Theory Dec. (2015) 79:601–613
DOI 10.1007/s11238-015-9485-0

Sub-models for interactive unawareness

Simon Grant · J. Jude Kline ·
Patrick O'Callaghan · John Quiggin

Published online: 6 February 2015

Abstract We propose a notion of a sub-model for each agent at each state in the Heifetz et al. (2006) model of interactive unawareness. Presuming that each agent is fully cognizant of his sub-model causes no difficulty and fully describes his knowledge and his beliefs about the knowledge and awareness of others. We use sub-models to motivate the HMS conditions on possibility correspondences.

Keywords Interactive unawareness · Sub-models · Events

1 Introduction

An innovative approach that allows for unawareness in interactive settings has been proposed by Heifetz et al. (2006), hereafter HMS. They follow the event-based approach of Aumann (1976), Aumann (1999).[1] A logic-based approach to capture awareness and unawareness was developed earlier by Fagin and Halpern (1988). Sub-

[1] Fagin et al. (1995) used the term event-based approach and compared it to the logic-based approach.

S. Grant
Research School of Economics, Australian National University,
Canberra, Australia

S. Grant · J. J. Kline (✉) · P. O'Callaghan · J. Quiggin
School of Economics, University of Queensland, Brisbane, Australia
e-mail: j.kline@uq.edu.au

S. Grant
e-mail: simon.grant@anu.edu.au

P. O'Callaghan
e-mail: p.ocallaghan@uq.edu.au

J. Quiggin
e-mail: j.quiggin@uq.edu.au

sequently, both Heifetz et al. (2008) and Halpern and Rego (2008) have shown that logic-based approaches to unawareness correspond to the HMS event-based approach.

While the event-based approach of HMS builds on that of Aumann, an important difference is that Aumann presumed in his set-up that the model is " commonly known" by all the agents. The " commonly known" appears here in (scare) quotes to emphasize that it refers to a claim about knowledge of the model as distinct from knowledge of an event which is described within the model. A natural question is whether or not an HMS model can be " commonly known," or even " fully known" by one agent. If it is " fully known" by an agent who is modeled as being unaware of some aspects, then the agent might arguably, by inspecting his model, become aware of those aspects.

A related difficulty arises from the structure of the agent's representation of the model. If this representation lacks structural features of a full model, then the agent can infer that she is unaware of aspects of the model of which other agents might be aware, even without any understanding of what those aspects might be.

This creates some tension between the unawareness the model intends to capture and the possible discovery by an agent of her unawareness via her reasoning within the model. In this paper, we resolve this tension by proposing a notion of a sub-model for each agent at each state. We argue that we can interpret each agent as being fully cognizant of her sub-model without compromising any agent's awareness level. That is, even if an agent fully explores her sub-model, she will never become aware of any aspect outside the awareness level she is prescribed within the (grand) model. Moreover, the agent's sub-model inherits all structural features of the full model that might be used to make inferences about the agent's own awareness.

HMS imposed a set of conditions on the possibility correspondences of the agents. We strip down the model to its bare essentials for an analysis of sub-models. We then use the sub-models to motivate the HMS conditions on the possibility correspondences.

The paper proceeds as follows. In Section 2, we present a generalized version of the HMS framework with minimal conditions on possibility correspondences and define the concept of a sub-model. We show that each sub-model is itself a model. In Section 3, we give necessary and sufficient conditions for HMS knowledge of an event to be an event. In Section 4, we use sub-models to motivate the full set of HMS conditions on possibility correspondences. In Section 5, we re-visit the Holmes and Watson Example of Galanis (2013) to explore our results on sub-models. We conclude in Section 6. Proofs appear in the appendix.

2 Models and sub-models of interactive unawareness

HMS introduced a framework for analyzing interactive unawareness as an analog to Aumann's event-based approach to knowledge. In models of knowledge with full awareness, the standard interpretation is that each agent knows each other agent's partition, and, moreover, the entire model can be taken to be commonly known. This interpretation, however, presents some difficulty in models of unawareness. If an agent is fully cognizant of the model, how can she possibly be unaware of any state or any event in the model?

We argue that our notion of a sub-model allows each agent to have full cognizance of a sub-model at each state of the world. Although these sub-models will in general differ across states and agents, we can keep the interpretation that an agent uses her (sub-)model at a particular state to reason about the knowledge and awareness not just of herself but also of the other agents. This provides a coherent interpretation of knowledge in the HMS framework.

We begin with a definition of a model of unawareness that places fewer restrictions on the possibility correspondences than those imposed by HMS.

Definition 1 (Model) A *model* is a quadruple $\mathcal{M} = (\mathcal{L}, r, N, \Pi)$ where:

M1 (***Base Spaces***): $\mathcal{L} \equiv (\mathcal{S}, \preceq)$ is
(M1.1) a complete lattice,[2] where
(M1.2) $\mathcal{S}$ is a non-empty collection of non-empty and disjoint *base spaces*. An element of $\mathcal{S}$ is denoted S and has typical element ω.
M2 (***Projections***): $r \equiv (r_S^{S'})_{S \preceq S'}$, where each $r_S^{S'} : S' \to S$ is a surjection satisfying:
(Identity map when $S = S'$) $r_S^S(\omega) = \omega$ for all $\omega \in S$ and $S \in \mathcal{S}$;
(Projections Commute) If $S \preceq S' \preceq S''$, then $r_S^{S''} = r_S^{S'} \circ r_{S'}^{S''}$.
M3 (***Agent Set***). $N = \{1, ..., \ell\}$ is a finite set of agents.
M4 (***Possibility Correspondences***): $\Pi = (\Pi_i)_{i \in N}$ where for each agent $i \in N$, $\Pi_i : \Sigma \to 2^{\Sigma} \backslash \varnothing$ is called the *possibility correspondence* of agent i, where Σ denotes the union of base spaces in $\mathcal{S}$.

We assume throughout that the possibility correspondences of a model satisfy the following coherency condition introduced by HMS.

C (*Confinement*): If $\omega \in S \in \mathcal{S}$, then: $\Pi_i(\omega) \subseteq S'$ for some $S' \preceq S$.

HMS view each base space $S \in \mathcal{S}$ as encoding everything that can be 'expressed' with the vocabulary of a particular 'language.' HMS interpret $S \preceq S'$ to mean that the language corresponding to the base-space S' is at least as expressive as the language corresponding to the base-space S. Note that $\mathcal{S}$ has a greatest base space (corresponding to the most expressive language with the richest vocabulary in the model) and a least base space (corresponding to the least expressive language with the most impoverished vocabulary). [3]

In light of this interpretation, we can in turn motivate C (confinement) as the requirement that all the states an individual considers possible at a given state ω can be expressed with the same vocabulary, namely the vocabulary of the language available to the individual at ω, which cannot be more expressive than the language in which ω is expressed. C seems to us to be fundamental for the coherence of a model and its consistency with the associated knowledge and awareness of the individuals therein.[4]

[2] A *complete lattice* is a pair $(\mathcal{S}, \preceq)$ where $\mathcal{S}$ is a set partially ordered by $\preceq$, and each subset B of $\mathcal{S}$ has an infimum and supremum in $\mathcal{S}$.

[3] A set B has a *greatest* (*least*) element under the partial order $\preceq$ iff there exists some $a \in B$ such that $b \preceq a (a \preceq b)$ for all $b \in B$. Note that since $\preceq$ is a partial order, the greatest (least) element is unique.

[4] Indeed the fact that it is designated by HMS (p83) as the "(0)" property for possibility correspondences suggests that they felt it was self-evident that this property should be satisfied.

Our aim is to define a notion of a sub-model for each agent at each state $\omega \in \Sigma$. Given a model $\mathcal{M}$ and a state $\omega \in \Sigma$, let $S(\Pi_i(\omega))$ denote the base space containing $\Pi_i(\omega)$. We define the sub-model $\mathcal{M}^{i,\omega}$ of agent $i \in N$ at $\omega \in \Sigma$ as follows.

Definition 2 (Sub-model) Given a model $\mathcal{M} = (\mathcal{L}, r, N, \Pi)$, and a state $\omega \in \Sigma$ the *sub-model* $\mathcal{M}^{i,\omega} = (\mathcal{L}^{i,\omega}, r^{i,\omega}, N^{i,\omega}, \Pi^{i,\omega})$ of agent $i \in N$ at ω is defined by:

R1 (Base Spaces) $\mathcal{L}^{i,\omega} \equiv (\mathcal{S}^{i,\omega}, \preceq^{i,\omega})$ where:
(i) $\mathcal{S}^{i,\omega} \equiv \{S' \in \mathcal{S} : S' \preceq S(\Pi_i(\omega))\}$;
(ii) $\preceq^{i,\omega} \equiv \preceq \cap (\mathcal{S}^{i,\omega} \times \mathcal{S}^{i,\omega})$;
R2 (Projections): $r^{i,\omega} \equiv (r_S^{S'})_{S \preceq^{i,\omega} S'}$;
R3 (Agent Set) $N^{i,\omega} \equiv N$;
R4 (Possibility Correspondences): $\Pi^{i,\omega} = (\Pi_j^{i,\omega})_{j \in N}$ where for each agent $j \in N$, $\Pi_j^{i,\omega}(\omega') \equiv \Pi_j(\omega')$ for all $\omega' \in \Sigma^{i,\omega} \equiv \cup\{S : S \in \mathcal{S}^{i,\omega}\}$.

The interpretation of a sub-model $\mathcal{M}^{i,\omega}$ within a given model $\mathcal{M}$ is as follows. When the actual state in the model $\mathcal{M}$ is ω, agent i is informed of the value of $\Pi_i(\omega)$ as a set of states and of the sub-model $\mathcal{M}^{i,\omega}$. Agent i may then use her sub-model $\mathcal{M}^{i,\omega}$ and her possibility set $\Pi_i(\omega)$ to reason about herself and the other agents.

Recall that under C it follows that for each $\omega \in \Sigma$ and each $i \in N$, there is a unique base space $S' \in \mathcal{S}$ which contains $\Pi_i(\omega)$. This ensures that each sub-model is itself a model.

Proposition 1 *(Sub-models are models) Fix a model $\mathcal{M}$. For any agent $i \in N$ and any state $\omega \in \Sigma$, the sub-model $\mathcal{M}^{i,\omega}$ given by Definition 2 is a model.*

3 Knowledge Events

Besides requiring that each sub-model be a model, we also want to ensure that knowledge, awareness, and unawareness are well defined events in the model and in each of its sub-models. In this section, we provide necessary and sufficient conditions for HMS's notion of knowledge (respectively, awareness, unawareness) of an event to be itself a well-defined event. Fix a model $\mathcal{M} = (\mathcal{L}, r, N, \Pi)$. For any $\omega \in S'$, and any $S \preceq S'$, HMS write ω_S for $r_S^{S'}(\omega)$, that is, ω_S is the projection of ω from its base space S' to the (less expressive) base space S. The set of extensions of a subset $B \subseteq S \in \mathcal{S}$ over more expressive base spaces is given by:[5]

$$B^{\uparrow} \equiv \bigcup_{S' \in \{S'' : S \preceq S''\}} \left(r_S^{S'}\right)^{-1}(B). \tag{1}$$

Definition 3 An *event* $(B^{\uparrow}, S)$ consists of a subset $B^{\uparrow}$ of Σ formed from a subset B of S according to (1).

[5] Here, $(r_S^{S'})^{-1}(B) = \{\omega \in S' : r_S^{S'}(\omega) \in B\}$.

We denote the set of events by $\mathcal{E}$. Let $(B^{\uparrow}, S) \in \mathcal{E}$ be an event where $B \neq \varnothing$. By the disjointedness of the base spaces assumed in M1.1, the base space S corresponding to $B^{\uparrow}$ is uniquely determined. Hence, we follow HMS and write $B^{\uparrow}$ instead of $(B^{\uparrow}, S)$ whenever $B \neq \varnothing$. While events defined in this way are clearly subsets of Σ, as HMS point out, even when $(B^{\uparrow}, S)$ is an event, its full complement $(\Sigma \backslash B^{\uparrow}, S)$ may not be. To deal with this problem, HMS use the relative complement $((S \backslash B)^{\uparrow}, S)$ of an event $(S, B^{\uparrow})$, which is itself an event.

Knowledge in HMS models follows Aumann (1976). For an event $(B^{\uparrow}, S)$, the subset $K_i(B^{\uparrow})$ of Ω is defined as:

$$K_i(B^{\uparrow}) \equiv \{\omega \in \Sigma : \Pi_i(\omega) \subseteq B^{\uparrow}\}. \tag{2}$$

The subset $K_i(B^{\uparrow})$ is regarded as the set of states in Σ where i knows the event $(B^{\uparrow}, S)$. Unawareness of an event is defined to be

$$U_i(B^{\uparrow}) \equiv \neg K_i(B^{\uparrow}) \cap \neg K_i(\neg K_i(B^{\uparrow})). \tag{3}$$

Finally, awareness of the event corresponds to the subset $A_i(B^{\uparrow}) \equiv \neg U_i(B^{\uparrow})$. As was shown in HMS, awareness can be expressed as

$$A_i(B^{\uparrow}) = K_i(B^{\uparrow}) \cup K_i\left(\neg K_i(B^{\uparrow})\right). \tag{4}$$

We would like to ensure that $(K_i(B^{\uparrow}), S)$ is an event whenever $(B^{\uparrow}, S)$ is an event. We formally define this property as

KE (*Knowledge Events*) $(K_i(B^{\uparrow}), S)$ is an event for all $S \in \mathcal{S}$ and $B \subseteq S$.

Since awareness and unawareness are defined in terms of knowledge, they become events once KE holds. HMS (Proposition 1, p. 84) showed that KE holds using C and other conditions on the possibility correspondences. Board et al (2011, Proposition 2.3, p.21) showed that some of those conditions can be dispensed with. In the following, we give a new condition and one of the HMS conditions, which in the presence of C are necessary and sufficient for KE:

NIA (*Non-increasing Awareness*) Let $S^{\ell} \preceq S$ and $\omega \in S$. If $\Pi_i(\omega) \subseteq S'$ and $\Pi_i(\omega_{S^{\ell}}) \subseteq S''$, then $S'' \preceq S'$.

NIA may be interpreted as requiring that the awareness of an agent does not increase as we move down the lattice. The other condition is the HMS condition that projections preserve knowledge.

PPK (*Projections Preserve Knowledge*): If $S \preceq S' \preceq S''$, $\omega \in S''$ and $\Pi_i(\omega) \subseteq S'$, then $(\Pi_i(\omega))_S = \Pi_i(\omega_S)$.[6]

Proposition 2 (KE is equivalent to PPK and NIA) *A model $\mathcal{M}$ satisfies KE if and only if it satisfies NIA and PPK.*

[6] Here, $(\Pi_i(\omega))_S = \{\omega'_S : \omega' \in \Pi_i(\omega)\}$.

When the model $\mathcal{M}$ satisfies KE, each sub-model $\mathcal{M}^{i,\omega}$ also satisfies KE. The proof of this result is straightforward and therefore omitted.

Proposition 3 (KE inherited) *Fix a model $\mathcal{M}$ satisfying KE. For any agent $i \in N$ and any state $\omega \in \Sigma$, the sub-model $\mathcal{M}^{i,\omega}$ given by Definition 2 is a model satisfying KE.*

Hence, knowledge, awareness, and unawareness are also well defined in each sub-model $\mathcal{M}^{i,\omega}$ of $\mathcal{M}$. We use $\Sigma^{i,\omega}, \uparrow^{i,\omega}$, and $\mathcal{E}^{i,\omega}$ to denote, respectively, the union of base spaces, the extensions over more expressive base spaces, and the set of events in $\mathcal{M}^{i,\omega}$. By R1 - R4, the following holds:

$$B^{\uparrow^{i,\omega}} = B^{\uparrow} \cap \Sigma^{i,\omega} \tag{5}$$

The events $\mathcal{E}^{i,\omega}$ in a sub-model $\mathcal{M}^{i,\omega}$ are tightly connected to the events in the original model that agent i is aware of at ω. Fix a model $\mathcal{M}$, an agent i and a state ω in that model. Let $\mathcal{E}_{A_i,\omega} \equiv \{(B^{\uparrow}, S) \in \mathcal{E} : \omega \in A_i(B^{\uparrow})\} \subseteq \mathcal{E}$ denote the set of events in $\mathcal{M}$ that agent i is aware of at ω.

Proposition 4 (Sub-model event consistency) *Fix a model $\mathcal{M}$ satisfying KE. For every agent $i \in N$ and every state $\omega \in \Sigma$, the sub-model $\mathcal{M}^{i,\omega}$ defined by Definition 2 satisfies the following:*

(a) (preservation of events) The function $f : \mathcal{E}_{A_i,\omega} \to \mathcal{E}^{i,\omega}$ defined by $f(B^{\uparrow}, S) = (B^{\uparrow} \cap \Sigma^{i,\omega}, S)$ is a bijection.
(b) (preservation of knowledge) $K_j^{i,\omega}(B^{\uparrow^{i,\omega}}) = K_j(B^{\uparrow}) \cap \Sigma^{i,\omega}$ for each $j \in N$ and each event $(B^{\uparrow^{i,\omega}}, S) \in \mathcal{E}^{i,\omega}$.

Part *(a)* shows that the set of events $\mathcal{E}^{i,\omega}$ in the sub-model $\mathcal{M}^{i,\omega}$ corresponds to the set of events $\mathcal{E}_{A_i,\omega}$ in $\mathcal{M}$ that i is aware of at ω. In other words, by restricting agent i to his sub-model $\mathcal{M}^{i,\omega}$, we formally restrict his attention to the events he is aware of at ω. Part *(b)* shows that knowledge is preserved in each sub-model. These results follow directly by Definition 2 of a sub-model.

4 Motivating Full HMS models by sub-models

In this section, we presume that each agent is sophisticated enough to compare her information $\Pi_i(\omega)$ with her sub-model $\mathcal{M}^{i,\omega}$ to determine if there is any inconsistency. If she finds any inconsistency, then she might doubt the validity of the information $\Pi_i(\omega)$ or her sub-model $\mathcal{M}^{i,\omega}$. The following property requires informational consistency between $\Pi_i(\omega)$ and $\mathcal{M}^{i,\omega}$. We do not speculate on how an agent might revise her sub-model or update her information if she detects some inconsistency. Rather we define the property of sub-model informational consistency and show that it corresponds to some conditions on the possibility correspondences of the original model $\mathcal{M}$.

Definition 4 (Sub-model Information Consistency [SMIC]) A model $\mathcal{M}$ satisfies sub-model information consistency (SMIC) if for each state $\omega \in \Sigma$ and for each agent $i \in N$, the sub-model $\mathcal{M}^{i,\omega}$ satisfies:

$$\Pi_i(\omega) = \{\omega' \in \Sigma^{i,\omega} : \Pi_i^{i,\omega}(\omega') = \Pi_i(\omega)\} \tag{6}$$

In the definition of SMIC, we have intentionally made reference to the agent's sub-model.[7] As noted above, we assume that the information received by agent i at ω comes in two parts: one part is her sub-model $\mathcal{M}^{i,\omega}$ and the other is her possibility set $\Pi_i(\omega)$. Although, in general, the agent need not be aware of ω, we presume that she has access to the contents of $\Pi_i(\omega)$. She can then use her possibility set $\Pi_i(\omega)$ to check if everything meshes with the states in her sub-model. For example, if she were to find some state ω' that is in $\Pi_i(\omega)$, but is not in her sub-model, then she would have encountered an inconsistency between the two parts of her information. More generally, if the set of states in her sub-model where she should receive the possibility set equivalent to $\Pi_i(\omega)$ is different to the possibility set $\Pi_i(\omega)$ she has received, then she regards her sub-model $\mathcal{M}^{i,\omega}$ to be informationally inconsistent with her possibility set $\Pi_i(\omega)$.

Here, we will characterize SMIC in terms of a new condition called *reflexivity given awareness* and the " stationarity" condition of HMS.

RGA (*Reflexivity Given Awareness*) If $\omega \in S$ and $\Pi_i(\omega) \subseteq S$ for some $S \in \mathcal{S}$, then $\omega \in \Pi_i(\omega)$.

ST (*Stationarity*): $\omega' \in \Pi_i(\omega)$ implies $\Pi_i(\omega') = \Pi_i(\omega)$.

We have the following proposition.

Proposition 5 (Characterization of SMIC) *A model $\mathcal{M}$ satisfies SMIC if and only if it satisfies RGA and ST.*

So far, we have motivated C, PPK, and ST among the conditions of HMS. There are two remaining conditions:

GR (*Generalized Reflexivity*): $\omega \in \Pi_i(\omega)^{\uparrow}$;
PPI (*Projections Preserve Ignorance*): If $\omega \in S'$ and $S \preceq S'$, then $\Pi_i(\omega)^{\uparrow} \subseteq \Pi_i(\omega_S)^{\uparrow}$.

Our condition NIA is a weakening of PPI, and RGA is a weakening of GR. However, if we require KE and SMIC, then we get the full set of HMS conditions.

Theorem 1 (KE and SMIC imply HMS) *A model $\mathcal{M}$ satisfies KE and SMIC if and only if it satisfies GR, ST, PPI, and PPK.*

[7] Since, by R4 in Definition 2, $\Pi_i^{i,\omega}(\omega') = \Pi_i(\omega')$, the expression (6) defining SMIC could have been written simply in terms of the possibility correpondence $\Pi_i(\cdot)$ of the full model.

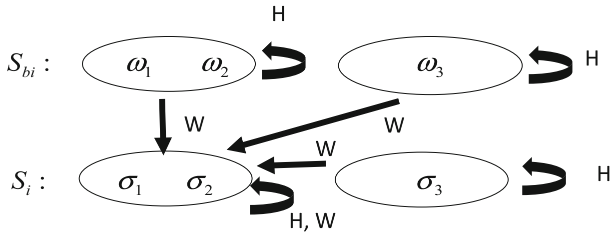

Fig. 1 Watson's False Knowledge

5 Holmes and Watson

In this section, we re-visit the example of Holmes and Watson presented by Galanis (2013). We use the example to demonstrate our results and to motivate further work with sub-models. The basic story goes as follows. Holmes realizes that if there was an intruder, the dog would have barked. Since the dog did not bark, there must have been no intruder. Watson, on the other hand, is unaware that the silence of the dog implies there was no intruder. Thus, he cannot conclude whether or not there was an intruder. Galanis argues that to handle this situation requires us to relax PPK. His model involves two base spaces, the top one with three states and the bottom one with two states.

We now provide an alternative version with one more state in the bottom base space that captures the situation and satisfies PPK and NIA. This model is depicted in Fig. 1. The top base space S_{bi} consists of three states and represents reasoning in the language built from the primitive propositions b (for bark) and i (for intruder). The first state, ω_1, corresponds to the situation where the dog barks and there is an intruder. The second state, ω_2, corresponds to the situation where the dog barks, and there is no intruder. The last state, ω_3, corresponds to the situation where the dog does not bark and there is no intruder. Since Holmes deduces that an intruder would have caused the dog to bark we can eliminate, as Galanis did, the state where there is an intruder and the dog does not bark. As HMS noted (pp. 85-86, Example 2 and Figure 2), and then formalized in their canonical model (Heifetz et al. (2008)), each state ω can be thought of as including the information of each individual at each state. Consequently, the state ω_3 can be thought of as including not only $\neg b$ (no barking) and $\neg i$ (no intruder), but also, the fact that Holmes knows there is no intruder. Similarly, the states ω_1 and ω_2 carry the information that Holmes does not know if there was an intruder.

The lower base space S_i represents reasoning in the language based only on the primitive proposition i. Unlike Galanis, however, we include a third state σ_3 which allows PPK to be satisfied. The projections are such that $r^{S_{bi}}_{S_i}(\omega_j) = \sigma_j$ for each $j = 1, 2, 3$. PPK states that when Holmes moves from the higher base space S_{bi} to the lower base space S_i, he should not become any more or less knowledgeable of things which he remains aware. In particular, when Holmes moves from S_{bi} to S_i, since he remains aware of i, he should not gain or lose any information about i. Having this third state allows this to hold. Holmes' possibility correspondence is given by: $\Pi_H(\omega_3) = \{\omega_3\}$ and $\Pi_H(\omega_j) = \{\omega_1, \omega_2\}$ for $j = 1, 2$; similarly, $\Pi_H(\sigma_3) = \{\sigma_3\}$

and $\Pi_H(\sigma_j) = \{\sigma_1, \sigma_2\}$ for $j = 1, 2$. Notice that, when Holmes moves from ω_j in S_{bi} to σ_j in S_i, his knowledge about i remains unchanged.

Watson, on the other hand, is uncertain about whether there was an intruder at every state in the model. That is, $\Pi_W(\omega) = \{\sigma_1, \sigma_2\}$ for all $\omega \in \Sigma$. This model satisfies NIA, PPK, and ST. However, it violates GR. In particular, $\omega_3 \notin \Pi_W(\omega_3)^{\uparrow}$. This violation, although understandable, has some implications. For example, consider the event consisting of all states in which there is no intruder. This event is $E = \{\sigma_2, \sigma_3, \omega_2, \omega_3\}$. The event that Holmes does not know E is thus $\neg K_H(E) = \{\sigma_1, \sigma_2, \omega_1, \omega_2\}$. Next, consider the event that Watson knows $\neg K_H(E)$, which is $K_W(\neg K_H(E)) = \Sigma$. In other words, Watson knows that Holmes does not know E at every state in the model. On the other hand, Holmes knows E at the state ω_3. Thus, Watson can be said to have false knowledge about Holmes' knowledge at ω_3.

Since this example satisfies PPK and NIA, it follows by Proposition 2 that it satisfies KE. But turning now to sub-model consistency, since it violates GR, it follows from Theorem 1 that SMIC fails. For an example of a violation of SMIC, consider the sub-model $\mathcal{M}^{W,\omega_3}$ of Watson at ω_3 which consists only of the base space S_i. Notice that, $\{\omega \in \Sigma^{W,\omega_3} : \Pi_W^{W,\omega_3}(\omega) = \Pi_W(\omega_3)\} = S_i \neq \Pi_W(\omega_3)$, which is a violation of SMIC. Given this violation, at the state ω_3, Watson might wonder why he has received the possibility set $\Pi_W(\omega_3) = \{\sigma_1, \sigma_2\}$ even though inspection of his sub-model suggests that σ_3 could also have happened. Faced with this situation, he might question the validity of his sub-model $\mathcal{M}^{W,\omega_3}$ or his possibility set $\Pi_W(\omega_3)$.

If we delve more deeply, we find that the problem arises because Watson has access to the *full* base space S_i. One way to resolve this problem would be to allow sub-models to have base spaces that are strict subsets of the corresponding base spaces in the full model. One might argue that Watson's sub-model should be restricted to $\{\sigma_1, \sigma_2\}$, similar to Galanis's treatment. In this way, unawareness may be captured not only by sub-lattices of the full model, but also by deletion of states within the base spaces. This approach goes beyond the scope of the current paper but may provide some new insights into unawareness. Such research may potentially be connected to the approaches taken by Halpern and Rego (2012) and Grant and Quiggin (2013) in modeling extensive games of unawareness that explicitly involved deletions of terminal histories (the analog of states in their setting).

6 Conclusions

The central idea of unawareness is that individuals do not consider all relevant possibilities and that they are therefore not aware of the full set of possible states of the world. The set of states they consider possible at any given state, based on the information available to them, are therefore defined on a limited set of states. The key idea in this paper is that of a sub-model $\mathcal{M}^{i,\omega}$ of the full model $\mathcal{M}$ held by agent i at ω, that is, a subset of the base spaces and information correspondences of $\mathcal{M}$ of which agent i is cognizant at ω. Our basic result, on which the rest of the paper is founded, is that for models $\mathcal{M}$ satisfying the confinement (C) property, all sub-models inherit the structural properties of $\mathcal{M}$ (including C). Hence, we can maintain the tradition of each agent reasoning within her own sub-model of which she is fully cognizant even in the face of unawareness.

This result provides a way of understanding the relationship between the perspective of agents in a model of differential awareness and that of the game theorist or outside analyst. In the Aumann model of common awareness, all agents are aware of the full model $\mathcal{M}$, just like the outside analyst (and this fact is common knowledge). In models of differential awareness, each agent has access to a restricted sub-model, but this sub-model ‘looks like’ a full model in all structural respects. In particular, the representation of other agents’ awareness in a given agent’s sub-model is structurally similar to the representation of agents’ awareness in the full model.[8]

With this result in hand, we have been able to clarify the relationship between knowledge and awareness in the HMS structure. The events of which agent i is aware at state ω in the full model $\mathcal{M}$ map neatly into the events in her sub-model $\mathcal{M}^{i,\omega}$. Hence, conditions ensuring that knowledge (and therefore awareness and unawareness) is represented by events in the full model are inherited by all sub-models. Along with the SMIC requirement that information structures in sub-models should be consistent with those of the full model, this is sufficient to motivate the full set of HMS conditions on possibility correspondences.

These results suggest ways of exploring a number of unresolved issues in the literature. As we have shown, the Holmes and Watson example of Galanis (2013) may be accommodated within the HMS framework (including the PPK condition) by allowing for false knowledge. False knowledge typically implies the exclusion of the true state from consideration. This suggests that the mapping from the full model of the most expressive base space available to an agent at a particular state in the full model to the sub-model considered by her at that state may exclude some elements of that most expressive base space (and the corresponding elements in less expressive base spaces). Potentially, this may allow for a unification of the HMS representation of unawareness with that of extensive-form representations involving the deletion of unconsidered terminal histories.

Acknowledgments We thank Spyros Galanis, Aviad Heifetz, Sander Heinsalu, and Burkhard Schipper for very constructive comments.

7 Appendix: Proofs

Proof of Proposition 1 First, we show that $\mathcal{M}^{i,\omega}$ satisfies M1. It follows by Definition 2.2. and M1.2 on $\mathcal{M}$ that M1.2 holds for $\mathcal{L}^{i,\omega} \equiv (\mathcal{S}^{i,\omega}, \preceq^{i,\omega})$. Let us see that M1.1 holds as well. Let $R \subseteq \mathcal{S}^{i,\omega}$. By C and Definition 2, there is a unique base space $S \in \mathcal{S}^{i,\omega}$ containing $\Pi_i(\omega)$ and $T \preceq S$ for all $T \in R$. By completeness of $\mathcal{L}$, the set R has a supremum, $\sup(R) \in \mathcal{S}$, and infimum, $\inf(R) \in \mathcal{S}$. By the definition of supremum, $\sup(R) \preceq S$. By Definition 2 , $\sup(R)$ is also the supremum of R in $\mathcal{S}^{i,\omega}$. Similarly, we find that $\inf(R)$ is also the infimum of R in $\mathcal{S}^{i,\omega}$. Thus, we have

[8] So, the ‘outside analyst’ might be a boundedly aware agent in some larger model.

shown that M1.2 holds. That $\mathcal{M}^{i,\omega}$ inherits M2, M3 and C follows from inspection of Definition 2. □

Proof of Proposition 2 (If-part): Let $(B^{\uparrow}, S) \in \mathcal{E}$. Then, $B \subseteq S$, and as such, $(\{\omega' \in S : \Pi_i(\omega') \subseteq B\}^{\uparrow}, S) \in \mathcal{E}$. We will show that $K_i(B^{\uparrow}) = \{\omega' \in S : \Pi_i(\omega') \subseteq B\}^{\uparrow}$. First, let us see that $K_i(B^{\uparrow}) \subseteq \{\omega' \in S : \Pi_i(\omega') \subseteq B\}^{\uparrow}$. Let $\omega \in K_i(B^{\uparrow})$. Then $\omega \in S''$ for some $S'' \in \mathcal{S}$, and by (2), $\Pi_i(\omega) \subseteq B^{\uparrow}$. By C, $\Pi_i(\omega) \subseteq S'$ for some $S' \in \mathcal{S}$ and $S' \preceq S''$. Since $\Pi_i(\omega) \subseteq B^{\uparrow} \cap S'$, and $B \subseteq S$, it follows that $S \preceq S'$ and $(\Pi_i(\omega))_S \subseteq B$. So, we have $S \preceq S' \preceq S''$, $\omega \in S''$ and $\Pi_i(\omega) \subseteq S'$. By PPK, $\Pi_i(\omega_S) = (\Pi_i(\omega))_S \subseteq B$. This implies that $\omega_S \in \{\omega' \in S : \Pi_i(\omega') \subseteq B\}$, which in turn implies $\omega \in \{\omega' \in S : \Pi_i(\omega') \subseteq B\}^{\uparrow}$.

Next, let's see that $\{\omega' \in S : \Pi_i(\omega') \subseteq B\}^{\uparrow} \subseteq K_i(B^{\uparrow})$. Let $\omega \in \{\omega' \in S : \Pi_i(\omega') \subseteq B\}^{\uparrow}$. Then, $\omega \in S'$ for some $S' \in \mathcal{S}$ with $S \preceq S'$ and $\Pi_i(\omega_S) \subseteq B \subseteq S$. By C, $\Pi_i(\omega) \subseteq S''$ for some $S'' \preceq S'$. By NIA, $S \preceq S''$. So, we have $S \preceq S'' \preceq S'$, $\omega \in S'$ and $\Pi_i(\omega) \subseteq S''$. It follows by PPK that $(\Pi_i(\omega))_S = \Pi_i(\omega_S) \subseteq B$. Since $S \preceq S''$ and $\Pi_i(\omega) \subseteq S''$, we find that $\Pi_i(\omega) \subseteq B^{\uparrow}$. Hence, $\omega \in K_i(B^{\uparrow})$.

(Only if-part): We prove the contrapositive, that is, if PPK or NIA fails, then KE fails. Observe that PPK can be broken into two conditions:

PPK$^{\supseteq}$ If $S \preceq S' \preceq S''$, $\omega \in S''$ and $\Pi_i(\omega) \subseteq S'$, the $(\Pi_i(\omega))_S \supseteq \Pi_i(\omega_S)$.
PPK$^{\subseteq}$ If $S \preceq S' \preceq S''$, $\omega \in S''$ and $\Pi_i(\omega) \subseteq S'$, then $(\Pi_i(\omega))_S \subseteq \Pi_i(\omega_S)$.

We break up the proof into three cases: Case 1: $\mathcal{M}$ violates PPK$^{\supseteq}$; Case 2: $\mathcal{M}$ satisfies PPK$^{\supseteq}$, but violates PPK$^{\subseteq}$; Case 3: $\mathcal{M}$ satisfies PPK, but violates NIA. For each case, we will construct an event $(E, S) \in \mathcal{E}$ and show that $(K_i(E), S) \notin \mathcal{E}$.

Case 1: $\mathcal{M}$ violates PPK$^{\supseteq}$.

In this case, we have $S \preceq S' \preceq S''$, $\omega \in S''$ and $\Pi_i(\omega) \subseteq S'$, but $\Pi_i(\omega_S) \nsubseteq (\Pi_i(\omega))_S$. Consider the event $(E, S) = (\Pi_i(\omega))_S^{\uparrow}, S) \in \mathcal{E}$. We will show that $(K_i((\Pi_i(\omega))_S^{\uparrow}), S) \notin \mathcal{E}$. Suppose, on the contrary, that $(K_i((\Pi_i(\omega))_S^{\uparrow}), S) \in \mathcal{E}$. Then, there must be some set B, such that $B \subseteq S$, and $B^{\uparrow} = K_i((\Pi_i(\omega))_S^{\uparrow})$. Since $\Pi_i(\omega) \subseteq (\Pi_i(\omega))_S^{\uparrow}$, it follows that $\omega \in K_i((\Pi_i(\omega))_S^{\uparrow}) = B^{\uparrow}$. Thus, $\omega_S \in B \subseteq B^{\uparrow} = K_i((\Pi_i(\omega))_S^{\uparrow})$. Since $\Pi_i(\omega_S) \nsubseteq (\Pi_i(\omega))_S$, it follows by C that $\omega_S \notin K_i((\Pi_i(\omega))_S^{\uparrow})$, which is a contradiction. Hence, we conclude that $(K_i((\Pi_i(\omega))_S^{\uparrow}), S) \notin \mathcal{E}$.

Case 2: $\mathcal{M}$ satisfies PPK$^{\supseteq}$, but violates PPK$^{\subseteq}$.

In this case, we have $S \preceq S' \preceq S''$, $\omega \in S''$ and $\Pi_i(\omega) \subseteq S'$, but $(\Pi_i(\omega))_S \nsubseteq \Pi_i(\omega_S)$. Consider the event $(E, S) = (\Pi_i(\omega_S)^{\uparrow}, S) \in \mathcal{E}$, which is an event since $\mathcal{M}$ satisfies PPK$^{\supseteq}$. We will show that $(K_i(\Pi_i(\omega_S)^{\uparrow}), S) \notin \mathcal{E}$. Suppose, on the contrary, that $(K_i(\Pi_i(\omega_S)^{\uparrow}), S) \in \mathcal{E}$. Then, there must be some set $B \subseteq S$, and $B^{\uparrow} = K_i(\Pi_i(\omega_S)^{\uparrow})$. Since $\Pi_i(\omega_S) \subseteq \Pi_i(\omega_S)^{\uparrow}$, $\omega_S \in K_i(\Pi_i(\omega_S)^{\uparrow}) = \{\hat{\omega} \in \Sigma : \Pi_i(\hat{\omega}) \subseteq \Pi_i(\omega_S)^{\uparrow}\}$. Since, by assumption, $B^{\uparrow} = K_i(\Pi_i(\omega_S)^{\uparrow})$, it follows that $\omega \in B^{\uparrow}$. However, since $(\Pi_i(\omega))_S \nsubseteq \Pi_i(\omega_S)$, $\omega \notin K_i(\Pi_i(\omega_S)^{\uparrow}) = B^{\uparrow}$, a contradiction. Hence, we conclude that$(K_i(\Pi_i(\omega_S)^{\uparrow}), S) \notin \mathcal{E}$.

Case 3: $\mathcal{M}$ satisfies PPK, but violates NIA.

In this case we have $\omega \in S$, $\Pi_i(\omega) \subseteq S'$, $S^\ell \preceq S$, and $\Pi_i(\omega_{S^\ell}) \subseteq S''$ but $S'' \npreceq S'$. Consider the event $(E, S'') = (\Pi_i(\omega_{S^\ell})^\uparrow, S'') \in \mathcal{E}$, and suppose that there is some set $B \subseteq S''$, and $B^\uparrow = K_i(\Pi_i(\omega_{S^\ell})^\uparrow)$. Since $\Pi_i(\omega_{S^\ell}) \subseteq \Pi_i(\omega_{S^\ell})^\uparrow$, $\omega_{S^\ell} \in K_i(\Pi_i(\omega_{S^\ell})^\uparrow) = B^\uparrow$. It follows that $\omega \in B^\uparrow$. However, since $\Pi_i(\omega) \subseteq S'$ and $S'' \npreceq S'$, $\Pi_i(\omega) \nsubseteq \Pi_i(\omega_{S^\ell})^\uparrow$. Hence, $\omega \notin K_i((\Pi_i(\omega_S)^\uparrow)$, a contradiction. Hence, we conclude that $(K_i(\Pi_i(\omega_S)^\uparrow), S) \notin \mathcal{E}$. □

Proof of Proposition 4 *(a)* First, we show that the function f is an injection. Suppose $(B^\uparrow, S)$ and $(C^\uparrow, S')$ are two distinct events in $\mathcal{E}_{A_i,\omega}$. Then, either $B^\uparrow \neq C^\uparrow$, or $S \neq S'$. In either case, $(B^{\uparrow^{i,\omega}}, S)$ is distinct from $(C^{\uparrow^{i,\omega}}, S')$. Now, let's see that f is a surjection. Let $(B^{\uparrow^{i,\omega}}, S) \in \mathcal{E}^{i,\omega}$. Then $S \in \mathcal{S}^{i,\omega}$, $(B^\uparrow, S) \in \mathcal{E}$ and $f(B^\uparrow, S) = (B^{\uparrow^{i,\omega}}, S)$. We need to show that $(B^\uparrow, S) \in \mathcal{E}_{A_i,\omega}$, which is equivalent to showing that $\omega \in A_i(B^\uparrow)$. Let S' denote the base space containing $\Pi_i(\omega)$. Since $S \in \mathcal{S}^{i,\omega}$ it follows from R1, that $S \preceq S'$. Hence, $\Pi_i(\omega) \subseteq S' \in S^\uparrow$. By (2), $\omega \in K_i(S^\uparrow)$.

(b) Let $(B^{\uparrow^{i,\omega}}, S)$ be an event in $\mathcal{E}^{i,\omega}$ and suppose that $\omega' \in K_j^{i,\omega}(B^{\uparrow^{i,\omega}})$. Then $\omega' \in \Sigma^{i,\omega} \subseteq \Sigma$ and $\Pi_j(\omega') \subseteq B^{\uparrow^{i,\omega}} \subseteq B^\uparrow$. Hence, $\omega' \in K_j(B^\uparrow) \cap \Sigma^{i,\omega}$. Conversely, suppose that $\omega' \in K_j(B^\uparrow) \cap \Sigma^{i,\omega}$. Then, $\Pi_j(\omega') \subseteq B^\uparrow$ and $\omega' \in \Sigma^{i,\omega}$. By C, $\Pi_j(\omega') \subseteq \Sigma^{i,\omega}$. Since $B^{\uparrow^{i,\omega}} = B^\uparrow \cap \Sigma^{i,\omega}$, we have $\Pi_j(\omega') \subseteq B^{\uparrow^{i,\omega}}$. By (2), $\omega' \in K_j^{i,\omega}(B^{\uparrow^{i,\omega}})$. □

Proof of Propostion 5 (If-part) Fix $\omega \in \Sigma$ and $i \in N$. First we show $\Pi_i(\omega) \subseteq \{\omega' \in \Sigma^{i,\omega} : \Pi_i(\omega') = \Pi_i(\omega)\}$. Let $\hat{\omega} \in \Pi_i(\omega)$ Then, $\hat{\omega} \in \Sigma^{i,\omega}$. By ST, $\Pi_i(\hat{\omega}) = \Pi_i(\omega)$. Hence, $\hat{\omega} \in \{\omega' \in \Sigma^{i,\omega} : \Pi_i(\omega') = \Pi_i(\omega)\}$. Next we show $\{\omega' \in \Sigma^{i,\omega} : \Pi_i(\omega') = \Pi_i(\omega)\} \subseteq \Pi_i(\omega)$. Let $\hat{\omega} \in \Sigma^{i,\omega}$ such that $\Pi_i(\hat{\omega}) = \Pi_i(\omega)$. By C, $\Pi_i(\hat{\omega}) \subseteq S(\hat{\omega})$. Hence, by RGA, $\hat{\omega} \in \Pi_i(\hat{\omega}) = \Pi_i(\omega)$.

(Only if-part) Suppose first that RGA fails. Then, for some $\omega \in \Sigma$ and $i \in N$, we have $\hat{\omega} \in \Sigma^{i,\omega}$ such that $\Pi_i(\hat{\omega}) \subseteq S(\hat{\omega})$, but $\hat{\omega} \notin \Pi_i(\hat{\omega})$. Then, $\hat{\omega} \in \Sigma^{i,\hat{\omega}}$, so $\hat{\omega} \in \{\omega' \in \Sigma^{i,\hat{\omega}} : \Pi_i(\omega') = \Pi_i(\hat{\omega})\}$. Hence, $\Pi_i(\hat{\omega}) \neq \{\omega' \in \Sigma^{i,\hat{\omega}} : \Pi_i(\omega') = \Pi_i(\hat{\omega})\}$, that is, $\mathcal{M}$ violates SMIC.

Finally, suppose that ST fails, i.e., there are $\omega, \hat{\omega} \in \Sigma$ such that $\hat{\omega} \in \Pi_i(\omega)$ but $\Pi_i(\hat{\omega}) \neq \Pi_i(\omega)$. Then $\hat{\omega} \notin \{\omega' \in \Sigma^{i,\omega} : \Pi_i(\omega') = \Pi_i(\omega)\}$. Hence, $\mathcal{M}$ violates SMIC. □

The proof of Theorem 1 is based on the following Lemma.

Lemma 1 *Fix a model $\mathcal{M}$ satisfying PPK. $\mathcal{M}$ satisfies RGA if and only if $\mathcal{M}$ satisfies GR.*

Proof The if-part is immediate from the definitions of RGA and GR. We prove the only if-part. Suppose $\omega \in \Sigma$. If $\Pi_i(\omega) \subseteq S(\omega)$, then by RGA, we have $\omega \in \Pi_i(\omega)$. Since $\Pi_i(\omega) \subseteq \Pi_i(\omega)^\uparrow$, GR holds. If, alternatively, $\Pi_i(\omega) \nsubseteq S(\omega)$, then by C, $\Pi_i(\omega) \subseteq S$ for some $S \prec S(\omega)$. Observe that we have $S \preceq S \preceq S(\omega)$, $\omega \in S(\omega)$ and $\Pi_i(\omega) \subseteq S$. Hence by PPK, $(\Pi_i(\omega))_S = \Pi_i(\omega_S)$. Since $(\Pi_i(\omega))_S \subseteq S$, we have $\Pi_i(\omega_S) \subseteq S$. Applying RGA to ω_S, we find $\omega_S \in \Pi_i(\omega_S)$. As observed already, $(\Pi_i(\omega))_S = \Pi_i(\omega_S)$, so $\omega_S \in (\Pi_i(\omega))_S$. But, since $\Pi_i(\omega) \subseteq S$, $(\Pi_i(\omega))_S = \Pi_i(\omega)$, we have $\omega_S \in \Pi_i(\omega)$ and $\omega \in \Pi_i(\omega)^\uparrow$, that is GR holds. □

Proof of Theorem 1 (Only if-part): PPK and NIA follow from KE and Proposition 2. GR and ST follow from SMIC and Proposition 5 and Lemma 1. It remains only to show PPI.

Let $\omega \in S''$ and $S \preceq S''$. We need to show that $\Pi_i(\omega)^{\uparrow} \subseteq \Pi_i(\omega_S)^{\uparrow}$. By C, $\Pi_i(\omega) \subseteq S^*$ for some $S^* \preceq S''$, and $\Pi_i(\omega_S) \subseteq S'$ for some $S' \preceq S$. By NIA, $S' \preceq S^*$. We will show the following two things:

(a) $\Pi_i(\omega)^{\uparrow} \subseteq \Pi_i(\omega_{S'})^{\uparrow}$;
(b) $\Pi_i(\omega_{S'}) = \Pi_i(\omega_S)$.

First, let's see (a). From our results above, we have $S' \preceq S^* \preceq S''$, $\omega \in S''$, and $\Pi_i(\omega) \subseteq S^*$. Hence, $\Pi_i(\omega)^{\uparrow} \subseteq (\Pi_i(\omega))^{\uparrow}_{S'}$. It follows by PPK$'$ that $(\Pi_i(\omega))_{S'} = \Pi_i(\omega_{S'})$, whence $\Pi_i(\omega)^{\uparrow} \subseteq \Pi_i(\omega_{S'})^{\uparrow}$ as required.

Next, let's see (b). Since $\Pi_i(\omega_S) \subseteq S'$, it follows by GR that $\omega_{S'} \in \Pi_i(\omega_S)$. Hence, by ST, $\Pi_i(\omega_{S'}) = \Pi_i(\omega_S)$ and (b) has been proved. It follows immediately from (a) and (b) that $\Pi_i(\omega)^{\uparrow} \subseteq \Pi_i(\omega_S)^{\uparrow}$, that is, we have proved PPI.

(If-part): As noted by HMS in their remark 2 on p.83, PPI and C imply NIA. Hence, KE follows from PPK and NIA using Proposition 2. SMIC follows from Proposition 5 and Lemma 1 using GR, ST, and PPK. □

References

Aumann, R. (1976). Agreeing to Disagree. *Annals of Statistics*, *4*, 1236–1239.

Aumman, R. (1999). Interactive Epistemology I: Knowledge. *International Jounral of Game Theory*, *28*, 263–300.

Board, O., Chung, K., & Schipper, B. (2011). Two models of unawareness: comparing the object-based and the subjective state approaches. *Synthese*, *179*, 13–34.

Dekel, E., Lipman, B., & Rustichini, A. (1998). Standard State-Space Models Preclude Unawareness. *Econometrica*, *66*, 159–173.

Fagin, R., Halpern, J., Moses, Y., & Vardi, M. (1995). *Reasoning about knowledge*. Cambridge: MIT.

Galanis, S. (2013). Unawareness of Theorems. *Economic Theory*, *52*, 41–73.

Grant, S., & Quiggin, J. (2013). Inductive Reasoning about Unawareness. *Economic Theory*, *54*, 717–755.

Halpern, J., & Rego, L. (2008). Interactive unawarenss revisited. *Games and Economic Behavior*, *62*, 232–262.

Halpern, J., & Rego, L. (2012). Generalized Solution Concepts in Games with Possibly Unaware agents. *International Journal of Game Theory*, *41*, 131–155.

Heifetz, A., Meier, M., & Schipper, B. (2006). Interactive Unawareness. *Journal of Economic Theory*, *130*, 78–94.

Heifetz, A., Meier, M., & Schipper, B. (2008). A Canonical Model for Interactive Unawareness. *Games and Economic Behavior*, *62*, 304–324.