

BONN ECON DISCUSSION PAPERS

Discussion Paper 23/2005

A Canonical Model for Interactive Unawareness

by

Aviad Heifetz, Martin Meier, Burkhard
C. Schipper

September 2005



Bonn Graduate School of Economics
Department of Economics
University of Bonn
Adenauerallee 24 - 42
D-53113 Bonn

The Bonn Graduate School of Economics is
sponsored by the

Deutsche Post  World Net
MAIL EXPRESS LOGISTICS FINANCE

A CANONICAL MODEL FOR INTERACTIVE UNAWARENESS*

Aviad Heifetz[†] Martin Meier[‡] Burkhard C. Schipper[§]

September 21, 2005

Abstract

Heifetz, Meier and Schipper (2005) introduced a generalized state-space model that allows for non-trivial unawareness among several individuals and strong properties of knowledge. We show that this generalized state-space model arises naturally if states consist of maximally consistent sets of formulas in an appropriate logical formulation.

JEL-Classifications: C70, C72, D80, D82.

Keywords: unawareness, awareness, knowledge, interactive epistemology, modal logic, lack of conception, bounded perception.

*We thank Joe Halpern for helpful comments. Martin acknowledges financial support from the European Union via a Marie-Curie Individual Fellowship (HPMF-2002-02073) during the preparation of the first draft of this paper and of the Spanish Ministerio de Educación y Ciencia via a Ramon y Cajal Fellowship (IAE-CSIC) and Research Grant SEJ2004-07861 while working on later versions. Burkhard acknowledges financial support by the DFG SFB/TR 15 and the Minerva Stiftung.

[†]The Economics and Management Department, The Open University of Israel. Email: aviadhe@openu.ac.il

[‡]Instituto de Análisis Económico - CSIC, Campus Universitat Autònoma de Barcelona. Email: martin.meier@uab.es

[§]Department of Economics, University of California, Davis. Email: bcschipper@ucdavis.edu

“5.6 Die Grenzen meiner Sprache bedeuten die Grenzen meiner Welt.”
“5.6 The limits of my language mean the limits of my world.”
(Ludwig Wittgenstein, *Tractatus Logico-Philosophicus*)

1 Introduction

Unawareness refers to lack of conception rather than to lack of information. In standard models of incomplete information, decision makers share the conception of the interactive decision problem but may have asymmetric information. While models of asymmetric information are common in game theory and economics, the modelling of unawareness in interactive decision making proves to be a tricky task. Modica and Rustichini (1994) and Dekel, Lipman and Rustichini (1998) showed that standard state-space models of asymmetric information preclude non-trivial forms of unawareness.

Heifetz, Meier and Schipper (2005) introduced a generalized state-space model that allows for non-trivial unawareness among several individuals and strong properties of knowledge. Such an *unawareness structure* involves as primitives a lattice of state-spaces ordered according to their strength of “expressive power,” and suitable projections among them. The possibility set of one individual in a state of one space may reside in a lower space, while the possibility set of a second individual in one of these possible states may reside in yet a lower space. The former phenomenon reflects the fact that the first individual is unaware of some aspects of reality, and conceives fewer dimensions of the situation than there actually are; the latter reflects the fact that the first individual considers as possible that the second individual is unaware of some aspects of which the first individual *is* aware. That’s how the model captures *mutual* beliefs about unawareness.

In this paper we substantiate this construction with logical building blocks. Such an investigation of foundations is indeed necessary. Any structure for modeling asymmetric cognition – even if it has the desired properties – begs the question whether the model *itself* should not be an object for further uncertainties of the individuals. One way to show that a model *is* comprehensive is to describe in minute detail the beliefs and mutual beliefs of all individuals in each state. If each relevant combination of such beliefs are described in some state of the model, then we are convinced that the phenomena we aim at modeling is captured by our construction.

For the case of knowledge and mutual knowledge, such a detailed substantiation of the beliefs in the standard partition model has been carried out by constructing the canonical model of all maximally-consistent sets of formulas in the S5 system of epistemic logic (Aumann, 1999). This is the standard propositional logic (with negations, conjunctions and disjunctions of formulas), augmented with a knowledge modality for each individual: For each formula φ and each individual i , there is also a formula $k_i\varphi$ (“individual i knows φ ”). The axioms of the system then specify (on top of the standard logical tautologies) that what an individual knows is true, and that the individual knows what she knows and what she ignores. The canonical construction in which the states are

the maximally-consistent sets of formulas in this logical system turns to have a natural partition structure, and the knowledge of events in this partition model reflects exactly the knowledge of formulas in the states. That’s how the canonical model is substantiated by the internal structure of its own states.

In this paper we aim at an analogous foundation for the case of unawareness. Two main differences arise.

First, the axiom system has to be amended: An individual knows that she doesn’t know some fact only if she is aware of that fact. Our axiom system manifests this feature, and is thus (equivalent to) a multi-person version of axiom systems proposed by Modica and Rustichini (1999) and Halpern (2001).

Second, in the lattice of spaces of an unawareness structure, only the states in the upper-most space are full descriptions of all aspects of reality relevant to the interaction among the individuals. States in lower spaces are subjective portraits of situations, in the mind of individuals who are unaware of some of these relevant dimensions. Therefore, to construct these subjective descriptions, one has to use a sub-language, with a proper subset of the primitive propositions of the logical syntax.

With these modifications done, we are able to accomplish our mission: The collections of maximally-consistent sets of formulas (across all sub-languages) constitute an unawareness structure (Theorem 1), in which knowledge of events reflects exactly the knowledge of formulas in states (Theorem 2).

This result entails an important corollary. Logicians are often interested whether an axiom system is sound and complete with respect to a given family of models. Soundness means that every theorem derivable from the axioms obtains in all states of all these models. Completeness means the reverse implication: Every formula that obtains in all states of all the models in the family is provable from the axioms of the system. Our results imply that the axiom system we devised is sound and complete with respect to the family of unawareness structures (Theorem 3).¹

The caveat here is that each space in the lattice of an unawareness structure is associated with a sub-language (determined by a subset of the primitive propositions), and that higher spaces in the lattice are associated with richer sub-languages. In a state which is a subjective description of reality in the mind of an individual who is unaware of a primitive proposition φ , neither φ nor $\neg\varphi$ would obtain. If one were to insist on a truth value for φ in such a state, it would have to be “mu” – a third value, distinct than “true” and “false”. This would lead to a three-valued logic, with an extended axiom system that accounts for the third truth value. In parallel work, Halpern and Rêgo (2005) have developed such three-valued sound and complete axiomatization for the family of our unawareness structures.

The article is organized as follows: In Section 2 we introduce the syntax and axioms. This is followed in Section 3 by the presentation of unawareness structures. In Section

¹In fact, even *strong* soundness and completeness are implied: A set of formulas Γ is consistent (that is, free of contradiction) if and only if there is a state in a model in which Γ obtains.

4 we construct the canonical model and prove its fundamental properties (Theorems 1 and 2). Soundness and Completeness are proved in Section 5 (Theorem 3). Some further properties of unawareness structures are collected in Appendix A.

2 Syntax and Axioms

Let X be the nonempty set of atomic formulas or primitive propositions, and let I be the set of individuals. We assume I to be nonempty, but otherwise no restriction is imposed (so I could be infinite, even uncountable).

The syntax is the usual multi-person modal syntax with knowledge modalities k_i , “negation” \neg , “and” \wedge and a constant \top for truth. As usual, \vee , \rightarrow , and \longleftrightarrow are abbreviations, defined in the usual way: if φ and ψ are formulas, then $\varphi \vee \psi := \neg(\neg\varphi \wedge \neg\psi)$, $(\varphi \rightarrow \psi) := (\neg\varphi \vee \psi)$, and $(\varphi \longleftrightarrow \psi) := (\varphi \rightarrow \psi) \wedge (\psi \rightarrow \varphi)$. Primitive propositions represent natural occurrences such as “tomorrow is payday” that are not themselves described either in terms of individuals knowing something, or as combinations of other natural occurrences using propositional connectives.

The *set of formulas* \mathcal{L} is the smallest set such that:

- \top is a formula,
- every $x \in X$ is a formula,
- if φ is a formula, then $\neg\varphi$ is a formula,
- if φ and ψ are formulas, then $\varphi \wedge \psi$ is a formula,
- if φ is a formula, then $k_i\varphi$ is a formula.

The *awareness* modality is defined by

$$a_i\varphi := k_i\varphi \vee k_i\neg k_i\varphi$$

and the *unawareness* modality is defined by

$$u_i\varphi := \neg a_i\varphi.$$

An individual’s awareness of a formula is defined as the individual knowing the formula or knowing that she doesn’t know the formula. An individual’s unawareness of a formula is defined as not being aware of the formula.

For a formula φ of our language, define $\text{Pr}(\varphi)$ to be the set of primitive propositions occurring in φ . More precise, define inductively:

$$\begin{aligned} \text{Pr}(\top) &:= \emptyset, \\ \text{Pr}(x) &:= \{x\}, \text{ for } x \in X, \\ \text{Pr}(\neg\varphi) &:= \text{Pr}(\varphi), \\ \text{Pr}(\varphi \wedge \psi) &:= \text{Pr}(\varphi) \cup \text{Pr}(\psi), \\ \text{Pr}(k_i\varphi) &:= \text{Pr}(\varphi). \end{aligned}$$

Let α be a subset of X . Then $\mathcal{L}_\alpha := \{\varphi \in \mathcal{L} \mid \text{Pr}(\varphi) \subseteq \alpha\}$. That is, \mathcal{L}_α is the sub-language of \mathcal{L} consisting of those formulas in which only primitive propositions in α occur.

An axiom is a formula assumed. An inference rule consists of hypotheses and the conclusion (e.g. see Modus Ponens below). An axiom system is a collection of axioms and inference rules. The axiom system $S5$ (see for example Chellas, 1980, pp. 14 or Fagin et al. 1995, p. 56) corresponds to standard partitional models usually applied in economics. This axiom system is weakened to the following system:

- All substitution instances of valid formulas of *Propositional Calculus* including the formula \top , (PC)

- the inference rule *Modus Ponens*:

$$\frac{\varphi, \varphi \rightarrow \psi}{\psi}, \quad (\text{MP})$$

- the *Axiom of Truth*:

$$k_i \varphi \rightarrow \varphi, \quad (\text{T})$$

- the *Axiom of Positive Introspection*:

$$k_i \varphi \rightarrow k_i k_i \varphi. \quad (4)$$

- the *Propositional Awareness axioms*:

$$\begin{aligned} 1. & \quad a_i \varphi \longleftrightarrow a_i \neg \varphi, \\ 2. & \quad a_i \varphi \wedge a_i \psi \longleftrightarrow a_i (\varphi \wedge \psi), \\ 3. & \quad a_i \varphi \longleftrightarrow a_i k_j \varphi, \text{ for } j \in I. \end{aligned} \quad (\text{PA})$$

- and the inference rule *RK-Inference*: For all natural numbers $n \geq 1$: If $\varphi_1, \varphi_2, \dots, \varphi_n$ and φ are formulas such that $\text{Pr}(\varphi) \subseteq \bigcup_{i=1}^n \text{Pr}(\varphi_i)$ then

$$\frac{\varphi_1 \wedge \varphi_2 \wedge \dots \wedge \varphi_n \rightarrow \varphi}{k_i \varphi_1 \wedge k_i \varphi_2 \wedge \dots \wedge k_i \varphi_n \rightarrow k_i \varphi}. \quad (\text{RK})$$

A *tautology* is a valid formula of Propositional Calculus.

Definition 1 *The set of theorems is the smallest set of formulas that contain all the axioms (that is all the substitution instances of valid formulas of Propositional Calculus, Truth, the Propositional Awareness Axioms and Axiom (4)) and that is closed under the inference rules Modus Ponens and RK-Inference.*

Definition 2 Let Γ be a set of formulas and φ a formula. A proof of φ from Γ is a finite sequence of formulas such that the last formula is φ and such that each formula is a formula in Γ , a theorem of the system or inferred from the previous formulas by Modus Ponens. If there is a proof of φ from Γ , then we write $\Gamma \vdash \varphi$. In particular, $\vdash \varphi$ means that φ is a theorem. If $\Gamma \vdash \varphi$, we say that Γ implies φ syntactically.

Definition 3 A set Γ of formulas is consistent if and only if there is no formula φ such that $\Gamma \vdash \varphi$ and $\Gamma \vdash \neg\varphi$. A set Γ of formulas is inconsistent, if it is not consistent.

RK-inference implies immediately (since $\varphi \wedge (\varphi \rightarrow \psi) \rightarrow \psi$ and $\varphi \wedge \psi \rightarrow (\varphi \wedge \psi)$ are instances of tautologies of the Propositional Calculus) that the following are theorems:

$$k_i\varphi \wedge k_i(\varphi \rightarrow \psi) \rightarrow k_i\psi \quad (\text{K})$$

and

$$k_i\varphi \wedge k_i\psi \rightarrow k_i(\varphi \wedge \psi). \quad (\text{C})$$

The derived inference rule (C), *Conjunction*, is implied by applying twice the inference rule Modus Ponens of the Propositional Calculus to instances of the tautology $p \rightarrow (q \rightarrow (p \wedge q))$:

$$\frac{\varphi, \psi}{\varphi \wedge \psi}.$$

By applying twice the inference rule Modus Ponens of the Propositional Calculus to instances of the tautology $(p \rightarrow q) \rightarrow ((q \rightarrow r) \rightarrow (p \rightarrow r))$, we obtain the derived inference rule *Implication*:

$$\frac{\varphi \rightarrow \psi, \psi \rightarrow \chi}{\varphi \rightarrow \chi}.$$

Since $k_i\neg k_i\neg k_i\varphi \rightarrow a_i\neg k_i\varphi$ is a theorem and since $a_i\neg k_i\varphi \rightarrow a_i\varphi$ follows from the propositional awareness axioms, $k_i\neg k_i\neg k_i\varphi \rightarrow a_i\varphi$ is a theorem:

Remark 1 The propositional awareness axioms imply the following weakening of the usual Axiom (5):²

$$u_i\varphi \rightarrow \neg k_i\neg k_i\neg k_i\varphi \quad (5')$$

or equivalently

$$k_i\neg k_i\neg k_i\varphi \rightarrow a_i\varphi. \quad (5'')$$

The positive introspection axiom implies introspection of awareness:

²We are grateful to Joe Halpern for pointing this out.

Proposition 1 *Axiom (4) implies*

$$a_i\varphi \rightarrow k_ia_i\varphi.$$

Proof. According to Axiom (4), $k_i\varphi \rightarrow k_ik_i\varphi$ and $k_i\neg k_i\varphi \rightarrow k_ik_i\neg k_i\varphi$ are theorems. Since $((p \rightarrow q) \wedge (r \rightarrow s)) \rightarrow ((p \vee r) \rightarrow (q \vee s))$ is a tautology, we have by Conjunction and Modus Ponens that $(k_i\varphi \vee k_i\neg k_i\varphi) \rightarrow (k_ik_i\varphi \vee k_ik_i\neg k_i\varphi)$ is a theorem. The formulas $k_i\varphi \rightarrow (k_i\varphi \vee k_i\neg k_i\varphi)$ and $k_i\neg k_i\varphi \rightarrow (k_i\varphi \vee k_i\neg k_i\varphi)$ are instances of tautologies. By RK $k_ik_i\varphi \rightarrow k_i(k_i\varphi \vee k_i\neg k_i\varphi)$ and $k_ik_i\neg k_i\varphi \rightarrow k_i(k_i\varphi \vee k_i\neg k_i\varphi)$ are theorems. Since $((q \rightarrow t) \wedge (s \rightarrow t)) \rightarrow ((q \vee s) \rightarrow t)$ is tautology, we have by Conjunction and Modus Ponens that $(k_ik_i\varphi \vee k_ik_i\neg k_i\varphi) \rightarrow k_i(k_i\varphi \vee k_i\neg k_i\varphi)$ is a theorem. By Implication, it follows that $(k_i\varphi \vee k_i\neg k_i\varphi) \rightarrow k_i(k_i\varphi \vee k_i\neg k_i\varphi)$ is a theorem. \square

Next, we formulate the Lemma on propositional awareness:

Lemma 1 *Let φ be a formula. Then, the following is a theorem:*

$$a_i\varphi \leftrightarrow \bigwedge_{x \in \text{Pr}(\varphi)} a_ix.$$

Proof. Clear by induction on the formation of the formulas using the propositional awareness axioms and the fact that the following is a tautology of the propositional calculus (for every natural number $n \geq 1$):

$$\left(\left(r \leftrightarrow \bigwedge_{i=1}^n q_i \right) \wedge \left(\bigwedge_{i=1}^n (q_i \leftrightarrow p_i) \right) \rightarrow \left(r \leftrightarrow \bigwedge_{i=1}^n p_i \right) \right).$$

\square

The Lemma says that individual i is aware of a formula if and only if she is aware of all primitive propositions in this formula.

The following proposition says that the following weaker form of the standard inference rule Necessitation obtains: Whenever an individual is aware of a theorem, then he knows that theorem:

Proposition 2 *If φ is a theorem, then $a_i\varphi \rightarrow k_i\varphi$ is theorem and hence $\{a_i\varphi\} \vdash k_i\varphi$.*

Proof. If φ is theorem then, since $\varphi \rightarrow (a_i\varphi \rightarrow \varphi)$ is an instance of a valid formula of PC, by Modus Ponens $a_i\varphi \rightarrow \varphi$ is also a theorem. By RK-Inference $k_ia_i\varphi \rightarrow k_i\varphi$ is a theorem. Since $a_i\varphi \rightarrow k_ia_i\varphi$ is a theorem, it follows by Implication that $a_i\varphi \rightarrow k_i\varphi$ is a theorem. \square

The following lemmata are standard and will be used later on:

Lemma 2 Let $\Gamma = \Gamma_1 \cup \Gamma_2$ be an inconsistent set of formulas, with Γ_1, Γ_2 both non-empty. Then there are finitely many formulas $\varphi_1, \dots, \varphi_n \in \Gamma_1$ and $\psi_1, \dots, \psi_m \in \Gamma_2$ such that $(\bigwedge_{i=1}^n \varphi_i) \rightarrow \neg(\bigwedge_{k=1}^m \psi_k)$ is a theorem.

Lemma 3 Let $\Gamma_1, \Gamma_2, \Gamma_3$ be nonempty sets of formulas, each closed under conjunctions, such that $\Gamma_1 \cup \Gamma_2 \cup \Gamma_3$ is inconsistent. Then there exist $\varphi \in \Gamma_1, \psi \in \Gamma_2$ and $\chi \in \Gamma_3$ such that $\vdash \varphi \wedge \psi \rightarrow \neg\chi$.

3 Unawareness Structures

In this section we recall the definition of unawareness structures in Heifetz, Meier and Schipper (2005).

Let $\mathcal{S} = \{S_\alpha\}_{\alpha \in \mathcal{A}}$ be a complete lattice of disjoint *state spaces*, with a partial order \preceq on \mathcal{S} . Denote by $\Sigma = \bigcup_{\alpha \in \mathcal{A}} S_\alpha$ the union of these spaces.

For every S and S' such that $S' \succeq S$, there is a surjective projection $r_{S'}^{S'} : S' \rightarrow S$, where r_S^S is the identity. Note that the cardinality of S is smaller than or equal to the cardinality of S' . We require the projections to commute: If $S'' \succeq S' \succeq S$ then $r_S^{S''} = r_S^{S'} \circ r_{S'}^{S''}$. If $s \in S'$, denote $s_S = r_S^{S'}(s)$. If $B \subseteq S'$, denote $B_S = \{s_S : s \in B\}$.

Denote $g(S) = \{S' : S' \succeq S\}$. For $B \subseteq S$, denote $B^\dagger = \bigcup_{S' \in g(S)} (r_{S'}^{S'})^{-1}(B)$.

An *event* is a pair (E, S) , where $E = B^\dagger$ with $B \subseteq S$, where $S \in \mathcal{S}$. B is called the *base* and S the *base-space* of (E, S) , denoted by $S(E)$. If $E \neq \emptyset$, then S is uniquely determined by E and, abusing notation, we write E for (E, S) . Otherwise, we write \emptyset^S for (\emptyset, S) . Note that not every subset of Σ is an event.

If (B^\dagger, S) is an event where $B \subseteq S$, the negation $\neg(B^\dagger, S)$ of (B^\dagger, S) is defined by $\neg(B^\dagger, S) := ((S \setminus B)^\dagger, S)$. Abusing notation, we write $\neg B^\dagger := (S \setminus B)^\dagger$. Note that by our notational convention, we have $\neg S^\dagger = \emptyset^S$ and $\neg \emptyset^S = S^\dagger$, for each space $S \in \mathcal{S}$. $\neg B^\dagger$ is typically a proper subset of the complement $\Sigma \setminus B^\dagger$. That is, $(S \setminus B)^\dagger \subsetneq \Sigma \setminus B^\dagger$. Thus our structure is not a standard state-space model in the sense of Dekel, Lipman, and Rustichini (1998).

If $\left\{ (B_\lambda^\dagger, S_\lambda) \right\}_{\lambda \in L}$ is a set of events (with $B_\lambda \subseteq S_\lambda$, for $\lambda \in L$), their conjunction $\bigwedge_{\lambda \in L} (B_\lambda^\dagger, S_\lambda)$ is defined by $\bigwedge_{\lambda \in L} (B_\lambda^\dagger, S_\lambda) := \left(\left(\bigcap_{\lambda \in L} B_\lambda^\dagger \right), \sup_{\lambda \in L} S_\lambda \right)$. Note, that since \mathcal{S} is a complete lattice, $\sup_{\lambda \in L} S_\lambda$ exists. If $S = \sup_{\lambda \in L} S_\lambda$, then we have $\left(\bigcap_{\lambda \in L} B_\lambda^\dagger \right) = \left(\bigcap_{\lambda \in L} \left((r_{S_\lambda}^{S_\lambda})^{-1}(B_\lambda) \right) \right)^\dagger$. Again, abusing notation, we write $\bigwedge_{\lambda \in L} B_\lambda^\dagger := \bigcap_{\lambda \in L} B_\lambda^\dagger$ (we will therefore use the conjunction symbol \wedge and the intersection symbol \cap interchangeably).

We define the relation \subseteq between events (E, S) and (F, S') , by $(E, S) \subseteq (F, S')$ if and only if $E \subseteq F$ as sets and $S' \preceq S$. If $E \neq \emptyset$, we have that $(E, S) \subseteq (F, S')$ if and only if $E \subseteq F$ as sets. Note however that for $E = \emptyset^S$ we have $(E, S) \subseteq (F, S')$ if and

only if $S' \preceq S$. Hence we can write $E \subseteq F$ instead of $(E, S) \subseteq (F, S')$ as long as we keep in mind that in the case of $E = \emptyset^S$ we have $\emptyset^S \subseteq F$ if and only if $S \succeq S(F)$. It follows from these definitions that for events E and F , $E \subseteq F$ is equivalent to $\neg F \subseteq \neg E$ only when E and F have the same base, i.e., $S(E) = S(F)$.

The disjunction of $\{B_\lambda^\uparrow\}_{\lambda \in L}$ is defined by the de Morgan law $\bigvee_{\lambda \in L} B_\lambda^\uparrow = \neg \left(\bigwedge_{\lambda \in L} \neg (B_\lambda^\uparrow) \right)$. Typically $\bigvee_{\lambda \in L} B_\lambda^\uparrow \subsetneq \bigcup_{\lambda \in L} B_\lambda^\uparrow$, and if all B_λ are nonempty we have that $\bigvee_{\lambda \in L} B_\lambda^\uparrow = \bigcup_{\lambda \in L} B_\lambda^\uparrow$ holds if and only if all the B_λ^\uparrow have the same base-space.

For each individual $i \in I$ there is a *possibility correspondence* $\Pi_i : \Sigma \rightarrow 2^\Sigma$ with the following properties:

0. *Confinedness*: If $s \in S$ then $\Pi_i(s) \subseteq S'$ for some $S' \preceq S$.
1. *Generalized Reflexivity*: $s \in \Pi_i^\uparrow(s)$ for every $s \in \Sigma$.³
2. *Stationarity*: $s' \in \Pi_i(s)$ implies $\Pi_i(s') = \Pi_i(s)$.
3. *Projections Preserve Awareness*: If $s \in S'$, $s \in \Pi_i(s)$ and $S \preceq S'$ then $s_S \in \Pi_i(s_S)$.
4. *Projections Preserve Ignorance*: If $s \in S'$ and $S \preceq S'$ then $\Pi_i^\uparrow(s) \subseteq \Pi_i^\uparrow(s_S)$.
5. *Projections Preserve Knowledge*: If $S \preceq S' \preceq S''$, $s \in S''$ and $\Pi_i(s) \subseteq S'$ then $(\Pi_i(s))_S = \Pi_i(s_S)$.

For 5., we could have assumed \supseteq and deduce $=$ from \supseteq , 3., and the other properties.

Remark 2 *Generalized Reflexivity implies that if $S' \preceq S$, $s \in S$ and $\Pi_i(s) \subseteq S'$, then $r_{S'}^S(s) \in \Pi_i(s)$. In particular, we have $\Pi_i(s) \neq \emptyset$, for all $s \in \Sigma$.*

Remark 3 *Property 4 and Confinedness imply that if $S' \preceq S$, $s \in S$ and $\Pi_i(s_{S'}) \subseteq S''$, then $\Pi_i(s) \subseteq S^*$ for some S^* with $S'' \preceq S^*$.*

Remark 4 *Property 5 and Confinedness imply Property 3.*

Generalized Reflexivity and Stationarity are the analogues of the partitional properties of the possibility correspondence in partitional information structures. In particular, Generalized Reflexivity will yield the truth property (that what an individual knows indeed obtains – property (iii) in Proposition 9 in Appendix A); Stationarity will guarantee the introspection properties (that an individual knows what she knows – property (iv) in Proposition 9 in Appendix A, and that an individual knows what she ignores provided she is aware of it – property 5. in Proposition 10 in Appendix A).

³Here and in what follows, we abuse notation slightly and write $\Pi_i^\uparrow(\omega)$ for $(\Pi_i(\omega))^\uparrow$.

Properties 3. to 5. guarantee the coherence of the knowledge and the awareness of individuals down the lattice structure. They compare the possibility sets of an individual in a state s and its projection s_S . The properties guarantee that, first, at the projected state s_S the individual knows nothing she does not know at s , and second, at the projected state s_S the individual is not aware of anything she is unaware of at s (Projections Preserve Ignorance). Third, at the projected state s_S the individual knows every event she knows at s , provided that this event is based in a space lower than or equal to S (Projections Preserve Knowledge). Fourth, at the projected state s_S the individual is aware of every event she is aware of at s , provided that this event is based in a space lower than or equal to S (Projections Preserve Awareness).

Definition 4 *The tuple*

$$\underline{\Sigma} := \left\langle (S_\alpha)_{\alpha \in \mathcal{A}}, (r_{S_\beta}^{S_\alpha})_{S_\beta \preceq S_\alpha}, (\Pi_i)_{i \in I} \right\rangle,$$

is called an unawareness structure for the set of individuals I .

Definition 5 *The knowledge operator of individual i on events E is defined, as usual, by*

$$K_i(E) := \{s \in \Sigma : \Pi_i(s) \subseteq E\},$$

if there is a state s such that $\Pi_i(s) \subseteq E$, and by

$$K_i(E) := \emptyset^{S(E)}$$

otherwise.

The following proposition is proved in Heifetz, Meier and Schipper (2005):

Proposition 3 *If E is an event, then $K_i(E)$ is an $S(E)$ -based event.*

Definition 6 *The unawareness operator of individual i from events to events is defined by*

$$U_i(E) = \neg K_i(E) \cap \neg K_i \neg K_i(E),$$

and the awareness operator is then naturally defined by

$$A_i(E) = \neg U_i(E).$$

This is the Modica-Rustichini (1999) definition. In particular, the Dekel-Lipman-Rustichini (1998) Plausibility requirement $U_i(E) \subseteq \neg K_i(E) \cap \neg K_i \neg K_i(E)$ is satisfied by this definition.

By Proposition 3 and the definition of the negation, we have

$$A_i(E) = K_i(E) \cup K_i \neg K_i(E).$$

For further properties of the unawareness structure see Appendix A.

4 The Canonical Unawareness Structure

Where do unawareness structures come from? What formal properties of knowledge and awareness do they capture? And can one model by unawareness structures *all* situations with such properties?

In this section we shall approach these questions by using the logical apparatus from section 2. “Properties” will be expressed by formulas of the syntax defined there; a “situation” will be a description of properties (a set of formulas) which is both *consistent* in the system (i.e. does not entail a contradiction), and *comprehensive* (i.e., for each potential property of knowledge and awareness, the description contains either the property or its negation).

We shall show that the collection of all such descriptions of situations does indeed constitute an unawareness structure (theorem 1 below). This unawareness structure is called the *canonical unawareness structure*. Any formula that belongs to some state of the canonical structure, and expresses knowledge or awareness, is mirrored accurately by a corresponding property of knowledge or awareness of *events* in the structure, a property that obtains in that state (theorem 2 below).

In short, the canonical unawareness structure consists of all the consistent and comprehensive descriptions of mutual knowledge and awareness, and these explicit descriptions are reflected by the knowledge and awareness operators on events in this structure. This substantiates the use of unawareness structure by providing an adequate foundation, and addresses the questions posed at the beginning of this section.

We now proceed with the formal definition of the canonical unawareness structure. Recall that X is the set of primitive propositions. For a subset $\alpha \subseteq X$, let Ω_α be the set of maximally consistent sets ω_α of formulas in the sub-language \mathcal{L}_α , that is, ω_α is consistent and for every formula $\varphi \in \mathcal{L}_\alpha \setminus \omega_\alpha$, the set $\omega_\alpha \cup \{\varphi\}$ is inconsistent. Let $\Omega = \bigcup_{\alpha \subseteq X} \Omega_\alpha$. We define $\Omega_\beta \preceq \Omega_\alpha$ whenever $\beta \subseteq \alpha$.

This makes $\{\Omega_\alpha\}_{\alpha \subseteq X}$ into a lattice of spaces. To complete its definition as *the canonical unawareness structure*, we proceed by defining the projections among the spaces and the possibility correspondences of individuals. To this end, we start with a pair of lemmas.

Lemma 4 *Let $\alpha \subseteq X$ and $\Gamma \subseteq \mathcal{L}_\alpha$. If Γ is a consistent subset of \mathcal{L}_α then it can be extended to a maximally consistent subset ω_α of \mathcal{L}_α .*

Proof. Clear by compactness and Zorn’s Lemma (note that by compactness the union of an increasing chain of consistent sets of formulas is consistent and an upper bound of that chain). \square

The following lemma can be proved along the lines of Chellas (1980, Theorem 2.18).

Lemma 5 *Let $\alpha \subseteq X$ and $\Gamma \subseteq \mathcal{L}_\alpha$. Then Γ is a maximally consistent subset of \mathcal{L}_α if and only if*

- (i) Γ is consistent, and
- (ii) for every formula $\varphi \in \mathcal{L}_\alpha$: $\varphi \in \Gamma$ or $\neg\varphi \in \Gamma$.

If $\Omega_\alpha \succeq \Omega_\beta$ (i.e., $\alpha \supseteq \beta$) we define a projection

$$r_\beta^\alpha : \Omega_\alpha \longrightarrow \Omega_\beta$$

by

$$r_\beta^\alpha(\omega) := \omega \cap \mathcal{L}_\beta.$$

Proposition 4 *The projection r_β^α is well-defined and surjective.*

Proof. By Lemma 5, r_β^α is well-defined, and by Lemma 4, r_β^α is surjective. \square

Remark 5 *If $\alpha \supseteq \beta \supseteq \gamma$ then $r_\gamma^\alpha = r_\gamma^\beta \circ r_\beta^\alpha$.*

For every formula φ , denote

$$[\varphi] := \{\omega \in \Omega : \varphi \in \omega\}.$$

The following lemma can be proved along the lines of Chellas (1980, Theorem 2.18).

Lemma 6 *Let $\omega \in \Omega_\alpha$ for some $\alpha \subseteq X$. Then ω is closed under inferences in the following sense:*

1. *If φ is a theorem such that $\varphi \in \mathcal{L}_\alpha$, then $\varphi \in \omega$.*
2. *If $\varphi \in \omega$ and $\varphi \rightarrow \psi$ is a theorem such that $\psi \in \mathcal{L}_\alpha$, then $\psi \in \omega$.*
3. *If $\varphi_1, \dots, \varphi_n \in \omega$, then $\bigwedge_{i=1}^n \varphi_i \in \omega$.*

Proposition 5 *$[\varphi]$ is an event.*

Proof. Consider $\alpha = \text{Pr}(\varphi)$ and $B := \{\omega_\alpha \in \Omega_\alpha \mid \varphi \in \omega_\alpha\}$. If $\omega \in \Omega$ such that $\varphi \in \omega$, then $\omega \in \Omega_\beta$ for some $\beta \supseteq \alpha$. Hence $\omega \cap \mathcal{L}_\alpha \in B$, that is $\omega \in B^\uparrow$, hence we have shown that $[\varphi] \subseteq B^\uparrow$. Conversely $\omega \in B^\uparrow$ implies that $\omega \cap \mathcal{L}_\alpha \in B$, therefore we have $\varphi \in \omega \cap \mathcal{L}_\alpha$, hence $\varphi \in \omega$, that is $\omega \in [\varphi]$. \square

Definition 7 For every $\omega \in \Omega$ and $i \in I$ define the possibility set

$$\Pi_i(\omega) = \left\{ \omega' \in \Omega : \text{For every formula } \varphi \begin{array}{l} (i) \ k_i\varphi \in \omega \text{ implies } \varphi \in \omega' \\ (ii) \ a_i\varphi \in \omega \text{ iff } (\varphi \in \omega' \text{ or } \neg\varphi \in \omega') \end{array} \right\}.$$

Definition 8 For every $\omega \in \Omega$ and $i \in I$ define $\alpha(\omega, i) := \{x \in X \mid a_i(x) \in \omega\}$.

Proposition 6 For every $\omega \in \Omega$ and $i \in I$ we have $\omega \cap \mathcal{L}_{\alpha(\omega, i)} \in \Pi_i(\omega)$.

Proof. By the Lemma on Propositional Awareness (Lemma 1) and Lemma 6, we have $a_i\varphi \in \omega$ iff $a_i(x) \in \omega$ for all $x \in \text{Pr}(\varphi)$. By maximal consistency of ω and Lemma 1, $\omega \cap \mathcal{L}_{\alpha(\omega, i)}$ has the property (ii). Note that $k_i\varphi \in \mathcal{L}_\alpha$ implies $a_i\varphi \in \mathcal{L}_\alpha$. Since $k_i\varphi \rightarrow a_i\varphi$ is theorem, $k_i\varphi \in \omega$ implies (by Lemma 6) $a_i\varphi \in \omega$. Since $k_i\varphi \rightarrow \varphi$ is a theorem, $k_i\varphi \in \omega$ implies (by Lemma 6) $\varphi \in \omega$, hence by (ii): $\varphi \in \omega \cap \mathcal{L}_{\alpha(\omega, i)}$ and $\omega \cap \mathcal{L}_{\alpha(\omega, i)}$ has property (i). Altogether, we have shown that $\omega \cap \mathcal{L}_{\alpha(\omega, i)} \in \Pi_i(\omega)$. \square

Theorem 1 For every $\omega \in \Omega$ and $i \in I$, $\Pi_i(\omega)$ is nonempty and satisfies the properties 0. - 5. of a possibility correspondence.

Proof. The nonemptiness follows from Proposition 6.

For 0., by (ii) and the Lemma on Propositional Awareness axioms we have $\Pi_i(\omega) \subseteq \Omega_{\alpha(\omega, i)}$.

Property 1. follows from $\omega \cap \mathcal{L}_{\alpha(\omega, i)} \in \Pi_i(\omega)$.

For 2., let $\omega' \in \Pi_i(\omega)$. We show first that $\Pi_i(\omega') \subseteq \Pi_i(\omega)$: Let $\omega'' \in \Pi_i(\omega')$. If $k_i\varphi \in \omega$ then, by the Axiom (4), we have, $k_i k_i\varphi \in \omega$, hence $k_i\varphi \in \omega'$ and therefore $\varphi \in \omega''$. If $a_i\varphi \in \omega$, then, we have $k_i a_i\varphi \in \omega$, therefore $a_i\varphi \in \omega'$ and hence $\varphi \in \omega''$ or $\neg\varphi \in \omega''$. Conversely, $(\varphi \in \omega'' \text{ or } \neg\varphi \in \omega'')$ implies $a_i\varphi \in \omega'$ and therefore $a_i a_i\varphi \in \omega$. From the Propositional Awareness Axioms, it follows that $a_i a_i\varphi \rightarrow a_i\varphi$ is a theorem, and hence by Modus Ponens, we have $a_i\varphi \in \omega$. We have now shown that $\Pi_i(\omega') \subseteq \Pi_i(\omega)$. For the reverse inclusion, let $\omega'' \in \Pi_i(\omega)$. First we have to show that if $k_i\varphi \in \omega'$ then $\varphi \in \omega''$. Indeed, if $k_i\varphi \in \omega'$, then $a_i k_i\varphi \in \omega$ and hence by Proposition 6 ($k_i\varphi \in \omega$ or $\neg k_i\varphi \in \omega$) and also $a_i\varphi \in \omega$ by Propositional Awareness. But $\neg k_i\varphi \in \omega$ together with $a_i\varphi \in \omega$ by Lemma 6 and PC would imply $k_i \neg k_i\varphi \in \omega$ and hence $\neg k_i\varphi \in \omega'$, a contradiction to the consistency of ω' . Therefore, we have $k_i\varphi \in \omega$ and hence $\varphi \in \omega''$. Next, if $a_i\varphi \in \omega'$, we have $a_i a_i\varphi \in \omega$, hence $a_i\varphi \in \omega$ and therefore $(\varphi \in \omega'' \text{ or } \neg\varphi \in \omega'')$. Conversely, $(\varphi \in \omega'' \text{ or } \neg\varphi \in \omega'')$ implies $a_i\varphi \in \omega$ and therefore, since $a_i\varphi \rightarrow k_i a_i\varphi$ is a theorem, we have $k_i a_i\varphi \in \omega$. But then we have $a_i\varphi \in \omega'$. Altogether we have shown that $\Pi_i(\omega) \subseteq \Pi_i(\omega')$.

Property 3. needs not to be proved since it follows from the other properties (see Remark 4).

For 4., if $k_i\varphi \in \omega \cap \mathcal{L}_\alpha = \omega_\alpha$ and $\omega' \in \Pi_i(\omega)$ then $\varphi \in \omega' \cap \mathcal{L}_\alpha$. Since $\omega' \in \Pi_i(\omega)$, we have that $a_i\varphi \in \omega \cap \mathcal{L}_\alpha$ if and only if $(\varphi \in \omega' \cap \mathcal{L}_\alpha \text{ or } \neg\varphi \in \omega' \cap \mathcal{L}_\alpha)$. Since we know that

$\omega' \cap \mathcal{L}_\alpha \in \Omega$, we have first that $\omega' \cap \mathcal{L}_\alpha \in \Pi_i(\omega_\alpha)$ and second that ω' is the inverse image of $\omega' \cap \mathcal{L}_\alpha$ under some projection.

For 5., let $\alpha \supseteq \beta \supseteq \gamma$, $\omega \in \Omega_\alpha$, $\Pi_i(\omega) \subseteq \Omega_\beta$. We have to show that $\omega' \cap \mathcal{L}_\gamma \in \Pi_i(\omega_\gamma) = \Pi_i(\omega \cap \mathcal{L}_\gamma)$. Indeed, if $k_i\varphi \in \omega \cap \mathcal{L}_\gamma$ then $\varphi \in \omega' \cap \mathcal{L}_\gamma$, and similarly if $a_i\varphi \in \omega \cap \mathcal{L}_\gamma$ then $(\varphi \in \omega' \cap \mathcal{L}_\gamma \text{ or } \neg\varphi \in \omega' \cap \mathcal{L}_\gamma)$. And if $(\varphi \in \omega' \cap \mathcal{L}_\gamma \text{ or } \neg\varphi \in \omega' \cap \mathcal{L}_\gamma)$ then $a_i\varphi \in \omega \cap \mathcal{L}_\gamma$.

Conversely, if $\omega'_\gamma \in \Pi_i(\omega_\gamma) = \Pi_i(\omega \cap \mathcal{L}_\gamma)$, then (i) when $k_i\varphi \in \omega \cap \mathcal{L}_\gamma$ we have $\varphi \in \omega'_\gamma$ and (ii) $a_i\varphi \in \omega \cap \mathcal{L}_\gamma$, $(\varphi \in \omega'_\gamma \text{ or } \neg\varphi \in \omega'_\gamma)$.

Let $\omega'_\gamma \in \Pi_i(\omega_\gamma) = \Pi_i(\omega \cap \mathcal{L}_\gamma)$. We have to show that there is a $\omega'_\beta \in \Pi_i(\omega)$ such that $(\omega'_\beta)_\gamma = \omega'_\gamma$. Consider the following set of formulas: $\Phi := \omega'_\gamma \cup \{k_i\varphi : k_i\varphi \in \omega\} \cup \{a_i\varphi : a_i\varphi \in \omega\}$. This set of formulas is contained in \mathcal{L}_β . If Φ is consistent, then it can be extended to a maximal consistent set of formulas Ψ in the language \mathcal{L}_β , and it is easy to see that $\Psi \in \Pi_i(\omega)$.

Hence it remains to show the consistency of the set Φ . Assume by contradiction that Φ is not consistent.

Note that $\{\varphi : k_i\varphi \in \omega\}$, $\{\psi : a_i\psi \in \omega\}$ and ω'_γ are all closed under conjunction, and that $(a_i \wedge_{i=1}^n \psi_i) \rightarrow \wedge_{i=1}^n a_i\psi_i$ (by the awareness axioms) is a theorem. By RK inference each $(k_i \wedge_{i=1}^m \varphi_i) \rightarrow k_j\varphi_j$ is a theorem and by Conjunction, Modus Ponens and the fact that $(\wedge_{i=1}^m (p \rightarrow q_i)) \rightarrow (p \rightarrow \wedge_{i=1}^m q_i)$ is a valid formula of PC, we infer that $(k_i \wedge_{i=1}^m \varphi_i) \rightarrow \wedge_{i=1}^m k_i\varphi_i$ is a theorem.

Hence, by Lemma 3 there are φ, ψ and χ such that $k_i\varphi \in \omega$, $a_i\psi \in \omega$ and $\chi \in \omega'_\gamma$ with $\vdash k_i\varphi \wedge a_i\psi \rightarrow \neg\chi$. Obviously, we have then (since $(p \rightarrow r) \rightarrow ((p \wedge q) \rightarrow r)$ is a tautology of PC) that $\vdash k_i\varphi \wedge a_i\psi \wedge a_i\chi \rightarrow \neg\chi$, and hence by RK-Inference that $\vdash k_i k_i\varphi \wedge k_i a_i\psi \wedge k_i a_i\chi \rightarrow k_i \neg\chi$. The formulas $k_i\varphi \rightarrow k_i k_i\varphi$, $a_i\psi \rightarrow k_i a_i\psi$ and $a_i\chi \rightarrow k_i a_i\chi$ are all theorems. By applying two times Conjunction, the fact that $((\wedge_{i=1}^n q_i) \rightarrow r) \wedge (\wedge_{i=1}^n (p_i \rightarrow q_i)) \rightarrow ((\wedge_{i=1}^n p_i) \rightarrow r)$ is a tautology of PC and by applying Modus Ponens we infer that $k_i\varphi \wedge a_i\psi \wedge a_i\chi \rightarrow k_i \neg\chi$ is a theorem.

But since $k_i\varphi, a_i\psi, a_i\chi \in \omega_\beta$ and since ω_β is a maximal consistent set of formulas in the language \mathcal{L}_β , it follows that $k_i(\neg\chi) \in \omega_\beta$. We have $k_i(\neg\chi) \in \omega_\gamma = \omega_\beta \cap \mathcal{L}_\gamma$ and hence $\neg\chi \in \omega'_\gamma$, a contradiction to $\chi \in \omega'_\gamma$. \square

By Remark 5, Proposition 4 and Theorem 1 it follows that

Corollary 1 *The tuple*

$$\underline{\Omega} := \left\langle (\Omega_\alpha)_{\alpha \subseteq X}, (r_\beta^\alpha)_{\beta \subseteq \alpha \subseteq X}, (\Pi_i)_{i \in I} \right\rangle,$$

is an unawareness structure for the set of individuals I .

The definition of the canonical unawareness structure is hence complete. We now proceed to show that the internal structure of its states is indeed reflected by operations on events in the structure. In particular, knowledge as expressed in syntactic terms within a state $(k_i\varphi)$ gets translated to knowledge of the corresponding event $(K_i[\varphi])$.

Theorem 2 For $\varphi \in \mathcal{L}$, we have

$$\begin{aligned} [\neg\varphi] &= \neg[\varphi], \\ [\varphi \wedge \psi] &= [\varphi] \cap [\psi], \\ [k_i\varphi] &= K_i[\varphi]. \end{aligned}$$

Proof. The first two items are straightforward.

For the third, suppose first that $\omega \in [k_i\varphi]$, i.e., $k_i\varphi \in \omega$. We have to show that $\Pi_i(\omega) \subseteq [\varphi]$. But this follows immediately from (i) in the definition of $\Pi_i(\omega)$.

For the reverse direction, suppose that $\omega \in K_i[\varphi]$, i.e., $\Pi_i(\omega) \subseteq [\varphi]$. In other words, $\forall \omega' \in \Pi_i(\omega)$ we have $\varphi \in \omega'$.

Suppose, by contradiction, that $k_i\varphi \notin \omega$. Since ω is maximally consistent in the sub-language built with all the primitive propositions in φ , we have $\neg k_i\varphi \in \omega$. Note that $a_i\varphi \in \omega$ by (ii). Since $a_i\varphi \rightarrow k_ia_i\varphi$ is a theorem in the language of ω , we have $k_ia_i\varphi \in \omega$. This implies $a_i\varphi \in \text{ken}_i(\omega)$, where $\text{ken}_i(\omega) := \{\psi : k_i\psi \in \omega\}$. Observe that $\{\psi : k_i\psi \in \omega\}$ is closed under conjunctions (this follows from the derived axiom (C)).

Now, if $\text{ken}_i(\omega) \cup \{\neg\varphi\}$ were inconsistent, then, by Lemma 3 and the fact that $\text{ken}_i(\omega)$ is closed under conjunctions, $\psi \rightarrow \varphi$ would be a theorem of the system for some $\psi \in \text{ken}_i(\omega)$. And then also $\psi \wedge a_i\varphi \rightarrow \varphi$ would be a theorem in the language of ω . By RK-Inference, this implies that $k_i\psi \wedge k_ia_i\varphi \rightarrow k_i\varphi$ is a theorem in the language of ω . But then we would have $k_i\varphi \in \omega$ in contrast with our assumption. Hence $\text{ken}_i(\omega) \cup \{\neg\varphi\}$ is consistent. But if $\text{ken}_i(\omega) \cup \{\neg\varphi\}$ is consistent, then we can extend $\text{ken}_i(\omega) \cup \{\neg\varphi\}$ to a maximally consistent ω' in the sub-language of the elements of $\Pi_i(\omega)$. By the definition of $\Pi_i(\omega)$, we have $\omega' \in \Pi_i(\omega)$, a contradiction to the assumption $\Pi_i(\omega) \subseteq [\varphi]$. \square

Corollary 2 For $\varphi \in \mathcal{L}$, we have $A_i[\varphi] = [a_i\varphi]$ and $U_i[\varphi] = [u_i\varphi]$.

5 Semantics, Soundness and Completeness

In this section we head to show how the canonical unawareness structure manifests that our axiom system is strongly sound and strongly complete with respect the family of unawareness structures. Strong soundness means that if a formula φ is provable from a set of formulas Γ (with the same set of atomic propositions as in φ), then φ obtains in all states of all unawareness structures in which Γ obtains. Strong completeness means the reverse implication: If φ obtains in all states of all unawareness structures in which such a Γ obtains, then φ is provable from Γ . Strong soundness and strong completeness is thus another way to formulate the sense in which the axiom system provides a foundation and substantiates the notion of an unawareness structure.

We now proceed with the formal definitions and analysis. If $\underline{\Sigma}$ is an unawareness structure, then, since the lattice is complete, there exists a state space $M(\underline{\Sigma})$, such that for all other state spaces S , we do have $S \preceq M(\underline{\Sigma})$. Note that for every event E it follows that $S(E) \preceq M(\underline{\Sigma})$.

We have to define what it means for a formula “to obtain” in a state. To this end we define below the notions of an evaluation function and a model.

For an unawareness structure $\underline{\Sigma}$ let \mathcal{E} be the collection of events. A function $v : X \longrightarrow \mathcal{E}$ is called an *evaluation function*. The set $v(x)$ is the event at which the primitive proposition x obtains.

Definition 9 *For a nonempty set X of primitive propositions, an unawareness model is a pair $\underline{\Sigma}^v := (\underline{\Sigma}, v)$, where $\underline{\Sigma}$ is an unawareness structure and v an evaluation function.*

If no confusion can arise, with some abuse of notation, we sometimes write $\underline{\Sigma}$ instead of $\underline{\Sigma}^v$.

Definition 10 *For a nonempty set X of primitive propositions and a set I of players, let $(\underline{\Sigma}, v)$ be an unawareness model, and let $s \in S$ for some $S \in \mathcal{S}$. Then we define by induction on the formation of the formulas in \mathcal{L} :*

- $(\underline{\Sigma}, s) \models \top$, for all $s \in \Sigma$,
- $(\underline{\Sigma}, s) \models x$, if $s \in v(x)$,
- $(\underline{\Sigma}, s) \models \varphi \wedge \psi$, if $s \in [\varphi] \cap [\psi]$,
- $(\underline{\Sigma}, s) \models \neg\varphi$, if $s \in \neg[\varphi]$,
- $(\underline{\Sigma}, s) \models k_i(\varphi)$, if $s \in K_i[\varphi]$,

where $[\psi] := \{s' \in \Sigma \mid (\underline{\Sigma}, s') \models \psi\}$, for every formula ψ .

Note that $[\psi]$ is an event, for every formula ψ . Note also that if φ is a formula, S' a state space such that $S([\varphi]) \not\subseteq S'$, and s a state in S' , then we have neither $(\underline{\Sigma}, s) \models \varphi$ nor $(\underline{\Sigma}, s) \models \neg\varphi$. However, for states s in the upmost state space $M(\underline{\Sigma})$, we always have either $(\underline{\Sigma}, s) \models \varphi$ or $(\underline{\Sigma}, s) \models \neg\varphi$, for every $\varphi \in \mathcal{L}$. We say that a formula φ is *defined in state s* in the unawareness model $\underline{\Sigma}$ if $s \in \bigcap_{x \in \text{Pr}(\varphi)} (v(x) \cup \neg v(x))$. It is easy to see that φ is defined in state s if and only if $s \in [\varphi] \vee \neg[\varphi]$.

Definition 11 *We say φ is true in state s , if $(\underline{\Sigma}, s) \models \varphi$. For a set of formulas Γ , we say that Γ is true in state s , if $(\underline{\Sigma}, s) \models \varphi$, for all $\varphi \in \Gamma$.*

Definition 12 *For $\Gamma \subseteq \mathcal{L}$, we say that Γ has a model if there is an unawareness model $\underline{\Sigma}$ and a state $s \in \Sigma$ such that $(\underline{\Sigma}, s) \models \Gamma$.*

The canonical structure becomes a model with the valuation function v that associates with every primitive proposition x the set of states to which x belongs.

Definition 13 For $x \in X$ define $v^\Omega(x) := \{\omega \in \Omega \mid x \in \omega\}$.

In the canonical model $\underline{\Omega}$, the properties of the canonical model make each state ω in it a model of all the formulas that ω contains.

Corollary 3 The pair $(\underline{\Omega}, v^\Omega)$ is an unawareness model such that for all $\varphi \in \mathcal{L}$:

$$(\underline{\Omega}, \omega) \models \varphi \text{ iff } \varphi \in \omega.$$

$(\underline{\Omega}, v^\Omega)$ is called the canonical unawareness model.

Proof. This follows directly from Corollary 1, Proposition 5, Definition 9, Theorem 2 and Definition 13. \square

Now we define a notion of semantic implication within an unawareness model.

Definition 14 Let $\underline{\Sigma}$ be an unawareness model. If Γ is a set of formulas and φ is a formula, we write $\Gamma \models_{\underline{\Sigma}} \varphi$ if for every state $s \in \bigcap_{x \in \text{Pr}(\varphi)} (v(x) \cup \neg v(x))$ we have that $(\underline{\Sigma}, s) \models \Gamma$ implies $(\underline{\Sigma}, s) \models \varphi$.

Definition 15 We write $\Gamma \models \varphi$ if for every unawareness model $\underline{\Sigma}$ we have $\Gamma \models_{\underline{\Sigma}} \varphi$. In this case, we say that Γ implies φ semantically. Accordingly, we write $\models \varphi$ if it is the case that $\emptyset \models \varphi$. We say that φ is valid, if $\models \varphi$.

It follows that a formula is valid if and only if it is true in all states in which it is defined.

Lemma 7 Let $\underline{\Sigma}$ be an unawareness model. Then $\Gamma \models_{\underline{\Sigma}} \varphi$ iff for all $s \in M(\underline{\Sigma})$: $(\underline{\Sigma}, s) \models \Gamma$ implies $(\underline{\Sigma}, s) \models \varphi$.

Proof. The “only if” is clear, since for all $s \in M(\underline{\Sigma})$ we do have $(\underline{\Sigma}, s) \models \varphi$ or $(\underline{\Sigma}, s) \models \neg\varphi$. For the other direction, suppose that $\Gamma \not\models_{\underline{\Sigma}} \varphi$ and that for all $s \in M(\underline{\Sigma})$: $(\underline{\Sigma}, s) \models \Gamma$ implies $(\underline{\Sigma}, s) \models \varphi$. Then there is a state space S and a state $s \in S$ such that $(\underline{\Sigma}, s) \models \Gamma$ and $(\underline{\Sigma}, s) \models \neg\varphi$. Since the projections are onto, there is a $w \in M(\underline{\Sigma})$ such that $w_S = s$. Since the set $[\psi]$ is an event, for every formula ψ , we have $(\underline{\Sigma}, w) \models \Gamma$ and $(\underline{\Sigma}, w) \models \neg\varphi$, a contradiction. \square

Definition 16 The system of axioms and inference rules is strongly sound (with respect to the class of unawareness models) if for every set of formulas Γ and every formula φ we have that $\Gamma \vdash \varphi$ implies $\Gamma \models \varphi$.

Soundness means that every formula that follows from axioms and inference rules, i.e. that is a theorem, is true in all states in all unawareness structures where this formula is defined. Strong soundness means that semantic implication follows from syntactic implication. Strong soundness is equivalent to that every set of formulas that is true in some state of some unawareness structure is consistent.

Proposition 7 *The system is strongly sound with respect to the class of unawareness models.*

Proof. To show strong soundness it is enough to show

1. that all axioms are valid formulas,
2. that the set of valid formulas is closed under RK-Inference, and
3. that for every state in every unawareness model the set of formulas that are true in that state is closed under Modus Ponens.

We first show 1:

To show that a formula φ is valid, by the above Lemma, it is enough to show that $(\underline{\Sigma}, s) \models \varphi$ for every unawareness model $\underline{\Sigma}$ and every $s \in M(\underline{\Sigma})$. It is clear (and standard) that if φ is a substitution instance of a valid formula, then we have $(\underline{\Sigma}, s) \models \varphi$ for every unawareness model $\underline{\Sigma}$ and every $s \in M(\underline{\Sigma})$. Using the definition of the relation \models , the validity of the Axiom of Truth follows from iii) of Proposition 9 and the validity of the Axiom of Positive Introspection is a consequence of iv) of Proposition 9. The validity of the first and the second Awareness Axiom follows from 6 and 7 of Proposition 10. By Proposition 11 (2.), $A_i(E) = A_i K_j(E)$. This implies, by the definition of the relation \models , the validity of the third Awareness Axiom.

Now we turn to 2: Let $\varphi_1, \varphi_2, \dots, \varphi_n$ and φ be formulas with $\text{Pr}(\varphi) \subseteq \bigcup_{i=1}^n \text{Pr}(\varphi_i)$ such that $\varphi_1 \wedge \varphi_2 \wedge \dots \wedge \varphi_n \rightarrow \varphi$ is a valid formula. Since $\text{Pr}(\varphi) \subseteq \bigcup_{i=1}^n \text{Pr}(\varphi_i)$, we have that φ is defined whenever $\varphi_1 \wedge \varphi_2 \wedge \dots \wedge \varphi_n$ is defined. Since $\varphi_1 \wedge \varphi_2 \wedge \dots \wedge \varphi_n \rightarrow \varphi$ is valid, we do have $[\varphi_1 \wedge \varphi_2 \wedge \dots \wedge \varphi_n] \subseteq [\varphi]$ in every unawareness model. By v) of Proposition 9, we do have $K_i([\varphi_1 \wedge \varphi_2 \wedge \dots \wedge \varphi_n]) \subseteq K_i([\varphi])$. By Theorem 2, the definition (of \models) and ii) of Proposition 9, we have $[k_i(\varphi_1) \wedge k_i(\varphi_2) \wedge \dots \wedge k_i(\varphi_n)] = K_i([\varphi_1]) \cap K_i([\varphi_2]) \cap \dots \cap K_i([\varphi_n]) = K_i[\varphi_1 \wedge \varphi_2 \wedge \dots \wedge \varphi_n] \subseteq K_i([\varphi]) = [k_i(\varphi)]$ in every unawareness model. This implies that $k_i\varphi_1 \wedge k_i\varphi_2 \wedge \dots \wedge k_i\varphi_n \rightarrow k_i\varphi$ is true whenever it is defined and hence that $k_i\varphi_1 \wedge k_i\varphi_2 \wedge \dots \wedge k_i\varphi_n \rightarrow k_i\varphi$ is valid.

Last we verify 3: Let $\underline{\Sigma}$ be an unawareness model, $s \in \Sigma$, $(\underline{\Sigma}, s) \models \varphi$, and $(\underline{\Sigma}, s) \models \varphi \rightarrow \psi$. We have to show that $(\underline{\Sigma}, s) \models \psi$. By the definition, we have $s \in [\varphi]$ and $s \in [\psi] \vee \neg[\varphi] \subseteq [\psi] \cup \neg[\varphi]$, hence $s \in [\psi]$ and therefore $(\underline{\Sigma}, s) \models \psi$. \square

Definition 17 *The system of axioms and inference rules is strongly complete (with respect to the class of unawareness models) if for every set of formulas Γ and every formula φ , we have that $\Gamma \models \varphi$ implies $\Gamma \vdash \varphi$.*

Completeness means that every formula that is true in all states in all spaces where it is defined is a theorem. Strong completeness means that syntactic implication follows from semantic implication. Strong completeness is equivalent to that every consistent set of formulas is true in some state in some unawareness structure.

Now, we can prove the strong characterization Theorem:

Theorem 3 (Strong soundness and strong completeness) *For $\Gamma \subseteq \mathcal{L}$ and $\varphi \in \mathcal{L}$, we have $\Gamma \vdash \varphi$ if and only if $\Gamma \models \varphi$. Furthermore, $\Gamma \vdash \varphi$ if and only if in the canonical unawareness model $\underline{\Omega}$, for every $\omega \in \Omega_\alpha$ such that $\text{Pr}(\varphi) \subseteq \alpha$, we have that $(\underline{\Omega}, \omega) \models \Gamma$ implies $(\underline{\Omega}, \omega) \models \varphi$.*

Proof. The “if” follows from soundness of our axioms and inference rules, i.e., Proposition 7.

For “only if” let $\Gamma \models \varphi$:

case 1.: Assume that $\Gamma \cup \{\neg\varphi\}$ is inconsistent. Then, by Lemma 2 there are $\psi_1, \dots, \psi_n \in \Gamma$ such that $\psi \wedge \dots \wedge \psi_n \rightarrow \neg(\neg\varphi)$ is a theorem. Hence, by Conjunction, Modus Ponens and Propositional Calculus, we have that $\Gamma \vdash \varphi$.

case 2.: Assume that $\Gamma \cup \{\neg\varphi\}$ is consistent. By Lemma 4 there is a maximal consistent set of formulas $\omega \subseteq \mathcal{L}$ with $\Gamma \cup \{\neg\varphi\} \subseteq \omega$. By definition, we have that $\omega \in \Omega_X$. By Corollary 3, for every formula $\chi \in \mathcal{L}$ we have $(\underline{\Omega}, \omega) \models \chi$ if and only if $\chi \in \omega$. It follows that $(\underline{\Omega}, \omega) \models \Gamma \cup \{\neg\varphi\}$, which is a contradiction to $\Gamma \models \varphi$. Hence we have that $\Gamma \cup \{\neg\varphi\}$ is inconsistent and therefore, by Case 1, $\Gamma \vdash \varphi$. \square

Corollary 4 *A set of formulas is consistent if and only if it has a model.*

Proof. Let Γ have a model $(\underline{\Sigma}, s)$. Since the set of states where a formula is true is an event, we can assume that $s \in M(\underline{\Sigma})$. If Γ is inconsistent, then $\Gamma \vdash \neg\varphi$ and $\Gamma \vdash \varphi$ for some formula φ . Since in s all formulas are defined, it follows by the strong soundness that $(\underline{\Sigma}, s) \models \neg\varphi$ and $(\underline{\Sigma}, s) \models \varphi$, which is impossible (by the definition of \models). Therefore Γ is consistent.

Let Γ be consistent. If Γ has no model, then by the definitions, we have that $\Gamma \models \neg\top$ and $\Gamma \models \top$. But then, by the strong completeness, we have $\Gamma \vdash \neg\top$ and $\Gamma \vdash \top$ and Γ is inconsistent, a contradiction. Hence, Γ has a model. \square

In the entire analysis above, the notion of a proof of φ from Γ allowed for using, along the proof, formulas with atomic propositions neither in φ nor in Γ . With theorem 3 in hand, we can now refine this notion, and show that theorem 3 obtains, in fact, even if the notion of proof does not involve formulas outside the sub-language of φ and Γ . To this end, we start with several definitions.

Definition 18 *Let Y be a set of primitive propositions. The set of \mathcal{L}_Y -theorems is the smallest set of formulas in \mathcal{L}_Y that contain all the axioms (that is all the substitution instances of valid formulas of Propositional Calculus, Truth, the Propositional Awareness Axioms and Axiom (4)) and that is closed under instances of the inference rules Modus Ponens and RK-Inference which mention only formulas in \mathcal{L}_Y .*

Definition 19 *Let Y be a set of primitive propositions, let Γ be a set of formulas in \mathcal{L}_Y and φ a formula in \mathcal{L}_Y . An \mathcal{L}_Y -proof of φ from Γ is a finite sequence of*

\mathcal{L}_Y —formulas such that the last formula is φ and such that each formula is a formula in Γ , an \mathcal{L}_Y —theorem or inferred from the previous formulas by an instance of Modus Ponens with formulas in \mathcal{L}_Y . If there is an \mathcal{L}_Y —proof of φ from Γ , then we write $\Gamma \vdash_{\mathcal{L}_Y} \varphi$. In particular, $\vdash_{\mathcal{L}_Y} \varphi$ means that φ is an \mathcal{L}_Y —theorem.

Definition 20 Let Y be a set of primitive propositions. A set of formulas $\Gamma \subseteq \mathcal{L}_Y$ is \mathcal{L}_Y —consistent if and only if there is no formula $\varphi \in \mathcal{L}_Y$ such that $\Gamma \vdash_{\mathcal{L}_Y} \varphi$ and $\Gamma \vdash_{\mathcal{L}_Y} \neg\varphi$. A set $\Gamma \subseteq \mathcal{L}_Y$ of formulas is \mathcal{L}_Y —inconsistent, if it is not \mathcal{L}_Y —consistent.

With these definitions, we have:

Proposition 8 Let $Y \supseteq Z$ be sets of primitive propositions and let $\Gamma \subseteq \mathcal{L}_Z$. Then Γ is \mathcal{L}_Z —consistent if and only if Γ is \mathcal{L}_Y —consistent.

Proof. Obviously, \mathcal{L}_Y —consistency of Γ implies \mathcal{L}_Z —consistency of Γ .

By Corollary 4 (applied to the case $X = Z$), the \mathcal{L}_Z —consistency of Γ implies that there is a unawareness structure $\underline{\Sigma}$, an evaluation function $v_Z : Z \rightarrow \mathcal{E}$ and a state $s \in \Sigma$ such that $(\underline{\Sigma}^{v_Z}, s) \models \Gamma$. Define $v_Y : Y \rightarrow \mathcal{E}$ by $v_Y(x) := v_Z(x)$ for $x \in Z$, and $v_Y(x)$ arbitrary for $x \in Y \setminus Z$. Then it is easy to see that $(\underline{\Sigma}^{v_Y}, s) \models \Gamma$. By Corollary 4 (applied to the case $X = Y$), it follows that Γ is \mathcal{L}_Y —consistent. \square

Corollary 5 Let $Y \subseteq X$ and let $\Gamma \subseteq \mathcal{L}_Y$ and $\varphi \in \mathcal{L}_Y$. Then we have $\Gamma \vdash_{\mathcal{L}_Y} \varphi$ if and only if $\Gamma \models \varphi$. Furthermore, this is the case if and only if in the canonical unawareness model $\underline{\Omega}$, for every $\omega \in \Omega_\alpha$ such that $\text{Pr}(\varphi) \subseteq \alpha$, we have that $(\underline{\Omega}, \omega) \models \Gamma$ implies $(\underline{\Omega}, \omega) \models \varphi$.

A Further Properties of the Unawareness Structure

The following proposition is proved in Heifetz, Meier and Schipper (2005):

Proposition 9 The Knowledge operator K_i has the following properties:

- (i) *Necessitation:* $K_i(\Sigma) = \Sigma$,
- (ii) *Conjunction:* $K_i(\bigcap_{\lambda \in L} E_\lambda) = \bigcap_{\lambda \in L} K_i(E_\lambda)$,
- (iii) *Truth:* $K_i(E) \subseteq E$,
- (iv) *Positive Introspection:* $K_i(E) \subseteq K_i K_i(E)$,
- (v) *Monotonicity:* $E \subseteq F$ implies $K_i(E) \subseteq K_i(F)$,
- (vi) $\neg K_i(E) \cap \neg K_i \neg K_i(E) \subseteq \neg K_i \neg K_i \neg K_i(E)$.

Proposition 9 says that the knowledge operator has all the strong properties of knowledge in partitioned information structures, except for the weakening (vi) of the negative introspection property. Negative introspection – the property $\neg K_i(E) \subseteq K_i \neg K_i(E)$ that when an individual does not know an event, she knows she does not know it – obtains only when the individual is also aware of the event (see property 5 of the next proposition).

The following proposition is proved in Heifetz, Meier and Schipper (2005):

Proposition 10 *The following properties of knowledge and awareness obtain:*

1. *KU Introspection:* $K_i U_i(E) = \emptyset^{S(E)}$,
2. *AU Introspection:* $U_i(E) = U_i U_i(E)$,
3. *Weak Necessitation:* $A_i(E) = K_i \left(S(E)^\uparrow \right)$,
4. *Strong Plausibility:* $U_i(E) = \bigcap_{n=1}^\infty (\neg K_i)^n(E)$,
5. *Weak Negative Introspection:* $\neg K_i(E) \cap A_i \neg K_i(E) = K_i \neg K_i(E)$,
6. *Symmetric:* $A_i(E) = A_i(\neg E)$,
7. *A-Conjunction:* $\bigcap_{\lambda \in L} A_i(E_\lambda) = A_i \left(\bigcap_{\lambda \in L} E_\lambda \right)$,
8. *AK-Self Reflection:* $A_i(E) = A_i K_i(E)$,
9. *AA-Self Reflection:* $A_i(E) = A_i A_i(E)$,
10. *A-Introspection:* $A_i(E) = K_i A_i(E)$.

Properties 1. to 4. have been proposed by Dekel, Lipman and Rustichini (1998), properties 6. to 9. by Modica and Rustichini (1999), and properties 5. to 9. by Fagin and Halpern (1988) and Halpern (2001). *A*-Introspection is the property that an individual is aware of an event if and only if she knows she is aware of it.

Define an alternative awareness operator,

$$\tilde{A}_i(E) := \{\omega \in \Sigma : \Pi_i(\omega) \subseteq S(E)^\uparrow\}$$

if there is a state ω such that $\Pi_i(\omega) \subseteq S(E)^\uparrow$, and by

$$\tilde{A}_i(E) = \emptyset^{S(E)}$$

otherwise. Then by weak necessitation, following remark is immediate:

Remark 6 *If E is an event, then $A_i(E) = \tilde{A}_i(E)$.*

Common knowledge can be defined in the usual way (see Aumann, 1999):

Definition 21 The “everybody knows” operator on events is defined by

$$K(E) = \bigcap_{i \in I} K_i(E).$$

The common knowledge operator on events is defined by

$$CK(E) = \bigcap_{n=1}^{\infty} K^n(E).$$

Analogously we can define common awareness.

Definition 22 The “everybody is aware” operator on events is defined by

$$A(E) = \bigcap_{i \in I} A_i(E),$$

and the common awareness operator on events by

$$CA(E) = \bigcap_{n=1}^{\infty} (A)^n(E).$$

The following proposition extends Remark 4 of Heifetz, Meier and Schipper (2005).

Proposition 11 Let E be an event. The following multi-agent properties obtain:

1. $A_i(E) = A_i A_j(E)$,
2. $A_i(E) = A_i K_j(E)$,
3. $K_i(E) \subseteq A_i K_j(E)$,
4. $A(E) = K(S(E)^\uparrow)$,
5. $A(E) = CA(E)$,
6. $K(E) \subseteq A(E)$,
7. $CK(E) \subseteq CA(E)$,
8. $CK(S(E)^\uparrow) \subseteq CA(E)$.

Proof. 1. By Proposition 3 and the definition of awareness operator, $S(E) = S(A_j(E))$. Hence by weak necessitation, $A_i(E) = A_i A_j(E)$.

2. By Proposition 3, $S(E) = S(K_j(E))$. Hence by weak necessitation, $A_i(E) = A_i K_j(E)$.

3. By definition of awareness, $K_i(E) \subseteq A_i(E)$. By 2., $K_i(E) \subseteq A_i(E) = A_i K_j(E)$.
4. By weak necessitation, $A(E) = \bigcap_{i \in I} A_i(E) = \bigcap_{i \in I} K_i(S(E)^\uparrow) = K(S(E)^\uparrow)$.
5. By 4. we have $A(E) = K(S(E)^\uparrow)$. By Proposition 3 $K_i(S(E)^\uparrow)$ is an $S(E)$ -based event for all $i \in I$. Hence $\bigcap_{i \in I} K_i(S(E)^\uparrow) = K(S(E)^\uparrow)$ is $S(E)$ -based. Hence $S(A(E)) = S(K(S(E)^\uparrow)) = S(E)$. Therefore $AA(E) = K(S(A(E))^\uparrow) = K(S(E)^\uparrow) = A(E)$. By induction it follows that $A^{n+1}(E) = AA^n(E) = AA(E) = A(E)$ for all $n = 1, 2, \dots$. Hence $CA(E) = \bigcap_{n=1,2,\dots} A^n(E) = A(E)$.
6. By definition, $s \in K(E)$ iff $s \in K_i(E)$ for all $i \in I$. Hence by definition of awareness, we have $s \in A_i(E)$ for all $i \in I$, and thus $s \in A(E)$.
7. By definition, $CK(E) \subseteq K(E)$. By 6., $CK(E) \subseteq A(E)$. By 5., $CK(E) \subseteq CA(E)$.
8. By definition, $CK(S(E)^\uparrow) \subseteq \bar{K}(S(E)^\uparrow)$. By 4., $CK(S(E)^\uparrow) \subseteq A(E)$. By 5., $CK(S(E)^\uparrow) \subseteq CA(E)$. \square

References

- [1] Aumann, Robert (1999). Interactive epistemology I: Knowledge, *International Journal of Game Theory* **28**, 263-300.
- [2] Chellas, Brian F. (1980). *Modal logic: An introduction*, Cambridge: Cambridge University Press.
- [3] Dekel, Eddie, Lipman, Bart and Aldo Rustichini (1998). Standard state-space models preclude unawareness, *Econometrica* **66**, 159-173.
- [4] Fagin, Ronald, Halpern, Joseph Y., Moses, Yoram and Moshe Y. Vardi (1995). *Reasoning about knowledge*, Cambridge, M.A.: MIT Press.
- [5] Fagin, Ronald and Joseph Y. Halpern (1988). Belief, awareness and limited reasoning, *Artificial Intelligence* **34**, 39-76.
- [6] Halpern, Joseph Y. (2001). Alternative semantics for unawareness, *Games and Economic Behavior* **37**, 321-339.
- [7] Halpern, Joseph Y. and Leandro Rêgo (2005). Interactive unawareness revisited, in *Theoretical Aspects of Rationality and Knowledge: Proc. Tenth Conference (TARK 2005)*; full version appears at <http://arxiv.org/abs/cs.AI/0509058>.
- [8] Heifetz, Aviad, Meier, Martin and Burkhard C. Schipper (2005). Interactive unawareness, *Journal of Economic Theory*, forthcoming.
- [9] Modica, Salvatore and Aldo Rustichini (1999). Unawareness and partitioned information structures, *Games and Economic Behavior* **27**, 265-298.

- [10] Modica, Salvatore and Aldo Rustichini (1994). Awareness and partitional information structures, *Theory and Decision* **37**, 107-124.
- [11] Wittgenstein, Ludwig (1922). *Tractatus logico-philosophicus*, London: Routledge.