Object-Based Unawareness II: Applications

Oliver Board and Kim-Sau Chung
University of Pittsburgh University of Minnesota

March 28, 2008

1 Introduction

This is a companion paper of Board and Chung (2007) (hereafter BC). BC propose a class of structures (called the OBU structures) which is capable of modelling an agent who is "not sure whether or not there exist things that he is unaware of" without the undesirable implication that the agent does not know what he is aware of.[1] BC also provide an axiomatization of those structures in first order modal logic. In this paper, we first prove some characterization results which provide precise description of the features of the OBU structures,[2] and then provide a couple of applications for these structures. The first application is to provide a rationale for a legal interpretive doctrine called *verba fortius accipiuntur contra proferentem*, which asks judges to resolve any ambiguity arising in a contract against the party who drafted the contract. We show that this doctrine systematically out-performs other doctrines when there is *persistent* asymmetric awareness between the contracting parties. The second application is to revisit a central result in Heifetz, Meier, and Schipper (2007) (hereafter HMS). HMS claim that "unawareness can be interpreted as a special form of delusion." However, surprisingly, they prove that the No-Trade Theorem continues to hold *despite* this existence of delusion. We show that, first of all, unawareness is *not* a special form of delusion. In particular, unawareness and non-delusion can co-exist in general, and HMS' model is a special case where they do not. Second of all, we show that HMS'

[1] Halpern and Rego (2006) propose an alternative model which can is also capable of modelling an agent who is "not sure whether or not there exist things that he is unaware of." However, in their model, whenever an agent has such uncertainty, he necessarily does not know what he is aware of, violating certain introspection axioms. See also Footnote 15.

[2] These are not the same as the soundness and completeness theorem proved in BC, but are rather set-theoretic completeness theorems in the spirit of Halpern (1999b).

model is a special case that satisfies a property called terminal-non-delusion, and it is this property that drives their result. Terminal-non-delusion is a property that can hold or fail independent of whether there is unawareness or not. Hence HMS' result has no bearing with unawareness at all.

2 The OBU Structures: A Review

We briefly review OBU structures here, and on our way we also prove some set-theoretic completeness theorems in the spirit of Halpern (1999b). The reader is referred to BC for an axiomatization of OBU structures in first order modal logic, as well as a review of the related literature.

2.1 OBU Frames

We start with OBU frames first. An *OBU frame* for n agents is a tuple $\langle W, O, \mathcal{I}_1, \ldots, \mathcal{I}_n, \mathcal{A}_1, \ldots, \mathcal{A}_n \rangle$, where:

- W is a set of states;
- O is a set of objects;
- $\mathcal{I}_i : W \to 2^W$ is an *information function* for agent i; and
- $\mathcal{A}_i : W \to 2^O$ is an *awareness function* for agent i.

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

In the standard model, events are subsets of the state space, corresponding (roughly) to the set of states in which some given proposition is true. In the our model, an event is an ordered pair (R, S), where $R \subseteq 2^W$ is a set of states and $S \subseteq 2^O$ is a set of objects; we call R the *reference* of the event (denoted $ref(R, S)$), corresponding (as before) to the set of states in which some proposition is true; and S is the *sense* of the event (denoted $sen(R, S)$), listing the set of objects referred to in the description of the proposition. (To give an example, the events representing the propositions "the dog barked" and "the dog barked and the cat either did or did not meow" have the same reference but difference senses.) We sometimes abuse notation and write (R, a) instead of $(R, \{a\})$, and (w, S) instead of $(\{w\}, S)$. We use $\mathcal{E}$ to denote the set of all events, with generic element E.

We now define three kinds of operators on events, corresponding to "not", "and", and "or".

$$\begin{aligned} \neg(R,S) &= (W\setminus R,S) \\ \wedge_j(R_j,S_j) &= (\cap_j R_j, \cup_j S_j) \\ \vee_j(R_j,S_j) &= (\cup_j R_j, \cup_j S_j) \end{aligned}$$

The negation of an event holds at precisely those states at which the event does not hold, but it refers to the same set of objects. The conjunction of two events holds only at those states at which *both* events hold, and it refers to both sets of objects. Finally, the disjunction of two events hold if either the first or the second holds, and it also refers to both sets of objects.

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

$$\mathsf{A}_i(R,S) = (\{w \mid S \subseteq \mathcal{A}_i(w)\}, S) \text{ (awareness)} \tag{1}$$

$$\mathsf{L}_i(R,S) = (\{w \mid \mathcal{I}_i(w) \subseteq R\}, S) \text{ (implicit knowledge)} \tag{2}$$

$$\mathsf{K}_i(R,S) = \mathsf{A}_i(R,S) \wedge \mathsf{L}_i(R,S) \text{ (explicit knowlege)} \tag{3}$$

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

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

A1 $\wedge_j \mathsf{A}(R,S_j) = \mathsf{A}(R,\cup_j S_j)$

A2 $\mathsf{A}(R,X) = \mathsf{A}(R',X)$ for all R, R'

A3 $\mathsf{A}(R,\varnothing) = (W,\varnothing)$

A4 $\mathsf{A}(R,X) = (R',X)$ for some R'

L1 $\mathsf{L}(W,O) = (W,O)$

L2 $\wedge_j \mathsf{L}(R_j, S) = \mathsf{L}(\cap_j R_j, S)$

L3 $\mathsf{L}(R,S) = (R', S)$ for some R'

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

The following results show that L1–L4 and A1–A4 also provide a precise characterization of awareness and (implicit) knowledge, respectively.

Proposition 1 *Suppose that A_i is defined as in (1). Then:*

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

Proof:

1. Straightforward.
2. Take some A'_i which satisfies A1–A4, and define $\mathcal{A}_i$ as follows: $a \in \mathcal{A}_i(w)$ iff $w \in ref \mathsf{A}'_i(W,a)$. We need to show that $\mathsf{A}'_i(R,S) = \mathsf{A}_i(R,S)$. We consider two cases:
   Case 1: $S \neq \varnothing$. Then

$$\begin{aligned}
\mathsf{A}'_i(R,S) &= \mathsf{A}'_i(W,S) \text{ (by A2)} \\
&= \wedge_{a\in S} \mathsf{A}'_i(W,a) \text{ (by A1)} \\
&= \wedge_{a\in S} (\{w \mid x \in \mathcal{A}_i(w)\}, a) \text{ (by A4 and the definition of } \mathcal{A}_i\text{)} \\
&= (\{w \mid S \subseteq \mathcal{A}_i(w)\}, S) \text{ (definition of } \wedge\text{)} \\
&= \mathsf{A}_i(R,S), \text{ as required.}
\end{aligned}$$

   Case 2: $S = \varnothing$. Then

$$\begin{aligned}
\mathsf{A}'_i(R,\varnothing) &= (W,\varnothing) \text{ (by A3)} \\
&= (\{w \in W \mid \varnothing \subseteq \mathcal{A}_i(w)\}, \varnothing) \\
&= \mathsf{A}_i(R,\varnothing), \text{ as required.}
\end{aligned}$$

■

Proposition 2 *Suppose that* L_i *is defined as in* (2). *Then:*

1. L_i *satisfies L1–L4; and*

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

PROOF:

1. Straightforward.

2. Take some L_i' which satisfies L1–L4, and define $\mathcal{I}$ as follows:

$$\mathcal{I}(w) = \{w' \mid w \in ref(\neg \mathsf{L}_i' \neg (w', O))\}.$$

Note that, by L4,

$$\{w' \mid w \in ref(\neg \mathsf{L}_i' \neg (w', O))\} = \{w' \mid w \in ref(\neg \mathsf{L}_i' \neg (w', S))\}$$

for all $S \subseteq O$, so $w' \in \mathcal{I}(w)$ iff $w \in ref(\neg \mathsf{L}_i' \neg (w', S))$, and hence

$$w' \notin \mathcal{I}(w) \text{ iff } w \in ref(\mathsf{L}_i' \neg (w', S)). \qquad (*)$$

We need to show that $\mathsf{L}_i'(R, S) = \mathsf{L}_i(R, S)$. We consider two cases:
Case 1: $R \neq W$. Then

$$\begin{aligned}
\mathsf{L}_i'(R, S) &= \mathsf{L}_i'(\cap_{w \notin R} W \setminus \{w\}, S) \\
&= \wedge_{w \notin R} \mathsf{L}_i'(W \setminus \{w\}, S) \text{ (by L2)} \\
&= \wedge_{w \notin R} \mathsf{L}_i' \neg (w, S) \text{ (definition of } \neg) \\
&= \wedge_{w \notin R} (\{w' \mid w \notin \mathcal{I}(w')\}, S) \text{ (by } (*) \text{ and L3)} \\
&= (\cap_{w \notin R} \{w' \mid w_1 \notin \mathcal{I}(w')\}, S) \text{ (definition of } \wedge) \\
&= (\{w' \mid \mathcal{I}(w') \subseteq R\}, S) \\
&= \mathsf{L}_i(R, S), \text{ as required.}
\end{aligned}$$

Case 2: $R = W$. Then $\mathsf{L}_i'(W, O) = (W, O)$ (by L1), so $\mathsf{L}_i'(W, S) = (W, S)$ (by L4). And $\mathsf{L}_i(W, S) = (\{w \mid \mathcal{I}_i(w) \subseteq W\}, S) = (W, S)$.

■

2.2 Introducing Properties

In economic applications, an important class of events consists of events that take the form of (for example) "a is tall," or more generally, "object a satisfies property p." So we introduce properties here.

Formally, properties can be represented as functions from objects to events: $p : O \to \mathcal{E}$ such that

$$p\,(a) = (R_a^p, S^p \cup \{x\}) \text{ for some } R_a^p \subseteq \mathcal{W} \text{ and some } S^p \subseteq O.$$

Intuitively, R_a^p is the set of states where object a possesses property p, and S^p is the set of objects referred to in the description of the property; for example, if p is the property "... is taller than Jim", then $S^p = \{Jim\}$. Note that S^p could be the empty set, for example if p is the property "... is tall". Let $\mathcal{P}$ denote the class of all properties.

Although the only primitives of the model are 1-place properties, we can build up n-place properties from n 1-place properties. Suppose we want to construct the 2-place property $taller\,(a, b)$, to be interpreted as "a is taller than b". We start with a family of 1-place properties $\{p_a : O \to \mathcal{E}\}_{a \in O}$, to be interpreted "$a$ is taller than ...". Define $f : O \to \mathcal{P}$ as $f\,(a) = p_a$. Then the two-place property $taller : O^2 \to \mathcal{E}$ is defined by $taller\,(a, b) = f\,(a)\,(b)$. Notice that, in particular, the sense of the event $taller(a, b)$ is $\{a, b\}$, because

$$sen(f(a)(b)) = S^{f(a)} \cup \{b\} = \{a\} \cup \{b\}.$$

We can also take negations, conjunctions, disjunctions, and implications of properties:

$$\begin{aligned}
\neg p &: O \to \mathcal{E} \text{ such that } (\neg p)\,(a) = \neg\,(p\,(a)) \\
p \wedge p' &: O \to \mathcal{E} \text{ such that } (p \wedge p')\,(a) = p\,(a) \wedge p'\,(a) \\
p \vee p' &: O \to \mathcal{E} \text{ such that } (p \vee p')(a) = p(a) \vee p'(a) \\
p \to p' &: O \to \mathcal{E} \text{ such that } (p \to p')(a) = (\neg p)(a) \vee p'(a)
\end{aligned}$$

A remark is warranted for the definition of negation. In daily English, we are accustomed to, for example, think of "short" as the negation of "tall." A more careful approach is to treat them as two different properties. In particular, "tall" should not be identified as "not short". This approach will allow us to handle even events that normally do not make sense to our ears; e.g., the event that "the coffee is tall." There is no state in which the coffee is tall, and there is no state in which the coffee is short, and in all states the coffee is not tall and not short.

2.3 Introducing Quantifiers

In economic applications, not only do we want to deal with events such as "a is tall" (represented by the event $tall(a)$, where $tall$ is a property) or "agent i knows that a is tall" (represented by the event $\mathsf{K}_i(tall(a))$), we also want to deal with events such as "agent i is not aware of all objects," or "agent i does not know whether or not he is aware of all objects." So we introduce (universal) quantifiers here.

As explained in BC, daily English admits multiple interpretations of the word "all", corresponding to different scopes implicit in the conversation: the "universe of objects" referred to by the word "all" can vary. We often freely switch back and forth among different interpretations, without making the scope explicit, and leaving it for the context to resolve the ambiguity. In a formal model, however, these different interpretations should be distinguished explicitly by different quantifiers. Two particular quantifiers that BC studies are possibilitist quantifier and actualist quantifier. The former has a scope that spans all possible objects, while the latter has a scope that spans only those objects that are "real," where "being real" is an exogenously given property. An example used by BC to illustrate the difference between the two quantifiers is as follows. Consider the following assertion: "Not all rights are included in the Bill of Rights—for example, it does not include the right to pick one's own fate." There can potentially be two reactions to this assertion. Some people would consider the example as valid, as it is indeed technically true that the Bill of Rights does not contain the right to pick one's own fate. Others would consider the example as invalid, as no one could be guaranteed the (unconstrained) right to pick pick his own fate, and hence the right to pick one's own fate is not a real right. The former reaction corresponds to the possibilitist interpretation of the word "all", whereas the latter corresponds to the actualist interpretation. In real life, most people have the latter reaction, meaning that we are likely to understand the word "all" in the above assertion as having a scope that spans only those objects that are real. Indeed, the actualist quantifier often plays a more important role in economic applications.

We introduce the possibilitist quantifier first, which is easier to understand, but less important in economic applications. For any property $p \in \mathcal{P}$, let $\overline{\mathsf{All}}\,p$ denote the event that "all objects satisfy property p," where "all" is interpreted in the possibilitist sense. Formally,

$\overline{\mathsf{All}}$ is a mapping from properties to events, such that

$$\overline{\mathsf{All}}\, p = (\cap_{a \in O} R^p_a, S^p).$$

We now introduce the actualist quantifier, which is our main concern. We first exogenously specify which objects are real in which state. Let $O_w \subseteq O$ be the set of objects that are real in state w. We then define a special property re ("... is real") in terms of the sets O_w:

$$re\,(a) = (\{w \mid a \in O_w\}\,, a)\,. \tag{4}$$

For any property $p \in \mathcal{P}$, recall that $re \to p$ is the property defined by $(re \to p)(a) = (\neg re)(a) \vee p(a)$. Let $\mathsf{All}\, p$ denote the event that "all objects satisfy property p", where "all" is interpreted in the actualist sense. Formally, All is a mapping from properties to events, such that

$$\mathsf{All}\, p = (\cap_{a \in O} R^{re \to p}_a, S^p). \tag{5}$$

Intuitively, $\mathsf{All}\, p$ holds at every state where all real objects possess property p; and the sense of $\mathsf{All}\, p$ is precisely the objects used to describe property p. It is easy to verify that the autualist quantifier satisfies the following properties:

All**1** $\mathsf{All}\ (\wedge_j p_j) = \wedge_j (\mathsf{All}\, p_j)$

All**2** if $w \in R^p_a$ for every $a \in O$, then $w \in ref(\mathsf{All}\, p)$

All**3** if $R^p_a = R^{p'}_a$ for every $a \in O$, then $ref(\mathsf{All}\, p) = ref(\mathsf{All}\, p')$

All**4** $sen(\mathsf{All}\, p) = S^p$

The following result shows that $\mathsf{All}\,1$–$\mathsf{All}\,4$ also provide a precise charactization of the actualist quantifier.

Proposition 3 *Suppose that* All *is defined as in* (4)–(5)*. Then:*

1. All *satisfies* $\mathsf{All}\,1$–$\mathsf{All}\,4$*; and*
2. *if* All' *is a mapping from properties to events which satisfies* $\mathsf{All}\,1$–$\mathsf{All}\,4$*, we can find a collection of real objects* $\{O_w\}_{w \in W}$ *such that* All' *and* All *coincide.*

PROOF:

1. Straightforward.

2. Take some All' which satisfies All1–All4. For any $w \in W$ and $a \in O$, construct the property p_{wa} such that:

$$p_{wa}(b) = \begin{cases} (W, b) & \text{if } b \neq a \\ (W \setminus \{w\}, b) & \text{if } b = a \end{cases}.$$

Observe for later use that, by All2, $W \setminus \{w\} \subseteq ref(\mathsf{All}'\, p_{wa})$, and hence, for any $R \subseteq W$,

$$\cap_{w \notin R} ref(\mathsf{All}'\, p_{wa}) = \{w | w \in ref(\mathsf{All}'\, p_{wa})\} \cup R. \tag{6}$$

We define $\{O_w\}_{w \in W}$ using these p_{wa}'s as follows:

$$O_w = \{a \mid w \notin ref\,(\mathsf{All}'\, p_{wa})\}\,.$$

These O_w's define the property re:

$$R_a^{re} = \{w \mid w \notin ref(\mathsf{All}'\, p_{wa})\}.$$

This property re, of course, in turn defines the operator All. We need to show that $\mathsf{All}' = \mathsf{All}$. Take an arbitrary property $\tilde{p}$. From All4, we have $sen(\mathsf{All}'\, \tilde{p}) = S^{\tilde{p}}$; and $sen(\mathsf{All}\, \tilde{p}) = S^{\tilde{p}}$ from the definition of All. It remains to show that $ref(\mathsf{All}'\, \tilde{p}) = ref(\mathsf{All}\, \tilde{p})$.

From $\tilde{p}$, construct another property $\hat{p}$ as follows:

$$\hat{p} := \wedge_{a \in O} \wedge_{w \notin R^{\tilde{p}_a}} \; p_{wa}.$$

We claim that $R_b^{\hat{p}} = R_b^{\tilde{p}}$ for every $b \in O$, and hence by All3, we have $ref(\mathsf{All}'\, \hat{p}) = ref(\mathsf{All}'\, \tilde{p})$. To prove this claim, notice that, for any $b \in O$,

$$\begin{aligned} R_b^{\hat{p}} &= \cap_{a \in O} \cap_{w \notin R_a^{\tilde{p}}} R_b^{p_{wa}} \\ &= (\cap_{\substack{a \neq b \\ w \notin R_a^{\tilde{p}}}} R_b^{p_{wa}}) \cap (\cap_{\substack{a = b \\ w \notin R_a^{\tilde{p}}}} R_b^{p_{wa}}) \\ &= (\cap_{\substack{a \neq b \\ w \notin R_a^{\tilde{p}}}} W) \cap (\cap_{w \notin R_b^{\tilde{p}}} W \setminus \{w\}) \\ &= R_b^{\tilde{p}}, \text{ as required.} \end{aligned}$$

Therefore, it suffices to prove that $ref(\mathsf{All}'\,\hat{p}) = ref(\mathsf{All}\,\tilde{p})$. By $\mathsf{All}1$, we have

$$\begin{aligned} ref(\mathsf{All}'\,\hat{p}) &= \cap_{a\in O} \cap_{w\notin R_a^{\tilde{p}}} ref(\mathsf{All}'\,p_{wa}) \\ &= \cap_{a\in O} (\{w \mid w \in ref(\mathsf{All}'\,p_{wa})\} \cup R_a^{\tilde{p}}) \quad \text{(by (6))} \\ &= \cap_{a\in O} (R_a^{\neg re} \cup R_a^{\tilde{p}}) \\ &= \cap_{a\in O} R_a^{\neg re \vee \tilde{p}} \\ &= \cap_{a\in O} R_a^{re \rightarrow \tilde{p}} \\ &= ref(\mathsf{All}\,\tilde{p}), \text{ as required.} \end{aligned}$$

■

An *OBU structure* is thus a tuple $\langle W, O, \{O_w\}_{w\in W}, \{\mathcal{I}_i\}_{i=1}^n, \{\mathcal{A}_i\}_{i=1}^n \rangle$.

3 OBU Structures with Probabilities

It is easy to introduce probabilistic beliefs into the OBU structures, although BC's axiomatization does not include this part. To introduce probabilistic beliefs, we first introduce implicit beliefs, which once again is a benchmark that serves as an intermediate tool to modelling what an agent actually believes. The relation between explicit beliefs (i.e., an agent's actual beliefs) and implicit beliefs is then analogous to the relation between explicit knowledge and implicit knowledge.

Start with an OBU structure $\langle W, O, \{O_w\}, \{\mathcal{I}_i\}, \{\mathcal{A}_i\} \rangle$. For simplicity let's assume that W is finite, so that we can avoid defining σ-algebras on W and hence ease notations. Augment the OBU structure with $\{q_i\}$, where each q_i is a probability assignment that associate with each state w a probability distribution on W satisfying $q_i(w)(\mathcal{I}_i(w)) = 1$. (an agent (implicitly) assigns probability 1 to those states that he consider possible when the true state is w). For any real number r, we introduce two belief operators for each agent, mapping any given event $E \in \mathcal{E}$ to the events that an agent implicitly and explicitly, respectively, believes that E happens with probability at least r:

$$\overline{\mathsf{B}}_i^r (R, S) = (\{w \mid q_i(w)(R) \geq r\}, S) \text{ (implicit belief)} \tag{7}$$

$$\mathsf{B}_i^r (R, S) = \mathsf{A}_i(R, S) \wedge \overline{\mathsf{B}}_i^r (R, S) \text{ (explicit belief)} \tag{8}$$

An *augmented OBU structure* is thus a tuple $\langle W, O, \{O_w\}, \{\mathcal{I}_i\}, \{\mathcal{A}_i\}, \{q_i\} \rangle$.

The common prior assumption is never a well-motivated assumption, even when there is no unawareness (see Morris (1995) and Gul (1998)). Nevertheless, for easy comparison with the previous literature in Section 5 , we introduce the assumption here as well. We say that an augmented OBU structure satisfies the *common prior assumption* if there exists a probability distribution q (without agent subscript) on W such that, whenever $q(\mathcal{I}_i(w)) > 0$, we have

$$q_i(w)(\cdot) = q(\cdot \mid \mathcal{I}_i(w)),$$

where $q(\cdot \mid \mathcal{I}_i(w))$ is the conditional probability distribution on W given $\mathcal{I}_i(w)$. When an augmented OBU structure satisfies the common prior assumption, we can represent it as the tuple $\langle W, O, \{O_w\}, \{\mathcal{I}_i\}, \{\mathcal{A}_i\}, q\rangle$, and simply call it an OBU structures with common prior.

In the subsequent sections, we shall study a couple applications of OBU structures with common priors.

4 The *verba fortius accipiuntur contra proferentem* Doctrine of Contract Interpretation

The *verba fortius accipiuntur contra proferentem* doctrine (or the cp doctrine in short), is a legal interpretive doctrine that asks judges to resolve any ambiguity[3] arising in a contract against the party who drafted the contract. In this section, we use OBU structures to formalize the rationale behind this doctrine. In particular, we show that the cp doctrine systematically out-performs other doctrines when there is *persistent* asymmetric awareness between the contracting parties.

Start with an OBU structure with common prior: $\langle W, O, \{O_w\}, \{\mathcal{I}_i\}, \{\mathcal{A}_i\}, q\rangle$. [4] There are two agents in the model. Agent 1 is a risk neutral insurer, and agent 2 is a risk averse insuree. We shall assume that agent 1 is female and agent 2 male. Absent any insurance contract between the agents, agent 1's income is 0 in every world, and agent 2's income is 0 in some worlds and 1 in other worlds. We can think of 0 income as the result of some negative income shock, which the risk averse agent 2 would like to insure against. Agent 1's utility is the same as his income, and agent 2's utility is $V(\cdot)$, which is strictly increasing and strictly concave in his income.

[3] "Ambiguity" is a loaded term in economics, and often refers to situations where decision makers entertain multiple prior probability distributions. Here, we are referring to a different kind of ambiguity, namely linguistic ambiguity.

[4] Given our distaste of the common prior assumption, the reader may wonder why we impose this assumption here. The common prior assumption allows us to state our results neatly. But we otherwise do not believe that the comparison between different doctrines depends on this assumption.

To make the set up as uncontroversial as possible, we make a couple of standard assumptions:[5]

1. Each agent i's $\mathcal{I}_i$ forms a partition of the state space W; i.e., $w \in \mathcal{I}_i(w)$ for every $w \in W$, and $w' \in \mathcal{I}_i(w)$ implies $\mathcal{I}_i(w') = \mathcal{I}_i(w)$. It means each agent i's implicit knowledge satisfies the axioms of truth, positive introspection, and negative introspection.

2. Each agent i implicitly knows what he is aware of; i.e., $w' \in \mathcal{I}_i(w)$ implies $\mathcal{A}_i(w') = \mathcal{A}_i(w)$.

We also make a couple of assumptions that are particularly natural to our current application.

1. It is natural to assume that agent 2 (the insuree) has a finer information partition then agent 1 (the insurer) has, meaning that agent 2 (implicitly) knows more about his likelihood of getting a negative income shock.

2. However, it is also natural to assume that agent 1, being an experienced insurer who has already dealt with many other insurees facing similar shocks, is aware of more objects than agent 2 is. It means $\mathcal{A}_2(w) \subset \mathcal{A}_1(w)$ for every w.

An example that satisfies the above assumptions is one where both agents' information partition is the coarest one (i.e., $\mathcal{I}_i(w) = W$ for every w and every i); and agent 1 is aware of every object (i.e., $\mathcal{A}_1(w) = O$ for every w), while agent 2 is aware of nothing except for his income (i.e., $\mathcal{A}_2(w) = \{o_1\}$, where o_1 is the object called "agent 2's income").[6] This example, although admittedly extreme, allows us to abstract away from the classical adverse selection problem, which is already well understood, and focus on the interaction between contractual ambiguity and asymmetric awareness. Therefore, we shall focus on this extreme case.

A rough outline of the contracting game's timing is as follows. There are 2 stages in the contracting phrase, which we shall refer to as *ex ante*. Then there is an *ex post* phrase, which contain the contract enforcement stage. In stage one, agent 1 proposes an insurance contract. The contract specifies a premium, a payment, and the condition C under which agent 1 has to pay the insurance payment to agent 2. A crucial assumption is that condition C has to be written with an exogenously given language, to be defined shortly, which does not allow agent 2 to directly refer to object o_1 (i.e., agent 2's income). Without this crucial

[5]See BC for discussion of these assumptions.

[6]A natural assumption which we do not make explicit in the text is that $o_1 \in O_w$ for every $w \in W$.

assumption, the insurance problem will be trivial. This assumption makes sense when, for example, agent 2's income is not verifiable and hence not contractible. In stage two, agent 2 either accepts or rejects the contract. If he accepts it, we move to stage three, the contract enforcement stage, where nature randomly picks a state according to the probability law q, and agent 1 has to pay agent 2 the insurance payment unless she can prove to a judge that condition C does not obtain.

4.1 Contracts and Interpretations

We now define the contractual language, which is very similar to the formal language of predicate logic. The building block of the language is a set called *vocabulary*, which contains the following three kinds of elements:

$a, b, c \ldots \in O \setminus \{o_1\}$ — the name of each object (except agent 2's income) is an element of the vocabulary

$P_1, P_2, \ldots$ — an exogenously given, non-empty list of predicates, each will later on be interpreted (by the judge) as corresponding to a certain property[7]

$\neg, \wedge, \vee$ — logical connectives

Intuitively, the contractual language is a collection of sentences, each is a finite string of letters (i.e., elements of the vocabulary) satisfying certain grammatic structure. We define this collection recursively as follows:

(i) for each object a and predicate P, $P(a)$ (to be interpreted as "object a satisfies the property called P") and $\neg P(a)$ are sentences;

(ii) if ϕ and ψ are sentences, then $\phi \wedge \psi$, and $\phi \vee \psi$ are sentences.

The contractual language, denoted by $\mathcal{L}$, is hence the minimal set that satisfies (i) and closed under (ii). A typical sentence in $\mathcal{L}$ looks like, for example, "the basement is flooded and the attic is burnt," where "basement" and "attic" are objects, and "... is flooded" and "... is burnt" are predicates.

To recap: an insurance contract is hence a triple (g, h, ϕ), where $g \in \mathbb{R}_+$ is the insurance premium that agent 2 pays agent 1 *ex ante*, $h \in \mathbb{R}_+$ is the insurance payment that agent 1 pays agent 2 *ex post*, and $\phi \in \mathcal{L}$ is a sentence that describes the condition C under which agent 1 pays h to agent 2.

[7] Without loss of generality, we assume that all these predicates are 1-place. See Section 2 for discussion.

Although a predicate P (in the vocabulary of the contractual language) is supposed to correspond to a specific property, whether an object satisfies that property or not is often ambiguous *ex post*. Does a basement with a leaky pipe satisfies the property of "being flooded"? Some people may say yes, while some say no. Without this kind of ambiguity, the cp doctrine would have been moot. So we now introduce this kind of ambiguity into our model.

It turns out that it is more convenient to think of this kind of ambiguity as ambiguity in the property that a predicate corresponds to. Formally, an *interpretation* is a mapping l from predicates to properties. We imagine that for every predicate P, there are two subpopulations in the society, each holds a different interpretation of P. Let l_1 and l_2 be the two interpretations in the society. We assume that these interpretations are exogenously given.

We can now formalize the cp doctrine. The cp doctrine asks the judge to resolve any ambiguity against the party who drafted the contract, which is agent 1 in this model. For example, if the contract requires that agent 1 pays the insurance payment when the basement is flooded, and if in the true state there is ambiguity in whether the house is flooded, then under the cp doctrine the judge is to rule that the house is indeed flooded. Formally, the cp doctrine d_{cp} is a mapping from $\mathcal{L}$ to events such that

d1 $d_{cp}(P(a)) = l_1(P)(a) \vee l_2(P)(a)$,

d2 $d_{cp}(\neg P(a)) = \neg(l_1(P)(a) \wedge l_2(P)(a))$,

d3 $d_{cp}(\phi \wedge \psi) = d_{cp}(\phi) \wedge d_{cp}(\psi)$, and

d4 $d_{cp}(\phi \vee \psi) = d_{cp}(\phi) \vee d_{cp}(\psi)$.

For sake of comparison, we set up a strawman and define the mirror image of the cp doctrine, and call it the anti-cp doctrine, which asks the judge to resolve any ambiguity in favor of agent 1. Formally, $d_{anti-cp}$ is the same as d_{cp}, except that d1 and d2 are replaced by

d1′ $d_{cp}(P(a)) = l_1(P)(a) \wedge l_2(P)(a)$, and

d2′ $d_{cp}(\neg P(a)) = \neg(l_1(P)(a) \vee l_2(P)(a))$.

Given any interpretive doctrine d of the judge, agent 1's problem in stage three (the contract enforcement stage) is to prove to the judge that condition C does not obtain, or equivalently that event $d(\phi)$ has not happened, where ϕ is the sentence that describes condition C. We assume that, *ex post*, once the true state w is realized, agent 2 would have all necessary evidences to prove whether or not any object a satisfies any property p, provided that object a is in the set of real objects O_w in state w. If $a \notin O_w$, then agent 2 cannot prove any statement regarding a, as there does not exist any real thing corresponding to the name a that agent 2 can point the judge to.

We assume that, ambiguity notwithstanding, the contractual language is rich enough so that full insurance is always possible, regardless of the interpretive doctrine. Formally, we assume the following.

Assumption 1 *For any $R \subseteq W$, there exists $\phi_R \in \mathcal{L}$ such that*

1. $d_{cp}(\phi_R) = d_{anti-cp}(\phi_R)$,
2. $ref(d_{cp}(\phi_R)) = R$, *and*
3. $sen(d_{cp}(\phi_R)) \subseteq O_w$ *for every* $w \notin R$.

Hence, if R is the set of states where agent 2 suffers an income shock, a contract (g, h, ϕ_R) with $h = 1$ will provide unambiguous full insurance for agent 2. In any state $w \notin R$, agent 2 would have no trouble proving that the (unambiguous) event described by ϕ_R has not happened. Assumption 1 also implicitly requires certain richness of the O_w's.

4.2 Benchmark: Symmetric Awareness

Finally, we need to explain how agent 2 makes his accept/reject decision in stage two. This can be tricky, as it depends on how agent 2's awareness changes after he reads the contract (which mentions objects that agent 2 was unaware of before he read it). Before we get into this, let's consider the benchmark case where there is symmetric awareness between the two agents; i.e., $\mathcal{A}_1(w) = \mathcal{A}_2(w) = O$ for every $w \in W$. In this case, agent 2 is aware of every object that agent 1 is aware of, and hence can perfectly distinguish different contracts written by agent 1. The solution concept of subgame perfect equilibrium is hence applicable here.

The following result demonstrates the importance of *asymmetric* awareness in rationizing the cp doctrine. When there is symmetric awareness, the choice interpretive doctrine is irrelevant.

Proposition 4 *Assume Assumption 1 and symmetric awareness. Under either the cp or anti-cp doctrine, in any subgame perfect equilibrium, agent 1 will offer a contract that fully insures agent 2, and agent 2 will accept it.*

PROOF: Let R be set of states where agent 2 suffers an income shock. Under either the cp or anti-cp doctrine, for any contract that does not fully insures agent 2, there exists another, full-insurance contract $(g, 1, \phi_R)$ for some appropriate insurance premium g such that agent 2 receives the same expected utility (and hence the two contracts are equally (un)acceptable to agent 2), while agent 1 receives a strictly higher expected utility. ■

Even if we drop Assumption 1, there is still no compelling argument for the cp doctrine. In general, the choice of interpretive doctrine will not be irrelevant. But for every example where the contracting parties prefer cp doctrine, there is an opposite example where they prefer the anti-cp doctrine. So making the cp doctrine the default interpretive doctrine does not save more contracting costs, as the contracting parties may need explicitly specify a different interpretive doctrine in their contracts half of the time.

4.3 Asymmetric Awareness

We now return to the case of asymmetric awareness; i.e., $\mathcal{A}_1(w) = O$ and $\mathcal{A}_2(w) = \{o_1\}$ for every $w \in W$. Here, an important modelling question to address is: how would agent 2's awareness changes after he reads the contract (which mentions objects that agent 2 was unaware of before he read it)?

If agent 2 was unaware of those objects because they slipped his mind, then it would be natural to assume that he becomes aware of them once he reads their name in a contract. If, instead, he was unaware of them because he genuinely had no idea what they were,[8] then it would be natural to assume that his awareness would not change even after reading the contract. In reality we will likely have both cases at the same time, which begs a richer model that distinguishes a slip-the-mind object from a genuinely-clueless object. But our current model is not rich enough for this purpose, so we will have to lump all objects into the same category. We consider both cases in turn.

4.3.1 The Slip-the-Mind Case

The slip-the-mind case is interesting because it is the only case considered by other authors so far.[9] But the next result says that this case is the same as the benchmark case of symmetric awareness, and hence the choice of interpretive doctrine is again irrelevant. A certain implicit assumption is responsible for this result, however, and it will prove inspiring

[8] Product warranties often mention that repair of certain specific parts is not covered, or is covered but not completely free, and we often have no idea what those parts are when we read and accept those warranties.

[9] See, for example, Filiz (2008).

for future research.

If agent 2 becomes aware of any object mentioned in a contract once he reads the contract, then, once again, he can perfectly distinguish different contracts written by agent 1. The solution concept of subgame perfect equilibrium is once again applicable.

Proposition 5 *Assume Assumption 1 and that agent 2 becomes aware of any object mentioned in a contract once he reads the contract. Under either the cp or anti-cp doctrine, in any subgame perfect equilibrium, agent 1 will offer a contract that fully insures agent 2, and agent 2 will accept it.*

The proof of Proposition 5 is exactly the same as that of Proposition 4. Even the discussion about what would happen if we relax Assumption 1 applies to the current case without change.

In case this result sounds counter-intuitive, let's point out one implicit assumption here. We implicitly assume that agent 2 assigns the same probability to the event $p(o_1)$, where p stands for the property "... drops," before and after he reads the names of some other objects in a contract. This is an extra assumption, because nothing in an OBU structure dictates this. An OBU structure is necessary static, and does not encode enough data regarding how an agent's knowledge, probabilistic belief, and awareness change when he is exposed to new information. An analogy is that Aumann's (static) partitional information structures do not encode enough data regarding how an agent's knowledge and probabilistic belief change when he receives unexpected information (such as when he observes an irrational move of an opponent who he originally thought was rational). Aumann's (static) partitional information structures are hence inadequate to study, for example, implications of common knowledge of rationality in dynamic games. The studies of these problems require richer models such as Samet (1996), Halpern (1999a), Battigalli and Siniscalchi (1999), and Board (2004). Similarily, satisfactory studies of dynamic games in the presense of unawareness require models that are richer than our OBU structures. Developing principled models of this kind should be a priority for future research.

4.3.2 The Clueless Case

If agent 2 does not become aware of an object after reading its name in a contract, his accept/reject decision cannot depend on the names mentioned in the contract. If he is to accept the contract $(g, h, P(a))$, he will also accept the contract $(g, h, P(b))$; and if he is to reject the contract $(g, h, P(a) \wedge \neg P'(b))$, he will also reject the contract $(g, h, P(b) \wedge \neg P'(c)$. Contracts written by agent 1 are in effect grouped into different equivalence classes, and agent 2's accept/reject decision has to apply uniformly to all contracts within the same equivalence class. This case hence resembles a game with imperfect information, and the solution concept of perfect Bayesian equilibrium is applicable.

A knee-jerk intuition may suggest that no contract will be accepted by agent 2, because he cannot understand fully any contract. This knee-jerk intuition, however, is at most half correct. It is easy to construct examples where, under either interpretive doctrines, in any perfect Bayesian equilibrium, agent 1 offers a contract that (unambiguously) fully insure agent 1, and agent 1 accepts it. For example, consider an example where $W = \{w_1, w_2\}$, $O_w = O = \{o_1, a\}$ for every w. There is only one predicate, P, in the contractual vocabulary, and according to both interpretations l_1 and l_2, object a satisfies the corresponding property in and only in state w_1. It is easy to verify that Assumption 1 is satisfied here. If w_1 is the state where agent 2 suffers an income shock, then he will only accept contracts of the form $(g, 1, P(\cdot))$ for low enough premium g, and agent 1 will offer a contract within this class in equilibrium. Full insurance is achieved. The analysis is similar if w_2 is the state where agent 2 suffers an income shock.

This kind of examples, however, rely on real object sets O_w's that are not rich enough (relative to the contractual language). Notice the subtle difference between rich O_w's and a rich O. For all real life examples, O is necessarily rich, as it is limited only by our imagination. For example, by combining a horn and a horse, we can imagine a unicorn, but it may not correspond to any real object. So a rich O does not imply rich O_w's. The issue here is whether O_w's are rich as well.

In most real life examples, however, we are accustomed to think of O_w's as rich as well, and hence the knee-jerk intuition remains appealing. The next result says that the knee-jerk intuition indeed is correct when O_w's are rich enough.

Assumption 2 *Let d denote the interpretive doctrine used by the judge. For any n-place sentence $\phi[a_1, \ldots, a_n] \in \mathcal{L}$, $n \geq 1$, there exist n distinct objects, $b_1, \ldots, b_n$, such that*

1. *$b_1, \ldots, b_n \in O_w \setminus \{o_1\}$ for evey $w \in W$, and*
2. *$ref(d(\phi[b_1/a_1, \ldots, b_n/a_n])) = \emptyset$, where $\phi[b_1/a_1, \ldots, b_n/a_n]$ is the same sentence as $\phi[a_1, \ldots, a_n]$ with each a_j replaced by b_j.*

Proposition 6 *Let d denote the interpretive doctrine used by the judge.*

1. *If Assumption 2 holds, then in any perfect Bayesian equilibrium, agent 2 receives no insurance.*

2. *If Assumption 2 does not hold, then there exists nonempty $R \subseteq W$ such that, if agent 2 suffers an income shock exactly in states in R, then there exists a perfect Bayesian equilibrium where agent 1 offers a contract that fully insures agent 2, and agent 2 accepts it.*

PROOF:

1. Suppose $(g, h, \phi[a_1, \ldots, a_n])$ is a contract that is both offered and accepted with positive probability in any equilibrium. Then $(g, h, \phi[b_1/a_1, \ldots, b_n/a_n])$, where the $\phi[b_1/a_1, \ldots, b_n/a_n]$ is defined as in Assumption 2, will also be accepted with positive probability. However, by Assumption 2, agent 1 can always prove that the event $d(\phi[b_1/a_n, \ldots, b_n/a_n])$ has not happened, and hence never needs to pay the insurance premium h. The fact that the original contract is offered with positive probability implies that agent 1 also never needs to pay the insurance premium under that contract. Hence agent 2 receives no insurance from it.

2. Let $\phi[a_1, \ldots, a_n]$ be the sentence that invalidates Assumption 2. For any n distinct objects $b_1, \ldots, b_n \in O \setminus \{o_1\}$, define $E(b_1, \ldots, b_n)$ to be the event

$$d(\phi[b_1/a_1, \ldots, b_n/a_n]) \vee \neg re(b_1) \vee \ldots \vee \neg re(b_n).$$

Let $(b_1^*, \ldots, b_n^*)$ be a solution of the following minimization problem:

$$\min_{\underbrace{b_1, \ldots, b_n}_{distinct} \in O \setminus \{o_1\}} q(ref(E(b_1, \ldots, b_n))),$$

where existence of a solution is guaranteed by the finiteness of W. Finally, define R to be $ref(E(b_1^*, \ldots, b_n^*))$. By assumption, R is nonempty. Then, if agent 2 suffers an income shock exactly in states in R, contracts of the form $(g, 1, \phi[b_1^*/a_1, \ldots, b_n^*/a_n])$ will fully insure agent 2, because agent 1 will fails to prove that condition C does not obtain exactly in states in R. Simple arguments then establish that there exists a perfect Bayesian equilibrium where agent 1 offers such a contract with an insurance premium g such that agent 2 is indifferent between accepting and rejecting, and agent 2 accepts the contract. The fact that $(b_1^*, \ldots, b_n^*)$ solves the above minimization problem implies that agent 1 cannot profitably deviate to other contracts within the equivalence class.

■

We can now formalize the benefit of the cp doctrine over, for example, the anti-cp doctrine: the cp doctrine minimizes the chance that Assumption 2 holds.

Proposition 7 *Whenever Assumption 1 holds under the cp doctrine, it will also hold under the anti-cp doctrine.*

PROOF: It suffices to observe that, for any $\phi \in \mathcal{L}$, $ref(d_{anti-cp}(\phi)) \subseteq ref(d_{cp}(\phi))$. ■

To summarize, the *verba fortius accipiuntur contra proferentem* doctrine of contract interpretation systematically out-performs other doctrines when there is *persistent* asymmetric awareness between the contracting parties.

5 Speculative Trade

This section revisits a central result in Heifetz, Meier, and Schipper (2007) (hereafter HMS). HMS claim that "unawareness can be interpreted as a special form of delusion." However, surprisingly, they prove that the No-Trade Theorem continues to hold *despite* this existence of delusion. We show that, first of all, unawareness is *not* a special form of delusion. In particular, unawareness and non-delusion can co-exist in general, and HMS' model is a special case where they do not. Second of all, we show that HMS' model is a special case that satisfies a property called terminal-non-delusion, and it is this property that drives their result. Terminal-non-delusion is a property that can hold or fail independent of whether there is unawareness or not. Hence HMS' result has no bearing with unawareness at all.[10]

5.1 Delusion and Unawareness

Let's recall that an OBU structure is a tuple $\langle W, O, \{O_w\}, \{\mathcal{I}_i\}, \{\mathcal{A}_i\}\rangle$. Given any OBU structure, the corresponding pair $\langle W, \{\mathcal{I}_i\}\rangle$ is called its *Kripke frame*, named after logician Saul Kripke. We say that a Kripke frame satisfies *non-delusion* if, for any world w and agent i, $w \in \mathcal{I}_i(w)$. We say that an OBU structure satisfies non-delusion if its corresponding Kripke frame satisfies non-delusion. Non-delusion is also known as reflexivity in the literature. It is

[10]This section does not imply that we endorse the common prior assumption. We do not, and we refer the reader to Morris (1995) and Gul (1998) for more on this subject. Nor does this section imply that we endorse the particular fomulation of the No Trade Theorem by HMS. We do not, and in particular we believe that any trade contract should instead be carefully modelled as written in certain contractual language which agents may or may not be able to comprehend. See our Section 4 for an example.

intimately related to the Truth Axiom, which says that if an agent knows a fact, then the fact is true (see BC for details).

It is easy to construct examples of OBU structures that satisfy non-delusion, and similarly for examples that violate it. Then why do HMS claim that "unawareness can be interpreted as a special form of delusion"? To understand their claim, let's consider the following two examples.

Example 1 *Consider an OBU structure with only one agent,* $W = \{w, w'\}$, $O = \{o_1, o_2\} = O_w$, $O_{w'} = \{o_1\}$, $\mathcal{A}(w) = \mathcal{A}(w') = \{o_1\}$. *The corresponding information function is* $\mathcal{I}(w) = \{world\}$ *and* $\mathcal{I}(w') = \{world'\}$.

Example 2 *Consider an OBU structure with only one agent,* $W = \{w, w'\}$, $O = \{o_1, o_2\} = O_w$, $O_{w'} = \{o_1\}$, $\mathcal{A}(w) = \mathcal{A}(w') = \{o_1\}$. *The corresponding information function is* $\mathcal{I}(w) = \{world'\} = \mathcal{I}(w')$.

In both examples, the agent is aware of only object o_1 but not object o_2. However, object o_2 is real only in world w (or one can think of it as physically exists only in world w), and hence there exists (in the actualist sense) something that the agent is unaware of only in world w. Therefore, in world w', the agent both implicitly and explicitly knows, and correctly so, that there does not exist anything that he is unaware of.

The two example differ, however, in $\mathcal{I}(w)$. As a result, the Kripke structure in Example 1 satisfies non-delusion, whereas that in Example 2 violates it. In particular, in Example 1, in world w, the agent (explicitly) knows, and correctly so, that there exists something that he is unaware of. On the contrary, in Example 2, in world w, the agent (explicitly) knows, but incorrectly so, that there does not exist anything that he is unaware of.

Example 2 is an example of OBU structures that satisfy two properties, both implicitly assumed in HMS: non-trivial unawareness, and denial of unawareness. Roughly speaking, we say that an OBU structure satisfies *non-trivial unawareness* if there exists a world w and an agent i such that, in world w, there exists (in the actualist sense) some objects that agent i is unaware of. Similarly, we say that an OBU structure satisfies *denial of unawareness* if for every world w and every agent i, in world w, agent i (explicitly) knows, not necessarily correctly, that there does not exist any object that he is unaware of. Formally, for any object o, choose an arbitrary event E_o such that $sen(E_o) = \{o\}$. The event $\mathsf{A}_i E_o$ is independent of how we choose this event E_o, and can be interpreted as the event that agent i is aware of object o. By varying across o, we can think of $\mathsf{A}_i E_.$ as a (1-place) property, and hence $\mathsf{All}\, \mathsf{A}_i E_.$ is an event, and corresponds to the fact that agent i is aware of all (in the actualist sense) objects. The event $\wedge_i \mathsf{All}\,\mathsf{A}_i E.$ hence corresponds to the fact that every agent is aware of all (in the actualist sense) objects. An OBU structure is said to satisfy *non-trivial unawareness* if

$$ref(\wedge_i \mathsf{All}\,\mathsf{A}_i E.) \neq W.$$

Similarly, the event $\wedge_i \mathsf{K}_i \mathsf{All}\,\mathsf{A}_i E.$ corresponds to the fact that every agent (explicitly) knows, not necessarily correctly, that he is aware of all (in the actualist sense) objects. An OBU structure is said to satisfy *denial of unawareness* if

$$ref(\wedge_i \mathsf{K}_i \mathsf{All}\,\mathsf{A}_i E.) = W.$$

HMS's claim that "unawareness can be interpreted as a special form of delusion" can now be formalized as a direct consequence of the implicit assumption of denial of unawareness.

Proposition 8 *If an OBU structure satisfies non-trivial unawareness and denial of unawareness, then it must violate non-delusion.*

PROOF: Pick any world $w \in W \setminus ref(\mathsf{All}\,\mathsf{A}_i E.)$, which can be done by the presumption of non-trivial unawareness. Since

$$w \in ref(\mathsf{K}_i \mathsf{All}\,\mathsf{A}_i E.) \subseteq ref(\mathsf{L}_i \mathsf{All}\,\mathsf{A}_i E.)$$

by the presumption of denial of unawareness, we have $\mathcal{I}_i(w) \subseteq ref(\mathsf{All}\ maths f A_i E.)$. Therefore $w \notin \mathcal{I}_i(w)$, violating non-delusion. ■

5.2 Terminal-Non-Delusion and Speculative Trade

HMS's formulation of the No Trade Theorem can be reproduced in our framework as follows. Given any OBU structure with a common prior, let $v : W \to \mathbb{R}^I$ be a function that satisfies $\sum_i v_i(w) = 0$ for every world w. The function v can be thought of as a trade contract that specifies the net monetary transfer to/from each agent in each world. Let F_i^v denote the event with empty sense ($sen(F_i^v) = \emptyset$)[11] and with its reference equal to the subset

[11] Empty sense is important. It guarantees that everyone — regardless of his awareness — can understand the trade contract and evaluate its profitability. But this also takes unawareness out of the picture, leaving the analysis with no bearing with unawareness at all. Again, see our disclaimer in Footnote 10.

of worlds in which agent i's conditional expection of v is strictly positive:[12]

$$ref(F_i^v) = \{w | \frac{\sum_{w' \in \mathcal{I}_i(w)} q(w')v_i(w')}{\sum_{w' \in \mathcal{I}_i(w)} q(w')} > 0\}.$$

Let F^v be the conjunction of these events: $F^v = \wedge_i F_i^v$. Let $\mathsf{K}^n F^v$ be recursively defined as $\wedge_i \mathsf{K}_i \mathsf{K}^{n-1} F^v$, with $\mathsf{K}^0 F^v = F^v$. Finally, define

$$\mathsf{CK} F^v := \wedge_{n \geq 1} \mathsf{K}^n F^v.$$

The No Trade Theorem result is said to obtain if $ref(\mathsf{CK}F^v) = \emptyset$ for every trade contract v.

As is apparent, whether or not the No Trade Theorem result obtains depends directly only on the $\langle W, \{\mathcal{I}_i\}, q\rangle$ part of any OBU structure with a common prior. It depends on unawareness only indirectly via channels such as Proposition 8. Let's call the $\langle W, \{\mathcal{I}_i\}, q\rangle$ part the *classical structure*. There are some well-known properties of the classical structure that would guarantee that the No Trade Theorem result obtains.

We say that a classical structure is *Euclidean* if $w' \in \mathcal{I}_i(w)$ and $w'' \in \mathcal{I}_i(w)$ imply $w'' \in \mathcal{I}_i(w')$. Euclideaness is intimately related to the Axiom of Negative Introspection, which says that every agent (implicitly) knows what he does not (implicitly) knows (see BC for details). We say that a classical structure satisfies *non-delusion* if its corresponding Kripke frame satisfies non-delusion. It is well-known that the No Trade Theorem result obtains if the classical structure is Euclidean and non-delusional. See, for example, Samet (1998) for a proof.

However, it is also well-known that Euclideaness alone, without non-delusion, does not suffice to guarantee the No Trade Theorem result. This makes HMS's central result very surprising, as it says that the No Trade Theorem result continues to obtain in their unawareness models in spite of their violation of non-delusion. To understand their result, we shall introduce a weaker property called terminal non-delusion.

We first generalize the notion of non-delusion to subspaces of W. For any subspace $W' \subseteq W$, we say that W' satisfies *non-delusion* if, for any world $w \in W'$ and agent i, $w \in \mathcal{I}_i(w) \subseteq W'$. For any subspace w', define[13]

$$D(W') = \{w \in W | \mathsf{I}_i(w) \neq \emptyset \text{ and } \mathsf{I}_i(w) \subseteq W' \text{ for some agent } i\}.$$

[12]If $\mathcal{I}_i(w) = \emptyset$, we should include w in the set on the right hand side as well.

[13]Our definition of OBU structures is general enough to allow $\mathsf{I}_i(w) = \emptyset$. But in most applications it makes sense to assume that $\mathsf{I}_i(w) \neq \emptyset$, which would simplify the equation below.

We say that a classical structure satisfies *terminal non-delusion* if there is a non-delusional subspace $W' \subseteq W$ such that $W = \cup_{n \geq 0} D^n(W')$, where $D^n(W')$ is defined recursively as $D(D^{n-1}(W'))$, and $D^0(W') = W$.

The classical structures in both Examples 1 and 2 satisfy terminal non-delusion, but only that in Example 1 satisfies non-delusion. More generally, if a classical structure satisfies non-delusion, it also satisfies terminal non-delusion, but not vice versa.

The next result says that non-delusion can be replaced by terminal non-delusion in order to obtain the No Trade Theorem result.

Proposition 9 *Suppose that the classical structure (of an OBU structure with a common prior) is Euclidean and terminally non-delusional. Then the No Trade Theorem result obtains.*

PROOF: Let W' be the non-delusional subspace in the presumption of terminal non-delusion. We prove by induction that

$$ref(\mathsf{CK}F^v) \cap D^n(W') = \emptyset$$

for every n, which immediately implies that the No Trade Theorem result obtains. For $n = 0$, this comes from the classical No Trade Theorem (see, for example, Samet (1998) for a proof). Suppose it is proved up to n. Consider any world w such that $\mathsf{I}_i(w) \neq \emptyset$ and $\mathsf{I}_i(w) \in D^n(W')$. Suppose $w \in ref(\mathsf{CK}F^v)$. Then $w \in ref(\mathsf{K}_i\mathsf{K}^m F^v)$ for every $m \geq 1$, and hence $\mathsf{I}_i(w) \subseteq ref(\mathsf{K}^m F^v)$ for every $m \geq 1$. Therefore $\mathsf{I}_i(w) \subseteq ref(\mathsf{CK}F^v)$. But then $ref(\mathsf{CK}F^v) \cap D^n(W') \supseteq \mathsf{I}_i \neq \emptyset$, a contradiction. So we have $ref(\mathsf{CK}F^v) \cap D^{n+1}(W') = \emptyset$ as well. ■

5.3 Terminal-Non-Delusion and Unawareness

Terminal non-delusion has no intrinsic bearing with unawareness. We have already seen two examples where the corresponding classical structure satisfies terminal non-delusion. We now present an example where it violates terminal non-delusion.

Example 3 *Consider an OBU structure with two agents,* $W = \{w, w'\}$*,* $O = \{o_1, o_2\} = O_w$*,* $O_{w'} = \{o_1\}$*,* $\mathcal{A}_i(w) = \mathcal{A}_i(w') = \{o_1\}$ *for both agent i. The corresponding information functions are* $\mathcal{I}_1(w) = \mathcal{I}_1(w') = \{w'\}$*, and* $\mathcal{I}_2(w) = \mathcal{I}_2(w') = \{w\}$*.*

In this example, agent 1 is a typical HMS agent — he lives in denial. He always thinks, sometimes incorrectly, that there is nothing that he is unaware of. On the contrary, agent 2 lives in paranoia. He always thinks, sometimes incorrectly, that there is something that he is unaware of. The classical structure in this example violates terminal non-delusion, although it is transitive and Euclidean. It is easy to see that the No Trade Theorem result does not obtain in this example.[14]

In order to understand HMS's surprising result, it is important to understand why the classical structures underlying their models always satisfy terminal non-delusion, and the key lies once again in their implicit assumption of denial of unawareness, which rules out Example 3. Before formalizing this, we need to introduce a few more concepts.

We say that an OBU structure satisfies LA-introspection if, in any world w and for any agent i, $w' \in \mathcal{I}_i(w)$ implies $\mathcal{A}_i(w') = \mathcal{A}_i(w)$. LA-introspection is intimately related to BC's **LA1** and **LA2** Axioms, which roughly say that every agent knows correctly what he is aware of (see BC for details).[15]

Now, consider an OBU structure that satisfies denial of unawareness. An agent in this structure knows, sometimes incorrectly, that there is nothing he is unaware of. The next property, weak non-delusion, says that: if his knowledge that "there is nothing he is unaware of" turns out to be correct, then his knowledge of any other facts is also correct. Formally, an OBU structure is said to satisfy *weak non-delusion* if, for any world w and agent i, $O_w \subseteq \mathcal{A}_i(w)$ implies $w \in \mathcal{I}_i(w)$. Although weak non-delusion looks like a peculiar property, it is actually weaker than what is implicitly assumed in HMS.

Finally, we say that an OBU structure satisfies Euclieaness and terminal non-delusion if its corresponding classical structure satisfies these properties respectively. The following result formalizes the tie between denial of unawareness and terminal non-delusion.

Proposition 10 *Consider an OBU structure that has a finite W and satisfies the property that, for any world w and agent i, $\mathcal{I}_i(w) \neq \emptyset$. Suppose this OBU structure further satisfies Euclieaness, denial of unawareness, LA-introspection, and weak non-delusion. Then it also satisfies terminal non-delusion.*

PROOF: Index every world $w \in W$ by the cardinality of O_w. This process allows for different worlds having the same index. For any two worlds, w and w', we say that *w points to w'* if there is an agent i such that $w \notin \mathcal{I}_i(w)$ and $w' \in \mathcal{I}_i(w)$. Suppose w points to w'.

[14] For example, consider a common prior which put equal weight on each world, and a trade contract which requires agent 1 to pay agent 2 in world w, and vice versa in world w'.

[15] As explained in Footnote 1, Halpern and Rego (2006) propose an alternative model which can is also capable of modelling an agent who is "not sure whether or not there exist things that he is unaware of." However, in their model, whenever an agent has such uncertainty, he necessarily does not know what he is aware of, violating the **LA1** and **LA2** Axioms.

Then, by denial of unawareness, LA-introspection, and weak non-delusion,

$$O_{w'} \subseteq \mathcal{A}_i(w') = \mathcal{A}_i(w) \subsetneq O_(w)$$

for some agent i. Therefore a world can only point to another world that has a strictly lower index. By finiteness of W, there exist worlds that do not point to any other worlds. Let W' be the collection of these worlds.

If w belongs to W', then $w \in \mathcal{I}_i(w)$ for any agent i, because by assumption $\mathcal{I}_i(w) \neq \emptyset$. Furthermore, by Euclideaness, $w' \in \mathcal{I}_i(w)$ implies $w' \in \mathcal{I}_i(w')$, and hence w' also does not point to any other worlds. Therefore W' is a non-delusional subspace.

If $W \neq W'$, then by finiteness of W, and by the observation that worlds can only point to worlds that have lower indices, there must exist worlds that point only to worlds inside W'. It is easy to verify that the union of these worlds and W' equals to $D(W')$. In general, if $W \neq D^n(W')$, $D^{n+1}(W')$ is a strict superset of $D^n(W')$. Therefore, by finitness again, $W = \cup_{n\geq 0} D^n(W')$. ■

6 Conclusion

[to be completed]

References

Battigalli, P., and M. Siniscalchi (1999): "Hierarchies of Conditional Beliefs and Interactive Epistemology in Dynamic Games," *Journal of Economic Theory*, 88, 188–230.

Board, O. (2004): "Dynamic Interactive Epistemology," *Games and Economic Behavior*, 49, 49–80.

Board, O., and K.-S. Chung (2007): "Object-Based Unawareness," mimeo, University of Pittsburgh and University of Minnesota.

Filiz, E. (2008): "Incorporating Unawareness into Contract Theory," mimeo, University of Maryland.

Gul, F. (1998): "A Comment on Aumann's Bayesian View," *Econometrica*, 66, 923–928.

Halpern, J. (1999a): "Hypothetical Knowledge and Counterfactual Reasoning," *International Journal of Game Theory*, 28, 315–330.

——— (1999b): "Set-Theoretic Completeness for Epistemic and Conditional Logic," *Annals of Mathematics and Artificial Intelligence*, 26, 1–27.

Halpern, J., and L. C. Rego (2006): "Reasoning About Knowledge of Unawareness," in *Proceedings of the Tenth International Conference on Principles of Knowledge Representation and Reasoning.*

Heifetz, A., M. Meier, and B. Schipper (2007): "Unawareness, Beliefs and Games," mimeo, UC-Davis.

Morris, S. (1995): "The Common Prior Assumption in Economic Theory," *Economics and Philosophy*, 11, 227–253.

Samet, D. (1996): "Hypothetical Knowledge and Games with Perfect Information," *Games and Economic Behavior*, 17, 230–251.

——— (1998): "Common Priors and Separation of Convex Sets," *Games and Economic Behavior*, 24, 172–174.