STANDARD STATE SPACE MODELS OF UNAWARENESS

PETER FRITZ AND HARVEY LEDERMAN

Abstract. The impossibility theorem of Dekel, Lipman and Rustichini [1998] has been thought to demonstrate that standard state-space models cannot be used to represent unawareness. We first show that Dekel, Lipman and Rustichini do not establish this claim. We then distinguish three notions of awareness, and argue that although one of them may not be adequately modeled using standard state spaces, it remains open whether standard state spaces can be used to provide models of the other two notions. In fact, standard space models of these forms of awareness are attractively simple. We illustrate this by describing a class of standard state space models which represent key features of awareness. We prove completeness and decidability for the logic of these models, show how propositional quantifiers can be added to our logic, and sketch how standard techniques from decision theory can be implemented in our models in a way which allows for speculative trade.

1. Introduction

Dekel, Lipman and Rustichini [1998, hereafter, "DLR"] claim to show that "standard state space models preclude unawareness". Their claim has achieved the status of orthodoxy, and has led to the development of a variety of novel, complex models of unawareness.[1] The first task of this paper is to clear the way for standard state space models of unawareness by showing that the formal result DLR present does not establish their headline conclusion. DLR informally motivate certain axioms concerning unawareness, but in their formal impossibility result, they rely on the claim that these axioms hold at all states in the model. As section 2 argues, the assumption that axioms hold at all states of the model is unwarranted; in fact, DLR themselves reject it. While DLR's formal results are valid, they are not sufficiently general to rule out standard state space models of unawareness. As we show, the impossibility results do not hold if one assumes only that DLR's explicit assumptions about unawareness hold at some "real" states, as opposed to at all states. Even strengthening those explicit assumptions considerably does not reinstate the results.

But this does not yet vindicate standard state space models of unawareness. Section 3 presents a novel impossibility result which uses widely shared assumptions

Date: April 24, 2015.

Thanks to Eddie Dekel, Bart Lipman and Timothy Williamson for very helpful discussions of drafts of this paper. Peter Fritz was supported by an AHRC/Scatcherd European Scholarship and the German Academic Exchange Service; Harvey Lederman acknowledges the support of the Clarendon Fund.

[1]See, e.g., [Schipper, 2014b, p. 2], Schipper [forthcoming], [Heifetz et al., 2006, p. 78], [Heifetz et al., 2008, p. 305], [Heifetz et al., 2013, p. 101], [Meier and Schipper, 2014, p. 220], [Karni and Vierø, 2013, p. 2790], [Li, 2009, pp. 977–978], [Heinsalu, 2012, p. 2454], [Heinsalu, 2014, p. 257], [Galanis, 2013, p. 42], [Sillari, 2008, p. 516], and Sillari [2006].

about awareness, and only requires the truth of these assumptions. In particular, the new impossibility theorem relies on the assumption that an agent who is aware of a conjunction is aware of its conjuncts. If awareness satisfies this assumption, then standard state space models may in fact preclude unawareness. But the assumption is not appropriate for all forms of awareness. We distinguish three notions of awareness, and suggest that two important ones do not satisfy this assumption, leaving open the possibility that they could be adequately modeled by standard state space models.

The remainder of the paper turns to a positive proposal. We describe a simple class of standard state space models which represent key features of awareness. In section 4, we establish completeness and decidability for the logic of these models. We also show that adding propositional quantifiers, a topic which has presented major difficulties for existing approaches to awareness, is straightforward in our standard state space models. In section 5, we present one way of implementing a choice-based approach to decision theory within these models, and show how non-trivial unawareness is consistent with speculative trade. Section 6 concludes.

2. DLR's Triviality Result

2.1. **Standard State Space Models.** Standard state space models for the knowledge and awareness of a single agent can be understood as certain tuples $\langle\Omega, k, a\rangle$. Ω is required to be a set, called the set of *states*, from which a set of events is derived by taking an *event* to be a set of states. k and a are functions on events, which represent the agent's knowledge and awareness, respectively: k maps each event E to the event $k(E)$ of the agent knowing E; a similarly takes each E to the event $a(E)$ of the agent being aware of E.

Such models are straightforwardly used to interpret a formal language in which one can talk about knowledge and awareness. Let L be such a language built up from proposition letters $p, q, \ldots$, using a unary negation operator $\neg$, a binary conjunction connective $\wedge$ and two unary operators K and A, respectively ascribing knowledge and awareness to the agent. Formulas of this language are interpreted relative to a model $M = \langle\Omega, k, a\rangle$ and a valuation function v which maps each proposition letter p to the event $v(p)$. This interpretation uses a function $[\![\cdot]\!]_{M,v}$ which maps each formula φ of L to the event expressed by φ in M, which can be understood as the set of states in which φ is true in M. To state the constraints on such a function let $-E = \Omega \backslash E$.

$$[\![p]\!]_{M,v} = v(p)$$
$$[\![\neg\varphi]\!]_{M,v} = -[\![\varphi]\!]_{M,v}$$
$$[\![\varphi \wedge \psi]\!]_{M,v} = [\![\varphi]\!]_{M,v} \cap [\![\psi]\!]_{M,v}$$
$$[\![K\varphi]\!]_{M,v} = k([\![\varphi]\!]_{M,v})$$
$$[\![A\varphi]\!]_{M,v} = a([\![\varphi]\!]_{M,v})$$

The agent being unaware of something can of course be understood as it not being the case that she is aware of it. We therefore syntactically use $U\varphi$ as an abbreviation for $\neg A\varphi$. Similarly, we introduce the other connectives of classical propositional logic as abbreviations, using $\vee$ for disjunction, $\rightarrow$ for material implication, $\leftrightarrow$ for bi-implication, and $\top$ and $\bot$ for an arbitrary tautology and contradiction, respectively. On the semantic side, we adopt the convention of writing fg for the composition of functions f and g, which allows us to write, e.g., $k - a(E)$ instead of $k(-(a(E)))$.

In order to express general constraints on these models, we say that a formula φ is *valid on* M if $[\![\varphi]\!]_{M,v} = \Omega$ for each valuation function v; this can be understood as requiring φ to be true in every state of M according to every valuation function. In order to limit this constraint to a particular state $\omega \in \Omega$, we say that φ is *valid in* ω if $\omega \in [\![\varphi]\!]_{M,v}$ for each valuation function v.

These models count as "standard" models in the sense of DLR. First, the events expressed by $A\varphi$ and $K\varphi$ are each a function of the event expressed by φ. (DLR call this "event-sufficiency".) Second, negation is interpreted as set-complement and conjunction as intersection, so that all tautologies of classical propositional logic, such as $p \vee \neg p$, are interpreted as the set of all states in every model. (DLR call this assumption "real states".)

2.2. **DLR on Standard State Space Models.** DLR introduce three constraints on awareness, which can be stated using the following three axioms:

Plausibility	$Up \to (\neg Kp \wedge \neg K \neg Kp)$
KU-Introspection	$\neg KUp$
AU-Introspection	$Up \to UUp$

Their constraints on knowledge can be stated using the following three axioms:

Necessitation:	$K\top$
Monotonicity:	$K(p \wedge q) \to (Kp \wedge Kq)$
Weak Necessitation:	$Kp \to K\top$

Their main results are then:

Theorem 1 (DLR)**.** *Let $M = \langle \Omega, k, a \rangle$ be a model on which Plausibility, KU-Introspection and AU-Introspection are valid.*

1(i): *If Necessitation is valid on M, then Ap is valid on M.*
1(ii): *If Monotonicity is valid on M, then $Kp \to Aq$ is valid on M.*
2: *If Weak Necessitation is valid on M, then $Kp \to Aq$ and $Ap \leftrightarrow Aq$ are valid on M.*

Our presentation of DLR's result differs in superficial respects from their original presentation. DLR do not present their constraints in terms of the validity of certain axioms. Thus, for example, instead of requiring KU-Introspection to be valid on M, they require $k - a(E) = \emptyset$ for all events E. However, it is a routine exercise to show that this condition is equivalent to the validity of our corresponding axiom. The same point holds for the other axioms. In short, our later models will not be escaping their triviality result by a sleight of hand which depends on this presentation. (Indeed, in an earlier working paper, we presented our models and DLR's axioms in DLR's setting.)

One reason for the variant presentation is that it will facilitate the later exposition. It also serves to demonstrate that standard state space models as discussed here are equivalent to what are now commonly known as neighborhood or Scott-Montague frames (see Scott [1970] and Montague [1970]). It is well known that given certain restrictions on the function interpreting knowledge, this function can be turned into a binary relation among states along the lines of those used by Kripke [1963] and Hintikka [1962]. This representation as a binary relation is, in turn, formally interchangeable with the "possibility correspondences" as introduced by Aumann [1976] (see also Aumann [1999]) and used throughout economic theory.[2]

[2]For this interchangeability, see, e.g. Fagin et al. [1995].

2.3. **Two Kinds of States.** In response to their triviality results, DLR suggest distinguishing informally between "real" states and "subjective" states. As we understand it, this distinction can be explained as follows. An epistemic model makes predictions about how an agent or group of agents will or would behave in particular situations. The model makes predictions about these situations by including states which represent them. The real states in a model are the states which represent situations the model is intended to describe. The model predicts an agent will behave a certain way in a particular situation just in case the agent behaves that way in the real state which represents that situation. The predictions of a single model are given by what holds at all its real states; the behavioral theory of a class of models is given by what holds at all real states in all its models.

A state in an epistemic model is subjective if it figures in the specification of what the agent knows or is aware of at some real state. According to this way of understanding real and subjective states, states may be both real and subjective. Suppose we wish to represent an agent who knows that a particular coin will be flipped, but who will not learn the outcome of this coin flip. If our model is intended to make predictions no matter how the coin lands, the subjective states needed to specify the agent's knowledge (heads and ignorant, tails and ignorant) will be exactly the real states; every state will be both real and subjective. But as DLR recognize, there is no reason to require all states to play both roles. In the earlier example, if we only wanted to make predictions about the situation in which the coin comes up heads, we would not count one of the states (tails and ignorant) as a real state; even so, to represent the agent's ignorance given heads, a state where the coin comes up tails would still have to be included as a subjective state. The point can also be illustrated with a less artificial example. Consider an analyst who wishes to model the interactions of agents who are rational, but who do not believe each other to be rational. To represent the beliefs of these agents, the analyst must include subjective states in which the agents are irrational. But although she includes these subjective states, the analyst has no intention of eliminating the claim that the agents are rational from the predictions of her theory. Rather, it is understood informally that these subjective states are not real; they do not represent situations the analyst aims to describe.

Put in the terms of our presentation, DLR's proposal is to allow subjective states in which the law of excluded middle, $p \vee \neg p$, may not be true. DLR have no intention of eliminating the law of excluded middle from the predictions of their theory. Rather, they introduce these subjective states to specify the agent's knowledge and unawareness at real states where classical logic holds. The theory of DLR's models is given by what holds at these classical real states, not by what holds at all states whatsoever. Still, since classical tautologies may fail in DLR's subjective states, their models violate the "real states" assumption, and so are not standard state space models.

But once we acknowledge that some states may not be intended to represent situations the analyst wishes to describe, a question arises: Why should one require DLR's three axioms on unawareness to be valid in *all* states? DLR do not argue for this assumption. At best, their arguments in favor of their axioms only motivate imposing these axioms at real states. These arguments provide no motivation for imposing the axioms on subjective states that are not real, since these states are merely included in the model to specify the agent's knowledge and unawareness at real states.

There is a general methodological principle in epistemic modeling that *axioms* are to be imposed at all states. But in the literature on awareness, following DLR, this methodological principle has long been abandoned. DLR's own non-standard models violate this requirement, as do the current leading proposals for representing awareness, for example that of Heifetz et al. [2006]. The subjective states in DLR's models include states in which logical axioms, including the law of excluded middle, do not hold. In DLR's models, as in the ones we will propose, even logical axioms are allowed to fail at some states. The only difference between our proposal and theirs concerns which axioms are allowed to fail. DLR preserve their own axioms at all states, and move to non-standard models in which classical propositional logic may fail at subjective states. We will preserve classical propositional logic at all states, and work with standard models in which DLR's axioms may fail at subjective states.

Still, one might ask: How should we *understand* a state where DLR's axioms are false? DLR interpret their own subjective states as "descriptions of possibilities as perceived by the agent" (p. 171). This interpretation does not seem appropriate for our models, in which DLR's axioms may fail at subjective states. But such metaphorical interpretations of these states are unnecessary. Subjective states where the axioms of awareness are invalid are simply to be understood in terms of the agent's knowledge and awareness at real states where the axioms are valid.

In fact, we can give a direct argument for not imposing DLR's axioms at all states, and in particular, for including states where KU-Introspection is invalid. First, by AU-Introspection, if an agent is unaware of p, then she must be unaware of being unaware of p. But then, by Plausibility, the agent does not know that she does not know that she is unaware of p. (Essentially, DLR already give this argument on p. 169.) In epistemic models, we generally represent an agent's not knowing q by including a state in which q is false. So, to allow for real states in which the agent does not know that she does not know that she is unaware of p, we must include subjective states in which the agent knows that she is unaware of p, and so violates KU-Introspection.[3]

To sum up: after rejecting standard state space models, DLR propose that we should use models in which the laws of logic fail at subjective states. They implement this proposal by countenancing states where propositional logic fails, so that their models are non-standard. But if we allow models in which the law of excluded middle may fail at subjective states, we must also consider models in which other axioms, including DLR's, may fail at subjective states. DLR's formal results only apply to standard state space models in which their axioms are imposed at all states; the results do not concern standard state space models in which the axioms are imposed only at real states. As a consequence, these formal results cannot provide a basis for the conclusion DLR draw: that standard state space models preclude unawareness.

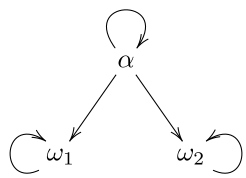
2.4. **Non-Triviality.** We have not argued against the validity of DLR's three axioms in real states – states representing the situations to be modeled. In our setting, we can formalize the distinction between real states and subjective states which are

[3]This argument differs from DLR's main proof of triviality, since it only assumes that the axioms hold at the real state.

not real, by only assuming the validity of the axioms in some subset of the states in a model. We can then ask: is assuming the validity of the three axioms in such distinguished real states enough to lead to triviality? The following example shows that it is not:

Theorem 2. *There is a model $M = \langle \Omega, k, a \rangle$, state $\alpha \in \Omega$ and event $E \subseteq \Omega$ such that Necessitation is valid on M, Plausibility, AU-Introspection and KU-Introspection are valid in α, and $\alpha \notin a(E)$.*

Proof. Let $\Omega = \{\alpha, \omega_1, \omega_2\}$. Define a binary (accessibility) relation R on Ω as follows:



R induces a possibility correspondence P such that $P(\sigma) = \{\tau : R\sigma\tau\}$. With P, define k and a such that for all $F \subseteq \Omega$:

$$k(F) = \{\sigma \in \Omega : P(\sigma) \subseteq F\}$$

$$a(F) = \begin{cases} \{\omega_2\} & \text{if } \omega_1 \in F \text{ and } \omega_2 \notin F \\ \Omega & \text{otherwise} \end{cases}$$

If $u(F)$ is defined as $-a(F)$, then we have $u(\{\omega_1\}) = u(\{\alpha, \omega_1\}) = \{\alpha, \omega_1\}$ and $u(F) = \emptyset$ for all other events F. It is routine to check that $M = \langle \Omega, k, a \rangle$, α and $E = \{\alpha, \omega_1\}$ witness the claim to be proven.[4] □

This shows that DLR's Theorem 1(i) cannot be extended to standard state space models in which DLR's three axioms are only required to be valid in real states. In fact, the model used in the above proof of Theorem 2 can also be used to show that DLR's other two results cannot be extended either. For 1(ii), note that Monotonicity is valid on M, and that $\alpha \in k(\Omega)$ although $\alpha \notin a(E)$. For 2, note first that Weak Necessitation is valid on M, and that $\alpha \in a(\Omega)$ but as before $\alpha \in k(\Omega)$ and $\alpha \notin a(E)$. More generally, any state in any model which satisfies both Necessitation and Monotonicity, in addition to DLR's three axioms, will be a counterexample not just to extensions of DLR's Theorem 1(i), but also to extensions of their Theorems 1(ii) and 2.

We conclude that none of DLR's three triviality results show that standard state space models preclude unawareness. One might wonder whether plausible strengthenings of the axioms on knowledge and unawareness allow us to reinstate the triviality results. We show first, that this cannot be achieved by strengthening their axioms governing knowledge, and, second, that it cannot be achieved by a particular strengthening of the axioms governing unawareness.

[4] As suggested in Modica and Rustichini [1999], one might consider an extension of Plausibility along the following lines: For each natural number n, let n-Plausibility be $Up \rightarrow (\neg K)^n p$, where $(\neg K)^n \varphi$ is defined inductively by the two clauses $(\neg K)^1 \varphi = \neg K \varphi$ and $(\neg K)^{n+1} \varphi = \neg K (\neg K)^n \varphi$. For each natural number n, n-Plausibility is valid in α.

2.5. **Non-Triviality under Stronger Assumptions on Knowledge.** The model used in Theorem 2 already validates a number of attractive axioms on knowledge, suggesting that strengthening DLR's constraints on knowledge is unlikely to yield an interesting triviality result. In particular, the following axioms are valid on the model:

Distribution:	$(Kp \wedge Kq) \to K(p \wedge q)$
Anti-Necessitation:	$\neg K\bot$
Reflexivity:	$Kp \to p$
Positive Introspection:	$Kp \to KKp$

If Negative Introspection (that is, $\neg Kp \to K\neg Kp$) were required to be valid on the model, the consequent of Plausibility could never be true. So Negative Introspection rules out non-trivial unawareness immediately (as was shown by Modica and Rustichini [1994]). Thus the list above comprises all the relevant standard principles usually considered for knowledge.

We can show more systematically that any strengthening of the axioms of knowledge which rules out unawareness does so trivially in the same way as Negative Introspection does. On the very mild assumption that the agent doesn't know the contradiction $\bot$, we can characterize the conditions under which a given model for the knowledge of an agent can be extended to an unawareness model in which the agent is unaware of a given p at a given point α in which DLR's three axioms are valid. Let a M' *extend* a *knowledge model* $M = \langle \Omega, k\rangle$ just in case $M' = \langle \Omega, k, a\rangle$ for some function $a : 2^\Omega \to 2^\Omega$.

Theorem 3. *Let $M = \langle \Omega, k\rangle$ be a knowledge model, $\alpha \in \Omega$ and $E \subseteq \Omega$ such that Anti-Necessitation is valid in α. M has an extension such that Plausibility, KU-Introspection and AU-Introspection are valid in α and $\alpha \notin a(E)$ if and only if*
(i) $\alpha \in -k(E) \cap -k - k(E)$, and
(ii) there is an event F such that $a \in F \cap -k(F) \cap -k - k(F)$.[5]

Proof. Assume first that (i) and (ii). Let $a : 2^\Omega \to 2^\Omega$ be defined so that $a(E) = a(F) = -F$, and $a(G) = \Omega$ for all other events G, and consider the model $M' = \langle \Omega, k, a\rangle$. It is routine to verify that Plausibility, KU-Introspection and AU-Introspection are valid in α, and $\alpha \notin a(E)$, appealing to Anti-Necessitation in the proof for KU-Introspection.

For the converse, note that (i) follows from the validity of Plausibility in α. For (ii), let $F = -a(E)$. Then $\alpha \in F$; by AU-Introspection, $\alpha \notin a(F)$; and so by Plausibility, $\alpha \in -k(F) \cap -k - k(F)$. □

In particular, as long as the constraints on knowledge allow for there to be an event E and a state $\alpha \in E$ such that $\alpha \in -k(E) \cap -k - k(E)$, standard state space models and DLR's three axioms will not preclude non-trivial unawareness.

2.6. **Non-Triviality under Stronger Assumptions on Awareness.** These results demonstrate that no plausible strengthening of the axioms governing knowledge will re-instate triviality. But what if we strengthen the axioms on awareness themselves? Working in standard state space models, we can think of the earlier observation that an agent who is unaware of p cannot know that she doesn't know that she is unaware of p as demonstrating that such an agent *must* fail to know

[5]This result also holds if we replace Plausibility by n-Plausibility, for all natural numbers n, and (i) and (ii) by the correspondingly iterated conditions.

KU-Introspection, i.e., $\neg KUp$. But what if the agent *is* aware of p? Can the agent know $\neg KUp$? One might consider strengthening DLR's axioms by additionally requiring that for all events of which the agent is aware, the agent knows all the relevant instances of the three axioms. Moreover, we could require that the agent knows this as well, and knows that she does, and so on. In the present single-agent case, in other words, we could require that the agent *commonly knows* the relevant instances of the axioms.

To investigate this suggestion formally, extend for the purposes of this section the language L by a unary operator CK for common knowledge. To define its interpretation on a model $M = \langle \Omega, k, a \rangle$, derive the following functions on events: $k^1(E) = k(E)$, $k^{n+1}(E) = kk^n(E)$, and $ck(E) = \bigcap_{1 \leq n < \omega} k^n(E)$.

$$[\![CK\varphi]\!]_{M,v} = ck([\![\varphi]\!]_{M,v})$$

The additional axioms are then:

CK-Plausibility	$Ap \to CK(Up \to (\neg Kp \wedge \neg K \neg Kp))$
CK-KU-Introspection	$Ap \to CK(\neg KUp)$
CK-AU-Introspection	$Ap \to CK(Up \to UUp)$

These additional axioms are also compatible with non-trivial unawareness. In fact, they are valid in state α of the example in the proof of Theorem 2. More generally, Theorem 3 can be extended straightforwardly to these three additional axioms, given the weak assumption that Necessitation and Anti-Necessitation are valid:

Theorem 4. *Let $M = \langle \Omega, k \rangle$ be a knowledge model in which Necessitation and Anti-Necessitation are valid, $\alpha \in \Omega$ and $E \subseteq \Omega$. M has an extension such that Plausibility, KU-Introspection, AU-Introspection, CK-Plausibility, CK-KU-Introspection and CK-AU-Introspection are valid in α and $\alpha \notin a(E)$ if and only if*

(i) $\alpha \in -k(E) \cap -k - k(E)$, and

(ii) there is an event F such that $\alpha \in F \cap -k(F) \cap -k - k(F)$.[6]

Proof. Assuming (i) and (ii), we define a as in the proof of Theorem 3, where it is noted that Plausibility, KU-Introspection and AU-Introspection are valid in α, and $\alpha \notin a(E)$. For the CK-conditions, consider any event G such that $\alpha \in a(G)$. Then by construction of a, $a(G) = \Omega$. Therefore $a(G) \cup H = \Omega$ for any event H, and so $\alpha \in ck(a(G) \cup H)$ by Necessitation, which establishes the validity of CK-Plausibility and CK-AU-Introspection in α. For CK-KU-Introspection, note that by Anti-Necessitation, $k - a(G) = \emptyset$, so $-k - a(G) = \Omega$, from which $\alpha \in ck(-k - a(G))$ follows again by Necessitation. For the converse direction, we establish (i) and (ii) as in the proof of Theorem 3. □

Not only does DLR's result fail to establish that standard state space models preclude unawareness, this conclusion can also not be reached by substantially strengthening their assumptions on knowledge and awareness along the lines discussed here.

3. Three Kinds of Awareness

3.1. **A New Triviality Result.** DLR's result had limited implications for state space models because it depended on the validity of their axioms at all states. One

[6] Again, we can extend this result to n-Plausibility for all n analogously to the extension in the previous footnote.

might thus wonder whether there is a triviality result which only uses the validity of axioms on awareness in real states, rather than their validity in all states. In fact, as we now show, widely accepted axioms on awareness *do* lead to triviality even if they are imposed only at real states. Later, we will argue that the new triviality result also does not rule out standard state space models. But we do think it poses a novel challenge to the use of standard state space models. The result uses the following two axioms:

AS $A\neg p \to Ap$
AC $A(p \wedge q) \to (Ap \wedge Aq)$

Awareness is widely assumed to satisfy both of these axioms; see, e.g., [Modica and Rustichini, 1999, pp. 274–275], [Halpern, 2001, p. 331] (axioms A1 and A2) and [Heifetz et al., 2008, p. 309] (axioms 1 and 2).

As the next theorem shows, these axioms lead straightforwardly to triviality.

Theorem 5. *Let $M = \langle \Omega, k, a \rangle$ be a model and $\alpha \in \Omega$ such that AS and AC are valid in α. Then $Ap \to Aq$ is valid in α.*

Proof. Consider any events E and F, and assume $\alpha \in a(E)$. Since $E = E \cap \Omega$, $\alpha \in a(E \cap \Omega)$, and so by AC, $\alpha \in a(\Omega)$. By AS, $\alpha \in a(\emptyset)$. Since $\emptyset = \emptyset \cap F$, $\alpha \in a(\emptyset \cap F)$, and so by AC, $\alpha \in a(F)$. □

The crucial difference between this and DLR's triviality result is that AS and AC are only assumed to be valid in a distinguished real state, for which it is shown that non-trivial unawareness in it is ruled out. This new triviality result is thus a more promising basis for an argument against standard state space models of awareness.

But does awareness really satisfy both AS and AC? In the following, we will focus in particular on AC, considering whether being aware of a conjunction entails being aware of its conjuncts.

3.2. **Attending vs Conceiving vs Processing.** In the literature on awareness, it is uncontentious that there is no single attitude of awareness; what is expressed by "aware" is a loose cluster of notions. This was noted at the very start of the literature, as witnessed by the lengthy discussions in Fagin and Halpern [1988]; another detailed discussion can be found in Schipper [forthcoming].[7] We will argue that at least some important notions of awareness do not satisfy AC (for others, as we will see, the situation is more complex). In order to do so, we start by making some rough distinctions between different notions of awareness.

We will distinguish between three notions of awareness, all of which are explicitly mentioned in the literature. These distinctions could be further subdivided, and no doubt there are additional categories which are not covered by the notions we'll discuss. But our three rough categories will be enough to illustrate how AC may fail, and thus, to vindicate the use of standard state space models for many applications. The three ways of understanding a claim of the form "The agent is aware of ..." are roughly as follows:

(i) The agent is attending to ...
(ii) The agent has the conceptual resources required to conceive of ...
(iii) The agent is able to process ...

[7]See also, for example, the distinction between syntactic "strong" awareness and semantic "weak" awareness in [Grossi and Velázquez-Quesada, 2009, p. 152–3].

We will introduce these notions – and distinguish between them – using various examples found in the literature.

Consider first attention. An influential example, which first appeared in Geanakoplos [1989], and which is discussed at length by DLR and in numerous other places in the subsequent literature, is based on the following quote from one of Arthur Conan Doyle's Sherlock Holmes stories [Doyle, 1901]:

> "'Is there any other point to which you would wish to draw my attention?'
> 'To the incident of the dog in the night-time.'
> 'The dog did nothing in the night-time.'
> 'That was the curious incident' remarked Sherlock Holmes."

Holmes's interlocutor is Inspector Gregory, a Scotland Yard detective. Before Holmes pointed out to Gregory that the dog did nothing in the night-time, Gregory was *unaware* of the dog doing nothing in the night-time. Gregory's state of unawareness is naturally understood as one of *inattention* – Holmes makes Gregory aware of the dog doing nothing in the night time in the sense of bringing this fact to his attention.

Gregory's failing to attend to the dog doing nothing in the night-time must be sharply distinguished from Gregory's not being able to conceive of the the dog doing nothing in the night-time. Before Holmes alerted Gregory to the dog doing nothing in the night-time, Gregory possessed the concepts required to entertain thoughts about the dog doing nothing in the night time. Contrast this with the following example for unawareness from [Fagin and Halpern, 1988, p. 40]:

> "How can someone say that he knows or doesn't know about p if p is a concept he is completely unaware of? One can imagine the puzzled frown of a Bantu tribesman's face when asked if he knows that personal computer prices are going down!"

The relevant state of unawareness in this example is not merely a matter of the agent failing to attend to the relevant event or subject matter. For example, if one is unaware in the sense of being unable to conceive of an event, it must be that one does not understand the words for those notions in any language. Contrast this with the case of Inspector Gregory. Gregory understands what Holmes says: he can conceive of the dog's doing nothing. But the purported example of inconceivability does not have this structure: the tribesman is supposed to be unable to think about computers using any of his conceptual resources, no matter what he attends to. The two notions of awareness – attending to versus being able to conceive of – are therefore clearly distinct.

The third notion of unawareness we want to single out is one which Fagin and Halpern [1988] (see also Halpern and Pucella [2011]) focus on; it can be understood as an attempt to deal with what is known as the "problem of logical omniscience" in epistemic logic. In standard state spaces, if two sentences φ and ψ are equivalent in classical propositional logic, then $K\varphi$ and $K\psi$ will be true in the same states. In particular, if $K(p \vee \neg p)$ is true in a given state, then so is $K\tau$ for any propositional tautology τ.[8] One of Fagin and Halpern's reasons for developing a logic of awareness is to obtain logics which do not have this property. They write:

> "The notion of awareness we use in this approach is open to a number of interpretations. One of them is that an agent is aware of a formula if he can compute whether or not it is true in a given situation within a certain time or space bound. This interpretation of awareness gives us a way of capturing resource-bounded reasoning in our model."

Being unaware of φ in the sense of not being able to process φ is clearly distinct from failing to attend to φ: although Gregory did not attend to the dog doing nothing in the night-time, he had no difficulties processing the sentence "the dog did nothing in the night-time." Not being able to process φ is also clearly distinct from not being able to conceive of φ: Gregory might not have been able to process an extremely complicated propositional tautology using only negation, conjunction and the sentence "the dog did nothing in the night-time", but he clearly possessed all the concepts required to entertain it.

The most prominent discussions of unawareness in the economics literature have tended to focus on the notion of inattention; see, e.g., Geanakoplos [1989], Modica and Rustichini [1994], Modica and Rustichini [1999], Dekel et al. [1998], Li [2009], or van Benthem and Velázquez-Quesada [2010]. A particularly often-discussed example is the following from [Heifetz et al., 2006, section 3]; see also Schipper [forthcoming]:

> "Consider an owner o of a firm and a potential buyer b. To make this example interesting, we assume that the agents' awareness differs. [...] 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."

Here, clearly, the point is not that the buyer lacks the concept of a lawsuit – to make the point vivid, he may himself be engaged in his own lawsuits in other parts of his life – nor is he incapable of processing the claim that there might be a lawsuit – after all, he can process this claim in the case of his own lawsuit. Rather, in making his investment decision, the buyer is not attending to the question of whether a lawsuit will arise.

Awareness as an ability to process has played a secondary role in the literature on awareness, even though it is important in the work of Halpern and his various co-authors, starting, as noted above, with Fagin and Halpern [1988]; see also, e.g.,

[8]In philosophy, the "problem of logical omniscience" is sometimes used to describe a different issue associated with "Stalnakerianism", the doctrine that the contents of propositional attitudes are sets of metaphysically possible worlds. (Stalnaker [1991], reprinted in Stalnaker [1999], seems to be the source of this confusion: he starts out by focusing on the *logical* problem, but shifts without warning on p. 247 of Stalnaker [1999] to the latter doctrine.) According to Stalnakerianism, there is only one metaphysically necessary proposition. Thus, if an agent knows a truth of logic such as, for example, that it is raining or it is not, she knows every metaphysically necessary truth, for example, that water is H_2O. But this conditional need not be true if we merely assume that the propositions (events) form a complete atomic Boolean algebra. In part the distinction between these "problems of logical omniscience" is just a terminological issue, although since Stalnakerianism brings with it problems which go well beyond logic, it seems to us better called "the problem of metaphysical omniscience".

Thijsse [1993], Halpern et al. [1994], Halpern and Pucella [2011]. Awareness as conceivability seems to have played no role at all, despite the early example in Fagin and Halpern [1988]. We will therefore mainly focus on attention and processing, although much of what we say about attention can also be applied to conceivability.[9]

3.3. **Sentences and Events.** Before returning to the question which notions of awareness satisfy the axiom AC, it will be helpful to apply our distinction between different notions of awareness to an issue which has provided a puzzle for research on awareness since its inception. The puzzle is: what are the objects of awareness?[10] Consider a claim such as "Gregory is not aware that the dog did nothing in the night-time". Does this express a relation of Gregory to the sentence "the dog did nothing in the night-time" or the event this sentence expresses? (Instead of "event", we might also be speaking of "content" or "proposition", and in the following, we will do so in cases where it is unnatural to speak of events – the three terms are to be understood as interchangeable theoretical terms.)

Awareness understood as an ability to process is naturally understood as expressing a relation towards a sentence. Part of what we say with a claim of the form "the agent can process . . . " is that the agent can figure out what the sentence ". . . " expresses. This can be illustrated using the following example: Let n be some large natural number. Reading "is (not) aware" as "can (not) process", there seems to be nothing problematic about asserting the following two sentences:

(1) "Holmes is aware that the dog did nothing in the night-time."
(2) "Holmes is not aware that not . . . not ($2n$ times) the dog did nothing in the night-time."

Assuming that the negation of the negation of a sentence expresses the same event as the original sentence, "the dog did nothing in the night-time" and "not . . . not ($2n$ times) the dog did nothing in the night-time" express the same event. If nevertheless both (1) and (2) are to be true, a natural conclusion is that "is aware" expresses a relation to a sentence rather than an event.

The matter is different with, e.g., attention. Supposing the Sherlock Holmes stories were about a French detective, and Grégoire were a monolingual French speaker, this Grégoire would not be able to process even the simple sentence "the dog did nothing in the night-time". But he would, nevertheless, be attending to the dog's doing nothing in the night-time. In the purest sense of attention,whether we correctly describe an agent as attending to an event does not depend on which sentence we use to express the event. This "purest" sense can sometimes take effort to isolate. Consider again the English-speaking Gregory. If we read awareness as expressing attention and assume the truth of (1), one naturally wants to accept (2) in some sense but to reject it in another: in some sense, Holmes clearly attends to

---

[9]Surprisingly, conceptual abilities are often appealed to when the notion of awareness is introduced. E.g., [Schipper, forthcoming, section 2.1] writes (see also the first sentence of his abstract):

> "While the precise connotations of all those uses of awareness are different, they have in common that the agent is able to conceive something. Being unaware means then that he lacks conception of something."

Given the variety of uses of "aware" we have found above, this seems too narrow.

[10]See Konolige [1986]. Related issues crop up throughout the literature under various guises, including the question whether a model theory must be "syntax-free"; see, e.g., Heifetz et al. [2006].

the issue at hand, but he also wouldn't put it in the terms in which (2) is phrased (involving a long string of negations).

What exactly we should make of this kind of puzzle is a vexed issue with a long history, going back at least to Frege [1892] and now making up a vast body of work (see, e.g, the references in McKay and Nelson [2014]). But much of the literature assumes that sentences express events which are not as finely individuated as the sentences used to express them, and that attitudes such as belief and knowledge are relations between agents and events of this kind. For example, the French Grégoire and the English Gregory can both believe that the dog did nothing in the night-time, even though they and their associates would express this belief using different sentences. In particular, in the technical literature, it is almost universally assumed that events expressed by sentences form a complete atomic Boolean algebra, and so, by a well-known representation theorem, can be modeled by the powerset of a set of states. The usual probabilistic models of belief in decision theory and game theory are just one example of this strategy; qualitative Hintikka-Kripke-Aumann models of belief and knowledge are another. It is of course controversial whether events (or: propositions) have this structure, but it should not be controversial that theorizing in terms of events with this structure has led to myriad important discoveries about the concepts represented in this way. In what follows, we will apply this methodology to awareness understood as attention, and as conceivability.[11]

It might be helpful to illustrate our understanding of the distinction between sentences and the events they express by explaining how we think of the so-called "problem of logical omniscience" in epistemic logic. As mentioned before, the alleged problem is that $K\varphi$ is true in all states for each propositional tautology φ. But aren't there complicated tautologies none of us know? It is true that only some claims of the form "The agent knows . . . " are appropriate to utter in English, even when ". . . " is filled in with a tautology. But a consequence of the assumption that events form a Boolean algebra is that sentences which are equivalent in propositional logic have identical contents. Since each propositional tautology expresses the same trivial event, there is nothing puzzling about ascribing knowledge of that trivial event to each agent. Once we think of events as forming a Boolean algebra, the problem of logical omniscience is not so much a problem about our epistemic models, but a question about why English draws distinctions of appropriateness among ways of expressing the same state of knowledge. That is, why can it be appropriate to say "Martina knows it's raining or it's not", but inappropriate to say "Martina knows it's not the case that it's both raining and not raining"? But *this* problem concerns speech patterns in English, and belongs to the study of natural language. As logicians interested in structural properties of belief, knowledge and awareness, we think such difficulties can be set aside for present purposes, given the attractive simplicity of models developed in the Boolean framework.

3.4. **Awareness of Conjunctions.** Let's return to the new triviality result introduced at the start of this section. As already advertised, we believe that the principle AC, which says that an agent who is aware of a conjunction is aware of its conjuncts, may be plausible for one notion of awareness, but it is not plausible

[11]For a sustained defense of this general approach, see Stalnaker [1984]. As noted in n. 8, Stalnaker's form of "Booleanism" – the doctrine that events or propositions form a complete atomic Boolean algebra – is a special one which relies on a substantive notion of metaphysical possibility. But Booleanism in its general form need not be committed on this count.

for the other two. In particular, the principle may apply to awareness as the ability to process, understood as a relation to sentences. It is natural to assume that an agent who is able to process a conjunction "... and —" will also be able to process "..." and "—". As noted already in [Fagin and Halpern, 1988, p. 54], even this may fail: an agent might be able to recognize that a very long sentence has the form $\varphi \wedge \neg\varphi$, and so be able to process it, although she is unable to process the complex φ on its own. Resolving this controversy may require distinguishing further among different notions of processing, and the appropriate resolution may depend on the intended application.

But however we think of the axiom as applied to the ability to process, AC is implausible for attention, understood for the moment as a relation to events, or perhaps more naturally in this context, propositions. Assume that after the conversation with Holmes quoted above, Gregory is alone thinking about the case, and attending to the event of the dog barking in the night (p). He is not, however, attending to the event of Holmes at that moment smoking a pipe (q). It is then natural to say that Gregory is also not attending to the conditional event that *if* Holmes is currently smoking a pipe *then* the dog barked in the night ($q \to p$). But notice that according to the coarse-grained Boolean theory of events, the event that the dog barked in the night (p) is identical to the event that if Holmes is smoking a pipe then the dog barked in the night, and the dog barked in the night ($(q \to p) \wedge p$). So if AC were valid, then since Gregory is attending to the event of the dog barking in the night, he would be attending to the event that if Holmes is smoking a pipe then the dog barked in the night. But by assumption Gregory is not attending to this last event. Thus it follows from the coarse-grained theory of events that AC must be rejected for attention.

A similar example can be given if we understand awareness as conceivability. Assume our agent does not have the conceptual resources to entertain the event of there being a black hole. According to the assumed coarse-grained theory of content, the event of there being a black hole and there being no black hole is identical to any event expressed by a propositional contradiction, such as the event of there being a sheep and there being no sheep. The agent might well have the conceptual resources to entertain the event of there being a sheep and there being no sheep, without having the conceptual resources to entertain the event of there being a black hole.

If we adopt the coarse-grained theory of events which is standard in epistemic models in a range of disciplines, AC appears to be inappropriate. It's worth noting, also, that if we were to give up this coarse-grained theory of events, it's unclear whether DLR's original triviality theorem would persist. But in any case the new triviality result with which we started this section also does not establish that standard state space models preclude unawareness understood as inattention or inability to conceive.

3.5. **Limitations of Current Non-Standard Models.** As we have seen so far, the only limitation of standard state space models is that they may not be able to model unawareness understood as being unable to process a sentence. Is this fact alone enough to warrant rejecting standard state space models? Shouldn't good models be capable of representing *all* notions of awareness?

We don't see why they should. The different notions of awareness have little in common. Perhaps most strikingly, processing is a relation to *sentences*, while attention and conceivability are relations to *events*. Since these two different classes of objects have fundamentally different structure – in particular, the first do not form a Boolean algebra,[12] while the second do – we see no reason why it should be possible to model all three notions of awareness using a single system.

This formal fact is reflected by powerful intuitions about the psychology of these states. As we saw above, examples can be given which sharply distinguish agents with limited processing power from those with limited attention, and similarly for the other pairs drawn from our three notions. So even if these attitudes *were* to have the same kind of object, it's unclear why we should expect them to have the same logic or formal representation, any more than we should expect regret and hope, knowledge and belief, or even desire and credence, to have the same formal representation.

Even setting aside these more general arguments, we can also argue that in not being able to model processing a sentence, standard state space models fare no worse than some of their much less tractable competitors. In particular, the benchmark models of Dekel et al. [1998] and Heifetz et al. [2006] *also* do not adequately model failures of processing.[13]

DLR do not tell us how to interpret Boolean connectives other than negation on their models, but since their models can be understood as interpreting formulas as true, false or neither relative to worlds, it is natural to expect these connectives to be interpreted according to one of the familiar three-valued sets of truth-tables, such as the so-called strong Kleene tables or Łukasiewicz tables – both fit their treatment of negation. No matter which of them is chosen, the de Morgan equivalences hold according them, which means, e.g., that $p \vee q$ and $\neg(\neg p \wedge \neg q)$ will always be interpreted as true at the same states and false at the same states; consequently $A(p \vee q)$ and $A\neg(\neg p \wedge \neg q)$ will be true at the same states. But surely an agent might be able to process $p \vee q$ without being able to process $\neg(\neg p \wedge \neg q)$: the addition of three truth-functional operators might push the complexity of the formula over the computational bounds within which the agent operates. If this example seems unsatisfying, more complicated examples can of course be chosen to make the difference in complexity more dramatic.[14]

Similarly, in the models of Heifetz et al. [2006], if Ap and Aq are true in a state, then so is $A\varphi$ for any propositional tautology φ containing only proposition letters p and q. But this φ may be arbitrarily complex; it could be so long that no human could read it in a single lifetime. Hence, these models are also incapable of modeling the notion of awareness as processing.

The inability of standard state space models to represent limited processing ability is no mark against standard state space models of attention and conceivability:

[12] When $\neg$ and $\wedge$ are understood respectively as mapping the formula φ to the formula $\neg\varphi$ and the formulas ϕ, ψ to the formula $\varphi \wedge \psi$, with the other operations derived as usual.

[13] The logic of local reasoning presented in [Fagin and Halpern, 1988, Section 6], as well as the algorithmic models of Halpern et al. [1994] and Pucella [2006] *do* represent failures of processing. But the economics literature has not seen substantial development of such models.

[14] See, e.g., [Priest, 2008, section 7.3] for a statement of the two sets of three-valued tables. More generally, assuming that Boolean connectives are truth-functional, there are only 3^n formulas up to truth-functional equivalence which can be constructed out of n proposition letters and Boolean connectives; thus no matter what kind of truth-tables are used to interpret the Boolean connectives, for any formula φ using n proposition letters, there will always be some arbitrarily complex formula using n proposition letters which is truth-functionally equivalent to φ. Of course, in such non-classical contexts, all connectives must be taken as primitive.

First, attention and the ability to conceive have objects which are not isomorphic to the objects of processing. Second, the attitudes appear to be psychologically distinct. Third, the most prominent models of unawareness in economics *also* cannot represent limited processing.

But these are negative claims. Standard state space models also have one crucial point in their favor: simplicity. The rest of the paper explores how the simplicity of standard state space models makes them a flexible and powerful tool for modeling what agents attend to and what they are able to conceive.

4. Partitional Models for Awareness

So far, we have shown that standard state-space models escape certain putative impossibility results for models of attention and conceivability. But this does not establish that standard space models can provide fertile models of these notions. In the remainder of the paper, we define, motivate, and examine a class of standard state space models for representing attention and the ability to conceive.

To show that our models generalize smoothly to the multi-agent case, from now on we use a language L_I parametrized to an arbitrary set of agent-indices I which is defined as the language L above, except that the operators A_i and K_i are indexed to $i \in I$. Models are consequently tuples of the form $\langle \Omega, k^i, a^i \rangle_{i \in I}$.

The models we will be working with are defined as follows:

Definition 1. $\langle \Omega, R^i, \approx^i \rangle_{i \in I}$ is a *partitional model* if Ω is a set and for each $i \in I$, R^i is a binary relation on Ω which is reflexive and transitive, and $\approx^i$ a function which maps each $\omega \in \Omega$ to an equivalence relation $\approx^i_\omega$ on Ω.

Here and in what follows, we make use of the fact that each equivalence relation corresponds to a unique partition, and *vice versa*; accordingly, we treat them as interchangeable.

Partitional models can be used to generate standard models in the following way: R^i specifies states of knowledge just as in Theorem 2. The idea behind $\approx^i$ is that the propositions the agent is aware of at ω are the events which are unions of sets of equivalence classes of $\approx^i_\omega$ (equivalently: unions of sets of cells of the induced partition). So for each $i \in I$, let R^i and $\approx^i$ determine functions k^i and a^i on events on Ω as follows:

$$k^i(E) = \{\sigma \in \Omega : P^i(\sigma) \subseteq E\}, \text{ where } P^i(\sigma) = \{\tau : R^i\sigma\tau\}$$
$$a^i(E) = \{\sigma \in \Omega : \text{for all } \rho \text{ and } \tau \text{ such that } \rho \approx^i_\sigma \tau,\ \rho \in E \text{ iff } \tau \in E\}$$

Let the standard model determined by a partitional model $\langle \Omega, R^i, \approx^i \rangle_{i \in I}$ be $\langle \Omega, k^i, a^i \rangle_{i \in I}$, with k^i and a^i as just defined. On such a standard model, L_I can be interpreted as above; obviously, this induces a way of interpreting L_I directly on partitional models.

We require R^i to be reflexive and transitive primarily to show that we can achieve consistency with a strong logic of knowledge, not because we believe that standard models of knowledge must satisfy these constraints.

4.1. The Attitude of Attention.

In order to motivate partitional models as models of limited attention, we suggest that attention in the sense we have been using the term should be understood as an attitude towards questions. There are many available formal approaches to modeling questions.[15] Each of these deserves

[15] For an overview, see Krifka [2011].

investigation, but for concreteness, we'll adopt a standard approach, representing a question as a partition of the state space (see Groenendijk and Stokhof [1984], building on Hamblin [1973] and Karttunen [1977]).[16] Although we think the attitude to questions is primary, we will follow the literature on awareness, in axiomatizing a notion of attention which has events as its objects. The relationship between this attitude to events and the attitude toward questions will be as follows: an agent attends to the question Q if and only if the agent attends to every partial answer to Q. Using partitions to model questions, partial answers are unions of sets of cells, corresponding to how standard models are derived from partitional models.[17]

We will assume for simplicity that agents don't fail to connect the different things they are attending to. In other words, if an agent is attending to the question Q, and attending to the question R, then she is also attending to the question $Q \sqcap R$, where $Q \sqcap R$ mathematically understood to be the meet of the two partitions, or the coarsest common refinement of both Q and R.[18] More generally, we will impose this constraint for arbitrary sets of questions. Furthermore, we assume that an agent attending to a given question Q thereby counts as attending to each question R of which Q is a refinement, and that they attend to at least one question (which might be the trivial question consisting of a single cell). It follows from these assumptions that the state of attention of an agent is uniquely characterized by the finest partition they attend to.

These assumptions can be seen as the interrogative analogue of the closure conditions imposed on Kripke frames for knowledge. Sets/events form a partial order (in fact, a Boolean algebra) under the subset relation, which corresponds to the relation of entailment on events. Partitions/questions form a partial order (in fact, a lattice) under the refinement relation, which corresponds to the relation of "entailment" on questions (see [Groenendijk and Stokhof, 1984, p. 16]). In our models, each agent is associated with a partition which is the "strongest" question they attend to, just as in Hintikka-Kripke-Aumann models each agent can be seen as associated with an event which is the strongest event they know. In both cases, the set of events the agent knows or the set of questions the agent attends to is exactly the principal filter generated by the relevant event or question.

These assumptions about the interrogative attitude of attention, together with our assumption that one attends to Q if and only if for every potential answer E to Q, one attends to E yields an exact semantics for attention. If the strongest question an agent attends to is Q, the agent attends to every event which is the union of a set of cells of a partition of which Q is a refinement, or, in other words, any event which is the union of a set of cells of Q. This, of course, is exactly the constraint imposed by partitional models.

---

[16]In philosophy Friedman [2013] has recently advocated that we take many rogative verbs to express "question-directed attitudes", which have questions as their objects.

[17]Kets [2014a] and Kets [2014b] develops Harsanyi type spaces in which not only agents' distributions, but also their σ-algebras may vary from type to type. The candidate algebras in Kets's models are those which coincide with the events definable at some finite order in the hierarchy of beliefs. In cases where these algebras are complete and atomic one can view Kets's models as a special case of partitional models, where the level-k agents are thought of as attending only to events definable in terms of level-k beliefs. Koralus [2014] also develops a theory of attention in which it has questions as its objects. Further discussion of references will be deferred until 4.9.

[18]For partitions Q and R, Q being a refinement of R – or R being at least as coarse as Q – is understood as each cell of R being a union of cells of Q.

In the above definition of partitional models we have imposed these basic constraints on the set of events attended to by an agent at *all* states. As brought out in the discussion of DLR's triviality argument, this may well be too strong, and the constraints may have to be restricted to distinguished states representing the way the world could be, in the relevant sense of "could". But for the moment, we opt for what might turn out to be overly strong constraints in order to obtain strong consistency results.

4.2. **Partitions for Conceivability.** We present two different ways of arguing for partitional models as models of awareness as conceivability. As with models for attention, the constraint on the awareness function might have to be restricted to states which represent the way the world could be.

The first argument is inspired by Fritz [unpublished b], which discusses and adapts models from [Stalnaker, 2012, Appendix A]. One's conceptual resources allow one to distinguish certain states from others. The relation among states of "not being distinguished by an agent's conceptual resources" is an equivalence relation: No state can be distinguished from itself (reflexivity). If a state x cannot be distinguished from a state y, then y cannot be distinguished from x (symmetry). Finally, assume that states x and z can be distinguished using certain conceptual resources. Then these resources allow one to draw a distinction p among states which applies to only one of x and z. Whether p applies to y or not, it either allows one to draw a distinction between x and y or between y and z. Thus, reasoning contrapositively, if x is indistinguishable from y and y is indistinguishable from z, then x is indistinguishable from z (transitivity).[19] This motivates the use of equivalence relations in specifying what an agent can conceive. If an event distinguishes between states which the agent can't distinguish, then the agent cannot conceive of the event. Conversely, if the agent can distinguish between the states which the event distinguishes between, she can conceive of it.

Another way of motivating the use of partitional models of conceivability is to assume that the agents to be modeled have the concept of negation and (infinitary) conjunction, so that the set of events they can conceive of are closed under complement and arbitrary intersection. This is mathematically equivalent to requiring that this set is derived from a equivalence relation as above.[20]

4.3. **An Example.** It will be useful to have a concrete partitional model before us, as a running example. The following model shows that there are non-trivial partitional models; for simplicity, a single-agent case is specified. In the initial presentation, we will not distinguish real states from states which are subjective but not real; from now until section 4.6, we will assume that all states are both real and subjective. Let $M = \langle \Omega, R, \approx \rangle$, with $\omega R \nu$ iff $\omega = 1$ or $\omega = \nu$, and $\approx$ given by the following equivalence classes:

$\approx_1$: $\{1\}, \{2, 3, 4\}$

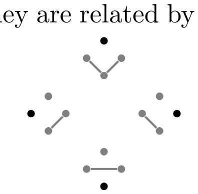
[19] In this respect, "not being distinguished by conceptual resources" differs from "not being distinguishable by an agent's perceptual faculties", since the latter is not transitive. See Williamson [1990].

[20] One might wonder at this claim given that there are complete atomless Boolean algebras, and every Boolean algebra can be represented as a field of sets. The conflict is only apparent, since a field of sets may be complete in the sense of having least upper bounds and greatest lower bounds of arbitrary sets of elements under the subset order without being closed under arbitrary union and intersection.

$\approx_2$: $\{1\}, \{2\}, \{3, 4\}$
$\approx_3$: $\{1\}, \{3\}, \{2, 4\}$
$\approx_4$: $\{1\}, \{4\}, \{2, 3\}$

Thus, at 1, the strongest event known by the agent is Ω, and at each other state n, it is $\{n\}$. At each state n, the events the agent is aware of are the events which don't distinguish between any states in $\Omega \backslash \{1, n\}$.

Drawing the four states in a circle, starting with 1 at the top and going clockwise, we can draw each equivalence relation in a similar smaller circle, connecting two states by a sequence of lines if they are related by the relevant equivalence relation:



This is a partitional model in which there is non-trivial unawareness at each state. We will appeal to it several times below in order to show the consistency of various constraints.

4.4. **Unawareness and Knowledge of Tautologies.** In most models for unawareness, an agent can only be unaware of something if she does not know all tautologies. For example, while DLR do not explicitly note this, their non-standard state space models impose the constraint that if an agent is unaware of something, she does not know that either she knows or does not know that she is unaware of it.[21] More generally, this is not surprising, since Plausibility, AS and AC are widely accepted, and they classically entail $Up \to \neg K \neg (p \wedge \neg p)$.

In contrast, partitional models validate $K\varphi$ for every propositional tautology φ. This shows that once we give up AC, we no longer have to assume that being unaware of something (in the sense of inattention or inconceivability) entails failing to know all tautologies.

4.5. **Axioms.** Given a class of models $\mathtt{C}$, a set of sentences $\Sigma \subseteq L_I$ is the logic of $\mathtt{C}$ if and only if Σ contains exactly those sentences which are valid on $\mathtt{C}$. Characterizing the logic of a class of models gives us a formal perspective from which to assess what assumptions our models encode about agents' knowledge and awareness.

Thus we may ask: What is the logic of partitional models? Standard techniques on completeness results in modal logic are easily adapted to obtain the following result.

[21] In DLR's proposed models, statements – the objects of knowledge and unawareness – can be neither true nor false at states. Therefore, they understand a statement S to be given by an arbitrary pair of disjoint sets of states, interpreted as the states where it is true, and the states where it is false. We write $k(S)$ or $a(S)$ for the statements that the agent knows or is aware of S, respectively, even though k and a must now be understood as functions on pairs of sets of states. Similarly, we still use $-$ for negation, and write $\cup$ for disjunction. DLR's version of KU-Introspection in this setting requires that for a given statement S, no state makes $k - a(S)$ true. Given the the very weak assumption that for a statement P, $P \cup -P$ is true just in case P or $-P$ is true, it follows that $k - a(S) \cup -k - a(S)$ and $-k - a(S)$ are true in the same states. DLR endorse, at least tentatively, a constraint they call "event-sufficiency", from which it follows that $k(k - a(S) \cup -k - a(S))$ and $k - k - a(S)$ are true in the same states as well. As we have seen above, if the agent is unaware of S in a state α, then by Plausibility and AU-Introspection, $k - k - a(S)$ is not true in α; it follows that $k(k - a(S) \cup -k - a(S))$ is not true in α either, as required.

Theorem 6. *A formula is valid on all partitional models if and only if it is derivable in the calculus with the following axiom schemas and rules:*

PL	*Any substitution instance of a theorem of propositional logic.*
K-K	$K_i(\varphi \to \psi) \to (K_i\varphi \to K_i\psi)$
K-T	$K_i\varphi \to \varphi$
K-4	$K_i\varphi \to K_iK_i\varphi$
A-Neg	$A_i\varphi \to A_i\neg\varphi$
A-M	$(A_i\varphi \wedge A_i\psi) \to A_i(\varphi \wedge \psi)$
A-N	$A_i\top$
K-RN	*From* $\vdash \varphi$ *infer* $\vdash K_i\varphi$
A-RE	*From* $\vdash \varphi \leftrightarrow \psi$ *infer* $\vdash A_i\varphi \leftrightarrow A_i\psi$

Moreover, the logic is decidable.

Proof. Since the formulas derivable in this calculus form a classical model logic in the sense of Segerberg [1971], we can apply the standard canonical model construction technique; in particular, consider the smallest canonical model (see [Chellas, 1980, chapter 9], especially p. 254). Consider any formula φ not provable in the above calculus, and let Γ be the set of subformulas of φ closed under Boolean combinations. A standard filtration of the canonical model through Γ produces a finite model in which φ is false. It is routine to prove that the neighborhood function for A_i associates with each state a field of sets; since the model is finite this field is generated by an equivalence relation, as required. The above filtration can be chosen in such a way as to preserve the transitivity of the relation for K_i; reflexivity is preserved by any filtration (see, e.g. [Chellas, 1980, chapter 3], especially p. 106, or Blackburn et al. [2001] p. 80).

The above argument also establishes that the logic thus axiomatized has the finite model property and so is decidable. □

4.6. **DLR Once More.** Consider again DLR's three axioms. Given our discussion above, it is natural to consider partitional models where DLR's axioms are required to be valid in some distinguished state:

Definition 2. $\langle \Omega, \alpha, R^i, \approx^i \rangle_{i \in I}$ is a *partitional DLR model* if $\langle \Omega, R^i, \approx^i \rangle_{i \in I}$ is a partitional model and Plausibility, KU-Introspection and AU-Introspection (for each $i \in I$) are valid in α.

We now show that DLR's triviality result cannot be revived in partitional models:

Theorem 7. *There is a partitional DLR model* $\langle \Omega, \alpha, R^i, \approx^i \rangle_{i \in I}$ *and an event* $E \subseteq \Omega$ *such that* $\alpha \in U(E)$.

Proof. Simply distinguish state 1 in the model presented in section 4.3. □

We conjecture that the logic of partitional DLR models can be axiomatized as follows:

Conjecture 1. *Add the following axioms to the theorems of the axiom system in Theorem 6 and close under modus ponens:*

P	$U_i\varphi \to (\neg K_i\varphi \wedge \neg K_i\neg K_i\varphi)$
AU	$U_i\varphi \to U_iU_i\varphi$

A formula is derivable in this calculus if and only if it is valid in every distinguished state of every partitional DLR model.

Note that $\neg KU\varphi$ can be derived using P, AU and K-T.

The present result shows that we can impose the DLR axioms without trivializing partitional models. But we confess to doubts about whether these axioms are appropriate. Just as with AC, once we understand more clearly the character of attention and conceivability, as well as the distinction between sentences and what they express, DLR's axioms become much less compelling. Consider first AU-Introspection, the principle that being unaware of something entails being unaware of being unaware of it. What if there is an event E such that due to inherent limitations in my human cognitive apparatus, I couldn't possibly attend to it or conceive of it? Then although in the relevant sense, I am unaware of it, the event of my being unaware of it might be the trivially true event, and so one I am always attending to.

A similar point holds for Plausibility. The contrapositive of Plausibility is $K\varphi \vee K\neg K\varphi \rightarrow A\varphi$. A consequence of this is that $K\varphi \rightarrow A\varphi$. But this principle is false for attention; agents can know things without attending to them. You know that there are more than four stars in the universe, but we doubt that you were attending to the question of how many stars there are prior to reading the previous clause. The case for conceivability requires a further assumption about knowledge, in particular, the claim that knowledge is closed under propositional logic, so that an agent who knows φ knows $\varphi \vee \psi$, for any ψ. Suppose that φ uses only concepts the agent possesses, while ψ does not, and that $\varphi \vee \psi$ is something the agent cannot conceive. Since φ is conceivable for the agent, she may know φ, and hence know $\varphi \vee \psi$ which, by assumption, she cannot conceive. Although this argument depends on an assumption some may wish to reject, it illustrates another way in which the coarse-grained conception of content may lead one to reject DLR's axioms.

4.7. **Propositional Quantification.** A challenge to some approaches to unawareness is to represent propositionally quantified statements. E.g., earlier models by Halpern made the claim that the agent knew she was unaware of something unsatisfiable (cf. Halpern and Rêgo [2009] and Halpern and Rêgo [2013]). In standard state space models such as ours, it is trivial to add propositional quantifiers without any such consequences. To do so, we write $v[E/p]$ for the valuation function which maps p to E and every other proposition letter q to $v(q)$:[22]

$$[\![\forall p\varphi]\!]_{M,v} = \bigcap_{E\subseteq\Omega}[\![\varphi]\!]_{M,v[E/p]}$$

To illustrate that these quantifiers behave just as one would expect, note that in state 1 of the example described in section 4.3, the agent knows that she is unaware of something without there being something that she knows to be unaware of: $K\exists pUp \wedge \neg\exists pKUp$ is true in this state.

4.8. **Closure and Automorphisms.** In this section, we discuss some further elaborations on partitional models, and show that they too are consistent with the DLR axioms.

In partitional models, what agents are aware of (attend to/can conceive) is closed under negation and conjunction. One might wonder whether we can also impose the constraints that what agents are aware of must be closed under awareness and

[22]See already Kripke [1959], and Fine [1970] for a more systematic development. See Fritz [unpublished a] for results on the complexity of propositional quantifiers in the related setting of Fritz [unpublished b].

knowledge. In other words, whether there are models on which the following axioms are valid:

A-4ij $A_i p \to A_i A_j p$
AK-4 $A_i p \to A_i K_j p$

We can think of these principles as appropriate for representing agents who are attending to questions formed by applying knowledge or awareness to questions they are attending to. Or in the case of conceivability, they represent agents who have the concepts of knowledge and unawareness.

To provide models which validate these principles we adapt the coherence constraint of Fritz [unpublished b].[23] The idea behind it is most easily described for awareness as conceivability, taking the equivalence relations of partitional models to represent a relation of indistinguishability using conceptual resources available to the relevant agent at the relevant state. Coherence requires that if two states are indistinguishable in this way, then there must be a way of permuting the state space in a way which preserves all structural facts about knowledge and awareness, as well as all the events which the relevant agent is aware of at the relevant state.

Let $M = \langle \Omega, R^i, \approx^i \rangle_{i \in I}$ be a partitional model. A permutation f of Ω is an automorphism of M if for all $i \in I$,

(i) for all $x, y, z \in \Omega$, $y \approx^i_x z$ iff $f(y) \approx^i_{f(x)} f(z)$, and
(ii) for all $x, y \in \Omega$, $R^i xy$ iff $R^i f(x) f(y)$.

A state $x \in \Omega$ *coheres* if for all $i \in I$ and $y, z \in \Omega$ such that $y \approx^i_x z$ there is an automorphism f of M such that $f(y) = z$ and $f \subseteq \approx^i_x$ (i.e., $\omega \approx^i_x \omega$ for all $\omega \in \Omega$). It's routine to verify that A-4ij and AK-4 are valid in any coherent state of a partitional model.

Once again, the model presented in section 4.3 demonstrates the satisfiability of this constraint: every state in this model is coherent. Since the model also satisfies the DLR axioms at state 1, the model also shows that even if we were to uphold the DLR axioms, imposing them together with coherence would not trivialize state space models of awareness.

One might of course wish to impose stronger constraints than coherence. For example, in the case of conceivability, it's plausible that an agent must be able to conceive of the strongest proposition she knows. (In the case of attention, this axiom is unattractive; clearly one is not always attending to the proposition which is the conjunction of everything one knows.) Formally, a state $x \in \Omega$ *strongly coheres* if for all $i \in I$ and $y, z \in \Omega$ such that $y \approx^i_x z$ there is an automorphism f witnessing coherence, and such that $R^i xy$ iff $R^i f(x) f(y)$. The model in section 4.3 also satisfies this constraint. Once again, we leave the study of the logic of coherent and strongly coherent models (and DLR models) to future work.

4.9. **Related Work.** To this point, we have mainly described the relationship of our models to the literature in economics and computer science. In this section, we discuss related work in linguistics and philosophy, with the aim of illustrating how partitional models could help to connect models of awareness with a range of other disciplines.

In the "question-under-discussion" framework for representing the information structure in a linguistic context (see especially Roberts [1996] now [2012]), the set

[23]The following notion of coherence differs importantly from that of Fritz [unpublished b] in that $\approx^i_x$ here need not relate x only to x.

of possible worlds which represents the informational context or "common ground" of a conversation (see Karttunen [1974], Stalnaker [1974, 2014]) is enriched by a distinguished partition which represents the "question under discussion". Although typically little is said about the source of this distinguished partition, it is natural to see it as derived from individuals' states of attention. (For example, in our setting, one might think of it as the finest common coarsening of the individuals' partitions.)

Yalcin [2011] develops the idea that agents have different belief-states relative to different partitions; in his terminology, agents' implicit belief states are "sensitive to" different questions, giving rise to different surface phenomena for what we might call "explicit beliefs".[24] But crucially Yalcin's agents are represented by *different* accessibility relations relative to different questions (partitions): the agents are thus "fragmented" or "compartmentalized" in the sense of Lewis [1982] and Stalnaker [1984]. These distinct "fragments" correspond to distinct total belief states, and there may be no single coherent belief state which represents their beliefs. (Interestingly, the idea of fragmentation was also briefly discussed from a formal perspective in [Fagin and Halpern, 1988, section 6].) Our models are more restrictive than his, since explicit knowledge at a given state is represented relative to a fixed partition.

These more restrictive models allow us to distinguish phenomena which arise due to limited attention from those which arise from "divided mind" phenomena. In particular, a number of authors (e.g. [Egan, 2008, p. 49-51], Elga and Rayo [2014]) have suggested that the hypothesis that agents suffer from fragmentation can be motivated by a need to explain failures of recall and failures of logical closure in beliefs or knowledge. But explicit knowledge (defined as $K^E\varphi \leftrightarrow (K\varphi \wedge A\varphi)$) can be used to model both of these phenomena. Failures of recall can be understood as deriving from different states of attention, while holding the unique belief state of the agent fixed (we'll say more about this in the next section). Similarly, explicit knowledge will not satisfy the axiom $K^E(\varphi \to \psi) \to (K^E\varphi \to K^E\psi)$, essentially because attention must fail AC, as we argued at length earlier.

While the more restrictive models described above promise to provide some new insight, we also think it worthwhile to explore more flexible models along Yalcin's lines (or those now found in [Yablo, 2014, Ch. 7]). Similarly, we think it worth exploring models in which the set of questions one attends to is itself "fragmented", that is, it does not form a principal filter on the lattice of questions.[25]

In philosophical discussions of the relationship between "full" or "qualitative" belief (sometimes also called "acceptance") and an agent's probability function, a

[24]In fact, Yalcin explicitly rejects reading the question-directed attitude in the vicinity as "attending to" [Yalcin, 2011, p. 314].

[25]In this connection, we note that our models cannot accommodate one of the phenomena which originally motivated Lewis [1982]. Lewis describes himself as in some contexts believing that Nassau Street runs East-West and in others believing that it runs North-South. Since we demand a coherent belief state, we cannot represent agents who believe p in one setting, and $\neg p$ in another. But Yalcin himself does not provide a systematic way of handling the example either, since it is natural to think that the "alternative semantic value" (i.e. a question) generated by "Nassau Street runs East-West" is the same as that of "Nassau Street runs North-South". The same point seems to apply to Yablo's proposal, since it is unclear how exactly the two claims above would introduce different "subject matters" (for the understanding of partitions as subject matters, see Lewis [1988b], and Lewis [1988a]). For further discussion of Yalcin, and for a different proposal for how such cases could be handled, see Goodman and Lederman [2015].

number of authors have considered views according to which whether an agent believes a proposition may be determined only relative to a partition of the probability space (e.g. Levi [1967], Lin and Kelly [2012], and Leitgeb [2014]).

5. Decision Theory

One advantage of standard state spaces is that we can build on a large body of work in decision theory which uses them. In this section, we discuss one way in which standard representation theorems can be applied to generate probabilities and utilities for agents represented by partitional models.[26]

In section 4.6 the example of the number of stars illustrated how one may believe and know things to which one is not attending. This kind of inattention may also affect choice. If Alice were asked about who will win the next presidential election, she might have definite odds at which she would accept bets on different candidates. But Alice may nevertheless not be attending to questions concerning the next presidency when she is making her decisions regarding various medium-term investments. When deciding between these different investment choices, it is plausible that Alice should not be represented as "using" the odds induced by her implicit choice-dispositions for acts which are distinguished only by how they depend on who the next president will be. Even though Alice's implicit preferences are represented by a probability distribution which distinguishes possibilities that differ only as regards who will become president, her explicit beliefs in the context of choice may fail to distinguish these possibilities.

For concreteness, we will develop our approach formally in the setting of Savage [1954]; since our models are standard state spaces, we could have used with *any* of the standard settings for the usual representation theorems.[27] Recall that in Savage's set-up S is a primitive set of states, X the space of outcomes, and $F = X^S$ the "acts", functions from states to outcomes. If preferences defined over acts satisfy certain axioms, the agent can be represented by a unique probability function on $\langle S, \mathcal{P}(S)\rangle$, and a utility function on X which is unique up to affine transformations.

To represent the agent's awareness in a choice context, we select a complete atomic sub-algebra of $\mathcal{P}(S)$, which we call $\mathcal{B}^C$, the agent's algebra in the context. The atoms of $\mathcal{B}^C$ are a partition of S, so we have a partitional model of unawareness. The distribution the agent "uses" in this context is given by letting $\mu^C(E) = \mu(E)$ for all $E \in \mathcal{B}^C$ and undefined otherwise. The events the agent "explicitly believes" in the context can then be defined as the events of which the agent is certain in μ^C. Contexts are required to be coherent in the sense that the agent is only presented with choices between acts that are measurable in $\mathcal{B}^C$.[28]

---

[26] We do not here attempt to back-form what the agent is aware of from her choice-dispositions, as Morris [1996, 1997] does for belief, and Schipper [2013, 2014a] does for awareness.

[27] e.g. von von Neumann and Morgenstern [1944], Savage [1954], Anscombe and Aumann [1963], Bolker-Jeffrey (Bolker [1967], Jeffrey [1983], Broome [1990]).

[28] An alternative approach, related to that of Karni and Vierø [2013] is also worth mentioning. On this approach, which goes most naturally with awareness as conceivability, the agent does not have preferences over acts which differ only on events of which she is unaware. We start from a state-space S already endowed with a partition $\mathcal{Q} \subseteq \mathcal{P}(S)$. We consider agents whose preferences are defined only over those acts which are constant on cells of the partition: this space of "conceivable" acts is thus $F^C = X^{\mathcal{Q}}$. The representation theorem goes through as before, since the cells of the partition simply behave as the states in the original result did. The result is a representation with a distribution over the space of what Savage [1954] called "small-worlds". The

We can generalize beyond Savage's setting (where the underlying algebra is the powerset algebra) to an arbitrary measurable space, as follows. An agent's state in a choice context is represented as $\langle S, \mathcal{B}, \mu, \mathcal{B}^C \rangle$, where $\langle S, \mathcal{B}, \mu \rangle$ is a measure space, and $\mathcal{B}^C$ is a complete atomic subalgebra of $\mathcal{B}$.

Here $\mathcal{B}^C$ can be used to parametrize "expanding" and more generally "changing" awareness, represented as transitions between different complete atomic subalgebras of $\mathcal{B}$.[29] Bayesian agents are typically thought to be able to "learn" in two ways: by conditionalizing or by altering the space of what they consider possible. In our setting, the first kind of learning operates as usual on μ itself; the second kind of learning alters the appropriate $\mathcal{B}^C$ and hence alters μ^C. Since different algebras will determine different *explicit* beliefs in different contexts, this changing awareness can also represent effects of limited attention such as framing effects or failures of recall.

An approach along these lines has already proven fruitful in epistemic game theory. Kets [2014a] and Kets [2014b] develops Harsanyi type spaces in which players' beliefs may be defined on different σ-algebras. If the algebras are taken as the events the agent is attending to, one may interpret these models as examples of agents who fail to attend to questions about the higher-order beliefs of others, and thus do not have *explicit* beliefs over events which can be defined only by the level-n beliefs of others for large enough n.

5.1. **Speculative Trade.** An important test of approaches to unawareness has been how they fare with speculative trade (Heifetz et al. [2013]). Building on the work of Aumann [1976], Milgrom and Stokey [1982] proved that agents with $S5$ knowledge – agents whose accessibility relation R^i is an equivalence relation – could not commonly know that they were engaging in speculative trade. This "no-trade" theorem is a kind of paradox, illustrating the extreme strength of $S5$ knowledge together with a common prior.[30] One aim of representing "bounded" agents such as those with limited attention is to escape such paradoxes. Accordingly, we now provide a partitional DLR model with a common prior in which speculative trade is possible.

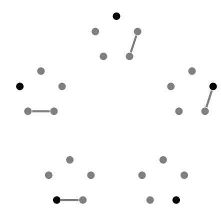
As is well known (see especially Geanakoplos [1989], Samet [1990], Rubinstein and Wolinsky [1990]), the "no-trade" theorem does not hold under weakenings of $S5$ for knowledge; in particular it does not hold if the accessibility relation R^i is merely transitive and reflexive, but does not form an equivalence relation. Modica and Rustichini [1994] argued that, since $S5$ was incompatible with Plausibility, it could not be used in conjunction with a representation of awareness. Still, the DLR axioms together with partitional awareness models impose substantial further constraints, which might be thought to rule out speculative trade. We now construct a partitional DLR model to show that speculative trade can still occur in the presence of DLR's axioms and nontrivial unawareness.

---

modeler (and perhaps other agents) may be able to conceive of events which the agent cannot, but as far as the agent is concerned, "states" are just the cells of her partition.

[29]In a number of theories, learning is partition relative. In Jeffrey conditionalization [Jeffrey, 1983] for example, a learning experience is a pair of a partition and weights on cells of the partition. Our approach differs by (1) taking unconditional probabilities also to be partition-relative, and (2) taking learning-by-refinement to be a distinct form of "update".

[30]For this perspective, see Morris [1995] and Lederman [2014].



Let the states be $W = \{1, 2, 3, 4, 5\}$ and the agents be Alice, A, and Bob B. The accessibility relations are defined so that: $1R^Ax$ iff $x \leqslant 3$; $5R^Ax$ iff $x \geqslant 3$ and otherwise wR^Ax iff $w = x$, while $R^B = W \times W$. The partitions of the agents are induced by $\approx^A_1=\approx^A_2= \{\{1\}, \{2, 3\}, \{4\}, \{5\}\}$; $\approx^A_4=\approx^A_5= \{\{1\}, \{2\}, \{3, 4\}, \{5\}\}$; and for all w, $\approx^B_w=\approx^A_3= \{\{1\}, \{2\}, \{3\}, \{4\}, \{5\}\}$. (Alice's is shown in the figure above.) The agents' common prior is the uniform one, and two acts f and g have utility as follows: $f(1) = f(5) = 1$, $f(2) = f(4) = 5$, $f(3) = 7$; $g(w) = 4$ for all $w \in W$. If the agents update by conditionalization on their implicit knowledge, then Alice will invariably maximize utility by choosing f (since in states $2, 3, 4$ she is certain it does better, and in states 1 and 5 she expects to gain $1/3 \cdot 1 + 1/3 \cdot 5 + 1/3 \cdot 7 > 4$). Bob meanwhile does not update at all, so that he strictly prefers g (since $4 > 2/5 \cdot 1 + 2/5 \cdot 5 + 1/5 \cdot 7$).

6. Conclusion

Standard state space models of awareness are at least as successful as current non-standard state space models. The non-standard models are more complicated, and it is unclear that this complexity affords any advantages in predictive strength or accuracy. Standard state space models of these phenomena promise to lead to a rich and rewarding theory, posing technical and conceptual challenges, and offering connections to related work by linguists, philosophers and logicians – as well as work on bounded reasoning elsewhere in economic theory.

Standard state space models are not a panacea for all issues related to bounded rationality. Some interesting notions of awareness as processing may satisfy the axiom AC; the triviality result in Section 3 demonstrates that for any such notion of processing, non-standard models *are* required. Even models such as those discussed in Dekel et al. [1998] and Heifetz et al. [2006], which we have shown to be inadequate as models of processing, may find other applications for notions of awareness not discussed here.

We emphasize that we are not dogmatic about the use of standard state space models, or the assumption of a coarse-grained theory of content. There may be good arguments to show that standard state space models can't model awareness. It may be that research into awareness can show that a coarse-grained theory of content, on which events form a complete atomic Boolean algebra, is incorrect. It may be that non-standard state space models can help to uncover arguments against this Boolean, coarse-grained theory. But so far, we have not seen any such arguments.

References

Francis J. Anscombe and Robert J. Aumann. A definition of subjective probability. *Annals of mathematical statistics*, pages 199–205, 1963.

Robert J. Aumann. Agreeing to disagree. *The Annals of Statistics*, 4(6):1236–1239, November 1976.

Robert J. Aumann. Interactive epistemology I: Knowledge. *International Journal of Game Theory*, 28:263–300, 1999.

Patrick Blackburn, Maarten de Rijke, and Yde Venema. *Modal Logic*. Cambridge University Press, 2001.

Ethan D. Bolker. A simultaneous axiomatization of utility and subjective probability. *Philosophy of Science*, pages 333–340, 1967.

John Broome. Bolker-Jeffrey expected utility theory and axiomatic utilitarianism. *The Review of Economic Studies*, 57(3):477–502, 1990.

Brian F. Chellas. *Modal Logic: An Introduction*. Cambridge: Cambridge University Press, 1980.

Eddie Dekel, Barton L. Lipman, and Aldo Rustichini. Standard state-space models preclude unawareness. *Econometrica*, 66(1):159–73, 1998.

Arthur Conan Doyle. The adventure of Silver Blaze. *The Strand Magazine*, 1901.

Andy Egan. Seeing and believing: perception, belief formation and the divided mind. *Philosophical Studies*, 140(1):47–63, 2008.

Adam Elga and Agustín Rayo. Fragmentation and information access. MS, November 2014.

Ronald Fagin and Joseph Y. Halpern. Belief, awareness, and limited reasoning. *Artificial Intelligence*, 34:39–76, 1988.

Ronald Fagin, Joseph Y. Halpern, Yoram Moses, and Moshe Y. Vardi. *Reasoning about Knowledge*. MIT Press, 1995.

Kit Fine. Propositional quantifiers in modal logic. *Theoria*, 36:336–346, 1970.

Gottlob Frege. Über Sinn und Bedeutung. *Zeitschrift für Philosophie und philosophische Kritik*, 100:25–50, 1892.

Jane Friedman. Question-directed attitudes. *Philosophical Perspectives*, 27:145–174, 2013.

Peter Fritz. Logics for propositional contingentism. unpublished a.

Peter Fritz. Propositional contingentism. unpublished b.

Spyros Galanis. Unawareness of theorems. *Economic Theory*, 52:41–73, 2013.

John Geanakoplos. Game theory without partitions, and applications to speculation and consensus. Technical report, Cowles Foundation Discussion Paper, 1989.

Jeremy Goodman and Harvey Lederman. Guise theory. MS, January 2015.

Jeroen Groenendijk and Martin Stokhof. *Studies on the Semantics of Questions and the Pragmatics of Answers*. PhD thesis, University of Amsterdam, 1984.

Davide Grossi and Fernando R. Velázquez-Quesada. Twelve angry men: A study on the fine-grain of announcements. In Xiangdong He, John Horty, and Eric Pacuit, editors, *Logic, Rationality, and Interaction*, pages 147–160. Berlin: Springer, 2009.

Joseph Y. Halpern. Alternative semantics for unawareness. *Games and Economic Behavior*, 37(2):321–339, 2001.

Joseph Y Halpern and Riccardo Pucella. Dealing with logical omniscience: Expressiveness and pragmatics. *Artificial Intelligence*, 175(1):220–235, 2011.

Joseph Y. Halpern and Leandro C. Rêgo. Reasoning about knowledge of unawareness. *Games and Economic Behavior*, 67(2):503–525, 2009.

Joseph Y. Halpern and Leandro C. Rêgo. Reasoning about knowledge of unawareness revisited. *Mathematical Social Sciences*, 65(2):73–84, 2013.

Joseph Y Halpern, Yoram Moses, and Moshe Y Vardi. Algorithmic knowledge. In *Proceedings of the 5th conference on Theoretical aspects of reasoning about knowledge*, pages 255–266. Morgan Kaufmann Publishers Inc., 1994.

C. L. Hamblin. Questions in Montague English. *Foundations of Language*, 10: 41–53, 1973.

Aviad Heifetz, Martin Meier, and Burkhard C. Schipper. Interactive unawareness. *Journal of Economic Theory*, 130(1):78–94, 2006.

Aviad Heifetz, Martin Meier, and Burkhard C. Schipper. A canonical model for interactive unawareness. *Games and Economic Behavior*, 62:304–324, 2008.

Aviad Heifetz, Martin Meier, and Burkhard C. Schipper. Unawareness, beliefs, and speculative trade. *Games and Economic Behavior*, 77:100–121, 2013.

Sander Heinsalu. Equivalence of the information structure with unawareness to the logic of awareness. *Journal of Economic Theory*, 147(6):2453–2468, 2012.

Sander Heinsalu. Universal type structures with unawareness. *Games and Economic Behavior*, 83:255–266, 2014.

Jaako Hintikka. *Knowledge and Belief.* Cornell University Press, 1962.

Richard C. Jeffrey. *The logic of decision.* Chicago: University of Chicago Press, second edition, 1983.

Edi Karni and Marie-Louise Vierø. "Reverse Bayesianism": A choice-based theory of growing awareness. *The American Economic Review*, 103(7):2790–2810, 2013.

Lauri Karttunen. Presupposition and linguistic context. *Theoretical linguistics*, 1 (1):181–194, 1974.

Lauri Karttunen. Syntax and semantics of questions. *Linguistics and Philosophy*, 1:3–44, 1977.

Willemien Kets. Bounded reasoning and higher-order uncertainty. Submitted MS, September 2014a.

Willemien Kets. Finite depth of reasoning and equilibrium play in games with incomplete information. Submitted MS, February 2014b.

Kurt Konolige. What awareness isn't: A sentential view of implicit and explicit belief. In Joseph Y. Halpern, editor, *Proceedings of the 1st Conference on Theoretical Aspects of Reasoning and Knowledge (TARK 1986)*, pages 241–250. Los Altos, CA: Morgan Kaufmann, 1986.

Philipp Koralus. The erotetic theory of attention: Questions, focus and distraction. *Mind & Language*, 29(1):26–50, 2014.

Manfred Krifka. Questions. In Klaus von Heusinger, Claudia Maienborn, and Paul Portner, editors, *Semantics. An International Handbook of Natural Language Meaning*, volume 2, pages 1742–1785. Berlin: Mouton de Gruyter, 2011.

Saul A. Kripke. A completeness theorem in modal logic. *Journal of Symbolic Logic*, 24:1–14, 1959.

Saul A. Kripke. Semantical analysis of modal logic I: Normal modal propositional calculi. *Zeitschrift für mathematische Logik und Grundlagen der Mathematik*, 9: 67–96, 1963.

Harvey Lederman. People with common priors can agree to disagree. *The Review of Symbolic Logic*, pages 1–35, 2014.

Hannes Leitgeb. The stability theory of belief. *Philosophical Review*, 123(2):131–171, 2014.

Isaac Levi. *Gambling with truth.* Knopf, 1967.

David Lewis. Logic for equivocators. *Noûs*, pages 431–441, 1982.

David Lewis. Relevant implication. *Theoria*, 54(3):161–174, 1988a.

David Lewis. Statements partly about observation. *Philosophical Papers*, 17(1): 1–31, 1988b.

Jing Li. Information structures with unawareness. *Journal of Economic Theory*, 144(3):977–993, 2009.

Hanti Lin and Kevin T. Kelly. A geo-logical solution to the lottery paradox, with applications to conditional logic. *Synthese*, 186(2):531–575, 2012.

Thomas McKay and Michael Nelson. Propositional attitude reports. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy.* Spring 2014 edition, 2014.

Martin Meier and Burkhard C. Schipper. Bayesian games with unawareness and unawareness perfection. *Economic Theory*, 56:219–249, 2014.

Paul Milgrom and Nancy Stokey. Information, trade and common knowledge. *Journal of Economic Theory*, 26(1):17–27, 1982.

Salvatore Modica and Aldo Rustichini. Awareness and partitional information structures. *Theory and Decision*, 37(1):107–124, 1994.

Salvatore Modica and Aldo Rustichini. Unawareness and partitional information structures. *Games and Economic Behavior*, 27(2):265–298, 1999.

Richard Montague. Universal grammar. *Theoria*, 36:373–398, 1970.

Stephen Morris. The common prior assumption in economic theory. *Economics and Philosophy*, 11(2):227–253, 1995.

Stephen Morris. The logic of belief and belief change: A decision theoretic approach. *Journal of Economic Theory*, 69(1):1–23, 1996.

Stephen Morris. Alternative definitions of knowledge. In Michael O. L. Bacharach, editor, *Epistemic logic and the theory of games and decisions*, pages 217–233. Kluwer Academic Publishers, 1997.

Graham Priest. *An Introduction to Non-Classical Logic: From If to Is.* Cambridge: Cambridge University Press, 2008.

Riccardo Pucella. Deductive algorithmic knowledge. *Journal of Logic and Computation*, 16(2):287–309, 2006.

Craige Roberts. Information structure in discourse: Towards an integrated formal theory of pragmatics. *Working Papers in Linguistics-Ohio State University Department of Linguistics*, pages 91–136, 1996.

Craige Roberts. Information structure in discourse: Towards an integrated formal theory of pragmatics. *Semantics and Pragmatics*, 5(6):1–69, 2012.

Ariel Rubinstein and Asher Wolinsky. On the logic of "agreeing to disagree" type results. *Journal of Economic Theory*, 51(1):184–193, 1990.

Dov Samet. Ignoring ignorance and agreeing to disagree. *Journal of Economic Theory*, 52(1):190–207, 1990.

Leonard J. Savage. *The Foundations of Statistics.* John Wiley and Sons, 1954.

Burkhard C. Schipper. Awareness-dependent subjective expected utility. *International Journal of Game Theory*, 42(3):725–753, 2013.

Burkhard C. Schipper. Preference-based unawareness. *Mathematical Social Sciences*, 70:34–41, 2014a.

Burkhard C. Schipper. Unawareness—a gentle introduction to both the literature and the special issue. *Mathematical Social Sciences*, 70:1–9, 2014b.

Burkhard C. Schipper. Awareness. In Hans van Ditmarsch, Joseph Y. Halpern, Wiebe van der Hoek, and Barteld Kooi, editors, *Handbook of Logics for Knowledge and Belief.* London: College Publications, forthcoming.

Dana Scott. Advice on modal logic. In Karel Lambert, editor, *Philosophical Problems in Logic: Some Recent Developments*, pages 143–173. Dordrecht: D. Reidel, 1970.

Krister Segerberg. *An Essay in Classical Modal Logic.* Number 13 in Filosofiska Studier. Uppsala: Uppsala Universitet, 1971.

Giacomo Sillari. Models of awareness. In Giacomo Bonanno, Wiebe van der Hoek, and Michael Wooldridge, editors, *Logic and the Foundations of Game and Decision Theory (* LOFT *7).* Amsterdam: Amsterdam University Press, 2006.

Giacomo Sillari. Quantified logic of awareness and impossible possible worlds. *The Review of Symbolic Logic*, 1:514–529, 2008.

Robert Stalnaker. Pragmatic presuppositions. New York University Press, New York, 1974.

Robert Stalnaker. *Inquiry.* Cambridge MA: MIT Press, 1984.

Robert Stalnaker. The problem of logical omniscience, I. *Synthese*, 89:425–440, 1991.

Robert Stalnaker. *Context and Content: Essays on Intentionality in Speech and Thought.* Oxford University Press, Oxford, 1999.

Robert Stalnaker. *Mere Possibilities.* Princeton: Princeton University Press, 2012.

Robert Stalnaker. *Context.* Oxford University Press, Oxford, 2014.

Elias Thijsse. On total awareness logics. In Maarten de Rijke, editor, *Diamonds and Defaults*, pages 309–347. Dordrecht: Kluwer, 1993.

Johan van Benthem and Fernando R. Velázquez-Quesada. The dynamics of awareness. *Synthese*, 177:245–267, 2010.

John von Neumann and Oskar Morgenstern. *Theory of Games and Economic Behavior.* Princeton: Princeton University Press, 1944.

Timothy Williamson. *Identity and discrimination.* John Wiley & Sons, 1990.

Stephen Yablo. *Aboutness.* Princeton University Press, 2014.

Seth Yalcin. Nonfactualism about epistemic modality. In Andy Egan and Brian Weatherson, editors, *Epistemic Modality.* Oxford University Press, 2011.