi'$ for all $t \in \mathcal{K}(s)$. But this, in turn, is true iff $(M,s) \models_{MR} X\varphi'$.

For part (b), given a GSM $M = (S, S', \rho, \mathcal{P}, \pi)$, where $S' = \cup_\Psi S'_\Psi$, define an awareness structure $M' = (S, \rho, \mathcal{K}, \mathcal{A})$, where $\mathcal{A}(s)$ is generated by the primitive propositions in Ψ if $\pi(s) \in S'_\Psi$ and $\mathcal{K}(s) = \{t : \pi(t) \in \mathcal{P}(s)\}$. Since $(s,t) \in \mathcal{K}$ and $\pi(s) \in S'_\Psi$ implies that $\pi(t) \in S'_\Psi$, it easily follows that M' is a propositionally determined awareness structure. Moreover, M' is reflexive (resp., reflexive and transitive, partitional) if M is.

I now prove by induction on the structure of φ that $(M,s) \models \varphi$ iff $(M',s) \models \varphi$ for all states $s \in S$. Again, the argument is immediate in the case that φ is a primitive proposition, conjunction, or negation. Finally, the argument in the case of $X\varphi'$ is essentially identical to that given in part (a). ▮

Theorem 4.1 shows that, in a precise sense, GSMs and propositionally determined awareness structures are equivalent, at least as far as the language $\mathcal{L}^X$ is concerned.[6] This equivalence does not extend to the language $\mathcal{L}^{XA}$ in general. In GSMs, $A\varphi$ is, by definition, equivalent to $X\varphi \vee X\neg X\varphi$. By Lemma 2.1, this equivalence holds as well in partitional awareness structures, but it does not hold in general. That is, in partitional awareness structures, Theorem 4.1 holds, not just for the language $\mathcal{L}^X$, but for the language $\mathcal{L}^{XA}$ as well. However, in general, Theorem 4.1 holds only for the language $\mathcal{L}^X$.

5. AXIOMATIZING AWARENESS STRUCTURES

The axioms that characterize awareness structures depend on three things: (1) the choice of language, (2) the assumptions made about the $\mathcal{K}$ operator,

[6]For simplicity, this theorem was proved for only four different assumptions about the possibility correspondence. However, the proof itself clearly shows that the equivalence continues for much more general assumptions about the possibility correspondence.

and (3) the assumptions made about awareness. I discuss axiomatizations with respect to some of the assumptions I have made in this paper. Some of the results, as noted below, are taken from [Fagin, Halpern, Moses, and Vardi 1995, Section 9.5].

Since the notion of axiomatization of a logic is somewhat different from the standard notion of axiomatizing an operator familiar in the economics literature, I now briefly review the relevant definitions. An *axiom system* AX consists of a collection of *axioms* and *inference rules.* An axiom is a formula, and an inference rule has the form "from $\varphi_1, \ldots, \varphi_k$ infer ψ," where $\varphi_1, \ldots, \varphi_k, \psi$ are formulas. Typically (and, in particular, in this paper), the axioms are all instances of *axiom schemes.* Thus, for example, an axiom scheme such as $K_i\varphi \Rightarrow \varphi$ defines an infinite collection of axioms, one for each choice of φ. A *proof* in AX consists of a sequence of formulas, each of which is either an axiom in AX or follows by an application of an inference rule. A proof is said to be a *proof of the formula* φ if the last formula in the proof is φ. We say φ is *provable in AX*, and write $\mathrm{AX} \vdash \varphi$, if there is a proof of φ in AX; similarly, we say that φ is *consistent with AX* if $\neg\varphi$ is not provable in AX.

Given a class $\mathcal{M}'$ of (awareness) structures, a formula φ is *valid in* $\mathcal{M}'$ if $(M, s) \models \varphi$ for every structure $M \in \mathcal{M}'$ and state s in M. An axiom system AX is said to be *sound* for a language $\mathcal{L}$ with respect to a class $\mathcal{M}'$ of structures if every formula in $\mathcal{L}$ provable in AX is valid with respect to every structure in $\mathcal{M}'$. The system AX is *complete* for $\mathcal{L}$ with respect to $\mathcal{M}'$ if every formula in $\mathcal{L}$ that is valid in $\mathcal{M}'$ is provable in AX. We think of AX as characterizing the class $\mathcal{M}'$ if it provides a sound and complete axiomatization of that class. Soundness and completeness provide a connection between the *syntactic* notion of provability and the *semantic* notion of validity.

5.1. Axiomatizing the language $\mathcal{L}^K$: a review

I begin with a review of standard axiomatizations for the language $\mathcal{L}^K$, since they form the basis of the axiomatizations with awareness.

Consider the following axioms:

Prop. All substitution instances of tautologies of propositional calculus
K. $(K\varphi \wedge K(\varphi \Rightarrow \psi)) \Rightarrow K\psi$, (Distribution Axiom)
T. $K\varphi \Rightarrow \varphi$, (Knowledge Axiom)
4. $K\varphi \Rightarrow KK\varphi$, (Positive Introspection Axiom)
5. $\neg K\varphi \Rightarrow K\neg K\varphi$, (Negative Introspection Axiom)
MP. From φ and $\varphi \Rightarrow \psi$ infer ψ (Modus ponens)
Gen. From φ infer $K\varphi$ (Knowledge Generalization)

The system with axioms and rules Prop, K, MP, and Gen has been called **K**. If we add T to **K**, we get the axiom system **T**; if we add 4 to **T**, we get **S4**; if we add 5 to **S4**, we get **S5**. (Many other systems can also be formed [Chellas 1980]; these are the four I focus on here.) The following result is well known (see, for example, [Chellas 1980; Fagin, Halpern, Moses, and Vardi 1995] for proofs.)

THEOREM 5.1. *For formulas in the language $\mathcal{L}^K$:*

*(a)***K** *is a sound and complete axiomatization with respect to $\mathcal{M}$,*
*(c)***T** *is a sound and complete axiomatization with respect to $\mathcal{M}^r$,*
*(c)***S4** *is a sound and complete axiomatization with respect to $\mathcal{M}^{rt}$,*
*(d)***S5** *is a sound and complete axiomatization with respect to $\mathcal{M}^{rst}$.*

While the axiomatizations above are the standard ones discussed in the literature, there are a number of equivalent formulations that will turn out to be useful for characterizing awareness. Consider the following rules of inference:

RK. From $\varphi_1 \wedge \ldots \wedge \varphi_n \Rightarrow \psi$ infer $K\varphi_1 \wedge \ldots \wedge K\varphi_n \Rightarrow K\psi$ $(n \geq 0)$. (The special case where $n = 0$ is just Gen.)
RK4. From $\varphi_1 \wedge \ldots \wedge \varphi_n \wedge K\sigma \Rightarrow \psi$ infer $K\varphi_1 \wedge \ldots \wedge K\varphi_n \wedge K\sigma \Rightarrow K\psi$ $(n \geq 0)$.
RK5. From $\varphi_1 \wedge \ldots \wedge \varphi_n \wedge K\sigma \wedge \neg K\tau \Rightarrow \psi$ infer $K\varphi_1 \wedge \ldots \wedge K\varphi_n \wedge K\sigma \wedge \neg K\tau \Rightarrow K\psi$ $(n \geq 0)$.

The following result is easy to prove (very similar results are proved in [Chellas 1980]).

PROPOSITION 5.1.

*(a)***K** *is equivalent to the system $\{Prop, MP, RK\}$.*
*(b)***T** *is equivalent to the system $\{Prop, MP, RK, T\}$.*
*(c)***S4** *is equivalent to the system $\{Prop, MP, RK4, T\}$.*
*(d)***S5** *is equivalent to the system $\{Prop, MP, RK5, T\}$.*

5.2. Axiomatizing $\mathcal{L}^{KXA}$

Axiomatizing $\mathcal{L}^{KXA}$ (with no assumptions on awareness) is easy.[7] The only axiom we need add to the basic axioms for K is

A0. $X\varphi \equiv K\varphi \wedge A\varphi$.

[7] Most of the material in this section is essentially in [Fagin, Halpern, Moses, and Vardi 1995, Section 9.5], although no theorems are stated there.

Axiomatizing the notions of awareness that I have been considering is also easy. Consider the following axioms.

A1. $A(\varphi \wedge \psi) \Rightarrow A\varphi \wedge A\psi$.
A2. $A\neg\varphi \Rightarrow A\varphi$.
A3. $AX\varphi \Rightarrow A\varphi$.
A4. $AA\varphi \Rightarrow A\varphi$.
A5. $AK\varphi \Rightarrow A\varphi$.
A6. $A\varphi \Rightarrow A\neg\varphi$
A7. $A\varphi \wedge A\psi \Rightarrow A(\varphi \wedge \psi)$.
A8. $A\varphi \Rightarrow AX\varphi$.
A9. $A\varphi \Rightarrow AA\varphi$.
A10. $A\varphi \Rightarrow AK\varphi$.
A11. $A\varphi \Rightarrow KA\varphi$.
A12. $\neg A\varphi \Rightarrow K\neg A\varphi$.

I said earlier that awareness being closed under subformulas corresponded to A1–A5, awareness being generated by primitive propositions corresponded to A1–A10, self-reflective agents were characterized by A9, and knowledge of awareness corresponded to A10 and A11. I now make this precise. Let $\mathcal{C}_1 = \{\text{A1}, \ldots, \text{A5}\}$, let $\mathcal{C}_2 = \mathcal{C}_1 \cup \{\text{A6}, \ldots, \text{A10}\}$, let $\mathcal{C}_3 = \{A9\}$, and let $\mathcal{C}_4 = \{A11, A12\}$. I use the superscripts *csub*, *gpp*, *sr*, and *ka* to restrict to awareness structures where awareness is closed under subformulas, awareness is generated by primitive propositions, the agent is self-reflective, and the agent knows which formulas he is aware of, respectively. Thus, for example, $\mathcal{M}^{rst,csub,ka}$ consists of partitional awareness structures where awareness is closed under subformulas and the agent knows which formulas he is aware of.

THEOREM 5.2. *Let $\mathcal{C}$ be a (possibly empty) subset of $\{\mathcal{C}_1, \mathcal{C}_2, \mathcal{C}_3, \mathcal{C}_4\}$ and let C be the corresponding subset of $\{$csub, gpp,sr,ka$\}$. For formulas in the language $\mathcal{L}^{KXA}$*

*(a)***K** $\cup\{A0\} \cup \mathcal{C}$ *is a sound and complete axiomatization with respect to* $\mathcal{M}^{C}$*,*

*(b)***T** $\cup\{A0\} \cup \mathcal{C}$ *is a sound and complete axiomatization with respect to* $\mathcal{M}^{r,C}$*,*

*(c)***S4** $\cup\{A0\} \cup \mathcal{C}$ *is a sound and complete axiomatization with respect to* $\mathcal{M}^{rt,C}$*,*

*(d)***S5** $\cup\{A0\} \cup \mathcal{C}$ *is a sound and complete axiomatization with respect to* $\mathcal{M}^{rst,C}$*.*

Proof. A straightforward modification of the proof of Theorem 5.1; I omit details here. ▌

5.3. Axiomatizing $\mathcal{L}^{XA}$

Perhaps the most interesting issues arise in providing an axiomatization for $\mathcal{L}^{XA}$. This forces us to consider the properties of explicit knowledge directly, without being able to take the indirect route through implicit knowledge and awareness provided by A0. The interplay between assumptions about awareness and axioms for explicit knowledge is far more significant here than for the language $\mathcal{L}^{KXA}$. To start with, consider propositionally determined awareness.

Let the axioms K_X, T_X, and 4_X be identical to K, T, and 4, respectively, except that K is replaced by X. We get 5_X and Gen_X by replacing K by X in 5 in Gen and, in addition, relativizing to awareness. More precisely:

5_X. $\neg X\varphi \wedge A\neg X\varphi \Rightarrow X\neg X\varphi$.
Gen_X. From φ infer $A\varphi \Rightarrow X\varphi$.

The analogous relativization to awareness is unnecessary in the case of K_X and 4_X, given the restriction to propositionally determined awareness. For example, in the case of K, we would have

K'_X. $X\varphi \wedge X(\varphi \Rightarrow \psi) \wedge A\psi \Rightarrow X\psi$.

However, if awareness is generated by primitive propositions, then $A\psi$ follows from $A(\varphi \wedge \psi)$ (which in turn follows from $X(\varphi \wedge \psi)$).

What happens to the axiom A0 that characterized X and the axioms A1–A12 that characterize propositionally determined awareness in the context of the language $\mathcal{L}^{XA}$, where there is no K operator? Clearly, A0 becomes

$\mathrm{A0}_X$. $X\varphi \Rightarrow A\varphi$.

A1–A4 and A6–A9 remain unchanged, since they do not mention the K operator, while A5 and A10 disappear; they are unnecessary, since K is not in the language. A11 changes to $\mathrm{A11}_X$, which is the result of replacing K by X. This is clearly sound in propositionally determined awareness structures. Finally, A12 disappears: if the agent is not aware of φ, then the agent cannot be aware of $\neg A\varphi$, and thus cannot explicitly know that he is unaware of φ.

My completeness proofs use one additional rule of inference, which says that certain formulas should be irrelevant in a proof.

Irr. If no primitive propositions in φ appear in ψ, then from $\neg A\varphi \Rightarrow \psi$ infer ψ.[8]

[8] This is an instance of a more general inference rule which is essentially the analogue of a well-known result in first-order logic called *Craig's Interpolation Lemma* [Craig 1957].

I strongly suspect that Irr is unnecessary and follows from the other axioms and inference rules I have stated. However, I have not been able to prove this.

Let $\mathbf{K}_X$ consist of Prop, MP, K_X, Gen_X, $\mathrm{A0}_X$, A1–A4, A6–A9, A11, and Irr; $\mathbf{T}_X$, $\mathrm{S4}_X$, and $\mathrm{S5}_X$ are obtained by successively adding T_X, 4_X, and 5_X to $\mathbf{K}_X$.

Theorem 5.3. *For formulas in the language $\mathcal{L}^{XA}$:*

(a)$\mathbf{K}_X$ is a sound and complete axiomatization with respect to $\mathcal{M}^{gpp,ka}$,
(c)$\mathbf{T}_X$ is a sound and complete axiomatization with respect to $\mathcal{M}^{r,gpp,ka}$,
(c)$\mathbf{S4}_X$ is a sound and complete axiomatization with respect to $\mathcal{M}^{rt,gpp,ka}$,
(d)$\mathbf{S5}_X$ is a sound and complete axiomatization with respect to $\mathcal{M}^{rst,gpp,ka}$.

Proof. I just sketch the proof of part (a) here, and then indicate how it can be modified to prove (b)–(d). For soundness, observe that all the axioms are valid and all the inference rules preserve validity. For example, in the case Gen_X, if φ is valid in $\mathcal{M}$, then it is easy to check that $A\varphi \Rightarrow X\varphi$ is valid. In the case of Irr, suppose that $\neg A\varphi \Rightarrow \psi$ is valid. Suppose that, by way of contradiction, ψ is not valid. Then there is a structure $M = (S, \rho, \mathcal{K}, \mathcal{A}) \in \mathcal{M}^{gpp,ka}$ and state s in M such that $(M, s) \models \neg\psi$. Now let $M' = (S, \rho, \mathcal{K}, \mathcal{A}')$ be the structure where all components agree with M except possibly $\mathcal{A}'$ and $\mathcal{A}'(t) = \mathcal{A}(t) - \{\theta :$ θ and φ have a primitive proposition in common$\}$. It is easy to check that $M' \in \mathcal{M}^{gpp,ka}$. Moreover, an easy induction on the structure of θ shows that for every formula σ that does not have any primitive propositions in common with φ and for every state $t \in S$, we have $(M, t) \models \sigma$ iff $(M', t) \models \sigma$. In particular, $(M', s) \models \neg\psi$. Moreover, by construction, $(M', s) \models \neg A\varphi$. But this contradicts the validity of $\neg A\varphi \Rightarrow \psi$. Thus, it must be the case that ψ is valid, as desired. Now an easy induction on the length of proofs shows that every formula φ provable in $\mathbf{K}_X$ is valid in $\mathcal{M}^{gpp,ka}$.

To prove completeness, it clearly suffices to show that every formula φ that is satisfiable in $\mathcal{M}^{gpp,ka}$ (i.e., every formula φ such that $M, s \models \varphi$ for some $M \in \mathcal{M}^{gpp,ka}$) is $\mathbf{K}_X$-consistent. The proof of this uses the idea of canonical structures. This approach was introduced in the economics literature by Aumann [1999], although the basic idea is well known in the

Craig's Interpolation Lemma says that if φ and ψ are formulas in first-order logic and $\varphi \Rightarrow \psi$ is provable, then there is a formula θ which involves only symbols that appear in both φ and ψ such that both $\varphi \Rightarrow \theta$ and $\theta \Rightarrow \psi$ are provable. It follows that if φ and ψ have no symbols in common and $\varphi \Rightarrow \psi$ is provable, then either $\neg\varphi$ is provable or ψ is provable. The analogue to the general Craig Interpolation Lemma holds for the language $\mathcal{L}^{KXA}$, but proving it is beyond the scope of this paper.

modal logic community, and seems to have been introduced independently by Kaplan [1966], Makinson [1966], and Lemmon and/or Scott [Lemmon 1977].

Given an axiom system AX, a formula φ is *AX-consistent* if $\neg\varphi$ is not provable in AX. A finite set of formulas is AX-consistent if their conjunction is; an infinite set of formulas is AX-consistent if every finite subset of its formulas is. Finally, F is a *maximal AX-consistent* set of formulas if F is AX-consistent and every strict superset of F is not AX-consistent.

Define the *canonical* $\mathbf{K}_X$ *model* $M^c = (S^c, \rho^c, \mathcal{K}^c, \mathcal{A}^c)$ as follows:

- S^c consists of all maximal $\mathbf{K}_X$-consistent sets of formulas (thus, each state in S^c is a set of formulas);
- $\rho^c(p) = \{s \in S^c : p \in s\}$ for each $p \in \Phi$;
- $(s,t) \in \mathcal{K}^c$ iff $\{\varphi : X\varphi \in s\} \cup \{Ap : Ap \in s, p \in \Phi\} \cup \{\neg Ap : \neg Ap \in s, p \in \Phi\} \subseteq t$;
- $\mathcal{A}(s) = \{\varphi : A\varphi \in s\}$.

Observe that it follows from A1–A4 and A6–A9 that awareness in M^c is generated by primitive propositions. Moreover, the construction of $\mathcal{K}^c$ and A6–A9 guarantees that the agent knows which formulas he is aware of in M^c. Thus, M^c is propositionally determined.

I now show that $\varphi \in s$ iff $(M^c, s) \models \varphi$, by induction on the structure of φ (where primitive propositions and formulas of the form $A\varphi'$ are treated as being "structurally simpler" than all other formulas). If φ is a primitive proposition or of the form $A\varphi'$, then this is immediate from the definition. If φ is a conjunction or a negation, it follows easily from standard properties of maximal consistent sets (in particular, the facts if F is a maximal consistent set, then $\varphi \wedge \psi \in F$ iff both $\varphi \in F$ and $\psi \in F$ and that $\neg\varphi \in F$ iff $\varphi \notin F$). Finally, suppose that φ is of the form $X\varphi'$.

If $X\varphi' \in s$, then since s is a maximal $\mathbf{K}_X$-consistent set, it follows from $\mathrm{A0}_X$ that $A\varphi' \in s$. Moreover, by construction of $\mathcal{K}$, if $(s,t) \in \mathcal{K}^c$, then $\varphi' \in t$. By the induction hypothesis, it follows that $(M^c, s) \models A\varphi'$ and $(M^c, t) \models \varphi'$ for all t such that $(s,t) \in \mathcal{K}^c$. Thus, $(M^c, s) \models X\varphi'$. Conversely, suppose that $(M^c, s) \models X\varphi'$. The key step in showing that $X\varphi' \in s$ is provided by the following claim.

Claim: Let $F_1 = \{\psi : X\psi \in s\}$, let $F_2 = \{Ap : Ap \in s, p \in \Phi\}$, and let $F_3 = \{\neg Ap : \neg Ap \in s, p \in \Phi\}$. If $X\varphi' \notin s$ and $A\varphi' \in s$, then $\{\neg\varphi'\} \cup F_1 \cup F_2 \cup F_3$ is $\mathbf{K}_X$-consistent.

Suppose that the claim is false. Then there are three finite set of formulas $F_1' \subseteq F_1$, $F_2' \subseteq F_2$, and $F_3' \subseteq F_3$ such that $\{\neg\varphi'\} \cup F_1' \cup F_2' \cup F_3'$ is $\mathbf{K}_X$-inconsistent. This means that

$$\mathbf{K}_X \vdash (\wedge F_1') \wedge (\wedge F_2') \wedge (\wedge F_3') \Rightarrow \varphi'.$$

Since $A\varphi' \in s$, if $\neg Ap \in F_3'$ then p does not appear in φ'. (For otherwise s would be inconsistent, using A6–A9.) Similarly, since $X\psi \in s$ (and hence $A\psi \in s$) for each formula $\psi \in F_1'$, if $\neg Ap \in F_3'$ then p does not appear in ψ. Of course, it is also immediate that if $\neg Ap \in F_3'$, then p does not appear in any formula in F_2'. By straightforward propositional reasoning (Prop and MP), it is immediate that $(\varphi \wedge \psi) \Rightarrow \theta$ is equivalent to $\psi \Rightarrow (\varphi \Rightarrow \theta)$. Thus, if $F_3' = \{\neg Ap_1, \ldots, \neg Ap_k\}$, we have that

$$\mathbf{K}_X \vdash \neg Ap_1 \Rightarrow (\neg Ap_2 \Rightarrow (\ldots \Rightarrow (\neg Ap_k \Rightarrow ((\wedge F_1') \wedge (\wedge F_2') \Rightarrow \varphi'))\ldots)).$$

Now applying Irr repeatedly, it follows that

$$\mathbf{K}_X \vdash (\wedge F_1') \wedge (\wedge F_2') \Rightarrow \varphi'.$$

Thus, by Gen_X, it follows that

$$\mathbf{K}_X \vdash A((\wedge F_1') \wedge (\wedge F_2') \Rightarrow \varphi') \Rightarrow X((\wedge F_1') \wedge (\wedge F_2') \wedge \Rightarrow \varphi'). \tag{1}$$

It is easy to see using A1–A4, A6–A9, and $\mathrm{A0}_X$ and the fact that all the formulas in F_2' have the form Ap for some $p \in \Phi$ that

$$\mathbf{K}_X \vdash ((\wedge_{\theta \in F_1'} X\theta) \wedge (\wedge F_2') \wedge A\varphi') \Rightarrow A((\wedge F_1') \wedge (\wedge F_2') \Rightarrow \varphi'). \tag{2}$$

Using $\mathrm{A11}_X$, it is immediate that

$$\mathbf{K}_X \vdash (\wedge F_2') \Rightarrow \wedge_{\theta \in F_2} X\theta. \tag{3}$$

Using axiom K_X, Prop, and MP, we get that

$$\mathbf{K}_X \vdash X((\wedge F_1') \wedge (\wedge F_2') \Rightarrow \varphi') \Rightarrow (X((\wedge F_1') \wedge (\wedge F_2')) \Rightarrow X\varphi'). \tag{4}$$

A standard argument using K_X and Gen_X shows that, for arbitrary formulas ψ_1 and ψ_2, we have

$$\mathbf{K}_X \vdash (X\psi_1 \wedge X\psi_2) \Leftrightarrow X(\psi_1 \wedge \psi_2). \tag{5}$$

Putting together (1)–(5), it follows that

$$\mathbf{K}_X \vdash ((\wedge_{\theta \in F_1'} X\theta) \wedge (\wedge F_2') \wedge A\varphi') \Rightarrow X\varphi'. \tag{6}$$

By construction, for each formula $\theta \in F_1'$, $X\theta \in s$. Moreover, each formula in F_2' is in s and, by assumption, $A\varphi' \in s$. Thus, since s is a maximal consistent set, it follows that $X\varphi' \in s$. But this contradicts the assumption that $X\varphi' \notin s$, and proves claim.

Now to prove completeness, suppose that $(M, s) \models X\varphi'$. This means that $(M, s) \models A\varphi'$, so by the induction hypothesis, $A\varphi' \in s$. Suppose, by way of contradiction, that $X\varphi' \notin s$. Then, by the claim, $\{\neg\varphi'\} \cup F_1 \cup F_2 \cup F_3$ is $\mathbf{K}_X$-consistent. Thus, there must be a maximal $\mathbf{K}_X$-consistent set $t \supseteq \{\neg\varphi'\} \cup F_1 \cup F_2 \cup F_3$. By construction, $(s, t) \in \mathcal{K}^c$. By the induction hypothesis, $(M^c, t) \models \neg\varphi'$. But this contradicts the assumption that $(M^c, s) \models X\varphi'$. Thus, it must be the case that $X\varphi' \in s$, as desired. This completes the completeness proof in the case if $\mathbf{K}_X$.

The modifications required to deal with $\mathcal{M}^r$, $\mathcal{M}^{rt}$, and $\mathcal{M}^{rst}$ are standard. Indeed, all that needs to be done to convert the completeness proof for $\mathbf{K}_X$ to one for $\mathbf{T}_X$ (resp., $\mathbf{S4}_X$, $\mathbf{S5}_X$) is to build the canonical model using maximal $\mathbf{T}_X$- (resp., $\mathbf{S4}_X$-, $\mathbf{S5}_X$-) consistent sets of formulas as states and to observe that the $\mathcal{K}^c$ relation must be reflexive (resp., reflexive and transitive, an equivalence relation). I leave the straightforward details to the reader. ▌

MR provide an axiomatization they call $\mathcal{U}$ for the language $\mathcal{L}^{XA}$ which they prove to be sound and complete for partitional GSMs. From Theorem 4.1, it follows immediately that $\mathbf{S5}_X$ and $\mathcal{U}$ are equivalent, so $\mathbf{S5}_X$ is a sound and complete axiomatization for partitional GSMs and U is a sound and complete axiomatization for partitional propositionally determined awareness structures. That is, $\mathbf{S5}_X$ provides an alternate axiomatization of (un)awareness in partitional GSMs. Note that it does *not* follow, however, that $\mathbf{K}_X$ (resp., $\mathbf{T}_X$, $\mathbf{S4}_X$) is a sound and complete axiomatization for the language $\mathcal{L}^{XA}$ for GSMs (resp., reflexive, reflexive and transitive) GSMs. This is because, as I observed earlier, the definition of A differs between GSMs and awareness structures in the non-partitional case.

What happens if we relax the assumption that awareness is propositionally determined? For ease of exposition, I continue to assume that the agent knows what formulas he is aware of. (It is possible to obtain axiomatizations if we drop this assumption, but they are uglier.) In this case, we need to use K'_X rather than K_X. However, it does not seem that K'_X and Gen_X suffice to get completeness. However, the following generalization of RK does suffice:

RK_X. From $\varphi_1 \wedge \ldots \wedge \varphi_n \Rightarrow \psi$ infer $X\varphi_1 \wedge \ldots \wedge X\varphi_n \wedge A\psi \Rightarrow X\psi$.

In the case of the K operator, Proposition 5.1 says that (in the presence of Prop and MP) RK is equivalent to K and Gen. With awareness, clearly Gen_X is just RK_X in the special case that $n = 0$ and K_X follows immediately from RK_X in the case that $n = 2$, since $(\varphi \wedge (\varphi \Rightarrow \psi)) \Rightarrow \psi$ is a propositional tautology, as is the fact that $\varphi_1 \wedge \varphi_2 \Rightarrow \psi$ is equivalent to $\varphi_1 \Rightarrow (\varphi_2 \Rightarrow \psi)$, for all choices of φ_1, φ_2, and ψ. However, RK_X does not seem to follow from in general from K_X and Gen_X. To understand why, consider why RK follows from K and Gen in the case that $n = 2$. Using Prop, from $\varphi_1 \wedge \varphi_2 \Rightarrow \psi$ we can conclude $\varphi_1 \Rightarrow (\varphi_2 \Rightarrow \psi)$. (The two formulas are propositionally equivalent.) Now applying Gen, we can conclude $K(\varphi_1 \Rightarrow (\varphi_2 \Rightarrow \psi))$. Now two applications of the axiom K together propositional reasoning (Prop and MP) given us $K\varphi_1 \Rightarrow K(\varphi_2 \Rightarrow \psi)$ and then $K\varphi_1 \Rightarrow (K\varphi_2 \Rightarrow K\psi)$, which is equivalent to $(K\varphi_1 \wedge K\varphi_2) \Rightarrow K\psi$. This reasoning breaks down if we replace K by X. The problem is that we cannot apply Gen_X to conclude $X(\varphi_1 \Rightarrow (\varphi_2 \Rightarrow \psi)$, since there is no reason to believe that $A(\varphi_1 \Rightarrow (\varphi_2 \Rightarrow \psi))$ holds, even if $X\varphi_1$ and $X\varphi_2$ and $A\psi$ hold. (We can draw this conclusion if awareness is propositionally generated and, indeed, in the presence of Prop and MP, RK_X is equivalent to K_X and Gen_X.)

We must generalize RK_X still further in the presence of knowledge of awareness, as follows:

RK_X^+. From $\varphi_1 \wedge \ldots \wedge \varphi_n \wedge A\tau_1 \wedge \ldots \wedge A\tau_k \wedge \neg A\tau_1' \wedge \ldots \wedge \neg A\tau_{k'}' \Rightarrow \psi$ infer $X\varphi_1 \wedge \ldots \wedge X\varphi_n \wedge A\tau_1 \wedge \ldots \wedge A\tau_k \wedge \neg A\tau_1' \wedge \ldots \wedge \neg A\tau_{k'}' \wedge A\psi \Rightarrow X\psi$.

Again, RK_X^+ seems strictly stronger that RK_X and the obvious extensions of A11 and A12, which are

$\mathrm{A11}_X'$. $A\varphi \wedge AA\varphi \Rightarrow XA\varphi$.
$\mathrm{A12}_X'$. $\neg A\varphi \wedge A\neg A\varphi \Rightarrow X\neg A\varphi$.

While $\mathrm{A11}_X'$ and $\mathrm{A12}_X'$ are easily seen to follow from RK_X^+ (in the presence of Prop and MP), the converse does not seem to hold.

In this spirit of RK_X^+, consider the following axioms and rules:

$\mathrm{RK4}_X^+$. From $\varphi_1 \wedge \ldots \wedge \varphi_n \wedge X\sigma_1 \wedge \ldots \wedge X\sigma_m \wedge A\tau_1 \wedge \ldots \wedge A\tau_k \wedge \neg A\tau_1' \wedge \ldots \wedge \neg A\tau_{k'}' \Rightarrow \psi$ infer $X\varphi_1 \wedge \ldots \wedge X\varphi_n \wedge X\sigma_1 \wedge \ldots \wedge X\sigma_m \wedge A\tau_1 \wedge \ldots \wedge A\tau_k \wedge \neg A\tau_1' \wedge \ldots \wedge \neg A\tau_{k'}' \wedge A\psi \Rightarrow X\psi$.

$\mathrm{RK5}_X$. From $\varphi_1 \wedge \ldots \wedge \varphi_n \wedge X\sigma_1 \wedge \ldots \wedge X\sigma_m \wedge A\tau_1 \wedge \ldots \wedge A\tau_k \wedge \neg A\tau_1' \wedge \ldots \wedge \neg A\tau_{k'}' \Rightarrow \psi$ infer $X\varphi_1 \wedge \ldots \wedge X\varphi_n \wedge X\sigma_1 \wedge \ldots \wedge X\sigma_m \wedge A\tau_1 \wedge \ldots \wedge A\tau_k \wedge \neg A\tau_1' \wedge \ldots \wedge \neg A\tau_{k'}' \wedge A\psi \Rightarrow X\psi$.

Let $\mathcal{C}_1' = \{\mathrm{A1}, \ldots, \mathrm{A4}\}$, let $\mathcal{C}_2' = \mathcal{C}_1' \cup \{\mathrm{A6}, \ldots, \mathrm{A9}\}$, and let $\mathcal{C}_3 = \{A9\}$.

THEOREM 5.4. *Let $\mathcal{C}$ be a (possibly empty) subset of $\{\mathcal{C}_1', \mathcal{C}_2', \mathcal{C}_3'\}$ and let C be the corresponding subset of $\{csub, gpp, sr\}$. For formulas in the language $\mathcal{L}^{XA}$*

(a)$\{$Prop, MP, RK_X^+, $A0_X\} \cup \mathcal{C}$ is a sound and complete axiomatization with respect to $\mathcal{M}^{C,ra}$.

(b)$\{$Prop, MP, RK_X^+, T_X, $A0_X\} \cup \mathcal{C}$ is a sound and complete axiomatization with respect to $\mathcal{M}^{r,C,ra}$,

(c)$\{$*Prop, MP,* $RK4^+_X$, T_X, $A0_X\} \cup \mathcal{C}$ *is a sound and complete axiomatization with respect to* $\mathcal{M}^{rt,C,ra}$*,*

(d)$\{$*Prop, MP,* $RK5^+_X$, T_X, $A0_X\} \cup \mathcal{C}$ *is a sound and complete axiomatization with respect to* $\mathcal{M}^{rst,C,ra}$*.*

Note that this means that we can replace K_X and Gen_X (resp., K_X, Gen_X, and 4_X; K_X, Gen_X, 4_X, and 5_X) by RK^+_X (resp., $\mathrm{RK4}^+_X$; $\mathrm{RK5}^5_X$) to get an axiomatization equivalent to $\mathbf{K}_X$ and $\mathbf{T}_X$ (resp., $\mathbf{S4}_X$; $\mathbf{S5}_X$).

6. DISCUSSION

I have shown that GSMs can be viewed as a special case of awareness structures in a precise sense. The advantage of thinking in terms of awareness structures is that we can then consider other notions of awareness, not just ones generated by primitive propositions. For example, as observed by FH, it is possible to consider more computationally oriented notions of awareness (see [Fagin, Halpern, Moses, and Vardi 1995, Chapter 10] for more discussion of this issue). The problem is then to come up with interesting notions of awareness that have enough structure to allow for interesting mathematical analysis. I believe it should also be possible to use awareness structures to allow for natural reasoning about awareness and lack of it (so that an agent can reason, for example, about the possibility that she is unaware of certain features that another may be aware of). I am currently working on modeling such reasoning.

References

Aumann, R. J. (1999). Interactive epistemology I: knowledge. *International Journal of Game Theory 28*(3), 263–301.

Chellas, B. F. (1980). *Modal Logic.* Cambridge, U.K.: Cambridge University Press.

Craig, W. (1957). Linear reasoning. A new form of the Herbrand-Gentzen Theorem. *Journal of Symbolic Logic 22*(3), 250–268.

Dekel, E., B. Lipman, and A. Rusticchini (1998). Recent developments in modeling unforeseen contingencies. *European Economic Review 42*, 523–542.

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

Fagin, R., J. Y. Halpern, Y. Moses, and M. Y. Vardi (1995). *Reasoning about Knowledge.* Cambridge, Mass.: MIT Press.

Geanakoplos, J. (1989). Game theory without partitions, and applications to speculation and consensus. Cowles Foundation Discussion Paper #914, Yale University.

Kaplan, D. (1966). Review of "A semantical analysis of modal logic I: normal modal propositional calculi". *Journal of Symbolic Logic 31*, 120–122.

Lemmon, E. J. (1977). *The "Lemmon Notes": An Introduction to Modal Logic*. Oxford, U.K.: Basil Blackwell. Written in collaboration with Dana Scott; edited by Krister Segerberg. American Philosophical Quarterly Monograph Series. Monograph No. 11.

Makinson, D. (1966). On some completeness theorems in modal logic. *Zeitschrift für Mathematische Logik und Grundlagen der Mathematik 12*, 379–384.

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

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