

INTERACTIVE UNAWARENESS*

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

February 2005

Abstract

The standard state-spaces of asymmetric information preclude non-trivial forms of unawareness (Modica and Rustichini, 1994, Dekel, Lipman and Rustichini, 1998). We introduce a generalized state-space model that allows for non-trivial unawareness among several individuals, and which satisfies strong properties of knowledge as well as all the desiderata on unawareness proposed this far in the literature.

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

Keywords: unawareness, awareness, knowledge, interactive epistemology, speculative trade, bounded perception.

*We thank Eddie Dekel, Christian Ewerhart, Simon Grant, Joe Halpern, Bart Lipman, Salvatore Modica, conference participants in TARK 2003, NASEM 2003, EEA 2003, Stony Brook 2003, ECCE 1, FUR 2004, GAMES 2004 and ESEM 2004 as well as seminar participants at Barcelona, Bonn, Brown, Heidelberg, Jerusalem, Madrid, Rice and Rochester for helpful comments on earlier drafts. Some parts of the material appeared without proofs in our paper “Multi-Person Unawareness” prepared for the conference TARK 2003.

[†]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; Financial support during the preparation of earlier drafts from the European Union via a Marie-Curie Individual Fellowship (HPMF-2002-02073) is gratefully acknowledged.

[§]Department of Economics, University of Bonn. Email: schipper@uni-bonn.de; Financial support from the Deutsche Forschungsgemeinschaft through SFB/TR 15 and the DAAD is gratefully acknowledged.

1 Introduction

It is hard to argue that decision makers are aware of all facts affecting the outcome of their decisions. Thus unawareness is a rather natural state of mind and its role merits investigation, especially in interactive decision making. Yet modeling unawareness proves to be a tricky task.

Geanakoplos (1989) suggested using non-partitional information structures to this effect. In such a model one can have states in which an individual doesn't know an event and is ignorant of her ignorance. However, Modica and Rustichini (1994) and Dekel, Lipman and Rustichini (1998) show that unawareness operators satisfying certain fundamental properties allow only for a trivial notion of unawareness in such structures. Namely, if an agent is unaware of anything, then he is unaware of everything and knows nothing. More generally, they showed that no standard information structure can capture adequately the notion of unawareness.¹

Modica and Rustichini (1999) suggest an enhanced structure in order to model unawareness of an individual. It consists of an "objective" space, describing the world with the full vocabulary, and a "subjective" space for the sub-vocabulary of which the agent is aware. When an individual is unaware of an event, the states she considers as possible belong to a subjective space in which this event cannot be described. Halpern (2001) offers an alternative formulation with one space but two different knowledge operators – implicit knowledge and explicit knowledge. Halpern (2001) proves that a particular kind of his awareness structures is equivalent to the Modica-Rustichini structure as a semantics for a modal syntax that includes both a knowledge and an awareness modality.

Both these approaches suffer from the following limitations. First, they involve an explicit use of the modal syntax *within* the semantic structures. This limits the audience that is capable of applying this machinery to specific problems. Just as the short paper by Aumann (1976) introduced to economists the partitional state-spaces as a logic-free tool to model knowledge, and was thus seminal to a large body of consecutive work in Economics, the analogue of such a presentation is still lacking for unawareness. Second, only one-person unawareness is treated explicitly both by Modica and Rustichini (1999) and Halpern (2001).

In an independent, parallel work, Li (2004) presents a set-theoretic version of a variant of the Modica and Rustichini (1999) model, and extends it also to the multi-person case. That extension involves an explicit specification of each player's state of mind about others' state of mind, about their state of mind about others, and so forth.

This complication exemplifies that unlike in the case of knowledge, in which the passage from the single-person case to the multi-person case involves no substantial

¹Ewerhart (2001) suggests a way to model unawareness in a standard information structure. However, in his modeling if an individual is unaware of an event then she believes its negation. While this property may be suitable for some aspect or view of unawareness, it is incompatible with all the other formal approaches cited here, as well as with the approach of the current contribution.

problems, the modeling of multi-person unawareness is more intricate. An individual i may be unaware of some issue, and may further be uncertain whether another individual j is aware of yet another issue (out of those issues of which i is aware). Furthermore, this uncertainty need not be correlated with the quality of i 's information about the issues of which she is aware. To model this appropriately, one needs an explicit *ordered structure of spaces*, where the possibility set of an individual in a state of one space may reside in another space, while the possibility set of a different individual in one of these possible states may reside in yet another space.

To wit, we consider a complete lattice of state-spaces accompanied by suitable projections among them. The partial order of spaces indicates the strength of their expressive power. The possibility set of an individual in a state of one space may reside in a less-expressive space. A crucial feature of the model is that it limits the subsets (of the union of all spaces) which are considered as events – those that can be “known” or be the object of awareness. The special structure of events is natural, in the sense that it is the same as that of subsets of states in which a particular proposition obtains – if states were to consist of maximally-consistent sets of propositions in an appropriate logical formulation.² In particular, in our setting the negation of an event is different from its set-theoretic complement. As a result, there are states that belong neither to an event nor to its negation. When the possibility set of an individual consists of such states, the individual is unaware of the event.

While our model of unawareness is presented in the following section, we apply interactive unawareness to an example of speculative trade in section 3. The so called “No-Trade-Theorems” (e.g., Milgrom and Stokey, 1982) show that when individuals know what they know and they are never certain of false statements, common knowledge of rationality precludes speculative trade. This is to be contrasted with e.g. the huge volume of daily trade in currency exchange, most of which is purely speculative. We show in a simple example that when combined with unawareness, these strong properties of knowledge and rationality *are* compatible with speculative trade. We conclude in section 4. All proofs are presented in the appendix.

2 Model

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

For every S and S' such that $S' \succeq S$ (“ S' is more expressive than S – states of S' describe situations with a richer vocabulary than states of S ”),³ there is a surjective projection $r_S^{S'} : S' \rightarrow S$, where r_S^S is the identity. (“ $r_S^{S'}(\omega)$ is the restriction of the

²We show this formally in a companion work (in preparation).

³Here and in what follows, phrases within quotation marks hint at intended interpretations, but we emphasize that these interpretations are not part of the definition of the set-theoretic structure.

description ω to the more limited vocabulary of S .”) 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 $\omega \in S'$, denote $\omega_S = r_S^{S'}(\omega)$. If $B \subseteq S'$, denote $B_S = \{\omega_S : \omega \in B\}$.

Denote by $g(S) = \{S' : S' \succeq S\}$ the set of spaces that are at least as expressive as S . For $B \subseteq S$, denote by $B^\dagger = \bigcup_{S' \in g(S)} (r_S^{S'})^{-1}(B)$ all the “extensions of descriptions in B to at least as expressive vocabularies.”

A subset E of Σ is an *event* if it is of the form B^\dagger for some $B \subseteq S$, where $S \in \mathcal{S}$. In such a case we call B the *basis* of the event E , and S the *base-space* of E , denoted by $S(E)$. Hence not every subset of Σ is an event.

If B^\dagger is an event where $B \subseteq S$, the negation $\neg B^\dagger$ of B^\dagger is defined by $(S \setminus B)^\dagger$. This is typically a proper subset of the complement $\Sigma \setminus B^\dagger$.

Intuitively, there may be states in which the description of an event E is both expressible and valid – these are the states in E ; there may be states in which this description is expressible but invalid – these are the states in $\neg E$; and there may be states in which neither this description nor its negation are expressible – these are the states in $\Sigma \setminus (E \cup \neg E) = \Sigma \setminus S(E)^\dagger$. Thus our structure is not a standard state-space model in the sense of Dekel, Lipman, and Rustichini (1998), since their “real states” assumption precludes such events.

If $B \neq \emptyset$ and $B \neq S$ for some $S \in \mathcal{S}$, then $\neg\neg B^\dagger = B^\dagger$, but otherwise it is not necessarily the case. To circumvent this, for each space $S \in \mathcal{S}$ we devise a distinct vacuous event \emptyset^S , and define $\neg S^\dagger = \emptyset^S$ and $\neg\emptyset^S = S^\dagger$. In fact, our notation of an event E is just the shorthand for (E, S) , where S is the base-space. No confusion arises if E is nonempty, otherwise we write \emptyset^S for (\emptyset, S) . The event \emptyset^S should be interpreted as a “logical contradiction phrased with the expressive power available in S .” 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. $S(E) = S(F)$.

If $\{B_\lambda^\dagger\}_{\lambda \in L}$ is a set of events (with $B_\lambda \subseteq S_\lambda$, for $\lambda \in L$), their conjunction $\bigwedge_{\lambda \in L} B_\lambda^\dagger$ is just the intersection $\bigcap_{\lambda \in L} B_\lambda^\dagger$ (we will therefore use the conjunction symbol \wedge and the intersection symbol \cap interchangeably). If $S = \sup_{\lambda \in L} S_\lambda$, then⁴ this conjunction is $\left(\bigcap_{\lambda \in L} \left((r_{S_\lambda}^S)^{-1}(B_\lambda)\right)\right)^\dagger$.

The disjunction of $\{B_\lambda^\dagger\}_{\lambda \in L}$ is defined by the de Morgan law $\bigvee_{\lambda \in L} B_\lambda^\dagger = \neg \left(\bigwedge_{\lambda \in L} \neg (B_\lambda^\dagger)\right)$. Typically $\bigvee_{\lambda \in L} B_\lambda^\dagger \subsetneq \bigcup_{\lambda \in L} B_\lambda^\dagger$, and $\bigvee_{\lambda \in L} B_\lambda^\dagger = \bigcup_{\lambda \in L} B_\lambda^\dagger$ holds if and only if all the B_λ^\dagger have the same base-space. Intuitively, if two events are described in distinct vocabularies, the disjunction of the events is expressible only in a vocabulary which is at least as rich as both vocabularies, but not necessarily in either vocabulary alone.

⁴Since \mathcal{S} is a *complete* lattice, $\sup_{\lambda \in L} S_\lambda$ exists.

Example 1. Let Φ be a set of facts. For $\alpha \subseteq \Phi$, let $S_\alpha = \{\omega : \omega = \{true, false\}^\alpha\}$. I.e., a state in S_α is a string indicating which facts in α are true and which are false. $S_\alpha \preceq S_{\alpha'}$ whenever $\alpha \subseteq \alpha'$. Consider for instance a set of three facts $\Phi = \{p, q, r\}$. For example, we write $\omega = (p, \neg q)$ for a state in $S_{\{p,q\}}$ in which the fact p is true and q is false. Clearly, we have $g(S_\emptyset) = \mathcal{S}$, $g(S_{\{p,q,r\}}) = \{S_{\{p,q,r\}}\}$ and e.g. $g(S_{\{r\}}) = \{S_{\{r\}}, S_{\{p,r\}}, S_{\{q,r\}}, S_{\{p,q,r\}}\}$. Figure 1 illustrates the state-spaces with the states. The projections are indicated by arrows (for clarity we do not consider in this figure any compositions of projections and the identity maps). Consider now the event that fact r is true “[r is true]”. The base-space is $S_{\{r\}}$, the basis of this event is $\{(r)\} \subset S_{\{r\}}$. Considering all extensions of $\{(r)\}$ we obtain the event

$$\{(r)\}^\uparrow = \{(r), (p, r), (\neg p, r), (q, r), (\neg q, r), (p, q, r), (p, \neg q, r), (\neg p, q, r), (\neg p, \neg q, r)\} = [r \text{ is true}].$$

This is the set of states in which fact r obtains. In Figure 1 the event [r is true] is indicated by the union of the dotted rectangles. The event that r is false [r is false] is the negation

$$\begin{aligned} \neg[r \text{ is true}] &= (S_{\{r\}} \setminus \{(r)\})^\uparrow = \\ &= \{(\neg r), (p, \neg r), (\neg p, \neg r), (q, \neg r), (\neg q, \neg r), (p, q, \neg r), (p, \neg q, \neg r), (\neg p, q, \neg r), (\neg p, \neg q, \neg r)\}. \end{aligned}$$

In Figure 1 it is indicated by the union of the grey rectangles. It becomes obvious that $[r \text{ is true}] \cup \neg[r \text{ is true}] \subsetneq \Sigma$. I.e., there are states such as (q) which belong neither to $[r \text{ is true}]$ nor $\neg[r \text{ is true}]$.

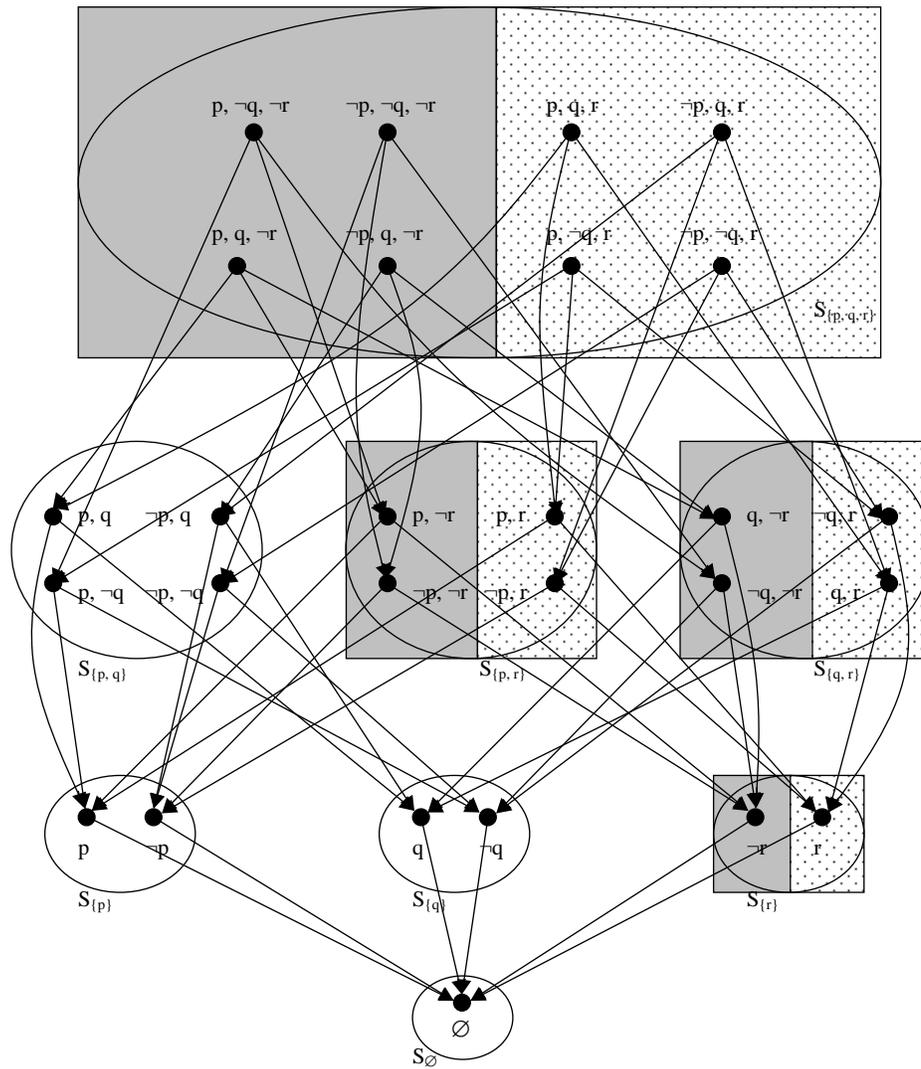
I is the set of individuals. For each individual $i \in I$ there is a possibility correspondence $\Pi_i : \Sigma \rightarrow 2^\Sigma \setminus \emptyset$ with the following properties:

0. Confinedness: If $\omega \in S$ then $\Pi_i(\omega) \subseteq S'$ for some $S' \preceq S$.
1. Generalized Reflexivity: $\omega \in \Pi_i^\uparrow(\omega)$ for every $\omega \in \Sigma$.⁵
2. Stationarity: $\omega' \in \Pi_i(\omega)$ implies $\Pi_i(\omega') = \Pi_i(\omega)$.
3. Projections Preserve Awareness: If $\omega \in S'$, $\omega \in \Pi_i(\omega)$ and $S \preceq S'$ then $\omega_S \in \Pi_i(\omega_S)$.
4. Projections Preserve Ignorance: If $\omega \in S'$ and $S \preceq S'$ then $\Pi_i^\uparrow(\omega) \subseteq \Pi_i^\uparrow(\omega_S)$.
5. Projections Preserve Knowledge: If $S \preceq S' \preceq S''$, $\omega \in S''$ and $\Pi_i(\omega) \subseteq S'$ then⁶ $(\Pi_i(\omega))_S = \Pi_i(\omega_S)$.

⁵Here and in what follows, we abuse notation slightly and write $\Pi_i^\uparrow(\omega)$ for $(\Pi_i(\omega))^\uparrow$.

⁶We could have assumed \supseteq and deduce $=$ from \supseteq , 3., and the other properties.

Figure 1: State-Spaces, Projections, and Event Structure in Example 1



Confinedness means that the states an individual considers as possible in a given state ω are all “described with the same vocabulary – the vocabulary available to the individual at ω .”

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 2); Stationarity will guarantee the introspection properties (that an individual knows what she knows – property (iv) in Proposition 2, and that an individual knows what she ignores provided she is aware of it – property 5. in Proposition 3).

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 ω and its projection ω_S , (“the restriction of the description ω to the more restricted vocabulary available in S ”). The properties guarantee that after this projection/restriction the individual learns nothing she did not know before, does not forget anything she knew (provided that it can be expressed with the restricted vocabulary available in S), and does not become aware of new facts, or unaware of facts of which she was aware (here again, provided that these facts can be expressed with the restricted vocabulary available in S).

Remark 1 *Property 1 implies that if $S' \preceq S$, $\omega \in S$ and $\Pi_i(\omega) \subseteq S'$, then $r_{S'}^S(\omega) \in \Pi_i(\omega)$.*

Remark 2 *Property 4 and Confinedness imply that if $S' \preceq S$, $\omega \in S$ and $\Pi_i(\omega_{S'}) \subseteq S''$, then $\Pi_i(\omega) \subseteq S^*$ for some S^* with $S'' \preceq S^*$.*

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

Definition 1 *The knowledge operator of individual i on events is defined, as usual, by*

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

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

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

otherwise.

Proposition 1 *If E is an event, then $K_i(E)$ is an $S(E)$ -based event.*

Proposition 2 *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 2 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 “everybody knows” operator on events is defined by

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

The common knowledge operator on events is defined by

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

The unawareness operator of individual i from events to events is now defined by⁷

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

and the awareness operator is then naturally defined by

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

By Proposition 1 and the definition of the negation, we have

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

Proposition 3 *The following properties of knowledge and awareness obtain:*

1. *KU Introspection*: $K_i U_i(E) = \emptyset^{S(E)}$

⁷This 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.

2. *AU Introspection*: $U_i(E) = U_i U_i(E)$
3. *Weak Necessitation*: $A_i(E) = K_i(S(E)^\uparrow)$
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. *Symmetry*: $A_i(\neg E) = A_i(E)$
7. *A-Conjunction*: $\bigcap_{\lambda \in L} A_i(E_\lambda) = A_i(\bigcap_{\lambda \in L} E_\lambda)$
8. *AK-Self Reflection*: $A_i K_i(E) = A_i(E)$
9. *AA-Self Reflection*: $A_i A_i(E) = A_i(E)$
10. *A-Introspection*: $K_i A_i(E) = A_i(E)$

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 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.

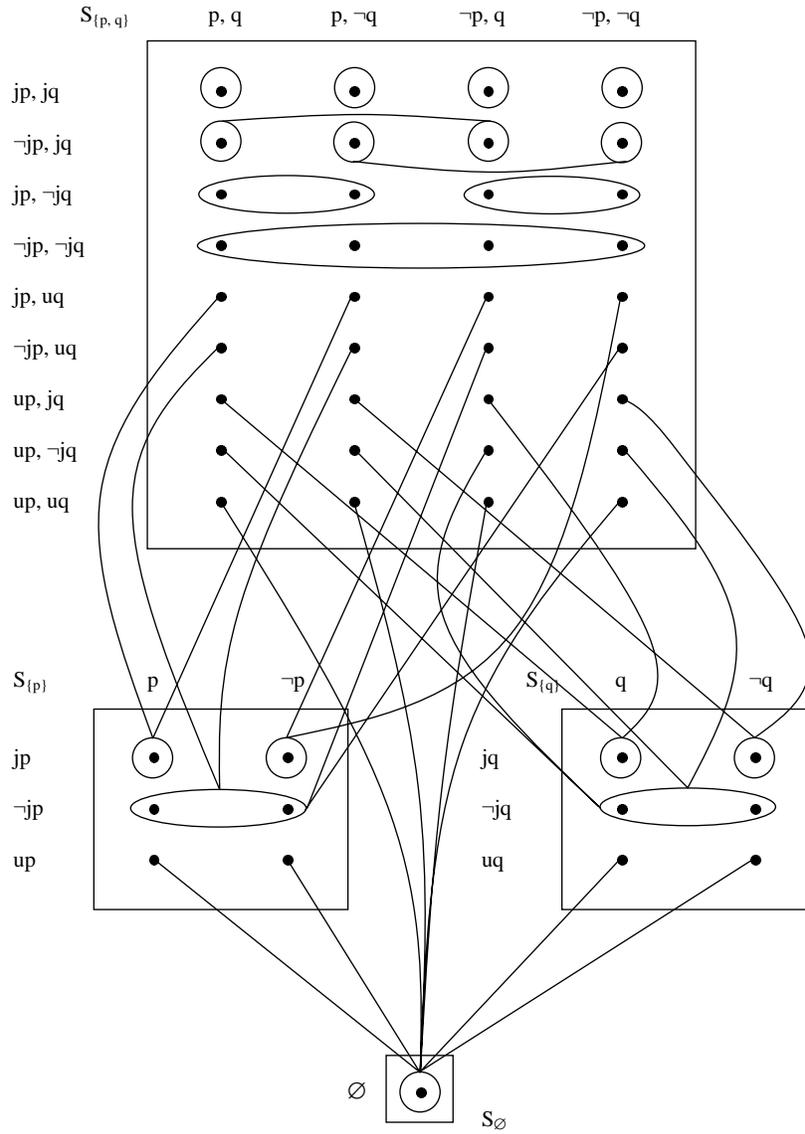
Remark 4 *Note that by Proposition 1 and Weak Necessitation, $A_i A_j(E) = A_i(E)$. In analogy with the “everybody knows” and the “common knowledge” operators we can define “everybody is aware” and “common awareness” operators. Note that by Proposition 1 and Weak Necessitation, when everybody is aware of an event E then everybody is also aware that everybody is aware of E . It then follows that the events “everybody is aware of E ” and “common awareness of E ” coincide.*

Remark 5 *Our unawareness operator is defined on events. However, this does not mean that we model unawareness of events only. Let an issue or question (e.g. “Is it raining?”) be such that it can be answered with a fact (“It is raining.”) or with the negation of the fact (“It is not raining.”). By symmetry, an individual is aware of an event if and only if she is aware of its negation. Thus, we model the awareness of questions and issues rather than just single events. Indeed, by weak necessitation, an individual is aware of an event if and only if she is aware of any event that can be expressed in the space with the same expressive power.*

Example 2. Consider a language with two basic propositions p, q and one individual with a knowledge modality k . Consider further the structure with four spaces $\mathcal{S} = \{S_{\{p,q\}}, S_{\{p\}}, S_{\{q\}}, S_\emptyset\}$ as indicated in Figure 2 by rectangles. To describe in a compact fashion the information of the individual in each state, we use the “knowing whether”⁸

⁸See Hart, Heifetz and Samet (1996).

Figure 2: One-person Awareness in Example 2



modality j defined by $jx \equiv kx \vee k\neg x$. For a proposition x , the proposition jx means “the individual knows whether x or $\neg x$ obtain.” The unawareness modality u is defined by $ux \equiv \neg kx \wedge \neg k\neg kx$.

For simplicity, each state is described by the basic propositions that hold in this state as well as by the propositions describing the information of the individual at that state. Thus we present in Figure 2 each state-space in a matrix-style. For example, the state (jp, jq, p, q) means that p and q obtain, and that the individual knows whether p and knows whether q . This of course implies that the individual knows p and q as reflected by the singleton possibility set. For each state ω , the possibility set $\Pi(\omega)$ of the individual is indicated by circles or ovals, some connected by lines. Other lines relate non-reflexive states (i.e., states ω such that $\omega \notin \Pi(\omega)$) to their possibility sets.

For a proposition x we denote by $[x]$ the set of states in which x obtains. Using the possibility correspondence Π in Figure 2 and the knowledge operator K from definition 1, we can build events such as $K[x]$, $\neg K[x]$, $K\neg K[x]$, $\neg K\neg K[x]$ and $U[x]$.

Negative introspection fails for non-reflexive states. To see this consider the event $[p]$, i.e., all states in which p obtains. It is easy to see that $(up, jq, p, q) \in \neg K[p]$. Since $(up, jq, p, q) \notin K\neg K[p]$, negative introspection fails. Moreover, also $K\left(S_{\{p,q\}}^\uparrow\right) = S_{\{p,q\}}^\uparrow$ fails since for instance $(up, jq, p, q) \in S_{\{p,q\}}^\uparrow$ but $(up, jq, p, q) \notin K\left(S_{\{p,q\}}^\uparrow\right)$. However, all the properties of Propositions 2 and 3 hold.

The example can also serve to highlight the difference between this model and the Generalized Standard Model (GSM) of Modica and Rustichini (1999) (which Halpern (2001) proves to be isomorphic to a particular kind of the Awareness Structures of Fagin and Halpern (1988)). In the GSM corresponding to this example, only the projections from the upper-most space $S_{\{p,q\}}$ to the other spaces $S_{\{p\}}, S_{\{q\}}, S_\emptyset$ would be defined, but not the projections among the lower spaces. More importantly, the states in the last rows of $S_{\{p\}}$ and $S_{\{q\}}$ (the two states in the row up of $S_{\{p\}}$ and the two states in the row uq of $S_{\{q\}}$) *do not exist in the corresponding GSM*. Indeed, these states do not belong to any possibility set of the individual in the states of the space $S_{\{p,q\}}$ of full descriptions of states of the world, and are hence redundant when the discussion is restricted to a single individual. However, it is exactly this kind of extra states that are needed in order to capture *interactive unawareness*, e.g. a situation in which one individual believes that another individual is unaware of something of which she herself *is* aware. This will become apparent in the following example, which explicitly features several individuals.

3 Example: Speculative Trade

Consider an owner o of a firm and a potential buyer b . To make this example interesting, we assume that the agents’ awareness differs. That is, we assume that there is a state such that the possibility sets of the agents reside in different spaces at that state. For

instance, the owner is aware that there might be a lawsuit $[l]$ involving the firm but he is unaware of a potential innovation or novelty $[n]$ enhancing the value of the firm. In contrast, the buyer is aware that there might be an innovation but unaware of the lawsuit.

Similarly to Example 2, Figure 3 presents the information structure graphically. The state spaces $\mathcal{S} = \{S_{\{n,l\}}, S_{\{n\}}, S_{\{l\}}, S_{\emptyset}\}$ are indicated by dotted rectangles. For instance, in space $S_{\{n\}}$ the event innovation $[n]$ can be expressed but not the event lawsuit $[l]$. As in Example 2, we use for convenience the “knowing whether” operator, j_b and j_o being the operator for the buyer and the owner, respectively. Then the ovals with horizontal lines indicate the possibility sets of the buyer, whereas the ones with vertical lines are those of the owner. A solid line connects a buyer’s non-reflexive state to its possibility set, whereas a dotted line corresponds to the owner.

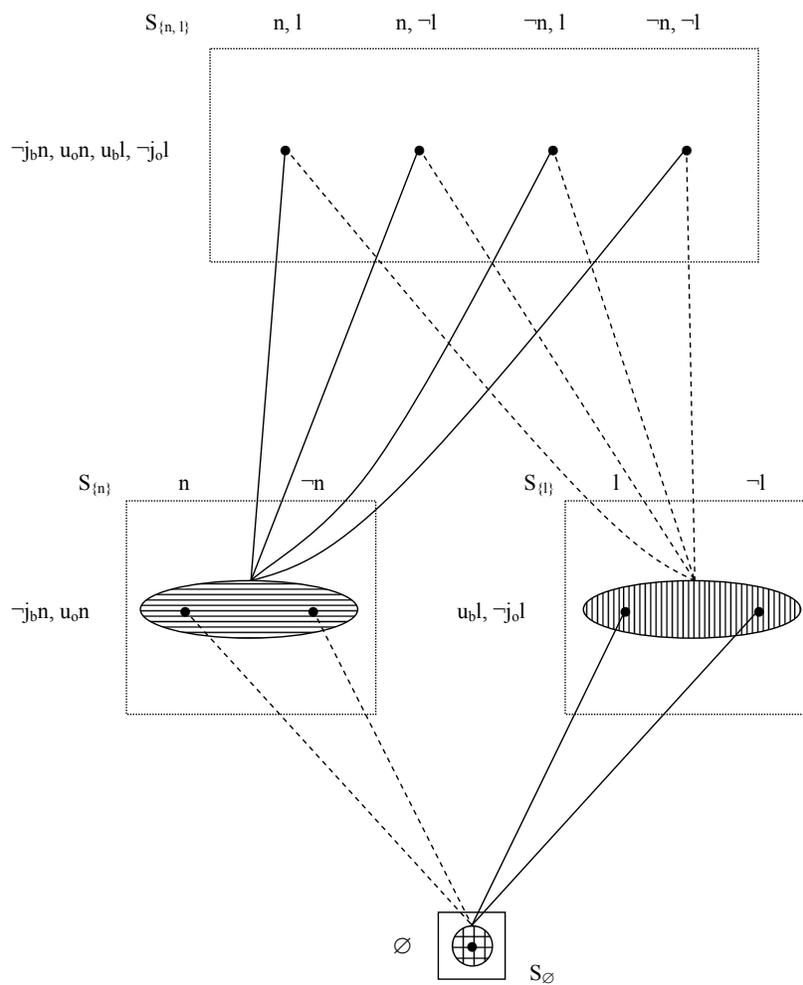
Consider for example the state $\omega = (\neg j_b n, u_o n, u_b l, \neg j_o l, n, l)$. At this state the buyer’s possibility set resides in $S_{\{n\}}$, whereas the owner’s one is in $S_{\{l\}}$. Hence the buyer is unaware of a lawsuit, $\omega \in U_b[l]$, and the owner of an innovation, $\omega \in U_o[n]$. The possibility sets are such that $\omega \in \neg K_b[n]$ but $\omega \in A_b[n]$ and similarly $\omega \in \neg K_o[l] \cap A_o[l]$.

Let the status quo value of the firm be 100 Taler. I.e., at the state (\emptyset) the value of the firm is 100 Taler. Suppose further, that if an innovation obtains, it raises the value of the firm by 10 Taler, whereas the implications of a lawsuit reduce the value by 10 Taler. Since at ω the buyer is aware of the event *innovation* $[n]$ but unaware of the event *lawsuit* $[l]$, the value of the firm to her is either 110 Taler in the event $[n]$ or 100 Taler if $\neg[n]$ obtains. At the same state, the owner values the firm at either 90 Taler in the event $[l]$ or at 100 Taler if $\neg[l]$ obtains.

We assume that agents are both rational in the sense of maximizing their respective payoffs, and that both agents know that. I.e., we introduce the mild assumption that if at all states an agent considers as possible the price is at least x , and in some of these states the price is strictly higher than x , then the agent strictly prefers to buy at the price x than not buying at x . Similarly, if at all states the agent considers as possible the price is at most x , and in some of these states the price is strictly lower than x , then the agent strictly prefers to sell at the price x than not selling at x . If, on the other hand, the price is exactly x in all the states that the agent considers as possible, then the agent is indifferent between trading or not at the price x . We will say that an agent is willing to trade at x if either she strictly prefers to trade at x or she is indifferent between trading or not at x .

Suppose now that the buyer offers to buy the firm from the owner for an amount of 100 Taler. Clearly, the buyer is willing to do that because she values the firm at 110 Taler (if $[n]$ obtains) or 100 Taler (if $\neg[n]$ obtains). Thus, she strictly prefers to buy at 100 Taler. The buyer also can expect that the owner is going to sell to her, since she believes the owner is unaware of an innovation that could enhance the value of the firm. In particular, she believes that the owner’s possibility set at $(\neg j_b n, u_o n, n)$ or $(\neg j_b n, u_o n, \neg n)$ resides in the space S_{\emptyset} , the owner’s valuation of the firm at state (\emptyset) being 100 Taler. Moreover,

Figure 3: Information Structure in the Example of Speculative Trade



the owner accepts the buyer's offer, since the former values the firm at 90 Taler (if $[l]$ obtains) or 100 Taler (if $\neg[l]$ obtains). He strictly prefers to sell at 100 Taler. To the owner, the buyer's offer is rational, since the owner believes that the buyer's possibility set at $(u_b l, \neg j_o l, l)$ or $(u_b l, \neg j_o l, \neg l)$ is in the space S_\emptyset , the buyer's valuation of the firm at state (\emptyset) being 100 Taler. So in this example, the agents trade, each expecting to make a strict positive gain and compensating the other with the status quo value.

Formally, in all states of the upper-most space $S_{\{n,l\}}$ both agents strictly prefer to trade at the price 100. Moreover, in all states of all spaces both agents are willing to trade at the price 100, and hence this fact is common knowledge among them. Thus, in all the states of $S_{\{n,l\}}$ there is both strict preference for trade and common knowledge of willingness to trade.

Such a state of affairs is *impossible to model in standard information structures* in which the knowledge operators K_i satisfy properties (i)-(v) of proposition 2 (i.e., all the properties of a partitioned information structure except, possibly, for the negative introspection property $\neg K_i(E) \subseteq K_i \neg K_i(E)$). Indeed, in a standard information structure Ω with possibility correspondences $(\Pi_i : \Omega \rightarrow 2^\Omega \setminus \emptyset)_{i=1,2}$ for the agents, there would be common knowledge at a state $\omega \in \Omega$ that both agents are willing to trade at the price x if and only if there would be a self-evident event $E \subseteq \Omega$ (i.e., satisfying $\Pi_i(\omega') \subseteq E$ for each $\omega' \in E$, $i = 1, 2$) with $\Pi_i(\omega) \subseteq E$ for $i = 1, 2$, such that both agents are willing to trade at x in all the states of E . The truth property $K_i(E) \subseteq E$ is equivalent to the property $\omega' \in \Pi_i(\omega')$ (reflexivity). This property would imply that the value in every $\omega' \in E$ is at least x (since the buyer is willing to buy at x in $\omega' \in E$, and ω' is one of the states the buyer considers as possible at ω'), and similarly the value in all the states of E is at most x , since the seller is willing to sell at x . It follows that the value would be exactly x in all the states $\omega' \in E$. But then, since $\Pi_i(\omega) \subseteq E$, it would not be the case that at ω each of the agents also *strictly prefers* to trade at x .

What would happen if we were to “flatten” the model, and consider the union of all states in all spaces of our unawareness model 3 as one state-space, while retaining the possibility correspondences? We would then get a standard non-partitional information model, in which reflexivity ($\omega \in \Pi_i(\omega)$) and hence the truth property ($K_i(E) \subseteq E$) fail. This may be interpreted as delusion on the part of the individuals – at some states they consider an entirely different set of states as possible.

On one hand, it is known that speculative trade is possible in such a model (Geanakoplos, 1989). However, it is also clear that the Dekel, Lipman and Rustichini (1998) critique would apply to the resulting model – it would only allow for a trivial notion of unawareness.

Our event structure rules out delusions – for every event E and every state $\omega \in E$, an individual does not “know” (believe, respectively) the negation $\neg E$, since $\Pi_i(\omega) \not\subseteq \neg E$. The individual's frame of mind at ω is still consistent with ω in the sense that she can not believe in facts that do not obtain at ω , but she may perceive less facts than

actually obtain at ω . Our unawareness structure is thus useful when we want to model *unawareness* as the driving force for an economic phenomenon, rather than *mistakes* in information processing.

4 Conclusion

Scientists were unaware of gravity until Newton conceived it. Mathematicians are now unaware of tomorrow’s proof-techniques for today’s long-standing conjectures. Some investors are unaware of financial market regularities that other investors exploit. In the large spectrum of such examples, unawareness is conceptually distinct from ignorance, incomplete information or faulty analysis: It has to do with the lack of conception, not the lack of knowledge.

To emphasize this distinction, we presented a tractable model of interactive unawareness, in which individuals are nevertheless introspective and non-deluded. Dekel, Lipman and Rustichini (1998) proved that no standard information space can truly capture the notion of unawareness. Accordingly, our model features an ordered set of spaces, with appropriate projections and inter-relations.

We interpret the order relation “ \preceq ” among spaces as ordering the expressive power or the richness of vocabulary with which states or situations are described. In a companion work we develop this idea formally. Starting with a multi-person epistemic logic with unawareness and a suitable axiom system, we show that the canonical structure built of the maximally-consistent sets of propositions in this system (for sub-languages corresponding to subsets of atomic propositions) is indeed an awareness structure as in section 2, each of whose states is a model for the propositions of which it consists.

Alternative (though less formal) interpretations of the order relation “ \preceq ” may depend on the motivation and reasons for unawareness. An individual may be unaware because of bounded perception or some form of resource boundedness. For instance, perception is studied (though less formally) in cognitive psychology. This literature suggests that perception is guided by mental models or categorization. A mental model is an individual representation of the world (Johnson-Laird, 1983). Mental models may differ in terms of comprehensiveness, motivating an order relation of expressive power. Categorization is suggested to guide a human’s perception by filtering observations (Goldstone and Kersten, 2002). Resource boundedness as source for incomplete knowledge of the relevant aspects of an individual’s environment was suggested by Simon (1955).

Reasoning takes time and effort, and it is computationally hard to find the best description. Thus if computational resources run out, individuals may arrive at different descriptions of the world, and in this sense may be unaware of the descriptions other people use. Such an argument is developed formally in Aragones, Gilboa, Postlewaite, and Schmeidler (2003).

We hope that our model will be helpful for developing applications of unawareness and bounded perception. Conceivable applications include the implications of unawareness to agreement, Dutch books, consumption behavior, emergence of novelty, insurance, inconceivable contingencies in (incomplete) contracting etc. This shall be left to future research.

A Appendix

A.1 Proof of Proposition 1

$K_i(E)$ is an event if there exists a space $S \in \mathcal{S}$ with a subset $B \subseteq S$ s.t. $B^\dagger = K_i(E)$.

Assume that $K_i(E)$ is non-empty. Choose $\omega \in K_i(E)$. We have $\omega \in K_i(E)$ iff $\Pi_i(\omega) \subseteq E$. By Generalized Reflexivity, it follows that $\omega \in E$. Since E is an event, there exists a unique base-space $S(E)$. It follows that $\omega_{S(E)} \in E$. Note that by Confinedness, $\Pi_i(\omega) \subseteq S$, for some $S \succeq S(E)$. Thus $(\Pi_i(\omega))_{S(E)}$ is defined. Moreover, $(\Pi_i(\omega))_{S(E)} \subseteq E \cap S(E)$. By Projections Preserve Knowledge, we have $\Pi(\omega_{S(E)}) \subseteq E$.

Define $B = \bigcup \{\Pi_i(\omega) : \Pi_i(\omega) \subseteq E \cap S(E)\}$. We first show that $B = K_i(E) \cap S(E)$. Note, that by the definition of B and by Stationarity, we have $B \subseteq K_i(E) \cap S(E)$. We also have $B \supseteq K_i(E) \cap S(E)$. Indeed, if $\omega \in K_i(E) \cap S(E)$, then by Confinedness $\Pi_i(\omega) \subseteq S$ for some $S \succeq S(E)$, and by Generalized Reflexivity $S(E) \succeq S$, implying together that $\Pi_i(\omega) \subseteq S(E)$. Therefore, by Generalized Reflexivity, $\omega \in \Pi_i(\omega)$. Since $\omega \in K_i(E)$, that is $\Pi_i(\omega) \subseteq E$, and since $\Pi_i(\omega) \subseteq S(E)$, we have $\omega \in \Pi_i(\omega) \subseteq E \cap S(E)$, that is $\omega \in B$.

We now have to show that $B^\dagger = K_i(E)$. Let $\omega \in B^\dagger$, that is $\omega \in S$ for some $S \succeq S(E)$ and $\omega_{S(E)} \in B$. By the definition of B , $\omega_{S(E)} \in \Pi_i(\omega')$ for some ω' such that $\Pi_i(\omega') \subseteq E \cap S(E)$. By Stationarity we therefore have $\Pi_i(\omega_{S(E)}) = \Pi_i(\omega') \subseteq B$. By Remark 2, it follows that $\Pi_i(\omega) \subseteq S'$, for some $S' \succeq S(E)$. Therefore $(\Pi_i(\omega))_{S(E)}$ is defined, and by Projections Preserve Knowledge, we have $(\Pi_i(\omega))_{S(E)} = \Pi_i(\omega_{S(E)}) \subseteq E$. Since E is an event, it follows that $(\Pi_i(\omega)) \subseteq E$ and hence $\omega \in K_i(E)$.

In the reverse direction, let $\omega \in K_i(E)$, that is $\Pi_i(\omega) \subseteq E$. By Confinedness, we have $\Pi_i(\omega) \subseteq S$, for some $S \succeq S(E)$, and by Generalized Reflexivity $\omega \in S'$ for some $S' \succeq S$. Hence $(\Pi_i(\omega))_{S(E)}$ is defined. Since E is a $S(E)$ -based event, we have $(\Pi_i(\omega))_{S(E)} \subseteq E \cap S(E)$. By Projections Preserve Knowledge, we have $\Pi_i(\omega_{S(E)}) = (\Pi_i(\omega))_{S(E)} \subseteq E \cap S(E)$, and therefore $\Pi_i(\omega_{S(E)}) \subseteq B$. By Generalized Reflexivity and the fact that $\omega_{S(E)} \in S(E)$, we have $\omega_{S(E)} \in \Pi_i(\omega_{S(E)}) \subseteq B$ and hence $\omega \in B^\dagger$.

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

A.2 Proof of Proposition 2

(i) $K_i(\Sigma) = \Sigma$ follows directly from the definition of K_i .

(ii) We have $\omega \in K_i\left(\bigcap_{\lambda \in L} E_\lambda\right)$ iff $\Pi_i(\omega) \subseteq \bigcap_{\lambda \in L} E_\lambda$ iff $\Pi_i(\omega) \subseteq E_\lambda$, for all $\lambda \in L$ iff $\omega \in K_i(E_\lambda)$, for all $\lambda \in L$ iff $\omega \in \bigcap_{\lambda \in L} K_i(E_\lambda)$.

(iii) Let $\omega \in K_i(E)$, that is $\Pi_i(\omega) \subseteq E$. Since E is an event, $\Pi_i^\dagger(\omega) \subseteq E$. By Generalized Reflexivity, $\omega \in \Pi_i^\dagger(\omega)$. Hence $\omega \in E$. In the case of $K_i(E) = \emptyset^{S(E)}$, we trivially have $K_i(E) \subseteq E$.

(iv) Let $\Pi_i(\omega) \subseteq E$ and $\omega' \in \Pi_i(\omega)$. We have to show that $\omega' \in K_i(E)$, that is $\Pi_i(\omega') \subseteq E$. But by Stationarity we have $\Pi_i(\omega') = \Pi_i(\omega) \subseteq E$. So we have shown that $K_i(E) \subseteq K_i K_i(E)$ in case $K_i(E)$ is not empty. If $K_i(E) = \emptyset^{S(E)}$, then $K_i(E) \subseteq K_i K_i(E)$ since by Proposition 1 $K_i K_i(E)$ is $S(E)$ -based.

(v) Monotonicity follows directly from the definition of K_i .

(vi) By the definition of the Unawareness operator and Strong Plausibility in Proposition 3, we have $\neg K_i(E) \cap \neg K_i \neg K_i(E) = U_i(E) = \bigcap_{n=1}^{\infty} (\neg K_i)^n(E) \subseteq \neg K_i \neg K_i \neg K_i(E)$. (Note that property (vi) of Proposition 2 will neither be used in the proof of Lemma 1, nor in the proof of Proposition 3.) \square

A.3 Proof of Proposition 3

Lemma 1 *Let E and F be events with the same base-space S . Then*

$$K_i(F \vee K_i(E)) = K_i(F \cup K_i(E)) = K_i(F) \cup K_i(E).$$

Proof of the Lemma: By Proposition 2, we have $K_i(K_i(E)) = K_i(E)$. Since, by Proposition 1 all the events in the lemma are S -based, \vee is equal to \cup , the set-theoretic union.

By the monotonicity of the K_i -operator, we have $K_i(F \cup K_i(E)) \supseteq K_i(F) \cup K_i(K_i(E)) = K_i(F) \cup K_i(E)$.

Conversely, let $\omega \in K_i(F \cup K_i(E))$.

1. case: Let $\omega' \in \Pi_i(\omega) \cap K_i(E)$. Since $\omega' \in K_i(E)$ it follows that $\Pi_i(\omega') \subseteq E$. But by Stationarity, we have $\Pi_i(\omega) = \Pi_i(\omega')$ and hence $\omega \in K_i(E)$.
2. case: $\Pi_i(\omega)$ and $K_i(E)$ are disjoint. Since $\omega \in K_i(F \cup K_i(E))$, we must have $\Pi_i(\omega) \subseteq F$ and hence $\omega \in K_i(F)$.

Thus we have shown that $K_i(F \cup K_i(E)) \subseteq K_i(F) \cup K_i(E)$. \square

Proof of Proposition 3: For convenience, the proof of the properties follows a different order than in the statement of the Proposition.

(1.) $K_i U_i(E) = K_i(\neg K_i(E) \cap \neg K_i \neg K_i(E)) = K_i \neg K_i(E) \cap K_i \neg K_i \neg K_i(E) \subseteq K_i \neg K_i(E) \cap \neg K_i \neg K_i(E) = \emptyset^{S(E)}$.

(3.) $A_i(E) = K_i(E) \cup K_i \neg K_i(E) = K_i(K_i(E)) \cup K_i(\neg K_i(E)) = K_i(S(E)^\uparrow)$, where the last equality follows from Lemma 1 above.

(2.) $U_i(E) = U_i U_i(E)$ is equivalent to $A_i U_i(E) = A_i(E)$. $A_i U_i(E) = K_i U_i(E) \cup K_i \neg K_i U_i(E)$. By KU-Introspection and Weak Necessitation, the last term is equal to $\emptyset^{S(E)} \cup K_i(\neg \emptyset^{S(E)}) = K_i(S(E)^\uparrow) = A_i(E)$. Thus $U_i(E) = U_i U_i(E)$.

(6.) Since $S(E) = S(\neg E)$, we have by Weak Necessitation that $A_i(\neg E) = A_i(E)$.

(5.) By Symmetry and the properties of the knowledge operator, $\neg K_i(E) \cap A_i \neg K_i(E) = \neg K_i(E) \cap A_i K_i(E) = \neg K_i(E) \cap (K_i K_i(E) \cup K_i \neg K_i K_i(E)) = (\neg K_i(E) \cap K_i K_i(E)) \cup (\neg K_i(E) \cap K_i \neg K_i K_i(E)) \subseteq (\neg K_i(E) \cap K_i(E)) \cup (\neg K_i(E) \cap K_i \neg K_i(E)) = \emptyset^{S(E)} \cup (\neg K_i(E) \cap K_i \neg K_i(E)) \subseteq K_i \neg K_i(E)$. The reverse direction follows by Truth and the definition of awareness.

(4.) By the definition of U_i we have $\bigcap_{n=1}^{\infty} (\neg K_i)^n(E) \subseteq \neg K_i(E) \cap \neg K_i \neg K_i(E) = U_i(E)$. It therefore remains to prove the reverse inclusion $U_i(E) \subseteq \bigcap_{n=1}^{\infty} (\neg K_i)^n(E)$, which, since the left-hand-side and the right-hand-side of the inclusion are both $S(E)$ -based events, is equivalent to $\bigvee_{n=1}^{\infty} K_i((\neg K_i)^{n-1}(E)) = \bigcup_{n=1}^{\infty} K_i((\neg K_i)^{n-1}(E)) \subseteq A_i(E)$ (Since, again, all the involved events are $S(E)$ -based, the disjunction and union operators coincide.)

The proof proceeds by induction. If $n = 1$, then $K_i((\neg K_i)^{1-1}(E)) = K_i(E) \subseteq A_i(E)$. If $n = 2$, then $K_i((\neg K_i)^{2-1}(E)) = K_i \neg K_i(E) \subseteq A_i(E)$.

For the induction step, we show that if $K_i((\neg K_i)^{n-1}(E)) \subseteq A_i(E)$, then $K_i((\neg K_i)^{n+1}(E)) \subseteq A_i(E)$. Set $F = (\neg K_i)^{n-1}(E)$. By Weak Negative Introspection, and the fact that all the events occurring here are $S(E)$ -based, we have $(\neg K_i)^2(F) \subseteq K_i(F) \cup U_i \neg K_i(F)$. By Monotonicity of the K_i -operator, and Lemma 1, it follows that $K_i[(\neg K_i)^2(F)] \subseteq K_i[K_i(F) \cup U_i \neg K_i(F)] = K_i(F) \cup K_i U_i(\neg K_i(E))$. Applying KU-Introspection, we obtain $K_i(F) \cup K_i U_i(\neg K_i(E)) = K_i(F) \cup \emptyset^{S(E)} = K_i(F)$. By the induction hypothesis, $K_i(F) \subseteq A_i(E)$.

(7.) By Weak Necessitation, Proposition 2, and the fact that $\bigcap_{\lambda \in L} (S(E_\lambda)^\uparrow) = S(\bigcap_{\lambda \in L} (E_\lambda)^\uparrow)$, we have $\bigcap_{\lambda \in L} A_i(E_\lambda) = \bigcap_{\lambda \in L} K_i(S(E_\lambda)^\uparrow) = K_i(\bigcap_{\lambda \in L} S(E_\lambda)^\uparrow) = K_i(S(\bigcap_{\lambda \in L} E_\lambda)^\uparrow) = A_i(\bigcap_{\lambda \in L} E_\lambda)$.

(8.) $A_i K_i(E) = K_i K_i(E) \cup K_i \neg K_i K_i(E)$. By Positive Introspection last term equals $K_i K_i(E) \cup K_i \neg K_i(E)$. Applying again Positive Introspection yields $K_i(E) \cup K_i \neg K_i(E) = A_i(E)$.

(9.) By Weak Necessitation, $A_i(E) = K_i(S(E)^\uparrow) = A_i(F)$, for any event F with $S(F) = S(E)$. Set $F = A_i(E)$. Hence $A_i(E) = A_i A_i(E)$.

(10.) By Weak Necessitation, we have $A_i(E) = K_i(S(E)^\uparrow)$. By (iii) and (iv) in Proposition 2, we have $K_i(S(E)^\uparrow) = K_i K_i(S(E)^\uparrow)$, and hence $A_i(E) = K_i A_i(E)$ obtains. \square

References

- [1] Aragonés, Enriqueta, Gilboa, Itzhak, Postlewaite, Andrew and David Schmeidler (2003). Fact-free learning, mimeo.
- [2] Aumann, Robert (1976). Agreeing to disagree, *Annals of Statistics* **4**, 1236-1239.
- [3] Dekel, Eddie, Lipman, Bart and Aldo Rustichini (1998). Standard state-space models preclude unawareness, *Econometrica* **66**, 159-173.
- [4] Ewerhart, Christian (2001). Heterogeneous awareness and the possibility of agreement, mimeo., University of Mannheim.
- [5] Fagin, Ronald and Joseph Y. Halpern (1988). Belief, awareness and limited reasoning, *Artificial Intelligence* **34**, 39-76.
- [6] Geanakoplos, John (1989). Games theory without partitions, and applications to speculation and consensus, Cowles Foundation Discussion Paper No. 914.
- [7] Goldstone, Robert and Alan Kersten (2003). Concepts and categorization, forthcoming in: Healy, A.F. and R.W. Proctor (eds.). *Comprehensive handbook of psychology, Vol. IV: Experimental psychology*, New York: Wiley, 599-621.
- [8] Halpern, Joseph Y. (2001). Alternative semantics for unawareness, *Games and Economic Behavior* **37**, 321-339.
- [9] Hart, Sergiu, Heifetz, Aviad and Dov Samet (1996). “Knowing whether”, “knowing that”, and the cardinality of state spaces, *Journal of Economic Theory* **70**, 249-256.
- [10] Johnson-Laird, P. N. (1983). *Mental models - Towards a cognitive science of language, inference and consciousness*, Cambridge et al.: Cambridge University Press.
- [11] Li, Jing (2004). Unawareness, mimeo., University of Pennsylvania.
- [12] Milgrom, Paul and Nancy Stokey (1982). Information, trade and common knowledge, *Journal of Economic Theory* **26**, 17-27.
- [13] Modica, Salvatore and Aldo Rustichini (1999). Unawareness and partitional information structures, *Games and Economic Behavior* **27**, 265-298.
- [14] Modica, Salvatore and Aldo Rustichini (1994). Awareness and partitional information structures, *Theory and Decision* **37**, 107-124.
- [15] Simon, Herbert (1955). A behavioral model of rational choice, *Quarterly Journal of Economics* **69**, 99-118.