Information Structures with Unawareness†

Jing Li

Department of Economics
University of Pennsylvania
3718 Locust Walk
Philadelphia, PA 19104
E-mail: jing.li@econ.upenn.edu

May, 2006

†I am indebted to Bart Lipman for his invaluable guidance throughout the course of this project. I am grateful to Larry Samuelson and Bill Sandholm for many inspiring discussions and advice. I also wish to thank Pierpaolo Battigalli, Eddie Dekel, Yossi Feinberg, Malcolm Forster and seminar participants at BU, CMU, Georgia State University, Northwestern, Penn, Penn State, Queen's, Toronto, UCLA, UIUC and UW-Madison for suggestions and comments. Financial support from Richard E. Stockwell fellowship is gratefully acknowledged.

Abstract

I construct a multi-agent state space model with unawareness following Aumann (1976). Dekel, Lipman and Rustichini (1998) show that standard state space models are incapable of representing unawareness. The model circumvents the impossibility result by endowing the agent a subjective state space that differs from the full state space when he has the unawareness problem. Information is modeled as a pair, consisting of both factual information and *awareness information*. The model exhibits nice properties parallel to those in the standard information partition model.

Keywords: unawareness, information, information partition, the state space models
JEL Classification: C70, C72, D80, D82, D83

> "There are things we know that we know. There are known unknowns - that is to say, there are things that we now know we don't know. But there are also unknown unknowns. There are things we do not know we don't know."
>
> Donald Rumsfeld, the U.S. Secretary of Defense

1 Introduction

A person is unaware of an event if he does not know it, and he does not know that he does not know it, and so on *ad infinitum*. In real life, *formulating* a decision problem, including recognizing all relevant uncertainties and available options, is at least as important as finding the solution to the formulated problem. Being unaware of some aspects of the situation is a common problem at this stage. For example, prior to the 911 attacks, most of us did not know that terrorists might use civilian aircraft as a weapon, and more importantly, we did not know that we did not know this. We were simply unaware of this possibility.

Unawareness plays an important role in economic life, especially through players' recognition of the possibility of their being unaware of *something*. One may not wish to commit to a seemingly attractive position in an alien environment due to the vast unknown unknowns. Contractual parties may opt to write incomplete contracts in order to preserve some flexibility to deal with contingencies of which they were unaware at the contractual date. See Dekel, Lipman and Rustichini (1998b) for more discussion.

These issues cannot be analyzed using the standard tools of economics. The prevailing model of uncertainty in economics is the standard information partition model: uncertainties are represented by a state space; information is represented by a partition over the state space; at each state, the agent is informed of the corresponding partition element. But then the agent cannot be unaware of anything: having an information partition is equivalent to having a knowledge hierarchy in which whenever the agent doesn't know something, he knows he doesn't know it (Bacharach 1985). Dekel, Lipman and Rustichini (1998a)(henceforth DLR) further show that the problem is fundamental: *any* model based on the standard state space specification, regardless of information being partitional or not, necessarily imposes either full awareness or full unawareness.

On the other hand, there are fruitful research in modeling unawareness using syntactic models, for example, Fagin and Halpern (1988), Halpern (2001), Modica and Rustichini (1994, 1999), to name a few. While this research has greatly improved our understanding of unawareness, the tools developed along this line are extraneous to many economists. Given the central role of decision-making under uncertainty in economics, a model that uses the familiar state space and information partition specifications, while highlighting the implication of unawareness on information processing is much desired. In this paper, I provide such a model.

The main idea is as follows. Fix the set of payoff-relevant uncertainties. One can think of them as a set of relevant questions. If the agent is unaware of a question, then a message reminding the agent of the question itself must be informative.[1] Such information is fundamentally different from the kind of factual information in the standard models. Modeling unawareness is equivalent to modeling such *awareness information*. On the other hand, one can only reason about things of which one is aware. Fixing a full state space and a full information partition representing the signal structure, if the agent is unaware of some uncertainties, his mind-set of reasoning must be represented by a less detailed model, where the uncertainties of which he is unaware are lacking. Therefore, I allow the agent to have (full-)state-contingent subjective models.

As an illustration, consider the following episode: Sherlock Holmes and Watson are investigating a crime. A horse has been stolen and the keeper was killed. From the narration of the local police, Holmes notices the dog in the stable did not bark that night and hence concludes that there was no intruder in the stable. Watson, on the other hand, although he also knows the dog did not bark – he himself mentioned this fact to Holmes – somehow does not come up with the inference that there was no intruder.

The feature I would like to capture in this story is the following. Watson is unaware of the possibility that there was no intruder, and hence fails to recognize the factual information "there was no intruder" contained in the message "the dog did not bark." Had someone asked Watson, "Could there have been an intruder in the stable that night?" He would have recognized his negligence and replied, "Of course not, the dog did not bark!"

The relevant question in this example is whether there was an intruder to the stable that night. Let $a = (a', \Delta)$, $b = (b', \Delta)$, where a' stands for "there was an intruder," b' stands for "there was no intruder" and Δ stands for "*cogito ergo sum*." The full state space is $\{a, b\}$. The dog barked in a and did not bark in b, yielding the full information partition $\{\{a\}, \{b\}\}$. However, in b, Watson is unaware of the possibility of no intruder, that is, his awareness signal does not include $\{a', b'\}$. To model this, I let Watson's mind-set be represented by the subjective state space $\{\Delta\}$, containing only specifications present in his awareness signal $\{\Delta\}$. Since the question of an intruder never occurs to Watson, the factual signal "the dog did not bark," or $\{b\}$, does not "ring a bell" in his mind. To Watson, the information he has is just the trivial partition of the subjective state space $\{\{\Delta\}\}$. As a consequence, Watson does not know $\{a, b\}$, and does not know that he does not know it.[2]

The question "Could there have been an intruder in the stable that night?" reveals the awareness information $\{a', b'\}$ to Watson. Now Watson adds the specification of

[1] The is without loss of generality. One could imagine situations where the agent has "partial awareness:" the agent is aware of the question but unaware of *some* answers. This model is capable of handling such situations, via proper rephrasing of the questions and answers. Li (2006b) discusses the issue in detail.

[2] Ely (1998) proposes a similar framework in the context of the Watson example. The observation is that unawareness of uncertainties causes unawareness of signals, thus the same signal structure induces different information partitions under different awareness. Ely considers an information structure that takes the form of state-contingent *partition* of the state space. Li (2006b) has more details.

whether there was an intruder to his subjective state space, updates it to the full state space $\{a, b\}$, and hence recognizes the information partition $\{\{a\}, \{b\}\}$, obtaining the knowledge "there was no intruder" as a result of simply being asked a question.

The model is a natural generalization of Aumann (1976) and exhibits nice properties parallel to those in the standard information partition model. In Aumann's model, an epistemic state specifies both resolutions of external uncertainties and the agent's knowledge. In this model, a full state specifies what the agent is or is not aware of, in addition to the resolution of external uncertainties and the agent's knowledge. Consequently, information takes the form of a pair, consisting of awareness signals, represented by (full)-state-dependent subjective state spaces, and factual signals, represented by an information partition over the full state space. The agent is said to know an event E if and only if he recognizes E given his awareness signal, and there is no uncertainty regarding E given his factual signal.

Each subjective state corresponds to an event in the full state space. Unawareness results in the agent's inability to imagine any scenario precisely. In this sense the subjective state space is incomplete with respect to the full state space in the sense of omitting "dimensions," not omitting "points." This suggests a clear distinction between probability zero and unawareness in this framework: if the agent is unaware of an event, then the event is beyond the agent's probability space and he is unable to assign any probability to it. It is worth pointing out that, since being unaware of an event results in being unaware of all possible factual signals regarding the event, unawareness introduces real dynamics into the model.

In a multi-agent environment, players could reason about what others are aware of as well as what they know, within the confines of their own mind-sets. Notice that player i could be unaware that player j is unaware of an event E of which i is aware himself. The resulting interactive knowledge hierarchies have rich but tractable structures. Unawareness has profound implications for common knowledge, which is a critical concept in economics. On the one hand, introducing the possibility of being unaware of an event obviously makes it harder for the players to arrive at common knowledge. On the other hand, common knowledge turns out to have surprisingly weak implications, namely only mutual knowledge. This is because, unawareness of a question results in unawareness of all factual signals concerning the question, and hence reduces the amount of uncertainties players reason about in this environment. I give conditions on the information structures under which one can bound common knowledge. I further show that if the information structures are sufficiently nice, then common knowledge of "E" can be fully characterized by adding to Aumann's classic characterization in the standard model a natural restriction that requires "common knowledge of awareness of E."

In an independently conceived work, Heifetz, Meier and Schipper (2004) propose a set-theoretic model by exploring a complete lattice of state spaces, ordered by "expressive power." The main difference between the two approaches is, while players' subjective perceptions of the environment (constrained by unawareness) are derived concepts in the current model, the subjective state spaces and subjective possibility sets are primitive notions in their model. Consequently, the current model focuses more heavily on the effects of introducing unawareness on the standard representation of uncertainty and information from an objective (the modeler's) perspective.

In connection to syntactic models in the literature, the current model could be viewed as a set-theoretic version of Fagin and Halpern (1988), in which the authors explore the standard Kripke structure with an additional modal operator of awareness. Modica and Rustichini (1999) explores a similar semi-set-theoretic, single-agent model which Halpern (2001) shows to be equivalent to a special case of the approach in Fagin and Halpern (1988).

The rest of the paper is organized as follows: Section 2 reviews the possibility correspondence models and DLR's impossibility results, highlighting the implicit assumption of full-awareness in the standard model and pointing to the parallel structure with the current model. Section 3 presents the model of unawareness, which I dub "the product model" for the use of the product structure of the state space. Section 4 characterizes the knowledge hierarchy with nontrivial unawareness. Section 5 extends the product model to multi-agent environment. Section 6 concludes. Proofs not found in the text are collected in the Appendix.

2 A Review of the Standard Model

The standard model, also known as the possibility correspondence model, consists of a state space Ω and a possibility correspondence $P : \Omega \to 2^{\Omega} \setminus \{\emptyset\}$. Each state $\omega \in \Omega$ completely specifies the resolution of all relevant uncertainties. An "event" in the ordinary usage of the term corresponds to a set of states in the model. For instance, the informal idea of the event that "there was an intruder" is formally taken to be the set of states where there was an intruder.

With this formulation, one can identify logical relations with set operations: set inclusion "$\subseteq$", set intersection "$\cap$", set union "$\cup$" and set complement (with respect to the state space) "$\setminus$" correspond to logical consequence "$\to$", conjunction "$\wedge$", disjunction "$\vee$" and negation "$\neg$" respectively.

The agent's information structure is represented by a possibility correspondence $P : \Omega \to 2^{\Omega} \setminus \{\emptyset\}$. P associates each state ω with a nonempty event $P(\omega)$, which is interpreted as the agent's information at ω. The idea is, at ω, the agent considers $P(\omega)$ to be the set of possible states. Note the information takes the form of an event, and hence is factual information.

Definition 1 *P induces an* **information partition** *over the state space if (1) for any $\omega \in \Omega$, $\omega \in P(\omega)$; and (2) for any $\omega, \omega' \in \Omega$, $\omega' \in P(\omega)$ implies $P(\omega') = P(\omega)$.*

If P induces an information partition, then (Ω, P) is an *information partition model*. Otherwise, it is a *non-partitional model*.

Knowledge is characterized as "truth in all possible states." Intuitively, if there was no intruder in every state Holmes considers possible, then for Holmes, there is no uncertainty left regarding this event, i.e., he knows "there was no intruder." Formally, for any event $E \subseteq \Omega$, define the knowledge operator $K : 2^\Omega \to 2^\Omega$ by

$$K(E) = \{\omega : P(\omega) \subseteq E\}$$

$K(E)$ is the set of states in which the agent knows E, and hence is interpreted as the event "the agent knows E." To see that it makes sense to interpret K as knowledge, consider the following properties. For any $E, F \subseteq \Omega$,

K1 *Necessitation*: $K(\Omega) = \Omega$

K2 *Monotonicity*: $E \subseteq F \Rightarrow K(E) \subseteq K(F)$

K3 *Conjunction*:[3] $K(E) \cap K(F) = K(E \cap F)$

K4 *The axiom of knowledge*: $K(E) \subseteq E$

K5 *The axiom of transparency*:[4] $K(E) \subseteq KK(E)$

K6 *The axiom of wisdom*: $\neg K(E) \subseteq K\neg K(E)$

A statement like "A is A" is universally true and the agent should know this. Indeed, a tautology is represented by the universal event Ω, and hence necessitation corresponds to knowledge of tautologies. Monotonicity says the agent is able to perform logical deductions. If E implies F and the agent knows E, then he knows F. Conjunction says the agent knows the events E and F if and only if he knows the event "E and F." K1-3 are basic properties of knowledge that make the characterization sensible. Without them, it is not clear what it means "to know" something. In contrast, the next three axioms K4-6 reflect the agent's *rationality* in information processing. The axiom of knowledge says the agent cannot know anything false. The axiom of transparency says whenever the agent knows something, he knows that he knows it. The axiom of wisdom says if the agent does not know something, he knows that he does not know it.

Theorem 1 *(Geanakoplos (1990)) In the possibility correspondence model* (Ω, P)*, knowledge satisfies K1-3. It satisfies K4-6 if* P *induces an information partition.*

It is obvious that the axiom of wisdom prevents an information partition model from having nontrivial unawareness. However, it is less obvious that the implicit assumption of full-awareness lies on the structure of the standard state space specification instead of the partitional information structure (Dekel, Lipman and Rustichini 1998a).

[3] Note that conjunction implies monotonicity.

[4] In places where there is no risk of confusion, I omit the parentheses when applying the operators.

DLR introduce an unawareness operator: $U : 2^\Omega \to 2^\Omega$, where $U(E)$ is the set of states where the agent is unaware of E, and hence is interpreted as the event "the agent is unaware of E." They consider three intuitive properties of unawareness: for any event $E \subseteq \Omega$,

U1 *Plausibility*: $U(E) \subseteq \neg K(E) \cap \neg K \neg K(E)$

U2 *AU introspection*: $U(E) \subseteq UU(E)$

U3 *KU introspection*: $KU(E) = \emptyset$

Plausibility says if one is unaware of something, then one does not know it, and does not know that one does not know it. AU introspection says if one is unaware of something, then one must be unaware of the possibility of being unaware of it. KU introspection says under no circumstances can one know exactly what one is unaware of.

DLR show that the combination of these three axioms implies one critical property of unawareness: whenever the agent is unaware of something, he must not know the state space. That is, U1-3 imply $U(E) \subseteq \neg K(\Omega)$ for any non-empty $E \subseteq \Omega$. But then adding necessitation or monotonicity eliminates nontrivial unawareness.

3 The Product Model

3.1 The primitives

I explore a product structure on the full state space. Intuitively, one can think of the set of payoff-relevant uncertainties as a set of questions, and each state specifies a complete collection of resolutions to these uncertainties, or answers to the questions, one for each. Therefore, without loss of generality, one can write the state space as the Cartesian product of the sets of answers.

Let $\mathcal{D}^* = \{D_i\}_{i \in Q}$,[5] where each D_i is the set of answers to question i and Q is an arbitrary index set for all relevant questions. Without loss of generality, I assume D_i is non-empty for all $i \in Q$. The collection $\mathcal{D}^*$ represents full awareness.

The full state space Ω^* is defined as the Cartesian product of all sets in the collection of full awareness information:[6]

$$\Omega^* = \prod_{i \in Q} D_i \equiv \times \mathcal{D}^*$$

Information is represented by a pair (W^*, P^*), dubbed *generalized information structure*, consisting of both awareness information and factual information. The novel

[5]To differentiate the "full model" from the potentially incomplete subjective models, I add a $*$ to elements in the full model for emphasis wherever I can.

[6]Despite of the syntactic flavor, the product structure does not impose real limitations. Li (2006b) shows this model can be constructed on arbitrary state spaces.

component is *the awareness function* $W^* : \Omega^* \to 2^{\mathcal{D}^*}$, associating each full state with a subset of $\mathcal{D}^*$. The interpretation is that at ω^*, the agent is aware of the uncertainties contained in $W^*(\omega^*)$. The second component is a full possibility correspondence $P^* : \Omega^* \to 2^{\Omega^*} \setminus \{\emptyset\}$, interpreted as the factual signal structure.[7]

I model the agent's state-dependent mind-sets using subjective state spaces and subjective factual information. At ω^*, the agent's subjective state space only contains those uncertainties of which the agent is aware, i.e. contained in $W^*(\omega^*)$, and hence is naturally defined as the Cartesian product of all sets in his awareness information:

$$\Omega(\omega^*) = \times W^*(\omega^*) \tag{3.1}$$

Every subjective state $\omega \in \Omega(\omega^*)$ leaves some questions (those corresponding to sets of answers not included in $W^*(\omega^*)$) unanswered, and thus is a "blurry" picture of the environment. Different subjective state spaces blur the full state space in different ways. Let $\mathcal{S} = \{\Omega = \times\mathcal{D} : \emptyset \neq \mathcal{D} \subseteq \mathcal{D}^*\}$ be the collection of all possible subjective state spaces. For any $\Omega \in \mathcal{S}, \Omega = \times\mathcal{D}$, let $\mathbb{P}^\Omega$ be the projection operator that yields the projection of points in Ω' on Ω, where Ω' is "finer" than Ω: $\Omega' = \times\mathcal{D}'$ for some $\mathcal{D}' \supseteq \mathcal{D}$.

The agent can only recognize the factual signal within his awareness constraints. For every $\omega^* \in \Omega^*$, the subjective possibility correspondence $P_{\omega^*} : \Omega(\omega^*) \to 2^{\Omega(\omega^*)} \setminus \{\emptyset\}$ is the projection of the full factual signal $P^*(\omega^*)$ on the corresponding subjective state space:[8,9]

$$P_{\omega^*}(\omega) = \mathbb{P}^{\Omega(\omega^*)}(P^*(\omega^*)) \quad \text{for all} \quad \omega \in \mathbb{P}^{\Omega(\omega^*)}(P^*(\omega^*)) \tag{3.2}$$

Let $s(\omega^*) = \mathbb{P}^{\Omega(\omega^*)}(\omega^*)$ denote the "subjective true state," i.e. the true state projected on the agent's mind-set. At ω^*, the agent considers subjective states in $P_{\omega^*}(s(\omega^*)) = \mathbb{P}^{\Omega(\omega^*)}(P^*(\omega^*))$ possible, which equals to $P^*(\omega^*)$ only if the agent is fully aware.

Definition 2 *The generalized information structure* (W^*, P^*) *is* **rational** *if it satisfies: for all* $\omega_1^*, \omega_2^* \in \Omega^*$, $\omega_1^* \in P^*(\omega_2^*) \Rightarrow (W^*(\omega_1^*), P^*(\omega_1^*)) = (W^*(\omega_2^*), P^*(\omega_2^*))$.

To see the connection with standard information partition model, I break it into the following two components:

---

[7]Although mathematically identical as the possibility correspondence P in the standard model, the interpretation of P^* is different. In the standard model (Ω, P), $P(\omega)$ is interpreted as the set of states *the agent considers possible at* ω; in the product model, $P^*(\omega^*)$ is interpreted as *the factual signal the agent receives at* ω^*. Thus P is subjective – from the agent's perspective – while P^* is objective. In fact the counterpart of P in the product model is the subjective factual information $P_{\omega^*}(s(\omega^*))$ – see below.

[8]The image of P_{ω^*} at subjective states the agent knows to be false is irrelevant for the knowledge hierarchy in single-agent environment, hence I leave it unspecified for now. For a full specification of P_{ω^*}, see Section 5.

[9]Definition (3.2) implicitly assumes P^* induces an information partition over Ω^*. But in general, modeling unawareness neither requires, nor implies a partitional structure of P^*. The details of a more general model allowing for non-partitional P^* is available upon request.

1. *Factual partition*: P^* induces an information partition over Ω^*;
2. *Rational awareness*: $\omega_1^* \in P^*(\omega_2^*) \Rightarrow W^*(\omega_1^*) = W^*(\omega_2^*)$.

The additional "rational awareness" condition says in states the agent receives different awareness signals, he receives different factual signals, too. The rationality condition generalizes the partitional structure of information in the standard model, inducing a "local" subjective information partition: the agent excludes all *subjective* states in which he has different *subjective* factual information.

Example 1. Suppose Charlie has an episodic hearing problem that causes him to hear a lot of noise when he experiences the problem, which prevents him from telling whether it rains outside. Suppose Charlie is never aware of the hearing problem.

This is modeled as follows. Let r, nr, p, np denote "it is raining," "it is not raining," "experiencing the hearing problem," "not experiencing the hearing problem" respectively.

$\mathcal{D}^* = \{\{r, nr\}, \{p, np\}\}$
$\Omega^* = \times\mathcal{D}^* = \{(r,p), (r,np), (nr,p), (nr,np)\}$
$W^*(\omega^*) = \{\{r, nr\}\}$ for all $\omega^* \in \Omega^*$;
P^* induces the full information partition $\{\{(r,p),(nr,p)\}, \{(r,np)\}, \{(nr,np)\}\}$

At (r,p), Charlie's full factual signal is the event $\{(r,p),(nr,p)\}$: the fact he cannot tell whether it rains indicates he has a hearing problem. However, being unaware of the hearing problem, Charlie only recognizes that he does not know whether it rains. This is reflected in his subjective model at (r,p):

$\Omega((r,p)) = \times\{\{r, nr\}\} = \{r, nr\}$
$P_{(r,p)}(s(r,p)) = P_{(r,p)}(r) = \mathbb{P}^{\{r,nr\}}(\{(r,p),(nr,p)\}) = \{r, nr\}$

3.2 The events

In the standard model where all information is factual, events only differ in the facts they convey. With the generalized information structure, events can differ in awareness, in facts, or in both. For instance, in the hearing problem example, the events "it rains, and there is a possibility that Charlie has a hearing problem" and "it rains," although expressing essentially the same factual information, are different events because they involve different levels of awareness.

Let E be a nonempty subset of some subjective state space. By construction, one can identify its space and hence the awareness information it carries. Let $\mathcal{D}_E$ denote the unique subset of $\mathcal{D}^*$ such that $E \subseteq \times\mathcal{D}_E$. I also include a collection of empty sets in the model, one for each subjective state space: for any state space $\Omega \in \mathcal{S}$, let $\emptyset_\Omega$ denote the empty set associated with Ω. Intuitively, this object is the empty set tagged with the awareness information. It behaves in the same way as the usual empty set, except that it is confined to its state space.[10,11] Finally, for convenience, I rule out the empty state space by requiring $\{\Delta\} \in \mathcal{D}^*$ and $\{\Delta\} \in W^*(\omega^*)$ for all $\omega^* \in \Omega^*$.[12]

The collection of events in this model is hence taken to be all conceivable events at all subjective state spaces:

$$\mathcal{E}^p = \{E \neq \emptyset : E \subseteq \Omega \text{ for some } \Omega \in \mathcal{S}\}$$

Note $\mathcal{E}^p$ includes the collection $\{\emptyset_\Omega\}_{\Omega \in \mathcal{S}}$.

Definition 3 *F is said to be an* **elaboration** *of E and E is said to be a* **reduction** *of F, if*

$$\mathcal{D}_F \supseteq \mathcal{D}_E \ \text{ and } \ F = \left\{\omega \in \times\mathcal{D}_F : \mathbb{P}^{\times\mathcal{D}_E}(\omega) \in E\right\}$$

Elaborations and reductions are events that contain the same factual information, but incorporate different levels of awareness. For example, $\{(r,p),(r,np)\}$ is an elaboration of $\{r\}$. For any $E \in \mathcal{E}^p$ and Ω satisfying $\mathcal{D}_E \subseteq \mathcal{D}_\Omega$, let E_Ω denote the elaboration of E in Ω. That is, E_Ω is the unique event satisfying $E_\Omega \subseteq \Omega$ and E is a reduction of E_Ω. In the above example, $\{(r,p),(r,np)\} = \{r\}_{\Omega^*}$.

Since the logical relations between events concern only facts, they are preserved by elaborations. Thus one can deal with logical relations between arbitrary subjective events E and F using their minimal elaborations that live in the same space, i.e. the space defined by $\times(\mathcal{D}_E \cup \mathcal{D}_F)$. This observation suggests the following definitions of *extended set relations and operations* on elements of $\mathcal{E}^p$ and their connections to logical relations:

Definition 4 Extended set relations and operations[13]

1. **Extended set inclusion** *(logical consequence): E is an extended (weak) subset of F, denoted by* $E \subseteq_* F$*, if* $E_{\Omega^*} \subseteq F_{\Omega^*}$*;*

2. **Extended set intersection** *(conjunction):* $E \cap_* F \equiv E_{\times(\mathcal{D}_E \cup \mathcal{D}_F)} \cap F_{\times(\mathcal{D}_E \cup \mathcal{D}_F)}$*;*

3. **Extended set union** *(disjunction):* $E \cup_* F \equiv E_{\times(\mathcal{D}_E \cup \mathcal{D}_F)} \cup F_{\times(\mathcal{D}_E \cup \mathcal{D}_F)}$.

[10] It is straightforward to incorporate this new object in set operations (for notational ease, I use the conventional symbols for these operations): for any state space Ω and any sets $E, F \neq \emptyset_\Omega$, $E, F \subseteq \Omega$, the set inclusion, intersection, union and complement notions are defined in the usual way, except that for disjoint E and F, $E \cap F = \emptyset_\Omega$ instead of $\emptyset$. In addition, for any $E \subseteq \Omega$, $\emptyset_\Omega \cup E = E$, $\emptyset_\Omega \cap E = \emptyset_\Omega$, $E \setminus \emptyset_\Omega = E$, and for any E such that $E \neq \emptyset$ and $E \subseteq \Omega$, $\emptyset_\Omega \subseteq E$.

[11] Another way to think about the multiple empty sets is that an event E in the model is actually a pair $(\times\mathcal{D}_E, E)$ where $E \subseteq \times\mathcal{D}_E$. The first object in the pair, $\times\mathcal{D}_E$, specifies the awareness in E while the second object represents the involved facts. In particular, $\emptyset_\Omega$ should be interpreted as the pair $(\Omega, \emptyset)$. Then the usual set inclusion can be extended to this space by letting $(\Omega, E) \subseteq (\Omega', F)$ if and only if $\Omega = \Omega'$ and $E \subseteq F$, and similarly for other set operations. For $E \neq \emptyset$, $\mathcal{D}_E$ is uniquely identified from it and hence is redundant.

[12] For simplicity, I omit Δ when there is no risk of confusion.

[13] Notice these relations and operations reduce to the usual ones for events from the same space.

Finally, the **negation** of a subjective event involves the same amount of awareness, and hence is identified with the set complement operation *with respect to the corresponding subjective state space*:

$$\neg E = \times\mathcal{D}_E \setminus E$$

3.3 The knowledge and unawareness operators

One can only reason about things of which one is aware. For any event E, only if there is some version of E – either E itself or some elaboration of it – in the agent's mind-set, can he reason about it. In addition, the agent's mind-set constrains his introspection of his own knowledge. Recall the hearing problem example. Charlie knows it rains only if it rains and he does not experience the hearing problem. The event "*Charlie* knows it rains" (from the modeler's perspective) is the singleton set $\{(r, np)\}$. On the other hand, at (r, np), since Charlie himself is unaware of the hearing problem, *from his perspective*, the event "*I* know it rains" is represented by the singleton set $\{r\}$ in the subjective state space. I call the former "objective" knowledge and the latter "subjective" knowledge. Since (objective) higher-order knowledge describes the agent's introspection, the involved lower-order knowledge should be subjective knowledge. For example, in the event "Charlie knows that *he knows it rains*," the first-order knowledge "he knows it rains" refers to Charlie's subjective knowledge "I know it rains."

Fix a full state ω^*, the agent's mind-set is modeled by $(\Omega(\omega^*), P_{\omega^*})$. Notice this is a standard information partition model when restricted to subjective states contained in $P_{\omega^*}(s(\omega^*))$. So subjective knowledge can be computed just as in the standard model. For notational ease, let $\emptyset_E \equiv \emptyset_{\times\mathcal{D}_E}$ denote the empty set that is associated with the state space of E. The *subjective knowledge operator* at ω^*, denoted by $\tilde{K}_{\omega^*}$, is defined by: for all $E \in \mathcal{E}^p$,

$$\tilde{K}_{\omega^*}(E) = \begin{cases} \{\omega \in \Omega(\omega^*) : P_{\omega^*}(\omega) \subseteq E_{\Omega(\omega^*)}\} & \text{if } \mathcal{D}_E \subseteq W^*(\omega^*) \\ \emptyset_E & \text{if } \mathcal{D}_E \nsubseteq W^*(\omega^*) \end{cases} \tag{3.3}$$

In words, the agent *subjectively* knows E if he is aware of E and E is true in all *subjective* states he considers possible. Here $E_{\Omega(\omega^*)}$ is the version of E in the agent's mind-set provided he is aware of E. If the agent is unaware of E, i.e. $\mathcal{D}_E \nsubseteq W^*(\omega^*)$, then the event $E_{\Omega(\omega^*)}$ is not defined and $\tilde{K}_{\omega^*}(E)$ is empty. To preserve the awareness information in E, I denote it by the empty set that contains the same awareness as E.

The subjective event $\tilde{K}_{\omega^*}(E)$ is interpreted as "knowledge of E in the agent's mind-set." Similar to the standard model, iteration of the subjective knowledge operator yields *subjective higher-order knowledge* in the agent's mind-set: "I know that I know $\cdots$ I know E."

$$\tilde{K}^n_{\omega^*}(E) = \tilde{K}_{\omega^*}(\tilde{K}^{n-1}_{\omega^*}(E)) \tag{3.4}$$

The *objective* description of the knowledge hierarchy, viewed from the modeler's perspective, is then obtained by putting together the relevant pieces of the subjective knowledge hierarchies. Formally, the *n-th order objective knowledge* is defined as:

$$K^n(E) = \left\{\omega^* \in \Omega^* : s(\omega^*) \in \tilde{K}^n_{\omega^*}(E)\right\} \tag{3.5}$$

To be consistent with the literature, I use the customary notation K^n, although it is defined through the iteration of the subjective knowledge operators instead of iteration of K itself. The latter is in fact not defined.[14]

The first-order knowledge can be directly derived from the generalized information structure without referring to the subjective models. When $n = 1$, equation (3.5) reduces to:

$$K(E) \quad = \quad \{\omega^* \in \Omega^* : \mathcal{D}_E \subseteq W^*(\omega^*),\ P^*(\omega^*) \subseteq E_{\Omega^*}\} \tag{3.6}$$

Since awareness information is built into every event, it is easy to define the *unawareness operator*:

$$U(E) = \{\omega^* \in \Omega^* : \mathcal{D}_E \nsubseteq W^*(\omega^*)\} \tag{3.7}$$

$U(E)$ is interpreted as the objective event "the agent is unaware of E," viewed from the modeler's perspective.

To see the connection with the standard model, notice that if the agent is fully aware in all full states, then he has the same subjective model – the full model – in all full states. The subjective knowledge operator $\tilde{K}$ then reduces to the usual K in the standard model and all subjective knowledge hierarchies become identical and coincide with the objective knowledge hierarchy. In that case the product model simply reduces to the standard model.

4 The Knowledge Hierarchy with Unawareness

Recall that in the standard information partition model (Ω, P), the agent's knowledge hierarchy is completely characterized at the first level: for any $E \subseteq \Omega$,

1. $K(E) = KK(E)$;

2. $\neg K(E) = K\neg K(E)$.

In words, given any event, the agent either knows it or does not know it, and he always knows whether he knows it. Natural generalization of the above characterization obtains in the product model.

Theorem 2 *In the product model (Ω^*, W^*, P^*), let (W^*, P^*) be rational. Then the agent's knowledge hierarchy satisfies: for any $E \in \mathcal{E}^p$,*

[14] Although $K(E)$ is a legitimate event, it is in general not an event the agent can conceive, unless he is fully aware in every full state, i.e. in the standard model. In general it does not make sense to ask whether the agent knows $K(E)$.

1. $U(E) = \neg K(E) \cap \neg K \neg K(E)$;

2. $K(E) = KK(E)$;

3. $\neg K(E) \cap \neg U(E) = K \neg K(E)$.

Theorem 2 says, given any event, the agent is either unaware of it, in which case he does not know it and does not know he does not know it; or he is aware of it, in which case he either knows it or does not know it, and always knows whether he knows it. Analogous to the standard model, the knowledge hierarchy is completely pinned down at the first level. In particular, this means as long as (W^*, P^*) is consistent and rational, the entire knowledge hierarchy can be derived from the pair (W^*, P^*) directly.

I prove theorem 2 via two Lemmas. The first lemma deals with the basic properties of knowledge and and unawareness without imposing rationality of (W^*, P^*). For all $E, F \in \mathcal{E}^p$,

U0* *Symmetry*: $U(E) = U(\neg E)$

U1$'$ *Strong plausibility*: $U(E) \subseteq \bigcap_{n=1}^{\infty} (\neg K)^n(E)$

U2* *AU introspection*:[15] $U(E) \subseteq (\neg K)^n U(E)$

U3$'$ *Weak KU introspection*: $U(E) \cap KU(E) = \emptyset_{\Omega^*}$

K1* *Subjective necessitation*: $\omega^* \in K(\Omega(\omega^*))$ for all $\omega^* \in \Omega^*$

K2* *Generalized monotonicity*:[16] $E \subseteq_* F,\ \mathcal{D}_E \supseteq \mathcal{D}_F \Rightarrow K(E) \subseteq K(F)$

K3* *Conjunction*:[17] $K(E) \cap K(F) = K(E \cap_* F)$

Symmetry is proposed by Modica and Rustichini (1999). It says that one is unaware of an event if and only if one is unaware of the negation of it. The other three unawareness properties correspond to the three axioms proposed by DLR. Strong plausibility strengthens DLR's plausibility axiom. Plausibility requires that whenever one is unaware of something, one does not know it and does not know that one does not

[15]Since awareness is embedded in the primitive of the model, namely the collection $\mathcal{D}^*$, I avoid defining an "unawareness of unawareness" operator directly.

The definition of the operators $K^n U$ is given in section 5. For the current purpose, it suffices to note that these events concern the agent's introspection of his own awareness, and hence can only be properly characterized by tracking the agent's *subjective knowledge* of the *subjective event* "I am unaware of E," which contains at least as much awareness information as E does.

[16]This property implies the agent is no longer logically omniscient: he knows the logical consequences of his knowledge *only if he is aware of them.*

[17]Parallel to the standard model, conjunction implies generalized monotonicity.

know it. I require such a lack of knowledge to be extended to an arbitrarily high order.[18] While KU introspection requires the agent never know exactly what he is unaware of, the weak KU introspection weakens it by allowing the agent to have false knowledge of his being unaware of a particular event.

$K1^* - 3^*$ are natural analogues of K1-3 in the context of nontrivial unawareness. Recall that necessitation says the agent knows all tautological statements: $K(\Omega^*) = \Omega^*$. However, while all theorems are tautologies, arguably Newton does not know the theory of general relativity because he is unaware of it. This is reflected in subjective necessitation, which says the agent knows all tautological statements *of which he is aware*.[19]

The essence of monotonicity is the intuitive notion that knowledge should be monotonic with respect to the information content of events. In the standard model, an event is more informative than another if and only if it conveys more facts. In the product model, an event is more informative than another if and only if it contains both more facts and *more awareness*. Alternatively, note monotonicity means the agent knows the logical consequences of his knowledge, while generalized monotonicity, which explicitly takes into account that the agent may not be fully aware, says the agent knows those logical consequences of his knowledge *of which he is aware*.

Lemma 3 *The product model* $\{\Omega^*, W^*, P^*\}$ *satisfies* $U0^*$, $U1'$, $U2^*$, $U3'$ *and* $K1^* - 3^*$.

To see the connection between lemma 3 and DLR's impossibility results, first notice that by (3.6),

$$K(\Omega^*) = \{\omega^* \in \Omega^* : \mathcal{D}^* \subseteq W^*(\omega^*), P^*(\omega^*) \subseteq \Omega^*\} = \{\omega^* \in \Omega^* : \mathcal{D}^* = W^*(\omega^*)\}$$

That is, $K(\Omega^*) = \Omega^* \Leftrightarrow W^*(\omega^*) = \mathcal{D}^*$ for all $\omega^* \in \Omega^*$. In other words, necessitation holds if and only if the agent is fully aware in every full state. But this implies $\mathcal{D}_E \subseteq W^*(\omega^*)$ for all $\omega^* \in \Omega^*$ and $E \in \mathcal{E}$, hence $U(E) = \emptyset_{\Omega^*}$, which is DLR's first impossibility result.

Secondly, observe that generalized monotonicity and monotonicity differ in that generalized monotonicity does not require knowledge to be monotonic when $\mathcal{D}_E \nsubseteq \mathcal{D}_F$, while in the standard model, one necessarily has $\mathcal{D}_E = \mathcal{D}_F$. Let $E \subseteq_* F$ and $K(E) \subseteq K(F)$. By (3.6), $\{\omega^* \in \Omega^* : \mathcal{D}_E \subseteq W^*(\omega^*)\} \subseteq \{\omega^* \in \Omega^* : \mathcal{D}_F \subseteq W^*(\omega^*)\}$. This implies $U(F) \subseteq U(E)$, which says whenever the agent is unaware of F, he is unaware of E. By (3.7), for any G such that $\mathcal{D}_G = \mathcal{D}_E$,

$$U(F) \subseteq U(G) = U(E)$$

[18] DLR consider the weaker property plausibility, which is sufficient for the negative result in which they are interested. However, when it comes to providing *positive* results in a model that deals with unawareness, strong plausibility, or an even stronger property that equates unawareness with the lack of knowledge of all orders which I discuss shortly, seems to be more interesting.

[19] Subjective necessitation is equivalent to the "weak necessitation" property DLR discussed in the context of propositional models.

That is, whenever the agent is unaware of F, he is unaware of any event that contains the same awareness information as E. By strong plausibility, he cannot know any of them, which is DLR's second impossibility result.

Lemma 3 does not require (W^*, P^*) to be rational. This is because analogous to the standard model, rationality in information processing mainly has implications for higher-order knowledge, while none of the above properties involves higher-order knowledge.

For all $E \in \mathcal{E}^p$,

U1* *UUU (Unawareness = unknown unknowns)*: $U(E) = \bigcap_{n=1}^{\infty} (\neg K)^n(E)$

U3* *KU introspection*: $KU(E) = \emptyset_{\Omega^*}$

K4.a^* *The axiom of knowledge I*: $K(E) \subseteq_* E$

*K*4.b^* *The axiom of knowledge II*: $K^n(E) \subseteq K^{n-1}(E)$

K5* *The axiom of transparency*: $K(E) \subseteq KK(E)$

K6* *The axiom of limited wisdom*: $\neg K(E) \cap \neg U(E) \subseteq K\neg K(E)$

Lemma 4 *The product model* (Ω^*, W^*, P^*) *satisfies* U1*, U3*, K4.a^*, K4.b^*, K5* *and* K6* *if* (W^*, P^*) *is rational.*

The axiom of limited wisdom extends the axiom of wisdom to an environment with unawareness by only requiring the agent to know that he does not know when he *is aware of the involved event*. UUU says the agent is unaware of an event *if and only if* he does not know it, he does not know that he does not know it, and so on. The extra strength added to strong plausibility is due to the axiom of limited wisdom: if the agent always knows that he does not know if he is aware of the event, then obviously the only circumstance where he does not know that he does not know is that he is unaware of it. Lastly, the axiom of knowledge says the agent can never have false knowledge, both with respect to the set of "natural" events $\mathcal{E}^p$ (K4.a^*), and in introspection of his own knowledge hierarchy (K4.b^*).[20] Such "non-delusion" property, combined with weak KU introspection, yields KU introspection.[21]

Remark 1. Information has more dramatic effects on the agent's knowledge hierarchy when the agent has nontrivial unawareness than when he does not. Upon receipt of new information, the agent updates his subjective state space as well as his subjective factual

[20] In the standard model, higher-order knowledge are defined through the iterations of the knowledge operator, and hence coincide with objective knowledge, i.e. "natural" events, K4 alone suffices and hence there is no need to split it into two separate properties.

[21] Strictly speaking, it is the non-delusion property in the subjective models, i.e. $\tilde{K}_{\omega^*}(E) \subseteq E$, where E could be subjective knowledge or unawareness, that delivers KU introspection.

information. Formally, given ω^*, let the agent's initial information be $(W_0^*(\omega^*), P_0^*(\omega^*))$. The agent has subjective factual information:

$$\mathbb{P}^{\times W_0^*(\omega^*)}[P_0^*(\omega^*)]$$

Upon receipt of new information $(W_1^*(\omega^*), P_1^*(\omega^*))$, the agent updates his subjective factual information to

$$\mathbb{P}^{\times [W_0^*(\omega^*) \cup W_1^*(\omega^*)]}[P_0^*(\omega^*) \cap P_1^*(\omega^*)]$$

As long as $W_1^*(\omega^*) \setminus W_0^*(\omega^*) \neq \emptyset$, the agent gains new knowledge.

In particular, if $P_0^*(\omega^*)$ is not an elaboration of $\mathbb{P}^{\times W_0^*(\omega^*)}[P_0^*(\omega^*)]$, that is, if it contains factual information about uncertainties beyond $W_0^*(\omega^*)$, then the agent could learn new facts from introspection of the first-period factual signal along. For example, in the Watson story, at b, Watson's initial information is the pair $(W_0^*(b), P_0^*(b)) = (\{\Delta\}, \{b\})$. His subjective factual information is $\mathbb{P}^{\{\Delta\}}\{b\} = \{\Delta\}$. Suppose Holmes asks Watson, "could there have been an intruder?" This question is represented by the signal pair $(W_1^*(b), P_1^*(b)) = (\{\{a', b'\}\}, \{a, b\})$. Now Watson updates his subjective state space and recognizes the factual information he has had all along but neglected: $\mathbb{P}^{\times(\{\Delta\} \cup \{a', b'\})}(\{b\} \cap \{a, b\}) = \{b\}$.

Remark 2. The model clearly differentiates assigning probability zero to an event and being unaware of it. In the hearing problem example, Charlie is unaware of the events "it rains and I have the hearing problem" and "it rains and I do not have the hearing problem," while he may assign positive probability to the event "it rains." In contrast, assigning zero probability to the former two events dictates that Charlie also assigns zero probability to the event "ir rains." This distinction is especially stark in dynamic environments. Upon receipt of new awareness information, the agent updates his probability space and may assign *any* probability to an event of which he is unaware before, while the events of which he assigns probability zero must still have probability zero.

5 The Multi-agent Model

In multi-agent environment, players reason about each other's awareness as well as knowledge, within the confines of their own awareness. To allow i to reason about j's mind-set, i's subjective model needs to include a specification of j's awareness signals at each *subjective state of i's*. Therefore, to model interactive mind-sets, I expand the product model to allow players to have *subjective product models.*

Let $N = \{1, \cdots, n\}$ be the set of players, and $\mathbf{W}^* = (W_1^*, \cdots, W_n^*)$ and $\mathbf{P}^* = (P_1^*, \cdots, P_n^*)$ denote the vector of awareness function and full possibility correspondences. Below I construct the subjective interactive models from the full model $(\Omega^*, \mathbf{W}^*, \mathbf{P}^*)$.

5.1 Deriving subjective models.

Let $W_j(\cdot|i_{\omega^*})$ denote *j's subjective awareness function* i ascribes to j in i's mind-set at ω^*. The symbol i_{ω^*} is to be understood as the pair (ω^*, i). For any $\omega \in \Omega_i(\omega^*)$,

$$W_j(\omega|i_{\omega^*}) = W_i^*(\omega^*) \cap [\bigcup_{\{\omega_1^*:\, \mathbb{P}^{\Omega_i(\omega^*)}(\omega_1^*)=\omega\}} W_j^*(\omega_1^*)] \tag{5.1}$$

Definition (5.1) says, j's awareness signal at i's subjective state $\omega \in \Omega_i(\omega^*)$ consists of any awareness information j receives in *some* full state underlying ω, and that i recognizes given i's own awareness constraints. The intersection captures that i cannot reason beyond his own mind-set, while the union captures that i cannot distinguish full states in $\{\omega_1^* : \mathbb{P}^{\Omega_i(\omega^*)}(\omega_1^*) = \omega\}$ due to his own unawareness.

The map $W_i(\cdot|i_{\omega^*})$ captures i's perception of his own awareness signal structure, given what i is aware of himself at ω^*. Rational awareness ensures $W_i(\omega|i_{\omega^*}) = W_i^*(\omega^*)$ for all $\omega \in P_{\omega^*}(s(\omega^*))$, that is, i knows his own awareness. Player i could be aware that he would have been unaware of some uncertainties in a counterfactual subjective state. For example, in a, Watson may be aware that he would have been unaware of the possibility of no intruder in b, he may even *know* it.

A complication in multi-agent environment is that i's knowledge at a counterfactual subjective state $\omega \in \Omega_i(\omega^*)$ matters if j considers (some version of) ω possible. Thus a fully specified subjective possibility correspondence is needed. It is intuitively appealing to take the projection of the full possibility correspondence, but what if unawareness affects what the agent *would have known* regarding things of which he is aware? Equivalently, the issue is, each subjective state may correspond to multiple full states where factual signals may differ, producing different projections in the subjective state space.

Recall the hearing problem example. At (r, np), Charlie knows it rains. It seems plausible to say that since he is unaware of the hearing problem which could cause him to be ignorant about the weather condition, in his mind-set he should consider that he would have known it does not rain had it not rained, the projection of the full factual signal he receives at (nr, np).

Therefore, it seems plausible to assume each player i has a set of "default" full states, one for each subjective state, that coincides with the current full state in the answers of questions of which i is unaware. Formally, for any $\omega^* \in \Omega^*$, let $u_i(\omega^*) = \mathbb{P}^{\times(\mathcal{D}^* \setminus W_i^*(\omega^*))}(\omega^*)$ denote the "default" resolution of the uncertainties of which i is unaware in his mind-set at ω^*. For any $\omega \in \Omega_i^*(\omega^*)$, let $\omega \times u_i(\omega^*)$ denote the full state where the uncertainties of which i is aware are resolved as in the subjective state ω, and those uncertainties of which i is unaware are resolved according to the "default."

Let $P_j(\cdot|i_{\omega^*}) : \Omega_i(\omega^*) \to 2^{\Omega_i(\omega^*)} \setminus \{\emptyset\}$ denote j's factual information structure in i's mind-set. Define:

$$P_j(\omega|i_{\omega^*}) = \begin{cases} \mathbb{P}^{\Omega_i(\omega^*)} P_j^*(\omega^*) & \text{for } \omega \in \mathbb{P}^{\Omega_i(\omega^*)} P_j^*(\omega^*) \\ \mathbb{P}^{\Omega_i(\omega^*)} P_j^*(\omega \times u_i(\omega^*)) & \text{otherwise.} \end{cases} \tag{5.2}$$

The tuple $(\Omega_i(\omega^*), W_j(\cdot|i_{\omega^*}), P_j(\cdot|i_{\omega^*}))$ describes the first-order interactive model capturing i's reasoning about j in his mind-set at ω^*. It is a product model itself, with $(\Omega_i(\omega^*), P_j(\cdot|i_{\omega^*}))$ being the "full" state space and "full" possibility correspondence. In sum, at ω^*, players' mind-sets are described by a collection of multi-agent product models, one for each agent:

$$(\Omega_i(\omega^*), \mathbf{W}(\cdot|i_{\omega^*}), \mathbf{P}(\cdot|i_{\omega^*}))_{i\in N}$$

where $\mathbf{W}(\cdot|i_{\omega^*}) = (W_1(\cdot|i_{\omega^*}), \cdots, W_n(\cdot|i_{\omega^*}))$ and $\mathbf{P}(\cdot|i_{\omega^*}) = (P_1(\cdot|i_{\omega^*}), \cdots, P_n(\cdot|i_{\omega^*}))$.

Each player i's mind-set is described by a multi-agent product model *of which i is fully aware*, based on which higher-order subjective models can be derived inductively.

Let $\Delta^1 = \{((\omega_1, i^1)) : \omega_1 \in \Omega^*, i^1 \in N\}$;

$\Delta^2 = \{q^1 + ((\omega_2, i^2)) : q^1 = ((\omega_1, i^1)) \in \Delta^1, \omega_2 \in \Omega_{i^1}(\omega_1), i^2 \in N\}$ where "+" denotes concatenation;

$\cdots$

$\Delta^{n+1} = \{q^n + ((\omega_{n+1}, i^{n+1})) : q^n \in \Delta^n, \omega_{n+1} \in \Omega_{i^n}(\omega_n|q^{n:n-1}), i_{n+1} \in N\}$ where $q^{n:n-1}$ denote the sequence consisting of the first $n-1$ terms of q^n; for ease of notation, when the sequence q^n is clear from the context, I simply write q^{n-1} in place of $q^{n:n-1}$.

The n-th order subjective model representing i^1's reasoning at ω_1 of i^2's reasoning at ω_2 of $\cdots$ of i^n's reasoning at ω_n about i^{n+1} is described by the product model:

$$(\Omega_{i^n}(\omega_n|q^{n-1}), W_{i^{n+1}}(\cdot|q^n), P_{i^{n+1}}(\cdot|q^n))$$

where:

$$\Omega_{i^n}(\omega_n|q^{n-1}) = \times W_{i^n}(\omega_n|q^{n-1}) \tag{5.3}$$

For notational ease, let $\Omega^m \equiv \Omega_{i^m}(\omega_m|q^{m-1}), m = 1, \cdots, n$ and q^{m-1} denote the first $m-1$ terms of the sequence q^n. Then for any $\omega \in \Omega^n$,

$$W_{i^{n+1}}(\omega|q^n) = W_{i^n}(\omega_n|q^{n-1}) \cap [\bigcup_{\{\omega'\in\Omega^{n-1}:\, \mathbb{P}^{\Omega^n}(\omega')=\omega\}} W_{i^{n+1}}(\omega'|q^{n-1})] \tag{5.4}$$

$$P_{i^{n+1}}(\omega|q^n) = \begin{cases} \mathbb{P}^{\Omega^n} P_{i^{n+1}}(\omega_n|q^{n-1}) & \text{for } \omega \in \mathbb{P}^{\Omega^n} P_{i^{n+1}}(\omega_n|q^{n-1}) \\ \mathbb{P}^{\Omega^n} P_{i^{n+1}}(\omega \times u_{i^n}(\omega_n|q^{n-1})|q^{n-1}) & \text{otherwise.} \end{cases} \tag{5.5}$$

where in (5.5), $u_{i^n}(\omega_n|q^{n-1}) = \mathbb{P}^{\times[W_{i^{n-1}}(\omega_{n-1}|q^{n-2})\setminus W_{i^n}(\omega_n|q^{n-1})]}(\omega_n)$ is the "default" resolution of those uncertainties of which i^n is unaware at ω_n (but i^{n-1} is aware), according to i^1's reasoning at ω_1 about i^2's reasoning at ω_2 about $\cdots$ about i^{n-1}'s reasoning at ω_{n-1} about i^n. Note for any $\omega \in \Omega^n$, $\omega \times u_{i^n}(\omega_n|q^{n-1})$ is the subjective state in Ω^{n-1} where the uncertainties of which i^n is aware are resolved as in ω and the uncertainties of which i^n is unaware are resolved according to the "default" value captured in $u_{i^n}(\omega_n|q^{n-1})$.

The tuple $(\Omega^*, \mathbf{W}^*, \mathbf{P}^*)$ along with the construction in (5.3)-(5.5) consists of the multi-agent product model.

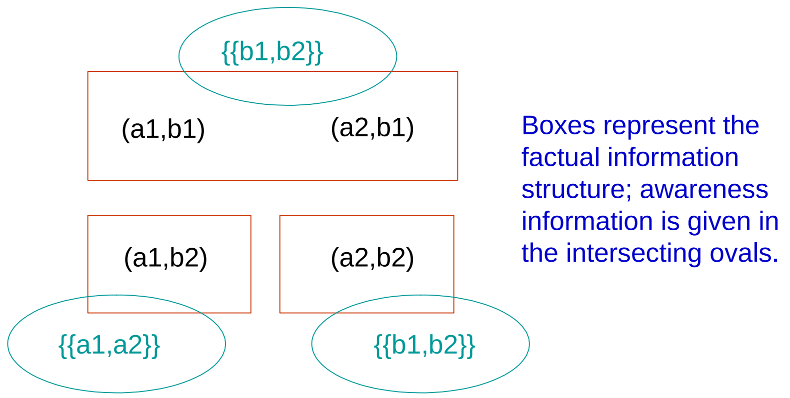

Figure 1: a full state space that is not rich.

5.2 Richness and rationality conditions.

I say $(\Omega^*, \mathbf{W}^*, \mathbf{P}^*)$ is *rational* if (W_i^*, P_i^*) is rational for all $i \in N$. Observe that i plays the role of the modeler in i's reasoning about j's knowledge and awareness, *subject to i's own unawareness, i.e. in i's subjective model.* To obtain a characterization of interactive knowledge hierarchies, I seek conditions to ensure $(W_j(\cdot|i_{\omega^*}), P_j(\cdot|i_{\omega^*}))$ to be rational for all $i, j \in N$ and $\omega^* \in \Omega^*$.

Definition 5 *The full state space Ω^* is* **rich** *if, for any $\omega_1^*, \omega_3^* \in \Omega^*$ such that $\mathbb{P}^\Omega(\omega_1^*) = \mathbb{P}^\Omega(\omega_3^*)$ for some $\Omega \in \mathcal{S}$, and $W_i(\omega_1^*) \neq W_i(\omega_3^*)$ for some $i \in N$, $\omega_2^* \in P_i^*(\omega_1^*)$ implies there exists some ω_4^* such that $\mathbb{P}^\Omega(\omega_2^*) = \mathbb{P}^\Omega(\omega_4^*)$ and $W_i(\omega_3^*) = W_i(\omega_4^*)$.*

Richness is a regularity condition. It says, if two full states ω_1^*, ω_3^* coincide in their answers to questions specified in Ω, and i has different awareness signals in them, then for any ω_2^* where i has the same factual signal as in ω_1^*, there must exist some ω_4^* that coincides with ω_2^* in their answers to questions specified in Ω, and where i's awareness signal coincides with that in ω_3^*. This condition requires the "coding" of generalized information to be sufficiently comprehensive: whenever answers to some question(s) (those not specified in Ω) affect *both* factual and awareness signals i receives, then such effects must be represented in the full state space symmetrically. Figure 1 illustrates the richness condition. In this example, since different answers to question b generate both different factual signal and different awareness signal in the two a_1 states, and the factual signal is the same in the two b_1 states, richness condition requires the full state space at least contain another a_2 state in which the awareness signal is question a, as in (a_1, b_2).

For any $q^n \in \Delta^n$ and $k < n$, let $q^{n\setminus k}$ denote the sequence obtained by removing the k-th term from q^n, i.e. $q^{n\setminus k} = ((\omega_1, i^1), \cdots, (\omega_{k-1}, i^{k-1}), (\omega_{k+1}, i^{k+1}), \cdots, (\omega_n, i^n))$. Let $s(\omega|q^n) \equiv \mathbb{P}^{\Omega^n}(\omega)$ denote the projection of ω on the interactive subjective state space Ω^n, i.e. $\Omega_{i^n}(\omega_n|q^{n-1})$.

Proposition 5 *Let the multi-agent product model* $(\Omega^*, \mathbf{W}^*, \mathbf{P}^*)$ *be rich and rational. Then for any* n*, and any* $q^n \in \Delta^n$ *satisfying* $\omega_k \in P_{i^k}(s(\omega_{k-1}|q^{k-1})|q^{k-1})$ *and* $i^k = i^{k-1}$ *for some* $k < n$*, the following holds for all* m *such that* $k < m \leq n$*:*

$$(\Omega_{i^m}(\omega_m|q^{m-1}), W_{i^{m+1}}(\cdot|q^m), P_{i^{m+1}}(\cdot|q^m)) = (\Omega_{i^m}(\omega_m|q^{(m-1)\setminus k}), W_{i^{m+1}}(\cdot|q^{m\setminus k}), P_{i^{m+1}}(\cdot|q^{m\setminus k}))$$

Proposition 5 says that, if an interactive reasoning involves a player i^{k-1} reasoning about his own model at a possible subjective state ω_k, then it leads to the same subjective interactive model as the one without such self-referential reasoning. In other words, a rich model yields a tractable hierarchy of interactive awareness under the construction (5.3)-(5.5): although agents may have different subjective models due to different awareness, everybody knows their own subjective model, everybody knows everybody knows their own subjective model, and so on.

To ensure the subjective possibility correspondence $P_j(\cdot|i_{\omega^*})$ to induce an information partition over $\Omega_i(\omega^*)$, I require the factual signals be decomposable into the conjunction of distinct signals, one for each relevant question.

Definition 6 *The possibility correspondence* $P^* : \Omega^* \to 2^{\Omega^*} \setminus \{\emptyset\}$ *satisfies* **product factual partition** *if it induces an information partition over* Ω^**, and that for all* $\omega^* \in \Omega^*$*,* $P^*(\omega^*)$ *can be written as a product set. That is, for every* ω^**, there exists a collection of partitions* $\{\pi^i_{\omega^*}\}_{i\in Q}$*, where* $\pi^i_{\omega^*}$ *is a partition over* $D_i \in \mathcal{D}^*$*, such that* $P^*(\omega^*) = \times_{i\in Q}\pi^i_{\omega}(\mathbb{P}^{\times\{D_i\}}(\omega^*))$*.*

It is worth mentioning that the decomposition may be state-dependent. All examples considered so far satisfy this condition.

Definition 7 *The multi-agent generalized information structure* $(\mathbf{W}^*, \mathbf{P}^*)$ *is* **interactively rational** *if* (W^*_i, P^*_i) *satisfies product factual partition and rational awareness for all* i*.*

The following key lemma says that in a rich and interactively rational product model, every derived interactive model is a product model with rational information structures, and hence the result on single-agent knowledge and unawareness carries to all interactive models, leading to Theorem 7.

Lemma 6 *Let* $(\Omega^*, \mathbf{W}^*, \mathbf{P}^*)$ *be rich and interactively rational. Then for any* n*, and any* $q^n \in \Delta^n$*,* $(W_{i^{n+1}}(\cdot|q^n), P_{i^{n+1}}(\cdot|q^n))$ *is rational.*

5.3 Interactive knowledge hierarchy.

Intuitively, i reasons about j in his subjective model for j, i.e. $(\Omega_i(\omega^*), W_j(\cdot|i_{\omega^*})P_j(\cdot|i_{\omega^*}))$, just like the modeler reasons about the agent in the single-agent case. The interactive knowledge hierarchy thus can be constructed recursively.

For example, the objective event "i knows j knows E," denoted by $K_iK_j(E)$, is obtained by tracking i's subjective knowledge "I know 'j knows E'," where "j knows E" is *j's knowledge of E in i's subjective model.* Let the latter be denoted by $\tilde{K}^j(E|i_{\omega^*})$. Analogous to the derivation of objective knowledge hierarchy in the single-agent case,

$$K_iK_j(E) \;=\; \left\{\omega^* \in \Omega^* : s(\omega^*|i_{\omega^*}) \in \tilde{K}^i_{\omega^*}(\tilde{K}^j(E|i_{\omega^*}))\right\} \tag{5.6}$$

where $\tilde{K}^j(E|i_{\omega^*}) = \left\{\omega \in \Omega_i(\omega^*) : P_j(\omega|i_{\omega^*}) \subseteq E_{\Omega_i(\omega^*)},\ W_j(\omega|i_{\omega^*}) \supseteq \mathcal{D}_E\right\}$ by (3.6).

Similarly, the objective event "i knows j is unaware of E," denoted by $K_iU_j(E)$, is defined by:

$$K_iU_j(E) \;=\; \left\{\omega^* \in \Omega^* : s(\omega^*|i_{\omega^*}) \in \tilde{K}^i_{\omega^*}(\tilde{U}^j(E|i_{\omega^*}))\right\} \tag{5.7}$$

where $\tilde{U}^j(E|i_{\omega^*}) = \{\omega \in \Omega_i(\omega^*) : \mathcal{D}_E \nsubseteq W_j(\omega|i_{\omega^*})\}$ by (3.7), interpreted as the event "j is unaware of E" *from i's perspective.*

In general, for any n and any $I = (i^1, \cdots, i^n) \in N^n$, the (objective) interactive knowledge "i^1 knows that i^2 knows $\cdots$ knows i^n knows E," denoted by $K_{i^1} \cdots K_{i^n}(E)$, is defined as follows. For any $\omega^* \in \Omega^*$, let $I^n_{\omega^*} = ((\omega_1, i^1), \cdots, (\omega_n, i^n)) \in \Delta^n$ denote the sequence where $\omega_1 = \omega^*, \omega_m = s(\omega^*|I^{m-1}_{\omega^*})$ for all $m = 2, \cdots, n$. For ease of notation, I use $I^m_{\omega^*}$ to denote $I^{n;m}_{\omega^*}$. Then, for any $E \in \mathcal{E}^p$, the subjective event "i^n knows E" in i^{n-1}'s subjective model ascribed by i^{n-2} ascribed by $\cdots$ by i^1 at ω^* is again derived from (3.6):

$$\tilde{K}^{i^n}(E|I^{n-1}_{\omega^*}) = \left\{\omega \in \Omega^{n-1} : P_{i^n}(\omega|I^{n-1}_{\omega^*}) \subseteq E_{\Omega^{n-1}}, W_{i^n}(\omega|I^{n-1}_{\omega^*}) \supseteq \mathcal{D}_E\right\} \tag{5.8}$$

Similarly, "i^n is unaware of E" in i^{n-1}'s subjective model ascribed by i^{n-2} ascribed by $\cdots$ by i^1 at ω^* is defined by:

$$\tilde{U}^{i^n}(E|I^{n-1}_{\omega^*}) = \left\{\omega \in \Omega^{n-1} : W_{i^n}(\omega|I^{n-1}_{\omega^*}) \nsupseteq \mathcal{D}_E\right\} \tag{5.9}$$

The objective higher-order interactive knowledge hierarchy is therefore defined by: for any $E \in \mathcal{E}^p$,

$$\begin{aligned} K_{i^1} \cdots K_{i^n}(E) \;=\; & \{\omega^* \in \Omega^* : s(\omega^*|I^1_{\omega^*}) \in \\ & \tilde{K}^{i^1}_{\omega^*}(\tilde{K}^{i^2}(\cdots(\tilde{K}^{i^n}(E|I^{n-1}_{\omega^*}))\cdots|I^1_{\omega^*}))\} \end{aligned} \tag{5.10}$$

$$\begin{aligned} K_{i^1} \cdots U_{i^n}(E) \;=\; & \{\omega^* \in \Omega^* : s(\omega^*|I^1_{\omega^*}) \in \\ & \tilde{K}^{i^1}_{\omega^*}(\tilde{K}^{i^2}(\cdots(\tilde{U}^{i^n}(E|I^{n-1}_{\omega^*}))\cdots|I^1_{\omega^*}))\} \end{aligned} \tag{5.11}$$

Theorem 7 *Let $(\Omega^*, \mathbf{W}^*, \mathbf{P}^*)$ be rich and interactively rational. Then, for any $E \in \mathcal{E}^p$,*

IK1 $i^k = i^{k-1}$ for some $1 < k \leq n \Rightarrow K_{i^1} \cdots K_{i^n}(E) = K_{i^1} \cdots K_{i^{k-1}} K_{i^{k+1}} K_{i^n}(E)$;

IK2 $K_i U_j(E) = \bigcap_{n=1}^{\infty} K_i(\neg K)_j^n(E)$;

IK3 $K_i K_j(E) \subseteq K_j(E) \cup [U_j(E) \cap K_i \neg U_j(E)]$.

$IK1$ says every agent knows his own knowledge, and everybody knows that everybody knows his own knowledge, and so on. $IK2$ says i knows j is unaware of E if and only if i knows the knowledge of E is lacking from j's knowledge hierarchy at all levels.

$IK3$ says two things: first, interactive knowledge could be "false," in the sense that it could be the case that i "knows" j knows E while j actually does not know E; second, the only situation this could happen is that j not only does not know E but also in fact is unaware of E; and that i "knows" j is aware of E. The interpretation is that i is actually unaware that j is unaware of E, and hence has wrong ideas about j's subjective factual signals. For example, suppose Holmes is unaware that Watson is unaware of the possibility of no intruder, then Holmes "knows" Watson knows there was no intruder. Intuitively, if i is aware of both E and j, then i would reason about j's reasoning about E *unless i is aware that j may be unaware of E*. The assumption that unawareness of unawareness is equivalent to being (falsely) aware of awareness is embedded in (5.1).[22] Notice it is possible for i to be aware that j may be unaware of E, while being unaware that k could be unaware of E, which introduces rich yet tractable interactive knowledge hierarchies, thanks to theorems 2 and 7.

5.4 Common knowledge.

An event E is common knowledge if everybody knows it, everybody knows everybody knows it, and so on. For any $i \in N$, let $I_i^m = (i^1, i^2, \cdots, i^m)$ be a sequence of players satisfying $i^1 = i$. For notational ease, let $K(E|I_i^m) = K_{i^1} \cdots K_{i^n}(E)$. Then the (objective) event "E is common knowledge," denote by $CK(E)$, is defined by:

$$CK(E) \equiv \cap_{i=1}^{n} \cap_{m=1}^{\infty} K(E|I_i^m) \tag{5.12}$$

The set $\cap_{m=1}^{\infty} K(E|I_i^m)$ represents the event "i knows E, i knows everybody knows E, i knows everybody knows everybody knows E, and so on." Intuitively, this event can be regarded as common knowledge of E *from i's perspective*, or i's *subjective common knowledge* of E, denoted by $CK_i(E)$. Rewrite (5.12) as:

$$CK(E) = \cap_{i=1}^{n} CK_i(E)$$

That is, an event is common knowledge if and only if it is common knowledge from everyone's perspective. In the standard model, since all subjective knowledge coincide with

[22] The presence of false interactive knowledge is consistent with the axiom of knowledge. This is because in the interactive knowledge "i knows j knows E," the event "j knows E" is a *subjective event that only exists in i's subjective model*, and hence axiom of knowledge has no bite. Rather, the problem is, i's subjective model is "wrong" because i fails to recognize j's lack of awareness signals, and hence "j knows E" in i's subjective model fails to reflect j's true knowledge.

objective knowledge, by axiom of knowledge, for any I_i^m, $K(E|I_i^m) \subseteq K(E|(i^2, \cdots, i^m))$, which implies $CK_i(E) \subseteq CK(E)$, and hence $CK(E) = \cap_{m=1}^{\infty} K(E|I^m) = CK_i(E)$ for all $i \in N$.

Aumann (1976) shows that in the standard model $(\Omega^*, \mathbf{P}^*)$, common knowledge has a simple and elegant characterization: for any $E \subseteq \Omega^*$,

$$CK(E) = \left\{\omega^* \in \Omega^* : \wedge_{j=1}^n P_j^*(\omega^*) \subseteq E\right\} \tag{5.13}$$

where $\wedge_{j=1}^n P_j^*$ denotes the meet of the information partition generated by the agents' possibility correspondences, and $\wedge_{j=1}^n P_j^*(\omega^*)$ represents the partition element containing ω^* in the meet.

This characterization obviously breaks down when unawareness is an issue. First, i may not know E even though his factual signal indicates E, due to his unawareness of E. Second, i may have "false" knowledge of j's knowledge due to i's unawareness of j's unawareness of E $(IK3)$. Suppose j is unaware of E, while i is unaware that j is unaware of E, then i could mistakenly consider E to be common knowledge. This is particularly relevant in situations where players have public factual signals. For example, suppose Holmes is unaware that Watson could be unaware of the possibility of no intruder, then from Holmes' perspective, the event "there was no intruder" is common knowledge, while it is not even mutual knowledge. From this aspect, (5.13) is too weak.

A natural candidate for the characterization of common knowledge in this environment is to add a "common awareness of E" clause to (5.13):

$$\underline{CK}(E) = \left\{\omega^* \in \Omega^* : \mathcal{D}_E \subseteq \cap_{j=1}^n \bigcap_{\omega_1^* \in \wedge_{j=1}^n P_j^*(\omega^*)} W_j^*(\omega_1^*),\ \wedge_{j=1}^n P_j^*(\omega^*) \subseteq E_{\Omega^*}\right\} \tag{5.14}$$

On the other hand, being unaware of others' unawareness may also result in being unaware of others' potential uncertainty, making it easier to achieve common knowledge. Recall the hearing problem example. Suppose it is raining, and Charlie does not have the hearing problem, i.e. the true state is (r, np). Suppose Dorothy is unaware that Charlie has the hearing problem, while she sees whether it rains outside, inducing a full information partition $\{\{(r,p),(r,np)\},\{(nr,p),(nr,np)\}\}$. It is easy to check that both of them know it rains, and know that both of them know it rains, and so on. However, Aumann's condition does not hold: at (r, np), Dorothy's full factual signal does not exclude (r, p), where Charlie does not know it rains. Intuitively, since Dorothy is unaware of Charlie's hearing problem, she is also unaware of how it affects whether Charlie hears the raining. To her, "it rains" is common knowledge between her and Charlie. In such situations, (5.13) is actually not necessary.

Theorem 8 *Let $(\Omega^*, \mathbf{W}^*, \mathbf{P}^*)$ be rich and interactively rational. Then for all $E \in \mathcal{E}^p$,*

$$\underline{CK}(E) \subseteq CK(E) \subseteq \cap_{i=1}^n K_i(E)$$

The prominent feature in the hearing problem example is the correlation between what *happens* along a dimension of which i is unaware and what j *knows about* and/or *whether* j *is aware of* E of which i is aware. It turns out the weak implication of common knowledge in this environment is precisely because of such correlation.

Definition 8 *The possibility correspondence* $P^* : \Omega^* \to 2^{\Omega^*} \setminus \{\emptyset\}$ *satisfies* **cylinder factual partition** *if it induces an information partition over* Ω^* *and* $P^*(\omega^*)$ *is a cylinder event for all* $\omega^* \in \Omega^*$. *That is, there exists a collection of partitions* $\{\pi^i\}_{i\in Q}$, π^i *is a partition over* $D_i \in \mathcal{D}^*$, *such that* $P^*(\omega^*) = \times_{i\in Q}\pi^i(\mathbb{P}^{\times\{D_i\}}(\omega^*))$.

Cylinder factual partition strengthens product factual partition by requiring the decomposition of distinct signals for each uncertainty be independent of the underlying full states, ruling out the correlation between answers to one question and factual signals about other questions.

Definition 9 *The awareness function* $W^* : \Omega^* \to 2^{\mathcal{D}^*} \setminus \{\emptyset\}$ *satisfies* **nice awareness** *if, for any* $\omega_1^*, \omega_2^* \in \Omega^*$,

$$[W^*(\omega_1^*) \triangle W^*(\omega_2^*)] \cap \left\{D \in \mathcal{D}^* : \mathbb{P}^{\times\{D\}}(\omega_1^*) = \mathbb{P}^{\times\{D\}}(\omega_2^*)\right\} = \emptyset$$

where $\triangle$ *denotes symmetric difference.*

In words, nice awareness requires that, if two full states coincide on the answers to a particular question, then either the agent is aware of the question in both states, or he is unaware of it in both states. This condition rules out the correlation between awareness of different questions, and hence rules out "unawareness of unawareness" in multi-agent environment.

Definition 10 $(\mathbf{W}^*, \mathbf{P}^*)$ *is* **strongly rational** *if it satisfies cylinder factual partition, rational awareness and nice awareness.*

Theorem 9 *Let* $(\Omega^*, \mathbf{W}^*, \mathbf{P}^*)$ *be rich and strongly rational. Then the formula (5.14) characterizes common knowledge, i.e. for all* $E \in \mathcal{E}^p$,

$$\underline{CK}(E) = CK(E)$$

6 Concluding remarks

In this paper, I construct a set-theoretic model of both single-agent and interactive information processing with unawareness. The main idea is that introducing unawareness necessitates examining a new type of information, i.e. awareness information, *in addition to* the usual factual information in the standard model. The resulting model is a natural generalization of the standard information partition model due to Aumann (1976), and hence well connecting to the existing literature. The model also sheds light on the formal distinction between unawareness and probability zero events: if the agent is unaware of an event, then the event is beyond the agent's probability space and he is unable to assign any probability to it. It follows that unawareness introduces real dynamics to the model, and that one should expect different behavior under unawareness as opposed to assigning probability zero in dynamic environments.

There are a number of exciting topics one can explore using this model. The anticipation of unforeseen contingencies has a significant impact in real-life decision processes. Many people keep some personal funds for unspecified emergencies. At a collective level, the scale can be quite impressive. For instance, "the City of New York's Five-Year Financial Plan" includes a "general reserve for *unforeseen contingencies* of \$42 million in FY 2001 and reserves of \$545 million in FY 2002."[23] However, decision-making under unforeseen contingencies is not well understood. The standard Savage framework assumes away unforeseen contingencies. Research in this field has been focusing on axiomatization of preferences over menus of items, which provides the dynamic structure intrinsically related to unforeseen contingencies. This model is suggestive on how a more direct approach exploring the generalized information structure might work. Contractual incompleteness is a particularly interesting and important economic phenomenon where anticipation of unforeseen contingencies seems to play an important role. It is not clear how to apply research in decision-making with unforeseen contingencies to the contractual environment without an explicit account for contractual parties' information structures. This model provides a simple tool for that purpose.

7 Appendix.

7.1 Proof of Lemma 3.

U0* *Symmetry*: $U(E) = U(\neg E)$

Follows from $\mathcal{D}_E = \mathcal{D}_{\neg E}$.

U1′ *Strong plausibility*: $U(E) \subseteq \bigcap_{n=1}^{\infty} (\neg K)^n(E)$

Let $\omega^* \in U(E)$. Then $\mathcal{D}_E \nsubseteq W(\omega^*)$. By 3.3, $\tilde{K}_{\omega^*}(E) = \emptyset_E$. By 3.5, $\omega^* \in \neg K(E)$. Note that $\mathcal{D}_{\neg\tilde{K}_{\omega^*}(E)} = \mathcal{D}_{\tilde{K}_{\omega^*}(E)} = \mathcal{D}_{\emptyset_E} = \mathcal{D}_E \nsubseteq W(\omega^*)$, now it follows $\tilde{K}_{\omega^*}\neg\tilde{K}_{\omega^*}(E) = \emptyset_E$ and $\omega^* \in \bar{K}_2(E)$. It is easy to see that $\mathcal{D}_{\neg\tilde{K}^{n-1}_{\omega^*}(E)} = \mathcal{D}_E$ for all n, which implies $\tilde{K}_{\omega^*}(\neg\tilde{K}^{n-1}_{\omega^*}(E)) = \emptyset_E$ for all n, and hence Thus, $\omega^* \in (\neg K)^n(E)$ for all n.

[23]The emphasis is mine. Source: "Review of the Mayor's Executive Budget for Fiscal Year 2002" by H. Carl McCall (State Comptroller), web address: http://www.osc.state.ny.us/osdc/rpt2002.pdf

U2**AU introspection*: $U(E) \subseteq (\neg K)^n U(E)$

Extending (5.7) to the general case in the obvious way, we have:

$$(\neg K)^n U(E) = \left\{\omega^* \in \Omega^* : s(\omega^*) \in (\neg \tilde{K}_{\omega^*})^n \tilde{U}_{\omega^*}(E)\right\}$$

where $\tilde{U}_{\omega^*}(E) = \{\omega \in \Omega(\omega^*) : \mathcal{D}_E \nsubseteq W_{\omega^*}(\omega)\}$ if $\mathcal{D}_E \subseteq W^*(\omega^*)$ and equals $\emptyset_E$ otherwise. This is interpreted as the subjective event "I am unaware of E" in the agent's subjective model at ω^*. Thus the subjective event $\tilde{U}_{\omega^*}(E)$ contains at least as much awareness information as event E itself: $\mathcal{D}_{\tilde{U}_{\omega^*}(E)} \supseteq \mathcal{D}_E$.

Now if $\omega^* \in U(E)$, then $\mathcal{D}_E \nsubseteq W^*(\omega^*)$, but then $\mathcal{D}_{\tilde{U}_{\omega^*}(E)} \nsubseteq W^*(\omega^*)$. But (3.3), $s(\omega^*) \in (\neg \tilde{K}_{\omega^*})^n \tilde{U}_{\omega^*}(E)$ for all n, and hence $\omega^* \in (\neg K)^n U(E)$.

U3$'$ *Weak KU introspection*: $U(E) \cap KU(E) = \emptyset_{\Omega^*}$

Observe $\neg KU(E) = \Omega^* \setminus KU(E)$ by (5.7), and hence the result follows from AU introspection.

K1**Subjective necessitation*: for all $\omega^* \in \Omega^*$, $\omega^* \in K(\Omega(\omega^*))$

For any $\omega^* \in \Omega^*$, $P_{\omega^*}(s(\omega^*)) = \mathbb{P}^{\Omega(\omega^*)}(P^*(\omega^*)) \subseteq \Omega(\omega^*)$, which implies $s(\omega^*) \in \tilde{K}_{\omega^*}(\Omega(\omega^*))$, and hence $\omega^* \in K(\Omega(\omega^*))$.

K2**Generalized monotonicity*: $E \subseteq_* F$, $\mathcal{D}_F \subseteq \mathcal{D}_E \Rightarrow K(E) \subseteq K(F)$

This is implied by conjunction. Take E and F such that $E \subseteq_* F$, $\mathcal{D}_F \subseteq \mathcal{D}_E$. By conjunction,

$$\begin{aligned} K(E) \cap K(F) &= K(E \cap_* F) \\ &= K(E \cap F_{\times \mathcal{D}_E}) \\ &= K(E) \end{aligned}$$

It follows $K(E) \subseteq K(F)$.

K3**Conjunction*: $K(E) \cap K(F) = K(E \cap_* F)$

Let $\omega^* \in K(E) \cap K(F)$. Then $\mathcal{D}_E \subseteq W^*(\omega^*)$, $\mathcal{D}_F \subseteq W^*(\omega^*)$ and $P^*(\omega^*) \subseteq E_{\Omega^*}$, $P^*(\omega^*) \subseteq F_{\Omega^*}$. Note that:

(1) $\mathcal{D}_E \subseteq W^*(\omega^*)$, $\mathcal{D}_F \subseteq W^*(\omega^*)$ if and only if $\mathcal{D}_E \cup \mathcal{D}_F \subseteq W^*(\omega^*)$, which implies $\mathcal{D}_{E \cap_* F} \subseteq W^*(\omega^*)$. Thus, the event $E \cap_* F$ has an elaboration in the space $\Omega(\omega^*)$.

(2) $P^*(\omega^*) \subseteq E_{\Omega^*}$, $P^*(\omega^*) \subseteq F_{\Omega^*}$ if and only if $P^*(\omega^*) \subseteq (E_{\Omega^*} \cap F_{\Omega^*})$;

Using the product structure, one has $E_{\Omega^*} = E_{\times(\mathcal{D}_E \cup \mathcal{D}_F)} \times (\mathcal{D} \setminus (\mathcal{D}_E \cup \mathcal{D}_F))$, $F_{\Omega^*} = F_{\times(\mathcal{D}_E \cup \mathcal{D}_F)} \times (\mathcal{D} \setminus (\mathcal{D}_E \cup \mathcal{D}_F))$. Therefore

$$\begin{aligned} E_{\Omega^*} \cap F_{\Omega^*} &= [E_{\times(\mathcal{D}_E \cup \mathcal{D}_F)} \times (\mathcal{D} \setminus (\mathcal{D}_E \cup \mathcal{D}_F))] \cap [F_{\times(\mathcal{D}_E \cup \mathcal{D}_F)} \times (\mathcal{D} \setminus (\mathcal{D}_E \cup \mathcal{D}_F))] \\ &= [E_{\times(\mathcal{D}_E \cup \mathcal{D}_F)} \cap F_{\times(\mathcal{D}_E \cup \mathcal{D}_F)}] \times (\mathcal{D} \setminus (\mathcal{D}_E \cup \mathcal{D}_F)) \\ &= (E \cap_* F)_{\Omega^*} \end{aligned}$$

Back to (2), one has $P^*(\omega^*) \subseteq (E \cap_* F)_{\Omega^*}$. Using (1), $P_{\omega^*}(s(\omega^*)) = \mathbb{P}^{\Omega(\omega^*)} P^*(\omega^*) \subseteq \mathbb{P}^{\Omega(\omega^*)}((E \cap_* F)_{\Omega^*}) = (E \cap_* F)_{\Omega(\omega^*)}$, hence $s(\omega^*) \in \tilde{K}_{\omega^*}((E \cap_* F)_{\Omega(\omega^*)})$, and hence $\omega^* \in K(E \cap_* F)$ □

7.2 Proof of Lemma 4.

U1* *UUU*: $U(E) = \bigcap_{n=1}^{\infty} (\neg K)^n(E)$

Strong plausibility gives ⇒; Applying De Morgan's law on the axiom of limited wisdom gives the other direction.

U3* *KU introspection*: $KU(E) = \emptyset_{\Omega^*}$

Suppose not. Let $\omega^* \in KU(E)$. By (5.7) and (3.3), this implies:
(1): $\mathcal{D}_{\tilde{U}_{\omega^*}(E)} \subseteq W^*(\omega^*)$; and
(2) $P_{\omega^*}(s(\omega^*)) \subseteq \tilde{U}_{\omega^*}(E)$.
By (1), $\mathcal{D}_E \subseteq W^*(\omega^*)$. By (5.1), $W_{\omega^*}(s(\omega^*)) = W^*(\omega^*)$, hence $\mathcal{D}_E \subseteq W_{\omega^*}(s(\omega^*))$, thus $s(\omega^*) \notin \tilde{U}_{\omega^*}(E)$, which obviously contradicts (2), because $s(\omega^*) \in P_{\omega^*}(s(\omega^*))$ by (3.2) and that $\omega^* \in P^*(\omega^*)$ (rational information).

Observe that when (W^*, P^*) is rational, the subjective model is a standard information partition model at the set of possible subjective states, and hence the standard results carry through.

K4.a^*, K4.b^* *The axiom of knowledge*: $K(E) \subseteq_* E$, $K^n(E) \subseteq K^{n-1}(E)$

Let $\omega^* \in K(E)$. By (3.6), $P^*(\omega^*) \subseteq E_{\Omega^*}$, but then since P^* induces an information partition over Ω^*, $\omega^* \in P^*(\omega^*)$; it follows that $\omega^* \in E_{\Omega^*}$, and hence $K(E) \subseteq E_{\Omega^*}$;

To see K4.b^*, let $\omega^* \in K^n(E)$. By (3.5), $s(\omega^*) \in \tilde{K}^n_{\omega^*}(E)$; by (3.4), $P_{\omega^*}(s(\omega^*)) \subseteq K^{n-1}_{\omega^*}(E)$. But then again $\omega^* \in P^*(\omega^*)$ implies $s(\omega^*) \in P_{\omega^*}(s(\omega^*)) \Rightarrow s(\omega^*) \in K^{n-1}_{\omega^*}(E)$, and hence $\omega^* \in K^{n-1}(E)$.

K5* *The axiom of transparency*: $K(E) \subseteq KK(E)$

Let $\omega^* \in K(E)$. It suffices to show $s(\omega^*) \in \tilde{K}_{\omega^*}\tilde{K}_{\omega^*}(E)$.

Since (W^*, P^*) is rational, $P_{\omega^*}(\omega) = P_{\omega^*}(s(\omega^*))$ for all $\omega \in P_{\omega^*}(s(\omega^*))$. Now

$$\begin{aligned} & s(\omega^*) \in \tilde{K}_{\omega^*}(E) \\ \Rightarrow\ & P_{\omega^*}(s(\omega^*)) \subseteq E_{\Omega(\omega^*)} \\ \Rightarrow\ & P_{\omega^*}(\omega) \subseteq E_{\Omega(\omega^*)} \\ \Rightarrow\ & \omega \in \tilde{K}_{\omega^*}(E) \text{ for all } \omega \in P_{\omega^*}(s(\omega^*)) \\ \Rightarrow\ & P_{\omega^*}(s(\omega^*)) \subseteq \tilde{K}_{\omega^*}(E) \\ \Rightarrow\ & s(\omega^*) \in \tilde{K}^2_{\omega^*}(E) \end{aligned}$$

K6* *The axiom of limited wisdom*: $\neg U(E) \cap \neg K(E) \subseteq K\neg K(E)$

Let $\omega^* \in \neg U(E) \cap \neg K(E)$. Then $\mathcal{D}_E \subseteq W^*(\omega^*)$, so we only need to show $P_{\omega^*}(s(\omega^*)) \subseteq \neg\tilde{K}_{\omega^*}(E)$. Let $\omega \in P_{\omega^*}(s(\omega^*))$. By (3.2), $P_{\omega^*}(\omega) = P_{\omega^*}(s(\omega^*))$, but $\omega^* \in \neg K(E) \Rightarrow P_{\omega^*}(s(\omega^*)) \nsubseteq E_{\Omega(\omega^*)}$, it follows $P_{\omega^*}(\omega) \nsubseteq E_{\Omega(\omega^*)} \Rightarrow \omega \in \neg\tilde{K}_{\omega^*}(E)$. □

7.3 Proof of Proposition 5.

Lemma 10 *Let $(\Omega^*, \mathbf{W}^*, \mathbf{P}^*)$ be rich, and let (W_i^*, P_i^*) satisfy rational awareness for all $i \in N$. Then for any n and any $q^n \in \Delta^n$, $(W_{i^{n+1}}(\cdot|q^n), P_{i^{n+1}}(\cdot|q^n))$ satisfies rational awareness.*

Proof. Given the recursive construction of higher-order subjective models, it suffices to prove the result for the case of $n = 1$. For notational ease, I write the first-order interactive model as $(\Omega_i(\omega^*), W_j(\cdot|i_{\omega^*}), P_j(\cdot|i_{\omega^*}))$ where $\omega^* \in \Omega^*$ is arbitrary. The goal is to show: for all $\omega \in \Omega_i(\omega^*), \omega' \in P_j(\omega|i_{\omega^*}) \Rightarrow W_j(\omega|i_{\omega^*}) = W_j(\omega'|i_{\omega^*})$.

By (5.2), $\omega' \in P_j(\omega|i_{\omega^*})$ implies that there exist two full states ω_1^*, ω_2^* such that $\mathbb{P}^{\Omega_i(\omega^*)}(\omega_1^*) = \omega$, $\mathbb{P}^{\Omega_i(\omega^*)}(\omega_2^*) = \omega'$, and that $\omega_2^* \in P_j^*(\omega_1^*)$. By hypothesis, (W_j^*, P_j^*) satisfies rational awareness, i.e. $W_j^*(\omega_1^*) = W_j^*(\omega_2^*)$.

Let $D \in W_j(\omega|i_{\omega^*}) = \bigcup_{\{\omega_0^*: \mathbb{P}^{\Omega_i(\omega^*)}(\omega_0^*) = \omega\}} W_i^*(\omega^*) \cap W_j^*(\omega_0^*)$. There are two cases.

1. $D \in W_j^*(\omega_1^*)$. Then $D \in W_j^*(\omega_2^*)$, and hence $D \in W_j(\omega'|i_{\omega^*})$ by (5.1);

2. $D \in W_j^*(\omega_3^*) \setminus W_j^*(\omega_2^*)$. Since $\mathbb{P}^{\Omega_i(\omega^*)}(\omega_1^*) = \mathbb{P}^{\Omega_i(\omega^*)}(\omega_3^*) = \omega$, by richness, there exists ω_4^* such that $\mathbb{P}^{\Omega_i(\omega^*)}(\omega_4^*) = \mathbb{P}^{\Omega_i(\omega^*)}(\omega_2^*) = \omega'$ and $W_j^*(\omega_3^*) = W_j^*(\omega_4^*)$. It follows $D \in W_j(\omega''|i_{\omega^*})$.

This proves $W_j(\omega|i_{\omega^*}) \subseteq W_j(\omega'|i_{\omega^*})$. The other direction is completely symmetric.

□

Proof for Proposition 5: It suffices to show

$$(\Omega_{i^{k-1}}(\omega_{k-1}|q^{k-2}), W_{i^{k+1}}(\cdot|q^{k-1}), P_{i^{k+1}}(\cdot|q^{k-1})) = (\Omega_{i^k}(\omega_k|q^{k-1}), W_{i^{k+1}}(\cdot|q^k), P_{i^{k+1}}(\cdot|q^k))$$

Since $\omega_k \in P_{i^k}(s(\omega_{k-1}|q^{k-1})|q^{k-1})$, by lemma (10), $W_{i^k}(\omega_k|q^{k-1}) = W_{i^k}(s(\omega_{k-1}|q^{k-1})|q^{k-1})$.

Let $G^{(k-1):(k-2)}(\omega) = \left\{\omega' \in \Omega_{i^{k-2}}(\omega_{k-2}|q^{k-3}) : \mathbb{P}^{\Omega_{i^{k-1}}(\omega_{k-1}|q^{k-2})}(\omega') = s(\omega_{k-1}|q^{k-1})\right\}$ denote the set of subjective states in i^{k-2}'s mind-set at ω_{k-2} ascribed by i^{k-3} at ω_{k-3} ascribed by $\cdots$ by i^1 at ω_1, that correspond to the subjective state ω in i^{k-1}'s mind-set by i^{k-2} at ω_{k-2} ascribed by $\cdots$ by i^1 at ω_1. By (5.4),

$$W_{i^k}(s(\omega_{k-1}|q^{k-1})|q^{k-1}) = W_{i^{k-1}}(\omega_{k-1}|q^{k-2}) \cap [\bigcup_{\omega' \in G^{(k-1):(k-2)}(s(\omega_{k-1}|q^{k-1}))} W_{i^k}(\omega'|q^{k-2})]$$

Since $q^n \in \Delta^n$, $\omega_{k-1} \in \Omega_{i^{k-2}}(\omega_{k-2}|q^{k-3})$. Notice $s(\omega_{k-1}|q^{k-1}) \in G^{(k-1):(k-2)}(s(\omega_{k-1}|q^{k-1}))$, therefore, we have:

$$W_{i^k}(\omega_{k-1}|q^{k-2}) \subseteq \bigcup_{\omega' \in G^{(k-1):(k-2)}(s(\omega_{k-1}|q^{k-1}))} W_{i^k}(\omega'|q^{k-2}) \tag{7.1}$$

Now $i^k = i^{k-1}$, thus $W_{i^k}(s(\omega_{k-1}|q^{k-1})|q^{k-1}) = W_{i^{k-1}}(\omega_{k-1}|q^{k-2})$, and hence:

$$\Omega_{i^k}(\omega_k|q^{k-1}) = \Omega_{i^{k-1}}(\omega_{k-1}|q^{k-2})$$

Let $\omega \in \Omega_{i^k}(\omega_k|q^{k-1})$. Now $G^{k:(k-1)}(\omega) = \left\{\omega' \in \Omega_{i^{k-1}}(\omega_{k-1}|q^{k-2}) : \mathbb{P}^{\Omega_{i^k}(\omega_k|q^{k-1})}(\omega') = \omega\right\}$, but since $\Omega_{i^k}(\omega_k|q^{k-1}) = \Omega_{i^{k-1}}(\omega_{k-1}|q^{k-2})$, $G^{k:(k-1)}(\omega) = \omega$.

Therefore, by (5.4),

$$\begin{aligned} W_{i^{k+1}}(\omega|q^k) &= W_{i^k}(\omega_k|q^{k-1}) \cap [\bigcup_{\omega' \in G^{(k):(k-1)}(\omega)} W_{i^{k+1}}(\omega'|q^{k-1})] \\ &= W_{i^{k-1}}(\omega_{k-1}|q^{k-2}) \cap W_{i^{k+1}}(\omega|q^{k-1}) \\ &\overset{\dagger}{=} W_{i^{k+1}}(\omega|q^{k-1}) \end{aligned}$$

where † follows from (7.1) and that, by (5.4),

$$W_{i^{k+1}}(\omega|q^{k-1}) = W_{i^{k-1}}(\omega_{k-1}|q^{k-2}) \cap [\bigcup_{\omega' \in G^{(k-1):(k-2)}(\omega)} W_{i^{k+1}}(\omega'|q^{k-2})]$$

Next consider $P_{i^{k+1}}(\cdot|q^k))$. By (5.5),

$$P_{i^{k+1}}(\omega|q^k) = \begin{cases} \mathbb{P}^{\Omega_{i^k}(\omega_k|q^{k-1})} P_{i^{k+1}}(\omega_k|q^{k-1}) & \text{for } \omega \in \mathbb{P}^{\Omega_{i^k}(\omega_k|q^{k-1})} P_{i^{k+1}}(\omega_k|q^{k-1}) \\ \mathbb{P}^{\Omega_{i^k}(\omega_k|q^{n-1})} P_{i^{k+1}}(\omega \times u_{i^k}(\omega_k|q^{k-1})|q^{k-1}) & \text{otherwise.} \end{cases}$$

Now note since $P_{i^{k+1}}(\cdot|q^{k-1})$ is defined on the interactive state space $\Omega_{i^{k-1}}(\omega_{k-1}|q^{k-2})$, but $\Omega_{i^{k-1}}(\omega_{k-1}|q^{k-2}) = \Omega_{i^k}(\omega_k|q^{k-1})$, this equation simply reduces to $P_{i^{k+1}}(\omega|q^k) = P_{i^{k+1}}(\omega|q^{k-1})$ which concludes the proof. □

7.4 Proof of Lemma 6.

The following intermediate result will be useful later.

Lemma 11 *For all $i, j \in N$, if P_j^* satisfies nice factual partition then, for any $\omega^* \in \Omega^*$ and all $\omega \in \Omega_i(\omega^*)$,*

$$P_j(\omega|i_{\omega^*}) = \mathbb{P}^{\Omega_i(\omega^*)} P_j^*(\omega \times u_i(\omega^*))$$

Proof. Let $\omega \in \mathbb{P}^{\Omega_i(\omega^*)} P_j^*(\omega^*)$. Only need to show $P_j^*(\omega \times u_i(\omega^*)) = P_j^*(\omega^*)$.

Since P_j^* induces an information partition, $\omega^* \in P_j^*(\omega^*)$. By definition of $u_i(\omega^*)$, $\omega^* = s(\omega^*|i_{\omega^*}) \times u_i(\omega^*)$. On the other hand, since P_j^* satisfies nice factual partition, it is a product set. It follows $\omega \times u_i(\omega^*) \in P_j^*(\omega^*) \Rightarrow P_j^*(\omega \times u_i(\omega^*)) = P_j^*(\omega^*)$. □

Now we are ready to prove Lemma 6:

Proof. By Lemma 10, $(W_{i^{n+1}}(\cdot|q^n), P_{i^{n+1}}(\cdot|q^n))$ satisfies rational awareness.

To see that $P_{i^{n+1}}(\cdot|q^n)$ induces an information partition over $\Omega_{i^n}(\omega_n|q^{n-1})$ for all q^n, it suffices to show $P_{i^{n+1}}(\cdot|q^n)$ satisfies nice factual partition for the case of $n = 1$. But this follows trivially from Lemma 11: the projection of an information partition is obviously an information partition of the corresponding state space; and projection preserves the product structure. □

7.5 Proof of Theorem 7.

Proof for IK1. First notice that for any ω^* and any I^n, $s(s(\omega^*|I^{n-2}(\omega^*))|I^{n-1}(\omega^*)) = s(\omega^*|I^{n-1}(\omega^*))$. By Lemma 6, the hypotheses for Proposition 5 are satisfied (take $\omega_k = s(\omega^*|I^{k-1}(\omega^*))$ and $\omega_{k-1} = s(\omega^*|I^{k-2}(\omega^*))$). For notational ease, I use Ω^k as the shorthand for $\Omega_{i^k}(s(\omega^*|I^{k-1}(\omega^*))|I^{k-1}(\omega^*))$. Thus for all $k < m < n$,

$$(\Omega^m, W_{i^{m+1}}(\cdot|I^m(\omega^*)), P_{i^{m+1}}(\cdot|I^m(\omega^*)))$$

$$=$$

$$(\Omega_{i^m}(s(\omega^*|I^{(m-1)\backslash k}(\omega^*))|I^{(m-1)\backslash k}(\omega^*)), W_{i^{m+1}}(\cdot|I^{m\backslash k}(\omega^*)), P_{i^{m+1}}(\cdot|I^{m\backslash k}(\omega^*)))$$

It follows $\tilde{K}^{i^m}(E|I^{m-1}(\omega^*)) = \tilde{K}^{i^m}(E|I^{(m-1)\backslash k}(\omega^*))$ for all m such that $k < m \leq n$. Thus, it suffices to show $\tilde{K}^{i^{k-1}}(E|I^{k-2}(\omega^*)) = \tilde{K}^{i^{k-1}}(\tilde{K}^{i^k}(E|I^{k-1}(\omega^*))|I^{k-2}(\omega^*))$.

For notational ease, let $i^k = i^{k-1} = j$, $E_{\Omega^{k-2}} = E_{k-2}$. Now by (5.8),

$$\tilde{K}^j(E|I^{k-2}(\omega^*)) = \left\{\omega \in \Omega^{k-2} : P_j(\cdot|I^{k-2}(\omega^*)) \subseteq E_{n-2}, W_j(\cdot|I^{k-2}(\omega^*)) \supseteq \mathcal{D}_E\right\}$$

This is the objective knowledge operator associated with the product model

$$(\Omega_{k-2}, W_j(\cdot|I^{k-2}(\omega^*)), P_j(\cdot|I^{k-2}(\omega^*)))$$

But by Lemma 6, $(W_j(\cdot|I^{k-2}(\omega^*)), P_j(\cdot|I^{k-2}(\omega^*)))$ is rational, and hence by Theorem 2, the iteration yields the same event:

$$\begin{aligned}\tilde{K}^j(E|I^{k-2}(\omega^*)) &= \tilde{K}^j(\tilde{K}^j(E|I^{k-2}(\omega^*))|I^{k-2}(\omega^*)) \\ &= \tilde{K}^j(E|I^{k-2}(\omega^*) \wedge (s(s(\omega^*|I^{k-2}(\omega^*))|I^{k-1}(\omega^*)), j)) \\ &= \tilde{K}^j(E|I^{k-2}(\omega^*) \wedge (s(\omega^*|I^{k-1}(\omega^*)), j)) \\ &= \tilde{K}^j(E|I^{k-1}(\omega^*))\end{aligned}$$

Use Theorem 2 again,

$$\begin{aligned}\tilde{K}^{i^{k-1}}(\tilde{K}^{i^k}(E|I^{k-1}(\omega^*))|I^{k-2}(\omega^*)) &= \tilde{K}^{i^{k-1}}(\tilde{K}^{i^k}(E|I^{k-2}(\omega^*))|I^{k-2}(\omega^*)) \\ &= \tilde{K}^{i^{k-1}}(E|I^{k-2}(\omega^*))\end{aligned}$$

□

Proof for IK2. For every ω^*, apply lemma 6 and Lemma 4 to the subjective model $(\Omega_i(\omega^*), W_j(\cdot|i_{\omega^*}), P_j(\cdot|i_{\omega^*}))$ to get $\tilde{U}_j(E|i_{\omega^*}) = \neg\tilde{K}^j(\cdots\tilde{(}K^j(E|I^{n-1}_{\omega^*}))\cdots|i_{\omega^*})$ for all n.□

Proof for IK3. Let $\omega^* \in K_iK_j(E)$. Then $s(\omega^*|i_{\omega^*}) \in \tilde{K}^i_{\omega^*}\tilde{K}^j(E|i_{\omega^*})$, and hence $s(\omega^*|i_{\omega^*}) \in \tilde{K}^j(E|i_{\omega^*})$. Thus the following holds:

$$P_j(s(\omega^*|i_{\omega^*})|i_{\omega^*}) \subseteq E_{\Omega_i(\omega^*)} \quad (7.2)$$
$$W_j(s(\omega^*|i_{\omega^*})|i_{\omega^*}) \supseteq \mathcal{D}_E \quad (7.3)$$

(7.3) implies $s(\omega^*|i_{\omega^*}) \in \neg\tilde{U}_j(E|i_{\omega^*}$. By (5.7), $\omega^* \in K_i\neg U_j(E)$;

By (5.2), $P_j(s(\omega^*|i_{\omega^*})|i_{\omega^*}) = \mathbb{P}^{\Omega_i(\omega^*)}P^*_j(\omega^*)$, and hence (7.2) implies $P^*_j(\omega^*) \subseteq E_{\Omega^*}$. Then if $\mathcal{D}_E \subseteq W^*_j(\omega^*)$ then $\omega^* \in K_j(E)$; if $\mathcal{D}_E \nsubseteq W^*_j(\omega^*)$ then $\omega^* \in U_j(E)$. □

7.6 Proof of Theorem 8.

Definition 11 *For any $\omega^* \in \Omega^*$, any n, any $i \in N$ and any $I^n_i = (i, i^2, \cdots, i^n)$, the sequence $r^n_i(\omega^*) = ((\omega^*, i), (\omega_2, i^2), \cdots, (\omega_n, i^n)) \in \Delta^n$ is* **relevant** *for I^n_i in ω^* if it satisfies: $\omega_2 \in P_i(s(\omega^*|i_{\omega^*})|i_{\omega^*})$, $\omega_k \in P_{i^{k-1}}(s(\omega_{k-1}|q^{k-2}_i)|q^{k-2}_i)$ for all $2 < k \leq n$, and $i, i^2, \cdots, i^k$ correspond to those in I^k_i.*

In words, $r^n_i(\omega^*)$ is a reasoning string where every player in the sequence I^n_i reasons about other's reasoning at a subjective state he himself considers possible. It is easy to see that if every player has full awareness, then the set of states involved in relevant sequences is just $\wedge^n_{j=1}P^*_j(\omega^*)$. With nontrivial unawareness, these are subjective states "reachable" from ω^* *subject to awareness constraints.*

Slightly abusing notation, let $\Omega^n_r = \Omega_{i^n}(s(\omega_n|r^{n-1}_i(\omega^*))|r^{n-1}_i(\omega^*))$ denote the subjective state space of i^n at subjective state $s(\omega_n|r^{n-1}_i(\omega^*))$ ascribed to him under $r^{n-1}_i(\omega^*)$.

The subscript r emphasizes this is a "relevant" subjective state space. Let $R(\omega|r_i^{n-1}(\omega^*))$ denote the set of reachable states from ω in the standard model $(\Omega_r^{n-1}, P_{i^n}(\cdot|r_i^{n-1}(\omega^*)))$, that is,

$$R(\omega|r_i^{n-1}(\omega^*)) = \wedge_{j=1}^{n} P_j(\omega|r_i^{n-1}(\omega^*))$$

Lemma 12 *Let $(\Omega^*, \mathbf{W}^*, \mathbf{P}^*)$ be rich and interactively rational. Let $E \in \mathcal{E}^p$ and $\omega^* \in \underline{CK}(E)$. Then, for any n, any $i \in N$ and any I_i^n, the following are true for all relevant $r_i^n(\omega^*)$ for I_i^n in ω^*, all j and all $\omega' \in R(s(\omega^*|r_i^{n-1}(\omega^*))|r_i^{n-1}(\omega^*))$,*

$$R(s(\omega^*|r_i^{n-1}(\omega^*))|r_i^{n-1}(\omega^*)) \subseteq E_{\Omega_r^{n-1}} \tag{7.4}$$

$$W_j(\omega'|r_i^{n-1}(\omega^*)) \supseteq \mathcal{D}_E \tag{7.5}$$

Intuitively, since the full possibility sets are product sets, the corresponding subjective possibility sets are product sets as well, and are obtained by projecting the full possibility sets onto the corresponding interactive state spaces. Since projection preserves set inclusion, (7.4) follows.

Proof. For $n = 2$:
Since $W_i^*(\omega^*) \supseteq \mathcal{D}_E$, $E_{\Omega_i(\omega^*)}$ is well-defined;
For all j, $P_j(s(\omega^*|i_{\omega^*})|i_{\omega^*}) = \mathbb{P}^{\Omega_i(\omega^*)} P_j^*(\omega^*)$, thus,

$$\begin{aligned} R(s(\omega^*|i_{\omega^*})|i_\omega^*) &= \wedge_{j=1}^{n} \mathbb{P}^{\Omega_i(\omega^*)} P_j^*(\omega^*) \\ &\subseteq \mathbb{P}^{\Omega_i(\omega^*)} \wedge_{j=1}^{n} P_j^*(\omega^*) \\ &\subseteq E_{\Omega_i(\omega^*)} \end{aligned}$$

Let $\omega' \in R(s(\omega^*|i_{\omega^*})|i_\omega^*)$. Since $R(s(\omega^*|i_{\omega^*})|i_\omega^*) = \mathbb{P}^{\Omega_i(\omega^*)} \wedge_{j=1}^{n} P_j^*(\omega^*)$, there exists $\omega_1^* \in \wedge_{j=1}^{n} P_j^*(\omega^*)$ such that $\mathbb{P}^{\Omega_i(\omega^*)}(\omega_1^*) = \omega'$. By hypothesis, $\mathcal{D}_E \subseteq W_j^*(\omega_1^*)$ for all j. Now,

$$\begin{aligned} W_j(\omega'|i_{\omega^*}) &= \bigcup_{\left\{\omega_2^*:\mathbb{P}^{\Omega_i(\omega^*)}(\omega_2^*)=\omega'\right\}} W_i^*(\omega^*) \cap W_j^*(\omega_2^*) \\ &\supseteq W_i^*(\omega^*) \cap W_j^*(\omega_1^*) \\ &\supseteq \mathcal{D}_E \end{aligned}$$

Suppose (7.4) and (7.5) hold for all $1 < n \leq k$. Observe that in the case of $n = k+1$, the interactive model $(\Omega_r^{n-1}, \mathbf{W}(\cdot|r_i^{n-1}(\omega^*)), \mathbf{P}(\cdot|r_i^{n-1}(\omega^*)))$ plays the role of the full model $(\Omega^*, \mathbf{W}, \mathbf{P})$ in the case of $n = 2$. Now observe this model is a multi-agent product model itself, and it satisfies all the properties of the full model by previous results (Lemma 11 and Lemma 6). Therefore the conclusion follows from the induction hypothesis. □

Now we are ready to prove the main theorem.

Proof. The second set inclusion is obvious.

To prove the first set inclusion, let $\omega^* \in \underline{CK}(E)$. By Lemma 12, in every interactive model at ω^*, the subjective factual signals satisfy the classic requirement that the meet is contained in the corresponding elaboration of E as in Aumann (1976). Therefore, the complication is to ensure in all relevant subjective states, every agent is aware of E, and there is higher-order knowledge of that.

It suffices to show that for all $i \in N$, all n and all $I_i^n = (i, i^2, \cdots, i^n)$,

$$\omega^* \in K_i K_{i^2} \cdots K_{i^n}(E) \tag{7.6}$$

By definition (5.10), (7.6) is equivalent to:

$$s(\omega^*|i_{\omega^*}) \in \tilde{K}^i_{\omega^*}(\tilde{K}^{i^2}(\cdots(\tilde{K}^{i^k}(E|I^{k-1}(\omega^*)))\cdots|i_{\omega^*}))$$

By definition (3.3), this is equivalent to:

$$\mathbb{P}^{\Omega_i(\omega^*)}P_i^*(\omega^*) \subseteq \tilde{K}^{i^2}(\cdots(\tilde{K}^{i^k}(E|I^{k-1}(\omega^*)))\cdots|i_{\omega^*})$$

By definition (5.8), the above amounts to: for all $\omega_2 \in P_i(s(\omega^*|i_{\omega^*})|i_{\omega^*})$,

$$P_{i^2}(\omega_2|i_{\omega^*}) \subseteq_* \tilde{K}^{i^3}(\cdots(\tilde{K}^{i^k}(E|I_i^{k-1}(\omega^*)))\cdots|I_i^2(\omega^*)) \tag{7.7}$$

$$W_{i^2}(\omega_2|i_{\omega^*}) \supseteq \mathcal{D}_E \tag{7.8}$$

Apply definition (5.8) on (7.7) recursively. It follows that (7.7) and (7.8) are equivalent to: for any relevant $r_i^{k-1}(\omega^*)$ under I_i^n and ω^*, and any $1 < h \le k$,

$$P_{i^k}(s(\omega_k|r_i^{k-1}(\omega^*))|r_i^{k-1}(\omega^*)) \subseteq_* E_{\Omega_r^{k-1}} \tag{7.9}$$

$$W_{i^h}(s(\omega_h|r_i^{h-1}(\omega^*))|r_i^{h-1}(\omega^*)) \supseteq \mathcal{D}_E \tag{7.10}$$

which follow from Lemma 12. □

7.7 Proof of Theorem 9.

Proof. Only need to prove necessity, i.e. $CK(E) \subseteq \underline{CK}(E)$. I show both clauses in the definition of $\underline{CK}(E)$ are necessary.

Let $\omega^* \notin \underline{CK}(E)$.

Case 1: Suppose there exists $\bar{\omega}^* \in \wedge_{j=1}^n P_j^*(\omega^*)$ such that $\bar{\omega}^* \notin E_{\Omega^*}$. Let $i \in N$. Suppose $\bar{\omega}^*$ is reachable from ω^* for i through the factual information sets of $i^2, \cdots, i^k$, that is,

$$\begin{aligned}
\bar{\omega}^* &\in P_{i^k}^*(\omega_k^*) \\
\omega_k^* &\in P_{i^{k-1}}^*(\omega_{k-1}^*) \\
&\cdots \\
\omega_2^* &\in P_i^*(\omega^*)
\end{aligned}$$

Since factual signals are cylinder events, the subjective possibility correspondence is independent of the default full states: $P_j(\omega|i_{\omega^*}) = \mathbb{P}^{\Omega(\omega^*)}P_j^*(\omega_1^*)$ for any ω_1^* satisfying $\mathbb{P}^{\Omega(\omega^*)}(\omega_1^*) = \omega$. In particular, this implies $\wedge_{j=1}^n \mathbb{P}^{\Omega_i(\omega^*)}P_j^*(\omega^*) = \mathbb{P}^{\Omega_i(\omega^*)} \wedge_{j=1}^n P_j^*(\omega^*)$: if a full state is reachable at ω^* in the full model (Ω^*, P^*), then in any subjective model, the corresponding subjective state is reachable, too.

It follows $r_i^k(\omega^*) = ((\omega^*, i), (s(\omega_2^*|i_{\omega^*}), i^2), \cdots, (s(\omega_k^*|r_i^{k-1}(\omega^*)), i^k))$ is a relevant sequence. But now since $\bar{\omega}^* \notin E_{\Omega^*}$, the corresponding projections satisfy $s(\bar{\omega}^*|r_i^{k-1}(\omega^*)) \notin E_{\Omega_r^{k-1}}$; since $s(\bar{\omega}^*|r_i^{k-1}(\omega^*)) \in P_{i^k}(s(\omega_k^*|r_i^{k-1}(\omega^*))|r_i^{k-1}(\omega^*))$, we have:

$$P_{i^k}(s(\omega_k^*|r_i^{k-1}(\omega^*))|r_i^{k-1}(\omega^*)) \not\subseteq E_{\Omega_r^{k-1}}$$

But we have shown in the proof of Theorem 8, the above is a sufficient condition for

$$s(\omega^*|i_{\omega^*}) \notin \tilde{K}^i_{\omega^*}(\tilde{K}^{i^2}(\cdots(\tilde{K}^{i^k}(E|r_i^{k-1}(\omega^*))\cdots)|i_{\omega^*})$$

which is equivalent to $\omega^* \notin K_i K_{i^2} \cdots K_{i^k}(E)$, which in turn implies $\omega^* \notin CK(E)$.

Case 2: suppose there exists $\bar{\omega}^* \in \wedge_{j=1}^n P_j^*(\omega^*)$ such that $\mathcal{D}_E \not\subseteq W_j^*(\bar{\omega}^*)$ for some $j \in N$. That is, at $\bar{\omega}^*$, j is unaware of E and hence does not know E. Under strongly rational information, this lack of knowledge of E is recognized by every player at ω^*, and hence generating enough uncertainty about k's knowledge of E to keep E from being common knowledge.

Suppose $\bar{\omega}^* \neq \omega^*$. (Otherwise $\omega^* \notin K_j(E)$ and the conclusion follows already.) Let $i \in N$ and let $\bar{\omega}^*$ be reachable from ω^* for i through the factual information sets of $i^2, \cdots, i^k$ as in case 1, thus the following sequence is relevant:

$$r_i^k(\omega^*) = ((\omega^*, i), (s(\omega_2^*|i_{\omega^*}), i^2), \cdots, (s(\omega_k^*|r_i^{k-1}(\omega^*)), i^k), (s(\bar{\omega}^*|r_i^k(\omega^*)), j))$$

Claim:

$$W_j(s(\bar{\omega}^*|r_i^k(\omega^*))|r_i^k(\omega^*)) = W_i^*(\omega^*) \cdots W_{i^k}^*(\omega_k^*) \cap W_j^*(\bar{\omega}^*) \tag{7.11}$$

Proof of the claim. Observe

$$W_{i^2}(s(\omega_2^*|i_{\omega^*})|i_{\omega^*}) \quad = \quad W_i^*(\omega^*) \cap [W_{i^2}^*(\omega_2^*) \cup_{\left\{\omega_0^*:\mathbb{P}^{\Omega_i(\omega^*)}(\omega_0^*)=s(\omega_2^*|i_{\omega^*})\right\}} W_{i^2}^*(\omega_0^*)]$$

Note $\left\{\omega_0^* : \mathbb{P}^{\Omega_i(\omega^*)}(\omega_0^*) = s(\omega_2^*|i_{\omega^*})\right\}$ is the set of full states that coincide with ω_2^* on $W_i(\omega^*)$. Thus by nice awareness, for all ω_0^* such that $\mathbb{P}^{\Omega_i(\omega^*)}(\omega_0^*) = s(\omega_2^*|i_{\omega^*})$,

$$[W_{i^2}^*(\omega_2^*) \triangle W_{i^2}^*(\omega_0^*)] \cap W_i^*(\omega^*) = \emptyset$$

It follows

$$W_{i^2}(s(\omega_2^*|i_{\omega^*})|i_{\omega^*}) = W_i^*(\omega^*) \cap W_{i^2}^*(\omega_2^*) \tag{7.12}$$

Let $B := \left\{\omega \in \Omega_i(\omega^*), \omega \neq s(\omega_3^*|i_{\omega^*}) : \mathbb{P}^{\Omega_r^2}(\omega) = s(\omega_3^*|r_i^2(\omega^*))\right\}$ denote the set of subjective states in i's subjective state space that are projected to the $s(\omega_3^*|r_i^2(\omega^*))$ in

i^2's subjective state space at $s(\omega_1^*|i_{\omega^*})$ ascribed by i at ω^*, and is not the projection of ω_3^* itself on $\Omega_i(\omega^*)$. Let $B^* := \left\{\omega_0^* : \mathbb{P}^{\Omega_i(\omega^*)}(\omega_0^*) = \omega \in B\right\}$ denote the set of full states that are projected to B. Using (7.12),

$$\begin{aligned} & W_{i^3}(s(\omega_3^*|r_i^2(\omega^*))|r_i^2(\omega^*)) \\ = \; & W_{i^2}(s(\omega_2^*|i_{\omega^*})|i_{\omega^*}) \cap [W_{i^3}(s(\omega_3^*|i_{\omega^*})|i_{\omega^*}) \cup_{\omega \in B} W_{i^3}(\omega|i_{\omega^*})] \\ = \; & W_i^*(\omega^*) \cap W_{i^2}^*(\omega_2^*) \cap [[W_i^*(\omega^*) \cap W_{i^3}^*(\omega_3^*)] \cup_{\omega_0^* \in B^*} [W_i(\omega^*) \cap W_{i^3}(\omega_0^*)]] \\ = \; & W_i^*(\omega^*) \cap W_{i^2}^*(\omega_2^*) \cap [W_{i^3}^*(\omega_3^*) \cup_{\omega_0^* \in B^*} W_{i^3}^*(\omega_0^*)] \end{aligned}$$

But since the projections of ω_3^* and all $\omega_0^* \in B^*$ coincide on $\Omega_r^2 = \times[W_i^*(\omega^*) \cap W_{i^2}^*(\omega_2^*)]$, by nice awareness,

$$[\bigcup_{\omega_0^* \in B^*} W_{i^3}^*(\omega_3^*) \triangle W_{i^3}^*(\omega_0^*)] \cap [W_i^*(\omega^*) \cap W_{i^2}^*(\omega_2^*)] = \emptyset$$

Therefore,

$$W_{i^3}(s(\omega_3^*|r_i^2(\omega^*))|r_i^2(\omega^*)) = W_i^*(\omega^*) \cap W_{i^2}^*(\omega_2^*) \cap W_{i^3}^*(\omega_3^*)$$

The conclusion follows from applying the arguments inductively. □

By (7.11), $\mathcal{D}_E \nsubseteq W_j(s(\omega_k^*|r_i^{k-1}(\omega^*))|r_i^{k-1}(\omega^*))$, which, as shown in the proof of Theorem 8, implies $s(\omega^*|j_{\omega^*}) \notin \tilde{K}_{\omega^*}^j(\tilde{K}^{i^2}(\cdots(\tilde{K}^i(E|r_j^{k-1}(\omega^*))\cdots)|j_{\omega^*})$ and hence $\omega^* \notin CK(E)$. This concludes the proof of Theorem 9. □

References

Aumann, Robert J., "Agreeing to Disagree," *Annals of Statistics*, 1976, *76* (4), 1236–1239.

Bacharach, Michael, "Some Extensions of a Claim of Aumann in an Axiomatic Model of Knowledge," *Journal of Economic Theory*, 1985, *37*, 167–90.

Dekel, Eddie, Barton L. Lipman, and Aldo Rustichini, "Standard State-Space Models Preclude Unawareness," *Econometrica*, 1998a, *66* (1), 159–173.

___ , ___ , and ___ , "Recent Developments in Modeling Unforeseen Contingencies," *European Economic Review*, 1998b, *42*, 523–542.

Ely, Jeffrey C., "A Note on Unawareness," 1998. Mimeo.

Fagin, Ronald and Joseph Y. Halpern, "Belief, Awareness and Limited Reasoning," *Artificial Intelligence*, 1988, *34*, 39–72.

Geanakoplos, John, "Game Theory without Partitions, and Applications to Speculation and Consensus," 1990. Cowles Foundation Discussion Paper No. 914.

Halpern, Joseph Y., "Alternative Semantics for Unawareness," *Games and Economic Behavior*, 2001, *37*, 321–339.

Heifetz, Aviad, Martin Meier, and Burkhard C. Schipper, "Interactive Unawareness," 2004. Forthcoming, Journal of Economic Theory.

Li, Jing, "Modeling Unawareness in Arbitrary State Space," 2006b. Working paper, University of Pennsylvania.

Modica, Salvatore and Aldo Rustichini, "Awareness and Partitional Information Structures," *Theory and Decision*, 1994, *37*, 107–124.

___ **and** ___ , "Unawareness and Partitional Information Structures," *Games and Economic Behavior*, 1999, *27*, 265–298.