# Two models of unawareness:
# Comparing the object-based and the
# subjective-state-space approaches*

Oliver J. Board
University of Pittsburgh

Kim-Sau Chung
University of Minnesota

Burkhard C. Schipper
University of California, Davis

March 19, 2009

## 1   Introduction

Over the past twenty years or so, a small but growing literature has emerged with the aim of modeling agents who are *unaware* of certain things. Early examples include Fagin and Halpern [4] and Modica and Rustichini [11]. More recently, a number of authors have begun to examine the implications of unawareness in economic theory (see e.g. Chung & Fortnow [3], Tirole [12], and Filiz-Ozbay [5]). In this paper we compare two different approaches to modeling unawareness: the *object-based* approach of Board & Chung [2] and the *subjective-state-space* approach of Heifetz *et al.* [8]. In particular, we show that subjective-state-space models (henceforth *HMS structures*) can be embedded within object-based models (henceforth *OBU structures*), demonstrating that the latter are at least as expressive. As long as certain restrictions are imposed on the form of the OBU structure, the embedding can also go the other way. A generalization of HMS structures (relaxing the partitional properties of knowledge) gives us a full converse. Given the rather different interpretations of each approach offered by their respective authors, we believe that these results may enhance our understanding of each.

# 2 Two models of unawareness

In what follows, we provide a brief presentation of OBU structures and HMS structures. More details, along with proofs of results can be found in Board & Chung [2] and Heifetz *et al.* [8, 9].

## 2.1 The object-based approach (Board & Chung [2])

An *OBU structure* is a tuple $\langle \mathsf{W}, \mathsf{O}, \{\mathcal{I}_i\}, \{\mathcal{A}_i\} \rangle$, where:

- $\mathsf{W}$ is a set of states;

- $\mathsf{O}$ is a set of objects;

- $\mathcal{I}_i : \mathsf{W} \to 2^{\mathsf{W}}$ is an *information function* for agent $i$; and

- $\mathcal{A}_i : \mathsf{W} \to 2^{\mathsf{O}}$ is an *awareness function* for agent $i$.

Intuitively, $\mathcal{I}_i(\mathsf{w})$ indicates the states that agent $i$ considers possible when the true state is $\mathsf{w}$, while $\mathcal{A}_i(\mathsf{w})$ indicates the objects she is aware of.[1]

In the standard information partition model familiar to economists, events are subsets of the state space, corresponding (roughly) to the set of states in which some given proposition is true. In our model, an event is an ordered pair $(\mathsf{R}, \mathsf{S})$, where $\mathsf{R} \subseteq \mathsf{W}$ is a set of states and $\mathsf{S} \subseteq \mathsf{O}$ is a set of objects; we call $\mathsf{R}$ the *reference* of the event, corresponding (as before) to the set of states in which some proposition is true; and $\mathsf{S}$ is the *sense* of the event, listing the set of objects referred to in the description of the proposition. (To give an example, the events representing the propositions "the dog barked" and "the dog barked and the cat either did or did not meow" have the same reference but difference senses.) We sometimes abuse notation and write $(\mathsf{R}, \mathsf{a})$ instead of $(\mathsf{R}, \{\mathsf{a}\})$, and $(\mathsf{w}, \mathsf{S})$ instead of $(\{\mathsf{w}\}, \mathsf{S})$. We use $\mathcal{E}$ to denote the set of all events, with generic element $\mathsf{E}$.

We now define two operators on events, corresponding to "not" and "and".

$$\neg(\mathsf{R}, \mathsf{S}) = (\mathsf{W} \setminus \mathsf{R}, \mathsf{S})$$
$$\wedge_j (\mathsf{R}_j, \mathsf{S}_j) = (\cap_j \mathsf{R}_j, \cup_j \mathsf{S}_j).$$

---

[1] In Board & Chung [2] OBU structures also include a subset $\mathsf{O}_{\mathsf{w}}$ of $\mathsf{O}$ for each state $\mathsf{w} \in \mathsf{W}$. The scope of quantifiers at a given state $\mathsf{w}$ includes only the objects in $\mathsf{O}_{\mathsf{w}}$ (the *real* objects), rather than all of the objects in $\mathsf{O}$. Allowing the $\mathsf{O}_{\mathsf{w}}$ sets to vary across states enables us to captures agents' uncertainty about whether they are aware of everything or not, e.g. "Peter is not sure whether he is aware of everything". The reader who is familiar with Board & Chung [2] can treat the OBU structures in this paper as special cases where $\mathsf{O}_{\mathsf{w}} = \mathsf{O}$ for all $\mathsf{w}$.

The negation of an event holds at precisely those states at which the event does not hold, but it refers to the same set of objects. The conjunction of several events holds only at those states at which *all* of those events hold, and it refers to each set of objects. It will often be convenient to use disjunction "or" as well, defined in terms of negation and conjunction as follows:

$$\vee_j(R_j, S_j) = \neg(\wedge_j \neg(R_j, S_j))$$
$$= (\cup_j R_j, \cup_j S_j).$$

We also introduce three modal operators for each agent, representing awareness, implicit knowledge, and explicit knowledge:

$$A_i(R, S) = (\{w \mid S \subseteq \mathcal{A}_i(w)\}, S) \qquad \text{(awareness)} \qquad (1)$$
$$L_i(R, S) = (\{w \mid \mathcal{I}_i(w) \subseteq R\}, S) \qquad \text{(implicit knowledge)} \qquad (2)$$
$$K_i(R, S) = A_i(R, S) \wedge L_i(R, S) \qquad \text{(explicit knowledge)} \qquad (3)$$

Intuitively, an agent is aware of an event at $w$ if she is aware of every object in the sense of the event; and the agent implicitly knows an event at state $w$ if the reference of the event includes every state she considers possible. However, implicit knowledge is not the same as explicit knowledge, and the latter is our ultimate concern. Implicit knowledge is merely a benchmark that serves as an intermediate step to modeling what an agent actually knows. Intuitively, an agent does not actually (i.e., explicitly) know an event unless he is aware of the event *and* he implicitly knows the event. Notice that $A_i$, $L_i$, and $K_i$ do not change the set of objects being referred to.

It is easy to verify that awareness and (implicit) knowledge satisfy the following properties (where we suppress the agent-subscripts):

**A1** $\wedge_j A(R, S_j) = A(R, \cup_j S_j)$

**A2** $A(R, X) = A(R', X)$ for all $R, R'$

**A3** $A(R, \varnothing) = (W, \varnothing)$

**A4** $A(R, X) = (R', X)$ for some $R'$

**L1** $L(W, O) = (W, O)$

**L2** $\wedge_j L(R_j, S) = L(\cap_j R_j, S)$

**L3** $L(R, S) = (R', S)$ for some $R'$

**L4** if $\mathsf{L(R,S)} = \mathsf{(R',S)}$ then $\mathsf{L(R,S')} = \mathsf{(R',S')}$

In Board & Chung [2] we show that the converse is also true: any awareness/knowledge operator satisfying these properties can be derived from some awareness/information function. Thus **A1**–**A4** and **L1**–**L4** provide a precise characterization of awareness and (implicit) knowledge, respectively.

**Proposition 1** *Suppose that* $\mathsf{A}_i$ *is defined as in* (1)*. Then:*

1. $\mathsf{A}_i$ *satisfies* **A1**–**A4***; and*

2. *if* $\mathsf{A}'_i$ *is an operator on events which satisfies* **A1**–**A4***, we can find an awareness function* $\mathcal{A}_i$ *such that* $\mathsf{A}'_i$ *and* $\mathsf{A}_i$ *coincide.*

**Proposition 2** *Suppose that* $\mathsf{L}_i$ *is defined as in* (2)*. Then:*

1. $\mathsf{L}_i$ *satisfies* **L1**–**L4***; and*

2. *if* $\mathsf{L}'_i$ *is an operator on events which satisfies* **L1**–**L4***, we can find an information function* $\mathcal{I}_i$ *such that* $\mathsf{L}'_i$ *and* $\mathsf{L}_i$ *coincide.*

**Aside: a formal language**

For the sake of transparency, and to aid interpretation, we now show how OBU structures can be used to provide truth conditions for a formal language, a version of first-order modal logic.[2] We start with a set of (unary) *predicates*, $P, Q, R, \ldots$, and an (infinite) set of *variables*, $x, y, z, \ldots$. Together with set of objects, $\mathsf{O}$, this generates a set $\Phi$ of *atomic formulas*, $P(\mathsf{a}), P(x), Q(\mathsf{a}), Q(x), \ldots$, where each predicate takes as its argument a single object or variable. Let $\mathcal{F}$ be the smallest set of *formulas* that satisfies the following conditions:

- if $\phi \in \Phi$, then $\phi \in \mathcal{F}$;

- if $\phi, \psi \in \mathcal{F}$, then $\neg\phi \in \mathcal{F}$ and $\phi \wedge \psi \in \mathcal{F}$;

- if $\phi \in \mathcal{F}$ and $x \in X$, then $\forall x \phi \in \mathcal{F}$;

- if $\phi \in \mathcal{F}$, then $\mathsf{L}_i\phi \in \mathcal{F}$ and $\mathsf{A}_i\alpha \in \mathcal{F}$ and $\mathsf{K}_i\alpha \in \mathcal{F}$ for each agent $i$.

---

[2]In Board & Chung [1], we offer a sound and complete axiomatization of this language.

Formulas should be read in the obvious way; for instance, $\forall x \mathtt{A}_i P(x)$ is to be read as "for every $x$, agent $i$ is aware that $x$ possesses property $P$." Notice, however, that it is hard to make sense of certain formulas: consider $P(x)$ as opposed to $P(\mathsf{a})$ or $\forall x P(x)$. Although it may be reasonable to claim that a specific object, $\mathsf{a}$, is $P$, or that every $x$ is $P$, the claim that $x$ is $P$ seems empty unless we specify which object variable $x$ stands for. In general, we say that a variable $x$ is *free* in a formula if it does fall under the scope of a quantifier $\forall x$, and define our language $\mathcal{L}$ to be the set of all formulas containing no free variables.[3] We use OBU structures to provide truth conditions only for formulas in $\mathcal{L}$, and not for formulas such as $P(x)$ that contain free variables.

Take an OBU structure $M = \langle \mathsf{W}, \mathsf{O}, \{\mathcal{I}_i\}, \{\mathcal{A}_i\} \rangle$, and augment it with an *assignment* $\pi(\mathsf{w})(P) \subseteq \mathsf{O}$ of objects to every predicate at every state (intuitively, $\pi(\mathsf{w})(P)$ is the set of objects that satisfy predicate $P$). If a formula $\phi \in \mathcal{L}$ is true at state $\mathsf{w}$ of OBU structure $M$ under assignment $\pi$, we write $(M, \mathsf{w}, \pi) \vDash P(\mathsf{a})$; $\vDash$ is defined inductively as follows:

$(M, \mathsf{w}, \pi) \vDash P(\mathsf{a})$ iff $\mathsf{a} \in \pi(\mathsf{w})(P)$;

$(M, \mathsf{w}, \pi) \vDash \neg\phi$ iff $(M, \mathsf{w}, \pi) \nvDash \phi$;

$(M, \mathsf{w}, \pi) \vDash \phi \wedge \psi$ iff $(M, \mathsf{w}, \pi) \vDash \phi$ and $(M, \mathsf{w}, \pi) \vDash \psi$;

$(M, \mathsf{w}, \pi) \vDash \forall x \phi$ iff $(M, \mathsf{w}, \pi) \vDash \phi[\mathsf{a} \backslash x]$ for every $a \in \mathsf{O}$ (where $\phi[\mathsf{a}\backslash x]$ is $\phi$ with all free occurrences of $x$ replaced with $\mathsf{a}$);[4]

$(M, \mathsf{w}, \pi) \vDash \mathtt{A}_i\phi$ iff $\mathsf{a} \in \mathcal{A}_i(\mathsf{w})$ for every object $\mathsf{a}$ in $\phi$;

$(M, \mathsf{w}, \pi) \vDash \mathtt{L}_i\phi$ iff $(M, \mathsf{w}', \pi) \vDash \phi$ for all $\mathsf{w}' \in \mathcal{I}_i(\mathsf{w})$;

$(M, \mathsf{w}, \pi) \vDash \mathtt{K}_i\phi$ iff $(M, \mathsf{w}, \pi) \vDash \mathtt{A}_i\phi$ and $(M, \mathsf{w}, \pi) \vDash \mathtt{L}_i\phi$.

Notice that there is a close connection between sentences of $\mathcal{L}$ and OBU events: for any given $\phi \in \mathcal{L}$, the reference of the corresponding OBU event is given by the set of states at which $\phi$ is true, while the sense is simply the set of objects in $\phi$

---

[3]More formally, we define inductively what it is for a variable to be free in $\phi \in \mathcal{F}$:

- if $\phi$ is an atomic formula of the form $P(x)$ where $x$ is a variable, then $x$ is free in $\phi$;
- $x$ is free in $\neg\phi$, $\mathtt{K}_i\phi$, $\mathtt{A}_i\phi$, and $\mathtt{L}_i\phi$ iff $x$ is free in $\phi$;
- $x$ is free in $\phi \wedge \psi$ iff $x$ is free in $\phi$ or $\psi$;
- $x$ is free in $\forall y \phi$ iff $x$ is free in $\phi$ and $x$ is different from $y$.

[4]In Board & Chung [2], OBU structures also specify a subset $\mathsf{O}_\mathsf{w}$ of $\mathsf{O}$ for every $\mathsf{w} \in \mathsf{W}$. In this more general case, we define $(M, \mathsf{w}, \pi) \vDash \forall x \phi$ iff $(M, \mathsf{w}, \pi) \vDash \phi[\mathsf{a}\backslash x]$ for every $\mathsf{a} \in \mathsf{O}_\mathsf{w}$.

To help understand how OBU structures work, consider the following simple example. There are two agents, 1 and 2, and two objects, a and b. There are two issues which are of (potential) interest to our agents: whether or not object a is $P$, and whether or not object b is $Q$ (to borrow a famous example, our agents might be interested in whether or not a dog barks, and whether or not a cat meows). Assume that Agent 1 is aware of object a and knows whether it is $P$ or not, but agent 2 is not and does not. On the other hand, agent 2 is aware of object b and whether it is $Q$ or not, while agent 1 is not and does not.

One OBU structure that could be used to reason about this situation is shown in Figure 1 below. There are two objects, $O = \{a, b\}$, and four states, $W = \{PQ, P\neg Q, \neg PQ, \neg P\neg Q\}$, where e.g. state $\neg PQ$ can be thought of as corresponding to the situation where a is not $P$ but b is $Q$.[5] Agent 1's (partitional) information function is given by $\mathcal{I}_1(PQ) = \mathcal{I}_1(P\neg Q) = \{PQ, P\neg Q\}$ and $\mathcal{I}_1(\neg PQ) = \mathcal{I}_1(\neg P\neg Q) = \{\neg PQ, \neg P\neg Q\}$, and is described by the dashed blue rectangles in Figure 1. Her awareness function is given $\mathcal{A}_1(w) = \{a\}$ for all $w \in W$, and is described by the dashed blue arrows. Similarly, agent 2's information function and awareness function are described by the solid red arrows.
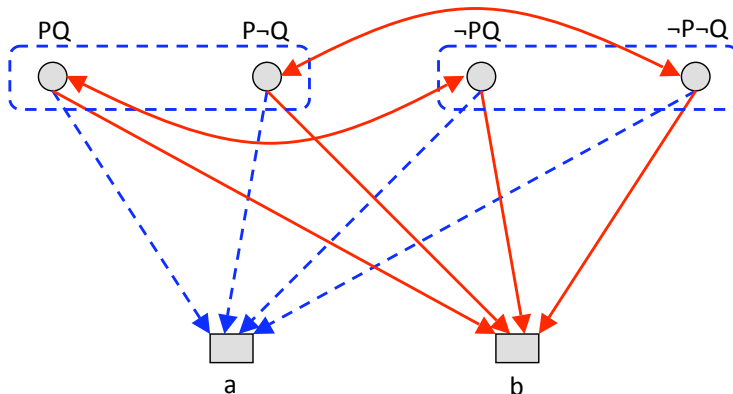


Figure 1: An OBU structure

Recall that in an OBU structure, an event is an ordered pair $(R, S)$, where $R \subseteq 2^W$ and $S \subseteq 2^O$. In this example, then, there are $2^4 \times 2^2 = 64$ distinct OBU events. For many of these, the interpretation is clear: for instance, the event $(\{PQ, P\neg Q\}, a)$ can be interpreted as "object a is $P$". But what about the event $(\{P\neg Q, \neg P\neg Q\}, \varnothing)$? This event holds when and only when object b is not $Q$, and yet neither of the objects, in particular not object b, are used to describe it. A natural translation of this event into English would be "nothing is $Q$" (indeed, with reference to language described above, the formula $\forall x \neg P(x)$

---

[5]For our current purposes, the labeling of the states serves merely to suggest a possible interpretation to the reader and to aid comprehension. This interpretation could be formalized as described above, by augmenting the OBU structure with an assignment. In Board & Chung [2] we show how to model properties of objects explicitly as mappings from objects to states, without reference to a formal language

would be true in precisely states $P\neg Q$ and $\neg P\neg Q$ according to the obvious assignment). In the current example, agent 1 does not know this event (implicitly or explicitly) in any of the states, although she does consider it possible—even though she is unaware of object b. Intuitively, she doesn't rule out the possibility that *something* is $Q$, even though she can't imagine what. When the true state is $PQ$, for example, this is captured by the fact that she considers both states $PQ$ and $P\neg Q$ to be possible. Note that is a modeling choice. We could have made different ones: In particular, we could have assumed that if an agent is not aware of anything which possesses a particular property, then that agent believes that nothing is $Q$; this would generate for agent 1 the (irreflexive) information function $\mathcal{I}_1(PQ) = \mathcal{I}_1(P\neg Q) = \{P\neg Q\}$. Alternatively, we could have assumed that if an agent is not aware of anything which possesses a particular property, then that agent believes that for sure *something* possesses that property; this would generate for agent 1 the (irreflexive) information function $\mathcal{I}_1(PQ) = \mathcal{I}_1(P\neg Q) = \{PQ\}$. Without a specific context in mind, none of these three modeling choices seems to us more appropriate than the others.

Tables 1 and 2 below describe the agents' awareness and knowledge of these two events, which we label $E$ and $F$ respectively.

|  | Agent 1 | Agent 2 |
|---|---|---|
| basic event | $E = \big(\{PQ, P\neg Q\}, a\big)$ | $E = \big(\{PQ, P\neg Q\}, a\big)$ |
| awareness | $A_1(E) = \big(W, a\big)$ | $A_2(E) = \big(\varnothing, a\big)$ |
| implicit knowledge | $L_1(E) = \big(\{PQ, P\neg Q\}, a\big)$ | $L_2(E) = \big(\varnothing, a\big)$ |
| explicit knowledge | $K_1(E) = \big(\{PQ, P\neg Q\}, a\big)$ | $K_2(E) = \big(\varnothing, a\big)$ |

Table 1: "object $a$ is $P$"

|  | Agent 1 | Agent 2 |
|---|---|---|
| basic event | $F = \big(\{P\neg Q, \neg P\neg Q\}, \varnothing\big)$ | $F = \big(\{P\neg Q, \neg P\neg Q\}, \varnothing\big)$ |
| awareness | $A_1(F) = \big(W, \varnothing\big)$ | $A_2(F) = \big(W, \varnothing\big)$ |
| implicit knowledge | $L_1(F) = \big(\varnothing, \varnothing\big)$ | $L_2(F) = \big(P\neg Q, \neg P\neg Q, \varnothing\big)$ |
| explicit knowledge | $K_1(F) = \big(\varnothing, \varnothing\big)$ | $K_2(F) = \big(P\neg Q, \neg P\neg Q, \varnothing\big)$ |

Table 2: "nothing is $Q$"

## 2.2   The subjective-state-space approach (Heifetz et al. [8, 9])

An *HMS structure*[6] is a tuple $\langle \mathcal{S}, \preceq, r, \Pi_i \rangle$. The first three components describe the event space. First, $\mathcal{S} = \{S_\alpha\}_{\alpha \in A}$ is a complete lattice of disjoint state spaces, partially ordered by $\preceq$. The intended interpretation of the ordering is that if $S \preceq S'$, then $S$ is less *expressive* than $S'$. $\Sigma = \bigcup_{\alpha \in A} S_\alpha$ is used to denote the set of all states.

Next, $r = \{r_S^{S'}\}_{S, S' \in \mathcal{S} \text{ with } S \preceq S'}$ is a set of surjective projections from each state space to every space that is (weakly) less expressive: if $w \in S'$, then $r_S^{S'}(w)$ is the restriction of the description of $w$ to the more limited vocabulary of $S$. These functions are required to commute (so that if $S \preceq S' \preceq S''$, then $r_S^{S''} = r_S^{S'} \circ r_{S'}^{S''}$), and $r_S^S$ is the identity function. The following notation will prove useful. Suppose $S \preceq S'$: then if $w \in S'$, let $w_S = r_S^{S'}(w)$; and if $B \subseteq S'$, let $B_S = \{w_S \mid w \in B\}$.

For $B \subseteq S$, denote by $B^\uparrow = \bigcup_{\{S' \mid S \preceq S'\}} \left(r_S^{S'}\right)^{-1}(B)$ the extension of $B$ to all more expressive vocabularies. Then $E \subseteq \Sigma$ is an *HMS event* if it is of the form $B^\uparrow$ for some $B \subseteq S$ and some $S \in \mathcal{S}$. $B$ is called the *basis* of $E$, and $S = S(E)$ the *base-space*.

If $B^\uparrow$ is an HMS event with basis $B \subseteq S$, its negation $\neg B^\uparrow$ is defined by $(S \setminus B)^\uparrow$. To handle the case where $B = S$, Heifetz *et al.* [8] introduce a distinct event $\varnothing^S$ ("a logical contradiction phrased with the expressive power available in $S$") for each $S \in \mathcal{S}$, and define $\neg S^\uparrow = \varnothing^S$ and $\neg \varnothing^S = S^\uparrow$.

The conjunction of a set of events $\left\{B_\lambda^\uparrow\right\}$ is simply the intersection:

$$\bigwedge_\lambda B_\lambda^\uparrow = \bigcap_\lambda B_\lambda^\uparrow,$$

while disjunction is defined from conjunction in the usual way:

$$\bigvee_\lambda B_\lambda^\uparrow = \neg \left( \bigwedge_\lambda \neg B_\lambda^\uparrow \right)$$

(note that the disjunction of a set of events is *not* equal to their union, except in the special case where they all have the same base-space).

The $\Pi_i$ functions are designed to capture the agents' knowledge and awareness. Agent $i$'s *possibility correspondence* $\Pi_i : \Sigma \to 2^\Sigma \setminus \varnothing$ is assumed to satisfy the following properties:

(0)  If $w \in S$ then $\Pi_i(w) \subseteq S'$ for some $S' \preceq S$.                    Confinement

(1)  $w \in \Pi_i^\uparrow(w)$ for every $w \in \Sigma$.                    Generalized Reflexivity

(2)  $w' \in \Pi_i(w)$ implies $\Pi_i(w') = \Pi_i(w)$                    Stationarity

---

[6]We follow the original notation and presentation as closely as possible in this section.

(3) If $w \in S'$, $w \in \Pi_i(w)$ and $S \preceq S'$ then $w_S \in \Pi_i(w_S)$     Projections Preserve Awareness

(4) If $w \in S'$ and $S \preceq S'$ then $\Pi_i^{\uparrow}(w) \subseteq \Pi_i^{\uparrow}(w_S)$     Projections Preserve Ignorance

(5) If $S \preceq S' \preceq S''$, $w \in S''$ and $\Pi_i(w) \subseteq S'$ then     Projections Preserve Knowledge
$(\Pi_i(w))_S = \Pi_i(w_S)$

Generalized reflexivity and stationarity are the analogues of the partitional properties of possibility correspondences in partitional information structures. Note, however, that $\Pi_i$ does not necessarily partition the state space $\Sigma$. In particular, there could be states in some space at which the possibility set of agent $i$ is in a different space. Nevertheless, all of the properties of knowledge associated with partitional possibility correspondences in standard state-space models are satisfied (see below). We also consider the case where only properties (0), (3), (4), and (5) are imposed, and refer to this larger class of models as *generalized HMS structures*.[7]

The knowledge operator can now be defined.

$$K_i(E) = \{w \in \Sigma \mid \Pi_i(w) \subseteq E\}$$

if there is a state $w$ such that $\Pi_i(w) \subseteq E$, or

$$K_i(E) = \varnothing^{S(E)}$$

otherwise.

Awareness is defined by:[8]

$$A_i(E) = \{w \in \Sigma \mid \Pi_i(w) \subseteq S(E)^{\uparrow}\}$$

if there is such a state $w$ such that $\Pi_i(w) \subseteq S(E)^{\uparrow}$, or

$$A_i(E) = \varnothing^{S(E)}$$

otherwise.

Heifetz *et al.* [8] show that in HMS structures, the knowledge operator satisfies the following properties:

---

[7]Halpern & Rego [7] also consider such a generalization. Heifetz *et al.* [10] introduce unawareness structures with probabilistic beliefs that are not required to satisfy generalized reflexivity.

[8]Note that Heifetz *et al.* [8] use a different definition of unawareness: $A_i(E) = K_i(E) \cup K_i(\neg K_i(E))$ (this definition was first adopted by Modica & Rustichini [11]). Although the two definitions coincide in HMS structures (see Halpern & Rego [7] or Heifetz *et al.* [9], Remark 6), the Modica & Rustichini definition is not appropriate for the case of generalized HMS structures.

(o) If $E$ is an event, then $K_i(E)$ is an $S(E)$-based event

(i) $K_i(\Sigma) = \Sigma$          Necessitation

(ii) $K_i\left(\bigcap_{\lambda \in L} E_\lambda\right) = \bigcap_{\lambda \in L} K_i(E_\lambda)$          Conjunction

(iii) $K_i(E) \subseteq E$          Truth

(iv) $K_i(E) \subseteq K_i K_i(E)$          Positive Introspection

(v) $E \subseteq F$ implies $K_i(E) \subseteq K_i(F)$          Monotonicity

(vi) $\neg K_i(E) \cap \neg K_i \neg K_i(E) \subseteq \neg K_i \neg K_i \neg K_i(E)$          Negative Non-Introspection

And Heifetz *et al.* [8, 9] show that in HMS structures the following properties hold for awareness and knowledge-awareness interaction. Letting $U_i(E)$ abbreviate $\neg A_i(E)$:

(vii) $A_i(E) = K_i(E) \cup K_i(\neg K_i(E))$          Plausibility

(viii) $K_i U_i(E) = \varnothing^{S(E)}$          *KU* Introspection

(ix) $U_i(E) = U_i U_i(E)$          *AU* Introspection

(x) $A_i(E) = K_i(S(E)^\uparrow)$          Weak Necessitation

(xi) $U_i(E) = \bigcap_{n=1}^{\infty} (\neg K_i)^n(E)$          Strong Plausibility

(xii) $\neg K_i(E) \cap A_i \neg K_i(E) = K_i \neg K_i(E)$          Weak Negative Introspection

(xiii) $A_i(\neg E) = A_i(E)$          Symmetry

(xiv) $\bigcap_{\lambda \in L} A_i(E_\lambda) = A_i\left(\bigcap_{\lambda \in L} E_\lambda\right)$          *A*-Conjunction

(xv) $A_i K_i(E) = A_i(E)$          *AK*-Self Reflection

(xvi) $A_i A_i(E) = A_i(E)$          *AA*-Self Reflection

(xvii) $K_i A_i(E) = A_i(E)$          *A*-Introspection

For generalized HMS structures, however, some of these properties may not hold.

**Proposition 3** *Consider a generalized HMS structure:*

1. *The knowledge operator $K_i$ satisfies property (o), Necessitation, Conjunction, and Monotonicity.*

2. *The awareness and knowledge operators, $A_i$ and $K_i$, satisfy Plausibility with "$\supseteq$", AU Introspection, Weak Necessitation, Strong Plausibility with "$\subseteq$", Symmetry, A-Conjunction, AK-Self Reflection, AA-Self Reflection, and A-Introspection with "$\subseteq$".*

In the appendix, we show that Negative Non-Introspection and Weak Negative Introspection may fail when Stationarity is violated (Example 1), and that $KU$-Introspection may fail when both Generalized Reflexivity and Stationarity are violated (Example 2).

To illustrate the mechanics of HMS structures, consider the example from the previous subsection. Recall that agent 1 is aware of object a, and knows whether or not it is $P$, while agent 2 is aware of object b, and knows whether or not it is $Q$. An HMS structure describing this situation shown in Figure 2 below.
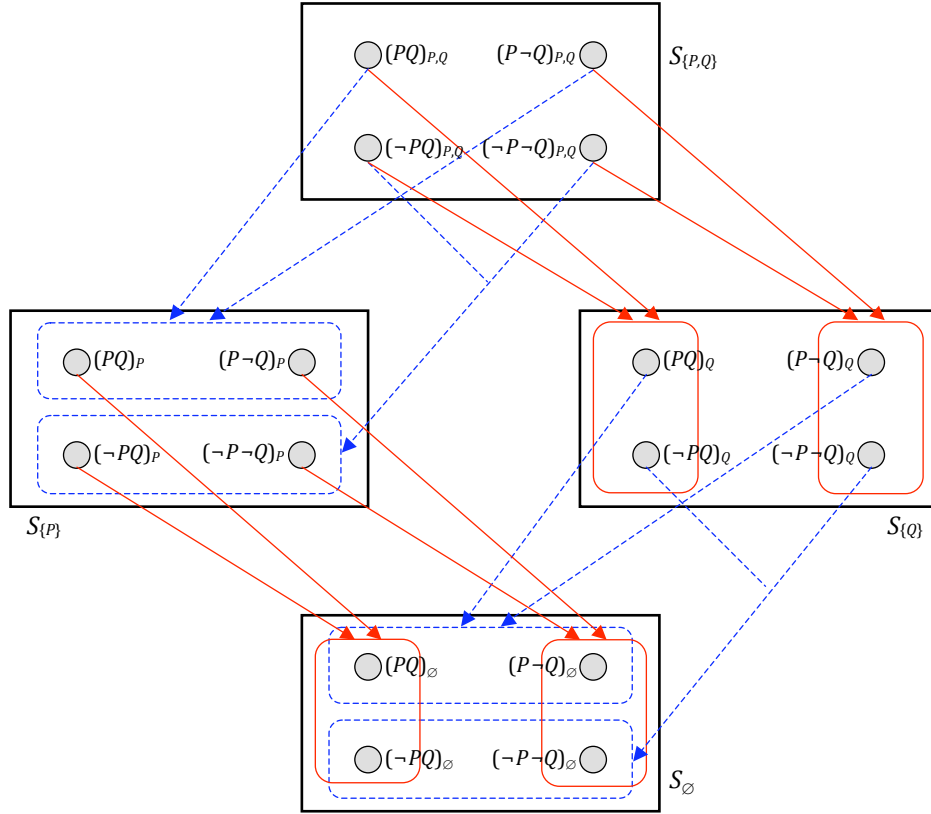


Figure 2: Example of an HMS structure

The four state spaces $\mathcal{S} = \{S_{\{P,Q\}}, S_{\{P\}}, S_{\{Q\}}, S_\varnothing\}$ (using the obvious notation) are indicated by solid black rectangles. The states are labeled according to whether or not a is $P$ and b is $Q$, and according to the expressiveness of the vocabulary. For example, at state $(\neg PQ)_Q \in S_{\{Q\}}$, object a is not $P$, but object b is $Q$, and the vocabulary of that state space is rich enough only to talk about whether or not object b is $Q$; and at state $(PQ)_\varnothing \in S_\varnothing$, a is $P$ and b is $Q$, although the vocabulary of that state space is not rich enough to talk about either

11

of these properties. The $r$ projection functions have been omitted to avoid clutter, but can be figured out from the descriptions of the states, so for example $r_{S_{\{P\}}}^{S_{\{P,Q\}}}(P\neg Q)_{P,Q} = (P\neg Q)_P$, $r_{S_{\{Q\}}}^{S_{\{P,Q\}}}(P\neg Q)_{P,Q} = (P\neg Q)_Q$, and $r_{S_\varnothing}^{S_{\{P,Q\}}}(P\neg Q)_{P,Q} = (P\neg Q)_\varnothing$.

The possibility correspondences for each agent are described by the colored arrows and ovals, dashed blue for agent 1 and solid red for agent 2. For a given state $w$, if $w$ is in an oval, then $\Pi_i(w)$ is given by all of the states in that oval; otherwise, if $\Pi_i(w)$ is given by the the states at which the arrow from $w$ points. For example, $\Pi_1((PQ)_{P,Q}) = \Pi_1((PQ)_P) = \{(PQ)_P, (P\neg Q)_P\}$, while $\Pi_1((PQ)_Q) = \Pi_1((PQ)_\varnothing) = \{(PQ)_\varnothing, (P\neg Q)_\varnothing\}$.

Recall that in an HMS structure, an event is a subset of some base space, along with all the states in more expressive state spaces that project onto a state in this subset. Thus in the HMS structure above, there are 64 distinct events (see Table 3 below).

| Base space | | | |
|---|---|---|---|
| $S_{\{P,Q\}}$ | $S_{\{P\}}$ | $S_{\{Q\}}$ | $S_\varnothing$ |
| $\varnothing^{S_{\{P,Q\}}}$ | $\varnothing^{S_{\{P\}}}$ | $\varnothing^{S_{\{Q\}}}$ | $\varnothing^{S_\varnothing}$ |
| $\{(PQ)_{P,Q}\}$ | $\{(PQ)_P, (PQ)_{P,Q}\}$ | $\{(PQ)_Q, (PQ)_{P,Q}\}$ | $\vdots$ |
| $\{(P\neg Q)_{P,Q}\}$ | $\{(P\neg Q)_P, (P\neg Q)_{P,Q}\}$ | $\{(P\neg Q)_Q, (P\neg Q)_{P,Q}\}$ | |
| $\{(\neg PQ)_{P,Q}\}$ | $\vdots$ | $\vdots$ | |
| $\{(\neg P\neg Q)_{P,Q}\}$ | | | |
| $\{(PQ)_{P,Q}, (P\neg Q)_{P,Q}\}$ | | | |
| $\vdots$ | | | $\Sigma$ |

Table 3: HMS events

In section 2.1 above, we considered an OBU event $\mathsf{E}$ corresponding to the sentence "object $\mathsf{a}$ is $P$", and an OBU event $\mathsf{F}$ corresponding to the sentence "nothing is $Q$". In the HMS structure just described, the equivalent events would be

$E = \{(PQ)_P, (P\neg Q)_P, (PQ)_{P,Q}, (P\neg Q)_{P,Q}\}$, and
$F = \{(P\neg Q)_\varnothing, (\neg P\neg Q)_\varnothing, (P\neg Q)_P, (\neg P\neg Q)_P, (P\neg Q)_Q, (\neg P\neg Q)_Q, (P\neg Q)_{P,Q}, (\neg P\neg Q)_{P,Q}\}$

We analyze the agent's awareness and knowledge of these events in the tables 4 and 5 below.

|  | Agent 1 | Agent 2 |
| --- | --- | --- |
| basic event | $E$ | $E$ |
| awareness | $A_1(E) = S_{\{P\}} \cup S_{\{P,Q\}}$ | $A_2(E) = \varnothing^{S_{\{P\}}}$ |
| knowledge | $K_1(E) = E$ | $K_2(E) = \varnothing^{S_{\{P\}}}$ |

Table 4: "object a is $P$"

|  | Agent 1 | Agent 2 |
| --- | --- | --- |
| basic event | $F$ | $F$ |
| awareness | $A_1(F) = \Sigma$ | $A_2(F) = \Sigma$ |
| knowledge | $K_1(F) = \varnothing^{S_\varnothing}$ | $K_2(F) = F$ |

Table 5: "nothing is $Q$"

# 3 Comparing the two models

We've shown how both an OBU structure (Figure 1) and an HMS structure (Figure 2) can be used to represent a simple story involving two agents and two propositions. Although the approaches are clearly different, it is natural to ask whether there are any equivalences between the two. In this section, we argue that *all of the relevant information encoded in the HMS structure is embedded within the OBU structure*, and *all of the relevant information encoded in the OBU structure is embedded within the HMS structure*.

Given the very different interpretations of their respective models provided by Board & Chung [2] and by Heifetz *et al.* [8], we believe these results add value to both types of model: in particular, they show that the HMS model can implicitly handle quantified events such as event $F$ above ("nothing is $Q$"); and that the unawareness analyzed by the OBU model, even though motivated in terms of objects, can be re-interpreted in terms of a proposition-based approach.

The obvious starting point when comparing the two structures depicted in figures 1 and 2 is to consider the event spaces in each case. Notice that there are 64 distinct events in each case. For each HMS event there is a corresponding OBU event, and for each OBU event there is a corresponding HMS event. This is the key to the embedding results. To illustrate the first half of this claim, consider HMS event $E$ ("object $a$ is $P$"). This corresponds to OBU event E. In general, for a given HMS event, the sense of the equivalent OBU event is determined by the base space of the HMS event, and the reference is determined by the states in the richest state space of the HMS structure. Of course showing this correspondence is not sufficient to demonstrate equivalence between the two models. We must also show

correspondence is preserved under negation, conjunction, knowledge and awareness. Tables 1 – 4 demonstrate that this is indeed the case for events $E$ and $F$. We now formalize the notion of an embedding, and prove our main result.

Let a *knowledge-awareness structure* (KA structure) consist of

1. a set of events, $\mathcal{E}$

2. a negation operator, $\neg : \mathcal{E} \to \mathcal{E}$

3. a conjunction operator, $\wedge : \mathcal{E} \times \mathcal{E} \to \mathcal{E}$

4. a knowledge operator for each agent $i$, $K_i : \mathcal{E} \to \mathcal{E}$

5. an awareness operator for each agent $i$, $A_i : \mathcal{E} \to \mathcal{E}$

(With slight abuse of terminology, in what follows we use "OBU structure" (or "HMS structure") to refer to the KA structure derived from a particular OBU (or HMS) structure, as well to refer to the original structure itself.)

Take two KA structures, $M^1 = (\mathcal{E}^1, \neg^1, \wedge^1, K_i^1, A_i^1)$ and $M^1 = (\mathcal{E}^2, \neg^2, \wedge^2, K_i^2, A_i^2)$. We say that $M^1$ can be *embedded in* $M^2$ if there is an injective function $f : \mathcal{E}^1 \to \mathcal{E}^2$ with the following properties

- $f(\neg^1 E) = \neg^2 f(E)$

- $f(E \wedge^1 F) = f(E) \wedge^2 f(F)$

- $f(K_i^1(E)) = K_i^2 f(E)$

- $f(A_i^1(E)) = A_i^2 f(E)$

Our main result says that generalized HMS structures can be embedded in OBU structures and *vice versa*.

**Theorem 1**

**(a)** *Every generalized HMS structure can be embedded in some OBU structure;*

**(b)** *Every OBU structure can be embedded in some generalized HMS structure.*

The need to consider generalized HMS structures to obtain this embedding result arises (roughly) because Heifetz *et al.* [8] impose "partitional" properties on their possibility correspondences (Generalized Reflexivity and Stationarity), while Board & Chung [2] make no such restrictions. This raises an obvious question: Do the embeddings preserve these partitional properties? To be more precise, can we embed every HMS structure into an OBU structure that is partitional in some appropriate sense, and *vice versa*? To answer this question, consider the following restrictions on an OBU structure:

1. reflexivity: $\mathsf{w} \in \mathcal{I}_i(\mathsf{w})$

2. stationarity: If $\mathsf{w}' \in \mathcal{I}_i(\mathsf{w})$ then $\mathcal{I}_i(\mathsf{w}') = \mathcal{I}_i(\mathsf{w})$

3. measurability: $\mathsf{w}' \in \mathcal{I}_i(\mathsf{w})$, then $\mathcal{A}_i(\mathsf{w}') = \mathcal{A}_i(\mathsf{w})$

The following proposition now provides a partial answer to our question:

**Proposition 4**

**(a)** *Every HMS structure can be embedded in some OBU satisfying reflexivity and station-arity;*

**(b)** *Every OBU structure satisfying reflexivity, stationarity and measurability can be embed-ded in some HMS structure.*

The difference between part (a) and the *quasi*-converse, (b), is the measurability condi-tion. Example 3 in the appendix shows that, with the embedding defined in the proof of Theorem 1(a), not every HMS structure can be embedded in an OBU structure satisfying reflexivity, stationarity *and* measurability. It is an open question whether measurability can be preserved with an alternative embedding. We can show that measurability of the OBU structure is required for part (b) of the Proposition: Example 4 in the appendix provides an OBU structure that satisfies reflexivity and stationarity but not measurability that cannot be embedded in an HMS structure with any embedding (of course, by Theorem 1(b), it can be embedded in some generalized HMS structure).

# 4   Conclusion

We have attempted to provide a direct comparison between two rather different approaches to modeling agents' unawareness: the object-based approach of Board & Chung [2] and the subjective-state-space approach of Heifetz *et al.* [8]. Our main result, Theorem 1, shows HMS structures can be embedded within OBU structures, and *vice versa.*

We believe that this result helps us understand both models better, adding value to each. In the case of OBU structures, the embedding result shows that it is valid to interpret the set of objects as propositions, so that (un)awareness of basic propositions provides the foun-dations for (un)awareness of more complex propositions, as is the case in most of the other related literature. On the other hand, we believe there may be some benefit in maintaining the distinction the objects and the properties they satisfy. Although the OBU structures described above derive an agent's unawareness of propositions from her unawareness of the objects described by those propositions, one can envisage an extension where unawareness of

properties is also modeled. A property-unawareness function could work (roughly) as follows: if an agent is unaware of a given property, then she would be unaware of any event containing one state but not another, where the two states could only be distinguished by whether or not various objects satisfied that property. Combining such a property-unawareness function with the object-unawareness function analyzed above would allow us to separate two kinds of unawareness: and agent could be unaware that "Yao Ming is tall" either because she has no idea who Yao Ming is or because she does not understand the concept of height.

The embedding results may also allow a re-interpretation of HMS structures. Note that both OBU structures and HMS structures can be used to provide semantics for formal languages that can be used to describe what agents know and what they are aware of. Both sets of authors provide details of how this can be done, along with sound and complete axiomatizations (Board & Chung [2] and Heifetz *et al.* [9]).[9] But while Heifetz *et al.* use propositional modal logic, Board & Chung use the considerably richer language of first-order modal logic, allowing us to separate objects from the properties they may or may not satisfy, and also enabling quantification over those objects. The proof of Theorem 1 (b) above suggests that HMS structures can be used to capture quantification, and therefore perhaps also to provide semantics for a first-order logic.

# A   Appendix

**Proof of Proposition 3.**

**Part 1:**

**(o)** Suppose that $E$ is an event. We need to show that there exists a basis $D \subseteq S(E)$ such that $D^\uparrow = K_i(E)$. Assume $K_i(E)$ is nonempty. Define $D = \{w \in S(E) \mid \Pi_i(w) \subseteq E\}$. By Confinement and the definition of $K_i$-operator, $D = K_i(E) \cap S(E)$.

We first show that $D^\uparrow \subseteq K_i(E)$. Let $w \in D^\uparrow$, $w \in S$ for some $S \succeq S(E)$. Then $w_{S(E)} \in D$. Hence, by definition of $D$ and Confinement, $\Pi_i(w_{S(E)}) \subseteq E \cap S(E)$. We claim that $\Pi_i(w) \subseteq E$. Since $w \in D^\uparrow$, $w \in S$ for $S \succeq S(E)$ then by Confinement $\Pi_i(w) \subseteq S'$ for some $S' \preceq S$. By Projections preserve ignorance, $\Pi_i^\uparrow(w) \subseteq \Pi_i^\uparrow(w_{S(E)})$. Hence $S' \succeq S(E)$. By Projections preserve knowledge, $(\Pi_i(w))_{S(E)} = \Pi_i(w_{S(E)})$. Since $\Pi_i(w_{S(E)}) \subseteq E \cap S(E)$, we have $(\Pi_i(w))_{S(E)} \subseteq E \cap S(E)$. Hence $\Pi_i(w) \subseteq E$. Thus by the definition of the $K_i$-operator, $w \in K_i(E)$.

Next, we show that $K_i(E) \subseteq D^\uparrow$. Let $w \in K_i(E)$. By definition of the $K_i$-operator, $\Pi_i(w) \subseteq E$. Let $w \in S'$. By Confinement, there exists space a $S \preceq S'$ such that $\Pi_i(w) \subseteq S$. Since $\Pi_i(w) \subseteq E$ we must have $S \succeq S(E)$. Since $\Pi_i(w) \subseteq E$, we have $(\Pi_i(w))_{S(E)} \subseteq$

---

[9]Halpern & Rego [7] provide two alternative axiomatizations of HMS structures

$E \cap S(E)$. By Projections preserve knowledge, $(\Pi_i(w))_{S(E)} = \Pi_i(w_{S(E)})$. Hence $\Pi_i(w_{S(E)}) \subseteq E \cap S(E)$. Therefore $w_{S(E)} \in D$ and thus $w \in D^\uparrow$.

Finally, if $K_i(E)$ is empty, then by the definition of the $K_i$-operator we have $K_i(E) = \varnothing^{S(E)}$.

**(i)**, **(ii)** and **(v)**. The proofs of Proposition 2 (i), (ii) and (v) in Heifetz *et al.* [8] apply respectively.

**Part 2:** For convenience, we prove the properties in a different order:

**(x)** Weak Necessitation: This follows directly from the definition of the awareness and knowledge operators.

**(ix)** AU-Introspection: $U_i(E) = U_i U_i(E)$. This is equivalent to $A_i(E) = A_i U_i(E)$. By Weak Necessitation and the definition of the unawareness operator, $A_i U_i(E) = K_i(S(U_i(E)^\uparrow)) = K_i(S(E)^\uparrow) = A_i(E)$.

**(vii)** Plausibility with "$\supseteq$": $U_i(E) \subseteq \neg K_i(E) \cap \neg K_i \neg K_i(E)$. This is equivalent to $A_i(E) \supseteq K_i(E) \cup K_i \neg K_i(E)$. By definition of the awareness operator, $K_i(E) \subseteq A_i(E)$. By Weak Necessitation and the definition of the awareness operator, $K_i \neg K_i(E) \subseteq A(\neg K_i(E)) = A_i(E)$. Hence the property follows.

**(x)** Strong Plausibility with "$\subseteq$": By definition of the awareness operator, $K_i(E) \subseteq A_i(E)$. By Weak Necessitation and the definition of the awareness operator,

$$K_i(\neg K_i)^n(E) \subseteq A_i((\neg K_i)^n(E)) = A_i(E)$$

for any $n = 1, \dots$. Hence the property follows.

**(xiii)**, **(xiv)** and **(xvi)**. The proofs of Proposition 3 (6), (7) and (9) in Heifetz *et al.* [8] apply respectively.

**(xiv)** AK-Self Reflection: $A_i K_i(E) = A_i(E)$. By Weak Necessitation and the definition of the knowledge operator, $A_i K_i(E) = K_i(S(K_i(E))^\uparrow) = K_i(S(E)^\uparrow) = A_i(E)$.

**(xvi)** A-Introspection with "$\subseteq$": $K_i A_i(E) \subseteq A_i(E)$. By the definition of the awareness operator, $K_i A_i(E) \subseteq A_i A_i(E) = A_i(E)$, where the last equality follows from AA-Self reflection. ∎

The following example shows that without Stationarity, HMS structures may fail to satisfy Negative Non-Introspection and Weak Negative Introspection.

**Example 1** *Consider a HMS structure with one space only, $S = \{w_1, w_2\}$. The possibility correspondence of the single agent is given by $\Pi(w_1) = \{w_1\}$ and $\Pi(w_2) = S$. Note that this specification violates Stationarity. Consider the event $E = \{w_1\}$. Then $K(E) = \{w_1\}$, $\neg K(E) = \{w_2\}$, $K\neg K(E) = \varnothing$, $\neg K\neg K(E) = S$, $K\neg K\neg K(E) = S$, $\neg K\neg K\neg K(E) = \varnothing$, and $A\neg K(E) = S$. Thus, $\neg K(E) \cap \neg K\neg K(E) = \{w_2\} \nsubseteq \neg K\neg K\neg K(E) = \varnothing$, violating Negative Non-Introspection.*

*Moreover, $\neg K(E) \cap A\neg K(E) = \{w_2\} \nsubseteq K\neg K(E) = \varnothing$, violating Weak Negative Intro-spection.*

The next example shows that without Generalized Reflexivity and Stationarity, HMS structures may fail to satisfy $KU$-Introspection. (Note that Heifetz *et al.* [10] show that $KU$-Introspection holds without Generalized Reflexivity.)

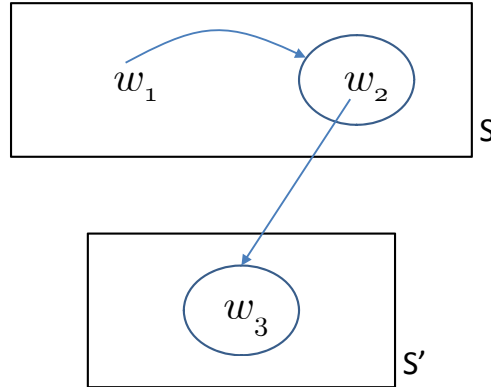**Example 2** *Consider the HMS structure shown in Figure 3. There is a totally ordered set of*



Figure 3: The failure of $KU$-Introspection

*two spaces, $S$ and $S'$. The possibility correspondence of the single agent is given by the ovals and arrows. Since $w_1 \notin \Pi^\uparrow(w_1)$, it violates Generalized Reflexivity. Moreover, $w_2 \in \Pi(w_1)$ but $\Pi(w_2) \neq \Pi(w_1)$. So the possibility correspondence violates Stationarity. Consider the event $E = \{w_1\}$. We have $A(E) = \{w_1\}$, $U(E) = \{w_2\}$, and $KU(E) = \{w_1\}$, violating $KU$-introspection.*

**Proof of Theorem 1.** As above, we use the symbols $\neg, \wedge, \mathsf{K}_i$, and $\mathsf{A}_i$ for OBU-negation, conjunction, knowledge and awareness; for the sake of clarity, we adopt the new symbols $\sim$ and $\curlywedge$ for HMS-negation and conjunction, but continue to use $K_i$ and $A_i$ for HMS-knowledge and awareness.

**(a)** We prove this direction of the embedding result result for a special class of generalized HMS structures (every example provided in Heifetz *et al.* [8] fits in this class). The extension to the class of all generalized HMS structures follows immediately from Lemma 1 below.

Say that a generalized HMS structure is *standard* if there is some set $\Phi$ and a labeling of the state spaces $\{S_\alpha\}_{\alpha \subseteq \Phi}$ such that $S_\alpha \preceq S_\beta$ if and only if $\alpha \subseteq \beta$.

From a standard generalized HMS model, construct an OBU model as follows:

- $\mathsf{W} = S_\Phi$

- $\mathsf{O} = \Phi$

- $\mathcal{I}_i(w) = \Pi_i^\uparrow(w) \cap S_\Phi$ (effectively, this projects the states the agent considers possible back to the richest state space, $S_\Phi$)

- $\mathcal{A}_i(w) = \alpha$, where $\alpha \subseteq \Phi$ satisfies $\Pi_i(w) \subseteq S_\alpha$ (note that by the HMS Confinement condition, this $\alpha$ is unique).

For any HMS-event $E$ with base space $S_\alpha$, we define $f(E) = (E \cap S_\Phi, \alpha)$. To see that $f$ is injective, note that two HMS-events $E$ and $F$ can differ only they have different bases; thus either (i) they have different base spaces, or (ii) their bases are different subsets of the same base space $S$. In the first case, the senses of $f(E)$ and $f(F)$ must differ, and in the second case their references must differ.

Next, let $E$ be an HMS-event, with base space $S_\alpha$ and basis $B$, and $F$ be an HMS-event with base space $S_\beta$ and basis $C$ (we ignore the cases where $B = \varnothing$ or $C = \varnothing$, which are straightforward). Then:

**Negation:** $f(\sim E) = f\left((S_\alpha \setminus B)^\uparrow\right) = (S_\Phi \setminus E, \alpha) = \neg(E \cap S_\Phi, \alpha) = \neg f(E)$

**Conjunction:** $f(E \curlywedge F) = f(E \cap F) = (E \cap F \cap S_\Phi, \alpha \cup \beta) = (E \cap S_\Phi, \alpha) \wedge (F \cap S_\Phi, \beta) = f(E) \wedge f(F)$

**Knowledge:**

$$f(K_i(E)) = f(\{w \in \Sigma \mid \Pi_i(w) \subseteq E\})$$
$$= f(D^\uparrow) \text{ for some } D \subseteq S_\alpha \text{ (by property (o) on page 10, since } S_\alpha = S(E))$$
$$= (X, \alpha), \text{ where } X = D^\uparrow \cap S_\Phi.$$

Take any $w \in X$. Then (i) $\Pi_i(w) \subseteq E$. The HMS Confinement conditions requires that $\Pi_i(w) \subseteq S_\gamma$ for some $\gamma \in \Phi$, and $\Pi_i(w) \subseteq E$ implies that $\alpha \subseteq \gamma$. Since $\mathcal{A}_i(w) = \gamma$ by construction, we have $\alpha \subseteq \mathcal{A}_i(w)$; and (ii) $\mathcal{I}_i(w) = \Pi_i^\uparrow(w) \cap S_\Phi \subseteq E \cap S_\Phi$, since $\Pi(w) \subseteq E$.

Next, for any $w \in S_\Phi$, suppose $\mathcal{I}_i(w) \subseteq E \cap S_\Phi$ and $\alpha \subseteq \mathcal{A}_i(w)$. Then $\Pi_i^\uparrow \cap S_\Phi \subseteq E \cap S_\Phi$ and $\Pi_i(w) \subseteq S_\gamma$ for some $\gamma \supseteq \alpha$; hence, $\Pi_i(w) \subseteq E$, and so $w \in D^\uparrow \cap S_\Phi = X$. Thus,

$$X = \{w \in S_\Phi \mid \mathcal{I}_i(w) \subseteq E \cap S_\Phi \text{ and } \alpha \subseteq \mathcal{A}_i(w)\},$$

and so $f(K_i(E)) = (X, \alpha) = \mathsf{K}_i f(E)$, as required.

**Awareness:** Note that by the definition of the awareness operator, we have by (x) Weak Necessitation, $A_i(E) = K_i(S(E)^\uparrow)$ for any event $E$. So it follows from property (o) of the knowledge operator that $A_i(E)$ has base space $S_\alpha = S(E)$. Thus $f(A_i(E)) = (X, \alpha)$, where $X = A_i(E) \cap S_\Phi$. Again by (x) Weak Necessitation and the proof of the knowledge part it follows that

$$
\begin{aligned}
X &= \{w \in S_\Phi | \mathcal{I}_i \subseteq S_\alpha^\uparrow \cap S_\Phi \text{ and } \alpha \subseteq \mathcal{A}_i(w)\} \\
&= \{w \in S_\Phi | \alpha \subseteq \mathcal{A}_i(w)\}
\end{aligned}
$$

and so $f(A_i(E)) = (X, \alpha) = \mathsf{A}_i f(E)$ as required.

**(b)** For the other direction, start with an OBU structure $\langle \mathsf{W}, \mathsf{O}, \{\mathcal{I}_i\}, \{\mathcal{A}_i\} \rangle$ and define a generalized HMS structure $\left\langle \left(\{S_\alpha\}_{\alpha \in \Phi}, \succeq\right), \left(r_\beta^\alpha\right)_{\beta \preceq \alpha}, (\Pi_i) \right\rangle$ as follows:

- $\Phi = 2^\mathsf{O}$. Define a partial order on $\Phi$ by set inclusion, i.e. $\alpha \succeq \beta$ if and only if $\alpha \supseteq \beta$. Since the set of all subsets is a complete lattice, so is $\Phi$.

- $S_\alpha = \mathsf{W}$ for all $\alpha \in \Phi$. That is, each space $S_\alpha$ is a copy of $\mathsf{W}$. Rename copies of $\mathsf{w} \in \mathsf{W}$ in $S_\alpha$ by $w_\alpha$. Spaces are disjoint. The order $\Phi$ can be extended to an order on the spaces. Hence, $\mathcal{S} = \{S_\alpha\}_{\alpha \in \Phi}$ is a complete lattice.

- Projections are defined in the obvious way by for $\alpha \succeq \beta$, $\alpha, \beta \in \Phi$, $r_\beta^\alpha(w_\alpha) = w_\beta$. It is straightforward to verify that indeed projections are surjective, commute and are the identity when domain and codomain coincide.

- For $w \in S_\alpha$,
$$\Pi_i(w) = \mathcal{I}_i\left(r_{\inf\{\alpha, \mathcal{A}_i(w)\}}^\alpha(w)\right).$$

Note that $\inf\{\alpha, \mathcal{A}_i(w)\}$ is well defined since $\Phi$ is a complete lattice. Confinement follows by construction. Projections Preserve Ignorance follows with equality from the construction. Projections Preserve Knowledge follows by construction. Projections Preserve Awareness follows from previous properties (see Remark 3 in Heifetz *et al.* [8]).

Thus we have shown that this construction indeed defines a generalized HMS structure. Next, define an embedding by for any OBU event $(B, \alpha) \in \mathcal{E}_{OBU}$, set

$$g(B, \alpha) = (B^\uparrow, S_\alpha)$$

where $B$ is the basis and $S_\alpha$ the base space of the corresponding event in the generalized HMS structure. First, note that $g$ is injective, since OBU events differ if and only if their references (bases) are different or their senses (base spaces) are different. Then:

**Negation:** $g(\neg(B, \alpha)) = g(\mathsf{W} \setminus B, \alpha) = ((S_\alpha \setminus B)^\uparrow, S_\alpha) = (\sim B^\uparrow, S_\alpha) = \sim g(B, \alpha)$

**Conjunction:** $g((B_1, \alpha_1) \wedge (B_2, \alpha_2)) = g(B_1 \cap B_2, \alpha_1 \cup \alpha_2) = (B_1^\uparrow \cap B_2^\uparrow, S_{\sup\{\alpha_1, \alpha_2\}}) = (B_1^\uparrow, \alpha_1) \curlywedge (B_2^\uparrow, \alpha_2) = g(B_1, \alpha_1) \curlywedge g(B_2, \alpha_2)$

**Knowledge:**

$$
\begin{aligned}
g(\mathsf{K}_\mathrm{i}(B, \alpha)) &= g((\{w \mid \alpha \subseteq \mathcal{A}_i(w)\}, \alpha) \wedge (\{w \mid \mathcal{I}_i(w) \subset B\}, \alpha)) \\
&= g((\{w \mid \alpha \subseteq \mathcal{A}_i(w)\}, \alpha)) \wedge g((\{w \mid \mathcal{I}_i(w) \subset B\}, \alpha)) \\
&= (\{w \in S_\alpha \mid \alpha \subseteq \mathcal{A}_i(w)\}^\uparrow, S_\alpha) \wedge (\{w \in S_\alpha \mid \mathcal{I}_i(w) \subseteq B\}^\uparrow, S_\alpha) \\
&= (\{w \in S_\alpha \mid \alpha \subseteq \mathcal{A}_i(w)\}^\uparrow \cap \{w \in S_\alpha \mid \mathcal{I}_i(w) \subseteq B\}, S_\alpha) \\
&= (\{w \in S_\alpha \mid \Pi_i(w) \subseteq B\}^\uparrow, S_\alpha) \\
&= (\{w \in \Sigma \mid \Pi_i(w) \subseteq (B^\uparrow, S_\alpha)\}, S_\alpha) \\
&= (K_i(B^\uparrow, S_\alpha), S_\alpha) \\
&= K_i(B^\uparrow, S_\alpha) \\
&= K_i(g(B, \alpha))
\end{aligned}
$$

**Awareness:**

$$
\begin{aligned}
g(\mathsf{A}_\mathrm{i}(B, \alpha)) &= g(\{w \mid \alpha \subseteq \mathcal{A}_i(w)\}, \alpha) \\
&= (\{w \in S_\alpha \mid \alpha \subseteq \mathcal{A}_i(w)\}^\uparrow, S_\alpha) \\
&= (\{w \in S_\alpha \mid \Pi_i(w) \subseteq S_\alpha\}^\uparrow, S_\alpha) \\
&= (\{w \in \Sigma \mid \Pi_i(w) \subseteq S_\alpha^\uparrow\}, S_\alpha) \\
&= (A_i(B^\uparrow, S_\alpha), S_\alpha) \\
&= A_i(B^\uparrow, S_\alpha) \\
&= A_i(g(B, \alpha))
\end{aligned}
$$

∎

**Proof of Proposition 4.**

**(a)** We use the same embedding as in the proof of Theorem 1(a). It remains to show that if the generalized HMS structure satisfies Generalized Reflexivity and Stationarity (i.e. if we started with an HMS structure), then the OBU structure thus defined satisfies reflexivity and stationarity.

**Generalized Reflexivity of $\Pi_i$ implies reflexivity of $\mathcal{I}_i$.** Suppose $w \in \Pi_i^\uparrow(w)$, with $\Pi_i(w) \subseteq S$. Consider two cases: Case (i) $S = S_\Phi$. Then by Confinement $w \in \Pi_i^\uparrow(w) \cap S_\Phi = \Pi_i(w)$. Thus $w \in \mathcal{I}_i(w)$ as required. Case (ii) $S \preceq S_\Phi$. Let $w \in S'$. By Confinement $S_\Phi \succeq S' \succeq S$. Consider any $w' \in (r_{S'}^{S_\Phi})^{-1}(w)$. Note that $w' \in \Pi_i^\uparrow(w)$ and $w'_{S'} = w$. By Projections Preserve Ignorance, $\Pi_i^\uparrow(w') \subseteq \Pi_i^\uparrow(w)$. Hence, by Confinement $\Pi_i(w') \subseteq S''$ for some $S'' \succeq S$. Note that $w'_S = w_S$. Hence by Projections preserve knowledge $(\Pi_i(w))_S = (\Pi_i(w'))_S$. Thus $w' \in \Pi_i^\uparrow(w') \cap S_\Phi$ and we conclude that $w' \in \mathcal{I}_i(w')$, as required.

**Stationarity of $\Pi_i$ implies stationarity of $\mathcal{I}_i$.** Suppose $w' \in \Pi_i(w)$ with $\Pi_i(w) \subseteq S$. By Stationarity, $\Pi_i(w') = \Pi_i(w)$. Let $w \in S^*$. By Confinement, $S^* \succeq S$. Consider any $w''' \in (r_S^{S_\Phi})^{-1}(w')$ and $w'' \in (r_{S^*}^{S_\Phi})^{-1}(w)$. Note that $w''' \in \Pi_i^\uparrow(w)$ and $w'''_S = w'$ and $w''_S = w_S$. By Projections Preserve Ignorance, $\Pi_i^\uparrow(w''') \subseteq \Pi_i^\uparrow(w')$ and $\Pi_i^\uparrow(w'') \subseteq \Pi_i^\uparrow(w)$. Hence, by Confinement $\Pi_i(w''') \subseteq S'''$ and $\Pi_i(w'') \subseteq S''$ for some $S''' \succeq S$ and $S'' \succeq S^* \succeq S$. By Projections Preserve Knowledge, $(\Pi_i(w'''))_S = \Pi_i(w')$ and $(\Pi_i(w''))_S = \Pi_i(w_S) = \Pi_i(w)$. Note that $w''' \in \Pi_i^\uparrow(w'') \cap S_\Phi$. Moreover, since we previously observed that $\Pi_i(w') = \Pi_i(w)$ by Stationarity, we must have $\Pi_i^\uparrow(w''') \cap S_\Phi = \Pi_i^\uparrow(w'') \cap S_\Phi$. Therefore for $w''' \in \mathcal{I}_i(w'')$ we have $\mathcal{I}_i(w''') = \mathcal{I}_i(w'')$, as required.

**(b)** We use the same embedding as in the proof of Theorem 1(b). It remains to show that if the OBU structure satisfies reflexivity, stationarity and measurability, then the embedding defines an HMS structure.

Recall that for $w \in S_\alpha$,

$$\Pi_i(w) = \mathcal{I}_i\left(r_{\inf\{\alpha, \mathcal{A}_i(w)\}}^\alpha(w)\right).$$

Note that if the OBU structure satisfies measurability, then Generalized Reflexivity and Stationarity follow from reflexivity and stationarity of $\mathcal{I}_i$ respectively. ∎

The following example shows that with the embedding considered in the proofs of Theorem 1(a) and Proposition 4(a), not every HMS structure can be embedded into some OBU structure satisfying reflexivity, stationarity *and* measurability.
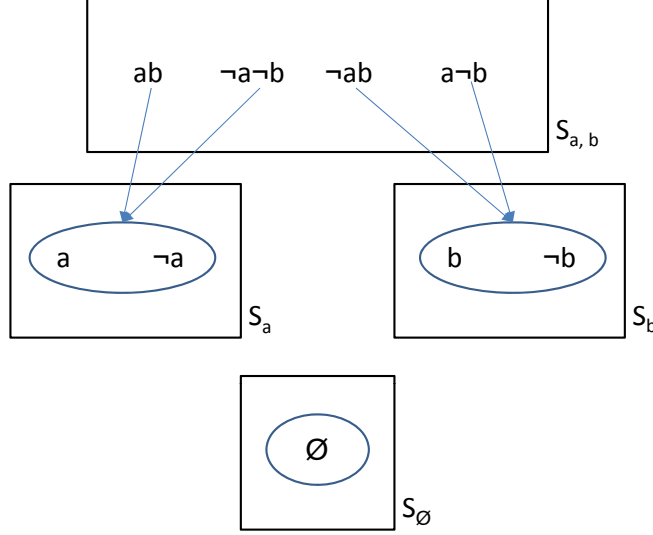
Figure 4: The failure of measurability

**Example 3** *Consider the HMS structure for one agent given in Figure 4. The possibility correspondence is given by the ovals and arrows on a lattice of four spaces. Notice that this is an HMS structure: In particular, Generalized Reflexivity and Stationarity are satisfied.*

*According to the embedding defined in Theorem 1(a), we have*

$$\mathcal{I}(ab) = \{a, \neg a\}^{\uparrow} \cap S_{a,b} = S_{a,b}$$
$$\mathcal{I}(\neg ab) = \{b, \neg b\}^{\uparrow} \cap S_{a,b} = S_{a,b}$$

*and*

$$\mathcal{A}(ab) = \{a\}$$
$$\mathcal{A}(\neg ab) = \{b\}$$

*So we have $(\neg ab) \in \mathcal{I}(ab)$ but $\mathcal{A}(ab) \neq \mathcal{A}(\neg ab)$. The measurability condition is not satisfied (though the OBU structure is reflexive and stationary).*

Example 4 shows that measurability is required for Proposition 4(b), i.e. that not every OBU structure satisfying reflexivity, stationarity but not measurability can be embedded into an HMS structure.

**Example 4** *Consider the following OBU structure for one agent:* $W = \{w_1, w_2\}$, $O = \{a, b\}$, $\mathcal{I}(w_1) = \mathcal{I}(w_2) = \{w_1, w_2\}$, $\mathcal{A}(w_1) = \{a, b\}$, *and* $\mathcal{A}(w_2) = \{a\}$. *Note that this OBU*

*structure satisfies both reflexivity and stationarity but does not satisfy measurability. Let event* $\mathsf{E} = (\mathsf{W}, \mathsf{O})$. *Then* $\mathsf{A}(\mathsf{E}) = (\mathsf{w}_1, \mathsf{O})$, *while* $\mathsf{K}(\mathsf{A}(\mathsf{E})) = (\varnothing, \mathsf{O})$. *Suppose this OBU structure can be embedded into some HMS structure, with embedding function* $f$. *Since* $f$ *is injective, we have* $f(\mathsf{A}(\mathsf{E})) \neq f(\mathsf{K}(\mathsf{A}(\mathsf{E})))$, *and hence* $A(f(\mathsf{E})) \neq K(A(f(\mathsf{E})))$. *This violates A-Introspection, yielding a contradiction.*

To extend the proof of Theorem 1(a) to the class of all generalized HMS structures, it suffices to show that every generalized HMS structure can be embedded in a standard generalized HMS structure.

**Lemma 1** *Every generalized HMS structure can be embedded in a standard generalized HMS structure.*

**Proof.** Before we start, we develop an alternative representation of generalized HMS structures that is more convenient for our purposes.

Fix an arbitrary generalized HMS structure $\langle \mathcal{S}, \preceq, r, \Pi_i \rangle$, where $\mathcal{S} = \{S_\alpha\}_{\alpha \in A}$ is a complete lattice of disjoint state spaces. Let $S_{\overline{\alpha}}$ and $S_{\underline{\alpha}}$ be the maximal and minimal state spaces, respectively. Let $X = S_{\overline{\alpha}}$. For any $\alpha$, define a class of subsets of $X$ as follows:

$$\mathcal{P}_\alpha := \{(r_\alpha^{\overline{\alpha}})^{-1}(w_\alpha) \mid w_\alpha \in S_\alpha\}.$$

We use $P_\alpha$ to denote a generic element of $\mathcal{P}_\alpha$. Since $r$ is surjective, $\mathcal{P}_\alpha$ is a partition of $X$. Moreover, there is a one-to-one correspondence between $\mathcal{P}_\alpha$ and $S_\alpha$. Let $f : \cup_\alpha S_\alpha \to \cup_\alpha \mathcal{P}_\alpha$ denote this one-to-one correspondence. The advantage of working with the $\mathcal{P}_\alpha$'s rather than the $S_\alpha$'s is that we can take joins and meets of partitions and create new partitions.

There are a few properties of this one-to-one correspondence that are worth pointing out. First, if $S_\alpha \preceq S_{\alpha'}$, then $\mathcal{P}_\alpha$ is a weakly coarser partition of $X$ than $\mathcal{P}_{\alpha'}$. Moreover, for any $w \in S_\alpha$ and $w' \in S_{\alpha'}$, $r_\alpha^{\alpha'}(w') = w$ if and only if $f(w') \subseteq f(w)$. Third, for any $\alpha$, $\mathcal{P}_\alpha = \bigvee \{\mathcal{P}_{\alpha'} \mid \alpha' \succeq \alpha\}$, where $\bigvee$ is the join operator on partitions.

Now, let's embed this arbitrary generalized HMS structure into a standard generalized HMS structure. Recall that an generalized HMS structure is *standard* if there is some set $\Phi$ and a labeling of the state spaces $\{\hat{S}_\beta\}_{\beta \subseteq \Phi}$ such that $\hat{S}_\beta \preceq \hat{S}_{\beta'}$ if and only if $\beta \subseteq \beta'$. Recall that $A$ is the index set in the original HMS structure, and $S_{\underline{\alpha}}$ is the minimal state space. Let $\Phi := A \setminus \{\underline{\alpha}\}$. Define a mapping $g : A \to 2^\Phi$ as follows: $g(\alpha) := \{\alpha' \in \Phi \mid \alpha' \preceq \alpha\}$. Let $B = g(A)$. Note that $g$ is a one-to-one correspondence between $A$ and $B$. Also note that, since $g(\underline{\alpha}) = \varnothing$ and $g(\overline{\alpha}) = \Phi$, we have $\varnothing, \Phi \in B$.

We now construct a class, $\{\mathcal{P}_\beta\}_{\beta \subseteq \Phi}$, of partitions of $X$. For any $\beta \subseteq \Phi$, define

$$\mathcal{P}_\beta := \bigvee \{\mathcal{P}_\alpha \mid g(\alpha) \supseteq \beta\}.$$

24

This construction has the following nice property: for any $\beta$ and $\beta'$ such that $\beta \subseteq \beta'$, we have $\{\mathcal{P}_\alpha \mid g(\alpha) \supseteq \beta\} \subseteq \{\mathcal{P}_\alpha \mid g(\alpha) \supseteq \beta'\}$, and hence $\mathcal{P}_\beta$ is a weakly coarser partition of $X$ than $\mathcal{P}_{\beta'}$. Moreover, For any $\alpha$ and $\beta$ such that $\beta = g(\alpha)$, we have $\mathcal{P}_\beta = \mathcal{P}_\alpha$. So there is a one-to-one correspondence between $\cup_{\alpha \in A} \mathcal{P}_\alpha$ and $\cup_{\beta \in B} \mathcal{P}_\beta$. Let's abuse notation and use $g$ to denote this one-to-one correspondence as well. Note that $g$ preserves both order and projections.

We use these $\mathcal{P}_\beta$'s as the state spaces in our standard generalized HMS structure. The partial ordering of these state spaces is the natural one, and so is the projection. It remains to define the possibility correspondences, $\hat{\Pi}_i$, in the standard generalized HMS structure. For any $\beta \in B$, and any $P_\beta \in \mathcal{P}_\beta$, the construction is straightforward: simply define $\hat{\Pi}_i(P_\beta) := g(f(\Pi_i(f^{-1}(g^{-1}(P_\beta)))))$. For any other $P_\beta$'s, the specification of $\Pi_i$ can be arbitrary. For example, we can simply define $\Pi_i(P_\beta) := \{P_\beta\}$.

That this standard generalized HMS structure embeds the original generalized HMS structure is obvious. ∎

# References

[1] Board, O.J. and K.-S. Chung (2007), "Object-based unawareness: axioms". Working paper, University of Minnesota, Minneapolis, MN.

[2] Board, O.J. and K.-S. Chung (2008), "Object-based unawareness: theory and applications". Working paper, University of Minnesota, Minneapolis, MN.

[3] Chung, K.-S. and L. Fortnow (2006), "Loopholes". Working paper, Northwestern University, Evanston, IL.

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

[5] Filiz-Ozbay, E. (2008), "Incorporating unawareness into contract theory". Working paper, University of Maryland.

[6] Halpern, J.Y. (1999), "Set-theoretic completeness for epistemic and conditional logic". *Annals of Mathematics and Artificial Intelligence* **26**, 1–27.

[7] Halpern, J.Y. & J.C. Rego (2008), "Interactive unawareness revisited". *Games and Economic Behavior* **62**, 232–262.

[8] Heifetz, A., M. Meier, and B.C. Schipper (2006), "Interactive unawareness". *Journal of Economic Theory* **130**, 78–94.

[9] Heifetz, A., M. Meier, and B.C. Schipper (2008), "A canonical model for interactive unawareness". *Games and Economic Behavior* **62**, 305–324.

[10] Heifetz, A., M. Meier, and B.C. Schipper (2009), "Unawareness, beliefs and speculative trade". Working paper, University of California, Davis.

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

[12] Tirole, J. (2008), "Cognition and incomplete contracts". *American Economic Review* forthcoming.