Bounded Rationality and Epistemic Blindspots

Abstract: Real-world agents do not know all consequences of what they know. But we are reluctant to say that a rational agent can fail to know some trivial consequence of what she knows. Since every consequence of what she knows can be reached via chains of trivial consequences of what she knows, we have a paradox. In this paper, I respond to the paradox in three stages. (i) I describe formal models which allow us to draw a distinction, at the level of content, between trivial (uninformative) and non-trivial (informative) inferences. (ii) I argue that agents can fail to know trivial consequences of what they know, but they can never do so determinately. Such cases are epistemic blindspots, and we are never in a position to assert that such-and-such constitutes a blindspot for agent *i*. (iii) I develop formal epistemic models on which the epistemic accessibility relations are vague. Given these models, we can show that epistemic blindspots always concern indeterminate cases of knowledge.

Keywords: Bounded rationality, logical omniscience, epistemic blindspots, problem of deduction, informative inference

1 Introduction

According to the possible worlds model of knowledge and belief, all agents know all logical consequences of what they know, and believe all logical consequences of what they believe. But real-world agents are not *logically omniscient* in this way. Deductive reasoning is often informative. We may know the premises (or have assumed them for the purposes of argument), and yet the conclusions we draw from them are often informative, and sometimes surprising.

Responses to the logical omniscience problem, taken as a technical problem for a particular semantics for knowledge and belief, are a dime a dozen. But the problem does not arise merely within a particular formal semantics: it is not an artefact of a particular theoretical framework. It is part of a broader problem of rationality quite generally. The problem concerns how deductive reasoning can be informative to us, given that the premises must guarantee the conclusion. In general, learning something, or gaining some new information, is a matter of having ruled out some possible scenarios. But in the case of a valid inference, there is no possibility of the premises being true and the conclusion being false, and hence nothing to be ruled out in learning that the premises entail the conclusion.

The problem is hard because all of the inference rules we typically work with are *trivial* and uninformative. Yet together (in a standard system) these rules are deductively complete, and hence any valid deduction is no more than the repeated application of these trivial, uninformative reasoning steps. As Dummett says,

> When we contemplate the simplest basic forms of inference, the gap between recognising the truth of the premisses and recognising that of the conclusion seems infinitesimal; but, when we contemplate the wealth and complexity of

> number-theoretic theorems which, by chains of such inferences, can be proved . . . we are struck by the difficulty of establishing them and the surprises they yield. (1978, 297)

So it will not help the problem merely to insist that any valid inference from *A* to *B* counts as informative. The task is to explain the difference between the trivial and the informative inferences, in such a way that chains of trivial inferences can be informative as a whole. Following Stalnaker (1984), I'll call this *the problem of deduction*.

The analogue of this problem in terms of knowledge is this. For any trivial consequence *A* of what agent *i* knows, it seems that *i* must also know that *A*. At least, there is strong rational resistance to describing an agent as failing to know a trivial consequence of what she knows (and similarly for what she believes). This rational resistance is built into our practice of knowledge and belief ascriptions. Yet it cannot be that agents know all trivial consequences of what they know, for otherwise they would know *all* consequences of what they know, and hence would be logically omniscient.

I will say that an agent has an *epistemic blindspot* when *A* is a trivial consequence of what she knows but she does not know that *A*.[1] The problem is then that agents who are not logically omniscient must suffer from epistemic blindspots, but we can never describe a particular blindspot without treating the agent in question as being irrational. I'll call this problem the *problem of rational knowledge*: how to reconcile the existence of epistemic blindspots with rational knowledge ascriptions. Both the problem of deduction and the problem of rational knowledge concern rational but not logically ideal agents. The problem in each case is one of describing rational agents, without treating them as being logically omniscient ideal reasoners. These are problems of *bounded rationality*.

My aim in this paper is as follows. After describing the problems in more detail (§2), I develop formal models which allow us to draw a distinction, at the level of content, between trivial (uninformative) and non-trivial (informative) inferences (§3 and §4). I then discuss how far this gets us with the problem of rational knowledge, and which aspects of the problem remain. I argue that we should understand the remaining aspects of the problem in terms of *rational unassertibility*, underpinned by vagueness. In §6, I develop formal epistemic models which incorporate these ideas. In this formal system, agents may be in epistemic blindspots, yet we can prove that such cases always concern indeterminate cases of knowledge.

2 Problems of Bounded Rationality

In this section, I discuss the problems of bounded rationality in more detail. This is how Stalnaker (1984) characterises the problem of deduction:

1. Sorensen (1984a;b) uses 'blindspot' to mean a true but inaccessible proposition. I am using 'epistemic blindspot' in a different (although related) sense.

> The problem is to explain how it is possible for the conclusion of a deductive argument to contain any information not already contained in the premises and, as a special case of this, how it is possible for a necessary truth to contain any information at all. An answer to this question is needed to explain how drawing deductive inferences can be a way of increasing one's knowledge, and how knowledge of necessary truths can be knowledge at all. (Stalnaker 1984, 25)

The problem is brought into sharp relief by considering *content* in terms of sets of possible worlds (Stalnaker 1976; 1984; Lewis 1986). For then, equivalent contents are identical, and hence no content can distinguish one valid deduction from another. Any content assigned to the premises of a valid inference already contains the content of the conclusion, and so there is no sense in which the move from premises to conclusion can be informative. But importantly, as Stalnaker goes onto say,

> The problem does not arise from any easily identifiable philosophical dogma which might be given up to avoid it. ... the conclusion really derives not from any substantive assumption about the source of knowledge, but from the abstract concept of content or information. The difficulty is, I think, that any way of conceiving of necessary truths as having content is at the same time a way of conceiving of them as contingent—as one way things could have been among others. This is, I think, because we do think of content and information in terms of alternative possibilities. Whether the source of my information is my senses, authority, or a faculty of intellectual intuition with access to a Platonic realm of abstract entities, its deliverances are not news unless they might have been different. (Stalnaker 1984, 25)

I agree with Stalnaker that the problem is quite general, and not the result of a particular semantic account. To highlight this fact, we can formulate a version of the general problem in terms of *closure conditions* on knowledge, without assuming the possible worlds account. Take a rule such as 'if agent i knows that A and that B, then she also knows that $A \wedge B$'. If this is rule is exceptionless, then knowledge is *closed* under *conjunction introduction*. If knowledge is closed under a deductively complete set of rules, it follows that each agent knows all consequences of what she knows: she is *logically omniscient*. This, it seems, is implausible. But if an agent is not logically omniscient, there is some inference rule ρ such that the agent's knowledge is not closed under ρ. Since each inference rule states a *trivial* inference, there is then some *trivial* consequence of the agent's knowledge that she does not know. This too seems implausible. So we have a dilemma: either we treat agents as logically omniscient, knowing all consequences of what they know, or we treat them as *logically ignorant*, failing to know some *trivial* consequence of what they know.[2] I will call the problem posed by this dilemma the *problem of rational knowledge*.

The problem arises regardless of whether we assume the possible worlds account of knowledge, according to which an agent agent i knows that A

2. The terminology 'logical omniscience *vs* logical ignorance' is from Ho (1995; 1997).

if and only if A is true according to all worlds *epistemically accessible* to i (Hintikka 1962; Lewis 1996). Yet it is instructive to consider the problem from the perspective of the possible worlds semantics. On that approach, an agent's knowledge is deductively closed if and only if the truths according to a world are deductively closed, for all worlds epistemically accessible to her. Hence to avoid logical omniscience (if we assume the standard possible worlds clause for knowledge ascriptions), there must be some world w accessible to the agent in question for which the truths are *not* deductively closed. Such worlds are often called *impossible* worlds (Hintikka 1975).

The strategy of dealing with logical omniscience by including impossible as well as possible worlds has proved a popular one. Hintikka (1975), Rantala (1975; 1982), Belnap (1977), Levesque (1984), Lakemeyer (1986; 1987) and Fagin et al. (1990) take this approach. And indeed, it is successful in delivering a worlds-based semantics on which agents are not modelled as knowing all consequences of what they know. But these approaches do not deal with the broader problems of bounded rationality. If the truths at some epistemically accessible world w fail some trivial inference rule ρ, then there are some truths according to w such that A follows trivially from these, but A is not true according to w. Then w seems to be not just impossible, but *trivially* impossible. Clearly, any trivially impossible situation is not an *epistemic* possibility, and so such worlds should not be counted as being epistemically accessible (from any world), for any rational agent.

The problem, therefore, is not one of providing a formal semantics which includes non-deductively-closed worlds. That task is relatively easy: we can use models of paraconsistent logic, or even arbitrary sets of sentences, as 'worlds'. The hard problem is to provide a notion of a world (or scenario, or situation) which is impossible, but not trivially so. Lewis (in complaining about paraconsistent logic) puts the point nicely:

> I'm increasingly convinced that I can and do reason about impossible situations. ... But I don't really understand how that works. Paraconsistent logic ... allows (a limited amount of) reasoning about *blatantly* impossible situations. Whereas what I find myself doing is reasoning about *subtly* impossible situations, and rejecting suppositions that lead fairly to blatant impossibilities. (Lewis 2004, 176)

A 'paraconsistent' world (or rather, a world or model used in the semantics of paraconsistent logic, on which a sentence may be both true and false) is an impossible world in the sense used above. But whenever such a world w is impossible, it is trivially impossible: some A is both true and false, according to w. But, as Lewis recognises, for the purposes of modelling rational but not ideal reasoners, we require *subtly* impossible worlds (or models of subtly inconsistent situations). Hintikka (1975) makes a similar point. He argues for a solution in terms of impossible worlds, and insists that impossible worlds must be 'subtly inconsistent' worlds (1975, 478) which 'look possible but which contain hidden contradictions' (1975, 476).

In the next section, I begin to develop formal models which aim at 'subtly inconsistent' worlds, and which may therefore be epistemically accessible for a rational agent.

3 Epistemic Space

An epistemically possible world is one which *seems* possible, and so one which can't be ruled out *easily* through *a priori* reasoning. As Hintikka says, epistemically possible but logically impossible worlds are ones which 'look possible but which contain hidden contradictions' (Hintikka 1975, 476). In previous work (reference omitted), I show how to construct an *epistemic space*, a structured domain of worlds, all of which are epistemically possible. My aim was to use an epistemic space to give an account of the content of deduction. On this account, some but not all valid deductions come out as being contentful and hence informative. The epistemic space approach allows us to make a distinction, at the level of content, between trivial and informative (non-trivial) valid inferences.

The epistemic space approach does not, in itself, solve the problem of bounded rationality. It is the beginning of a solution, in allowing us to draw a distinction between trivial and informative inferences. But it does not yet provide a response to the problem of rational knowledge discussed above. In this section, I'll give a brief overview of how an epistemic space is constructed, show how it helps with the problem of deduction (§4), and highlight aspects of the problem of rational knowledge which remain (§5). Then, in §6, I'll extend the approach to deal with the problem of knowledge.

In terms of worlds, the dilemma we face is this. If truth at a world is deductively closed, then the worlds we get are all logically possible worlds (with the exception of the *trivial* world, according to which everything is true), and we are back to the standard possible worlds account. Call such a space of worlds *ideal epistemic space*. It gives us idealised contents, in the sense that agents are treated as being logically omniscient. If truth at some world w is not deductively closed, on the other hand, then w looks to be trivially impossible. Call any space involving such worlds *trivial epistemic space*. It allows us to represent agents who are not logically omniscient, but it does not do justice to an agent's *rationality*. Rational agents reject trivially impossible scenarios as epistemic possibilities.

What we require is a *rational* (hence non-trivial) but *non-ideal* epistemic space. To construct one, we being with *very fine-grained worlds*. These worlds are fine-grained in the sense that, for any sets of sentences Γ and Δ, there is a fine-grained world w which represents that it is the case that A for each $A \in \Gamma$ and which represents that it is not the case that B for each $B \in \Delta$. Rather than closing what is true (and what is false) according to a world under deductive consequence, we use inference rules to assign a *rank* to each world. In (reference omitted), I do this by imposing a graph-structure on worlds, and defining the rank of a world w in terms of the graphs, built in a certain way, which have w as their root. To save space, I'll tell a much shorter story here. For the details I skip over (and for the 'official' way of doing things), see (reference omitted).

To assign a rank to each of our fine-grained worlds, we first need to fix on a system of proof rules with certain properties. In (reference omitted), I give a system of *sequent rules*. For the uninitiated, *sequents* have the form $\Gamma \vdash \Delta$, where Γ and Δ may be sequences, multisets or sets of sentences. I'll assume they are sets. Rules manipulate such sequents, and may have either one or two sequents as premises (or *upper sequents*) and a single sequent as conclusions (or *lower sequents*). A standard sequent calculus for classical logic employs left and right logical rules for each connective (see, e.g., Buss 1998), plus a set of *structural rules*.[3] I adopt a slightly non-standard system, in which all sentences appearing in the lower sequent of a rule must appear in the upper sequent(s) too. This, together with the fact that Γ and Δ in a sequent are sets, allows us to live without structural rules other than the *identity* rule:[4]

$$\frac{}{\Gamma, A \vdash A, \Delta}\ [\textsc{id}]$$

The logical rules for '¬' and '∨', for example, are:

$$\frac{\Gamma, \neg A \vdash A, \Delta}{\Gamma, \neg A \vdash \Delta}\ [\neg\textsc{l}] \qquad \frac{\Gamma, A \vdash \neg A, \Delta}{\Gamma \vdash \neg A, \Delta}\ [\neg\textsc{r}]$$

$$\frac{\Gamma, A \vee B, A \vdash \Delta \qquad \Gamma, A \vee B, B \vdash \Delta}{\Gamma, A \vee B \vdash \Delta}\ [\vee\textsc{l}] \qquad \frac{\Gamma \vdash A, B, A \vee B, \Delta}{\Gamma \vdash A \vee B, \Delta}\ [\vee\textsc{r}]$$

A proof of a sequent $\Gamma \vdash \Delta$ is a tree of sequents whose root is $\Gamma \vdash \Delta$ and whose leaves are all instances of ID. In practice, proofs are constructed bottom-up, beginning with the sequent to be proved and working upwards, applying the rules from lower sequent to upper sequent(s).[5]

To assign a rank to a world w, suppose that Γ is the set of truths according to w and Δ is the set of falsehoods according to w. Then if w is refutable on logical grounds alone—that is, if w is implicitly inconsistent in what it represents as being true and false—then the sequent $\Gamma \vdash \Delta$ is valid in our system.[6] That is, there is a tree built from the proof rules, with $\Gamma \vdash \Delta$ at its root and an instance of ID at each of its leaves (which makes the tree a *closed* tree). Say that such a tree is *associated with* w. The *size* of a proof tree is the number of its non-leaf nodes. The rank of w, $\#w$, is then the size of the smallest closed tree associated with w, if there is one, and ω (the first limit ordinal) otherwise.

Intuitively, the lower the rank of a world w, the easier it is to rule out through logical reasoning. 'Ruling out a world' amounts to deriving an *explicit*

3. Structural rules allow us to manipulate sequents into the right form for applying logical rules, e.g., by changing the order of sentences in a sequent (the *exchange* rules) or deleting duplicates (the *contraction* rules).

4. Here, 'Γ, A' is shorthand for '$\Gamma \cup \{A\}$'.

5. This system is equivalent to the standard presentation: any sequent derivable in one system is derivable in the other. Hence the new system is sound and complete with respect to the truth-table semantics.

6. This follows from the fact that the sequent rules are complete with respect to the classical semantics.

contradiction from it, given what is (explicitly) true and what is false according to that world. A world should count as epistemically possible only if it isn't easy to rule it out in this way. So I assume that w is epistemically possible if and only if $\#w$ is not small. Of course, it's an indeterminate matter which numbers are small and which are not. My approach in (reference omitted) did not specify a way to resolve the indeterminacy. Rather, we can use whichever we think is in general the best semantics for vague terms. Thus, if you're an epistemicist, then you already think that some some particular integer n is the largest small number. You can then treat all worlds for which $\#w > n$ as the epistemically possible ones. Other approaches to vagueness will work differently. On any approach, for some worlds w it will be indeterminate whether w is epistemically possible. This is as it should be.[7] Nevertheless, we can think of each integer as a *sharpening* of 'epistemically possible world', and reason accordingly, relative to some sharpening or other.

With this in mind, consider a structure $E = \langle W, V^+, V^- \rangle$, where W is a set of fine-grained worlds and V^+ and V^- are labelling functions, telling us what is true and what is false according to a world, respectively. The rank of E is the smallest $\#w$ such that $w \in W$, that is, $\min\{\#w \mid w \in W\}$. E counts as a *genuine epistemic space* if and only if E's rank is not small, that is, if and only if all worlds in W are epistemically possible. I'll leave this as an informal notion. Again, we can think of each integer as giving a sharpening of 'epistemic space', qua set of all epistemically possible worlds, and reason accordingly. In what follows, I won't develop a formal semantics for reasoning about the indeterminacy of epistemic space. Instead, I'll work with particular sharpenings (all of which I'll call 'epistemic spaces'), and leave which are the admissible sharpenings as an informal notion.

4 Epistemic Entailment

Let E be a sharpening of epistemic space $\langle W, V^+, V^- \rangle$. A *pointed epistemic space* is a pair $\langle E, w \rangle$, where $w \in W$. The rank of $\langle E, w \rangle$ is the rank of E. Say that $E, w \Vdash A$ iff $A \in V^+w$ in E, and $E, w \dashV A$ iff $A \in V^-w$ in E. For a pointed space E, we define $E \Vdash \Gamma$ iff $E \Vdash A$ for each $A \in \Gamma$, and $E \dashV \Gamma$ iff $E \dashV A$ for at least one $A \in \Gamma$. We then define a notion of entailment, which I'll call *epistemic n-entailment*, as follows.

Definition 1 (Epistemic n-entailment) *For any integer $n \in \mathbb{N} \cup \{\omega\}$, a set of premises Γ epistemically n-entails A, $\Gamma \vDash^e_n A$, if and only if, for all pointed epistemic spaces E of rank $r > n$, $E \Vdash \Gamma$ only if $E \not\dashV A$.*

As a definition of entailment, this definition is rather unusual: the clause has '$E \not\dashV A$' where we would usually have '$E \Vdash A$'. Epistemic entailment is not about truth-preservation; rather, an epistemic entailment holds when the truth of the premises guarantee falsity avoidance for the conclusion.

7. If we could fix a determinate threshold for a world to count as being an epistemic possibility, we would have artificially sharpened 'epistemic possibility'.

Theorem 1 $\vDash_n^e$ *has the following properties, for all* $n \in \mathbb{N} \cup \{\omega\}$*:*

(a) $\vDash_n^e \subset \vDash_{n+1}^e$ *(that is, if* $\Gamma \vDash_n^e A$*, then* $\Gamma \vDash_{n+1}^e A$*).*

(b) $\vDash_n^e$ *is monotonic: if* $\Gamma \vDash_n^e A$ *and* $\Gamma \subseteq \Delta$ *then* $\Delta \vDash_n^e A$.

(c) $\Gamma \vDash_n^e A$ *only if* Γ *classically entails* A.

(d) $\vDash_n^e$ *is reflexive and non-symmetrical.*

(e) $\Gamma \vDash_0^e A$ *if and only if* $A \in \Gamma$.

(f) Except for $n = 0$, $\vDash_n^e$ *is non-transitive and does not satisfy cut: it is not the case that if* $\Gamma \vDash_n^e A$ *and* $\Gamma, A \vDash_n^e B$ *then* $\Gamma \vDash_n^e B$.

Proof: For (a), suppose that $\Gamma \vDash_n^e A$ and that E is a space of rank $r > n+1$ s.t. $E \Vdash \Gamma$. Then for all spaces E' of rank $r > n$ s.t. $E' \Vdash \Gamma$, $E' \not\dashv\!\vdash A$. Hence $E \not\dashv\!\vdash A$, and so $\Gamma \vDash_{n+1}^e A$. For (b), suppose $\Gamma \vDash_n^e A$, $\Gamma \subseteq \Delta$ and $E \Vdash \Delta$. Then $E \Vdash \Gamma$ and so, by definition, $E \not\dashv\!\vdash A$. Hence $\Delta \vDash_n^e A$. For (c), suppose Γ does not classically entail A. Then there is no closed tree in our sequent system with $\Gamma \vdash A$ at its root. Let $E = \langle \{w\}, V^+, V^- \rangle$ where $V^+w = \Gamma$ and $V^-w = \{A\}$, so that $E \Vdash \Gamma$ and $E \dashv\!\vdash A$. Since $\#w = \omega$, E has rank ω and hence $\Gamma \not\vDash_n^e$ for any $n \leq \omega$. (c) follows by contraposition.

For (d), first suppose that E has rank $r > 0$. Then there is no $w \in W$ in E such that $A \in V^+w \cap V^-w$ (since for any such world, $\#w = 0$). Hence if $E \Vdash A$ then $E \not\dashv\!\vdash A$, and so $A \vDash_0^e A$. Then, by (b), if $A \in \Gamma$ then $\Gamma \vDash_0^e A$ and, by (a), $\Gamma \vDash_n^e A$ for any $n \in \mathbb{N}$, hence each $\vDash_n^e$ is reflexive. Since classical entailment is non-symmetrical, (c) entails that each $\vDash_n^e$ is non-symmetrical. For (e), the 'if' follows from reflexivity. For the 'only if', suppose $A \notin \Gamma$ and let $E = \langle \{w\}, V^+, V^- \rangle$ where $V^+w = \Gamma$ and $V^-w = \{A\}$. Then $E \Vdash \Gamma$ and $E \not\Vdash A$. Since $A \notin \Gamma$, $\#w > 0$, hence E has a rank $r > 0$, hence $\Gamma \not\vDash_0^e A$.

For (f), note that Γ_0^e trivially satisfies cut. If $\Gamma \vDash_0^e A$ and $\Gamma, A \vDash_0^e B$ then given (e), $A \in \Gamma$, hence $\Gamma \cup \{A\} = \Gamma$ and so $\Gamma \vDash_0^e B$. Now suppose $n > 0$ and let $\Gamma_n = \{p_1, p_1 \to p_2, \ldots, p_n \to p_{n+1}\}$. It is easy to see that $\Gamma \vDash_n^e p_{n+1}$ and $\Gamma, p_{n+1} \vDash_n^e p_1 \wedge p_{n+1}$, but $\Gamma \not\vDash_n^e p_1 \wedge p_{n+1}$. For a counterexample, let $E = \langle \{w\}, V^+, V^- \rangle$ where $V^+w = \Gamma$ and $V^-w = \{p_1 \wedge p_{n+1}\}$. We have $E \Vdash \Gamma$ and $E \dashv\!\vdash p_1 \wedge p_{n+1}$. Since $\#w = n+1$ and so E has rank $n+1$, it follows that $\Gamma \not\vDash_n^e p_1 \wedge p_{n+1}$. Hence, for $n > 0$, $\vDash_0^e$ does not satisfy cut. Since cut is a form of transitivity, the argument that $\vDash_0^e$ is non-transitive is similar. ∎

So long as n is not too small, the epistemic n-entailments include all the inferences usually called trivial (on the use of 'trivial' that does not apply to all valid inferences). For example, we have:

$$A \wedge B \vDash_1^e A \qquad A, B \vDash_1^e A \wedge B$$
$$A \vDash_1^e A \vee B \qquad A \vee B, \neg A \vDash_2^e B$$
$$A \to B, A \vDash_1^e B \qquad A \to B, \neg B \vDash_3^e \neg A$$
$$\neg(A \wedge B) \vDash_5^e \neg A \vee \neg B \qquad \neg(A \vee B) \vDash_5^e \neg A \wedge \neg B$$

and so on.[8] Above, I said that we can think of each n within some admissible range as a sharpening of 'epistemic space'. In this range, the epistemic n-entailments give us sharpenings of 'trivial inference'. Thus, an inference from Γ to A is *determinately* trivial iff $\Gamma \vDash^e_n A$ for all n in the range, determinately non-trivial iff $\Gamma \nvDash^e_n A$ for all n in the range, and it is indeterminate whether the inference is trivial otherwise. (As I said above, I will not incorporate this machinery into our formal semantics: I will leave is as an informal notion.)

We can also use the notions defined above to say *what* the content of a valid deduction is, so that deductions concerning trivial inferences are contentless, whereas those concerning non-trivial inferences have content and hence are treated as being informative. The content of a valid deduction from Γ to A, relative to an epistemic space $E = \langle W, V^+, V^- \rangle$ of rank n, is the set of worlds $w \in W$ such that $\Gamma \subseteq V^+w$ but $A \in V^-$. According to such worlds, the premises are true but the conclusion is false. These are the worlds ruled out by an agent who performs the deduction from Γ to A. Now suppose $\Gamma \vDash^e_n A$. Then, for every world $w \in W$ such that $\Gamma \subseteq V^+w$, $A \notin V^-$, and hence the content of the deduction from Γ to A, relative to E, is empty. In this way, epistemic n-entailments guarantee that the corresponding deductions have no content (that is, they have content $\varnothing$), in all models of rank n. Then, given our informal understanding of which range of integers give admissible sharpenings of 'genuine' epistemic space, we can say which deductions are determinately informative, which are determinately uninformative and which are indeterminate between informative and trivial.

I've given a way to differentiate between trivial and non-trivial valid inferences, and to assign contents to deductions such that contentful (and hence informative) inferences are treated as non-trivial, whereas trivial ones are always contentless. On this account, some but not all valid inferences count as contentful and hence informative, as it should be.

5 The Problem Bites Back

We have epistemic spaces which, for some (informally specified) range of integers, tell us which worlds count as epistemic possibilities, and which can be used to differentiate between trivial and informative valid deductions. So it might seem a simple matter to develop epistemic spaces into models of agents' epistemic states. We would do this by adding an epistemic accessibility relation R_i for each agent i, and then give the usual clauses for 'agent i knows that A'. Since it is indeterminate which worlds are epistemically possible, it may be indeterminate whether R_i holds between worlds w and u. (More precisely, there may be admissible sharpenings on which R_iwu, and others on which $\neg R_iwu$.) Consequently, it may be indeterminate whether agent i knows that A. In this way, we allow that it may be indeterminate *what* i knows.

8. These listed entailments give the minimal n-value for which the epistemic n-entailment holds. For example, $A \vee B, \neg A \nvDash^e_1 B$. To see why, consider the space $E = \langle \{w\}, V^+, V^- \rangle$ such that $V^+w = \{A \vee B, \neg A\}$ and $V^- = \{B\}$. Then $\#w = 1$ and hence the pointed space $\langle E, w \rangle$ is of rank 1.

This feature, it might be thought, makes inroads into the problem of rational knowledge. We worried about epistemic blindspots, in which an agent does not know some trivial consequence of what she knows. If what an agent knows can be indeterminate, then perhaps we can avoid this problem. Things are not nearly so simple, however. The problem of rational knowledge bites back. In order to build epistemic spaces, I made use of very fine-grained worlds w, for which V^+w (the truths according to w) and V^-w (the falsehoods according to w) are arbitrary sets of sentences. These sets need not be logically structured in any way whatsoever. There are worlds which are radically 'gappy', in the sense that they say nothing whatsoever about the truth or falsity of some sentence A. Moreover, some world w may say nothing about A *explicitly*, but might say something about A *implicitly*, in the sense that the truths (or the falsehoods) according to w either entail A or entail $\neg A$. This can happen in a very stark way: w might say that $A \wedge B$ is true, for example, without saying anything explicitly about A. This opens the possibility of an agent knowing that $A \wedge B$ but not that A, according to our model. But, as noted above, we then seem to be treating the agent as being *irrational*, and not merely as a non-ideal reasoner.

Perhaps, you might think, the correct response is to allow a small amount of logical closure on worlds. But no good: either w represents that A and that B whenever it represents that $A \wedge B$, or it doesn't. If it does, then the world is fully closed with respect to conjunction elimination; if it doesn't, then there is some instance of $A \wedge B$ true according to w, such that one of the conjuncts isn't true according to w. Similarly for all other closure conditions. So again, either all agents are modelled as being logically omniscient, or else some agent does not know some trivial consequence of what she knows.

Instead, perhaps the correct response is to 'fill in' incomplete worlds as far as possible. On this approach, whenever w says nothing explicitly about A, we add either A or $\neg A$ to the set of truths according to w, or either $\neg A$ or A to the set of falsehoods according to w. That is, we insist that $V^+w \cup V^-w$ is a maximal set, containing either A or $\neg A$ for every A. Call such worlds *complete*. Perhaps as a result of adding the new sentences, w will become epistemically impossible. But no problem, so long as we recalculate the value of $\#w$ and update our opinion on whether the structure E containing w counts as an admissible sharpening of 'epistemic space'.

If we make this move, however, we soon run into a serious and irredeemable problem. Trivially, for any maximal consistent set X and any $A \notin X$, $X \cup \{A\}$ is trivially inconsistent. Thus any maximal inconsistent set is trivially inconsistent, containing $A, \neg A$ for some A. Now suppose that w is a complete but logically impossible world, so that $V^+x \vdash V^-w$ is a valid sequent. Then $V^+w \cup \{\neg A \mid A \in V^-w\}$ is both complete and inconsistent, hence trivially inconsistent, and hence either $\#w = 0$ or $\#w = 1$.[9] But determinately, no such world is epistemically possible (since 0 and 1 are determinately small numbers), and hence no such world appears in a genuine epistemic space. So, if we insist that all worlds be

9. If there is any $A \in V^+w \cap V^-w$, then $\#w = 0$. Otherwise, there is some $A, \neg A \in V^+w$ or $A, \neg A \in V^-w$, in which case, $\#w = 1$.

complete, then any impossible world immediately becomes trivially impossible and so is rejected from epistemic space. Epistemic space will then contain just the logically possible worlds, and we are back to ideal epistemic space, which treats all agents as being logically omniscient.

This problem is essentially the one discussed by Bjerring (2012), in response to (reference omitted). The results that Bjerring establishes show that, if we want to have an epistemic space which contains some logically impossible worlds but no trivially impossible worlds, then all those logically impossible worlds must also be incomplete worlds. But for an incomplete world w, there is some trivial consequence of the truths according to w which is not itself a truth according to w. Hence, when we build our model of knowledge from such epistemic spaces, we will have agents who do not know some trivial consequence of what they know, according to the model. Indeed, every logically non-omniscient agent in the model will be like this. Every such agent will have epistemic blindspots.

We cannot avoid the conclusion that epistemic space contains highly incomplete worlds. Explicit incompleteness, unlike explicit inconsistency, is not a reason for ruling a world out as a genuine epistemic possibility. We judge whether a world is an epistemic possibility based on what it represents, not in terms of what it does not represent. Hence, if what a world represents as being the case seems possible, then that world is an epistemically possible world, regardless of whether it is complete. But then, we must accept that agents suffer from epistemic blindspots. There is no avoiding this conclusion. Having accepted that agents suffer from epistemic blindspots, however, we need to explain why a particular blindspot can never be rationally ascribed to an agent. We cannot rationally claim that agent i does not know that either it is snowing or it is not snowing here right now, for example. The problem is to explain why this is so.

The best response available is as follows. It is never assertible that such-and-such constitutes an epistemic blindspot for agent i because it is never a *determinate* matter such-and-such constitutes a blindspot for agent i. Whenever an epistemic blindspot arises for an agent, it is always indeterminate that the agent is in that blindspot. And when A is indeterminate in truth, we may not rationally assert that A. This is the response I will adopt here. I claim that, although agents suffer from epistemic blindspots, these blindspots always concern indeterminate cases of knowledge. There is then no particular instance of an epistemic blindspot that is determinately a blindspot. In the next section, I build formal epistemic models which incorporate this idea. Before that, I need to say more on how taking epistemic blindspots to be indeterminate cases of knowledge helps with the problem of rational knowledge.

Suppose guests at a party are invited to take some number of sweets from a box. The greedy guests are those who take many sweets, but there is no precise number we can give to answer the question, 'what is the minimum number of sweets a guest could take, and thereby be greedy?' If n is the correct answer, then it is indeterminate that n is the correct answer and hence 'n is the correct answer' is not something we can genuinely know.[10] Had things (including the

10. Philosophical theories of vagueness disagree on whether any particular number in fact answers

meanings of words like 'greedy') been ever so slightly different, $n+1$ or $n-1$ rather than n might easily have been the correct answer. Given that we cannot epistemically discriminate between these three possibilities, we cannot *know* that n is the correct answer (even assuming it is).[11] As a consequence, we may not rationally *assert* the answer to be n. For one should assert only what one knows, or at least, what one has good reason to believe. In asserting 'n is the minimum number of sweets a guest could take and thereby be greedy', one is thereby claiming some evidence in favour of n in answer to the question, but one cannot have such evidence. In all cases of vagueness, we simply aren't at liberty to assert precisely in borderline cases, even if we happen (purely by chance) to hit on the truth.

This phenomena of *unassertibility at the borderline* helps us explain the rational prohibition on asserting that agent i does not know some trivial consequence of her knowledge, given that such cases are always indeterminate cases. Similarly, we can explain why we cannot rationally assert that agent i does not know that either it is snowing or it is not snowing here right now (for example). Given my overall picture, I am committed to the existence of such instances of knowledge failure, but they are always indeterminate instances. Whenever A is a trivial truth, either agent i knows that A or else it is indeterminate whether i knows that A: it is never determinate that i does not know that A. More generally, if A follows easily from what agent i determinately knows, then either i also knows that A, or else it is indeterminate whether i knows that A. There is never a case in which i determinately knows such-and-such, from which it easily follows that A, such that it is determinate that i does not know that A. In the next section, I explain how to account for this idea within the framework so far established.

6 Epistemic Models

In this section, I develop formal epistemic models. These models take an epistemic space E and add accessibility relations between worlds. The target concept of epistemic accessibility is a vague one. The intuitive idea I want to capture is this. If world w is putatively epistemically accessible to agent i, and w says that $A \wedge B$ but says nothing about A, then it is indeterminate whether that world, or a world like it but which also says that A, is accessible to i. Yet we do not want to say that it is indeterminate whether w is accessible for *every* incomplete world w. It is determinate that agent i does not know all consequences of what she knows, and hence some incomplete worlds must be determinately accessible to i. To avoid this consequence, I will build models relative to a fixed sharpening of 'epistemic accessibility'. Truth in a model will equate to truth on this sharpening. That sharpening will also generate a fixed set of alternative sharpenings. *Determinate* truth will equate to truth on all of those sharpenings. Now for the details.

---

the question. But all theories agree with this conditional.

11. We *can* know that the answer is in the vicinity of n: we know that taking, say, 3 sweets isn't greedy, whereas taking 1,000 is. The point is that we can have *inexact* knowledge only (Williamson 1992).

First, we expand the object language $\mathcal{L}$ (over a set of primitive sentences $\mathcal{P}$) to include a determinacy operator '$\vartriangle$', where '$\vartriangle A$' abbreviates 'it is determinate that A is the case'. We can introduce an indeterminacy operator '$\triangledown$' by definition:

$$\triangledown A \;=_{\text{df}}\; \neg\vartriangle A \wedge \neg\vartriangle\neg A$$

with '$\triangledown A$' read as 'it is indeterminate whether A is the case'. In the model theory, I will work with epistemic *projection functions* f_i, rather than accessibility relations R_i, where $f_i w$ gives the set of worlds that are epistemically accessible from world w for agent i. Given an accessibility relation R_i, we set $f_i w = \{u \mid R_i wu\}$ and, given a projection function, we can define $R_i wu$ iff $u \in f_i w$. Working with projection functions is merely a convenient notational change. Relative to $\mathcal{L}$ (which, I am assuming, is fixed throughout), we define models and related notions as follows.

Definition 2 (Epistemic model) *An* epistemic model *for k agents is a tuple $M = \langle W^P, W^I, V^+, V^-, f_1, \ldots, f_k\rangle$, where W^P and W^I are sets of worlds (thought of as sets of possible and impossible worlds, respectively), V^+, V^- are as above, and each f_i is an epistemic projection function. Let $W^\cup = W^P \cup W^I$. The rank of M is* $\min\{\#w \mid w \in W^\cup\}$.

Given any accessibility projection function f_i in M and any sentence $A \in \mathcal{L}$, we define f_i^A, the A-variant of f_i, as follows:

Definition 3 (A-variant of f_i)

$$f_i^A w = \begin{cases} (f_i w \cap \{w \mid A \in V^+ w\}) \cup (f_i w \cap W^P) & \textit{if } f_i w \subseteq \{w \mid A \notin V^- w\} \\ f_i w & \textit{otherwise} \end{cases}$$

Let $f_i^{\mathcal{L}} = \{f_i\} \cup \{f_i^A \mid A \in \mathcal{L}\}$.

Definition 4 (Alternative sequences) *For an epistemic model M as above, let* $\alpha_M = \{\langle g_1 \cdots g_k\rangle \mid g_i \in f_i^{\mathcal{L}}, i \leq k\}$.

I will use the notation '$\vec{g}$' to denote sequences of functions, and '$\vec{g}^i$' to denote the ith function of the sequence $\vec{g}$. Thus, if $\vec{g} = \langle f_i \cdots f_k\rangle$ then $\vec{g}^i = f_i$. We can think of each alternative sequence $\vec{g} \in \alpha_M$ as a sharpening of 'epistemic accessibility' for our k agents. On this account, models are defined relative to a particular sharpening $\langle f_1 \cdots f_k\rangle$, from which all the others in α_M are generated. We think of $\langle f_1 \cdots f_k\rangle$ as the sharpening that gets things right: it tells us what's true (simpliciter), whereas what's *determinately* true is a matter of what is true on *all* alternatives in α_M. This will allow us to give a classical definition of truth, on which $M \Vdash A$ or $M \Vdash \neg A$ for all (pointed) models M.[12]

12. Alternatively, we could give a supervaluationist-style treatment by defining models relative to a set of alternative sequences, and define truth (simpliciter) as truth on all alternatives. Then it will always be the case that $M \Vdash A \vee \neg A$ but not always the case that $M \Vdash A$ or $M \Vdash \neg A$. Theorem 3, the main result of the paper, holds on either approach.

Definition 5 (**$\vec{g}$-truth and $\vec{g}$-falsity**) *Given an epistemic model M as above and an alternative sequence $\vec{g} \in \alpha_M$, we define $\vec{g}$-relative truth and falsity in M, $\Vdash_{\vec{g}}$ and $\dashv\Vert_{\vec{g}}$, as follows. (M is implicit in each clause.) For possible worlds $w \in \mathbb{W}^P$:*

$$
\begin{array}{rcl}
w \Vdash_{\vec{g}} p & \textit{iff} & p \in V^+ w \\
w \Vdash_{\vec{g}} \neg A & \textit{iff} & w \nVdash_{\vec{g}} A \\
w \Vdash_{\vec{g}} A \wedge B & \textit{iff} & w \Vdash_{\vec{g}} A \textit{ and } w \Vdash_{\vec{g}} B \\
w \Vdash_{\vec{g}} A \vee B & \textit{iff} & w \Vdash_{\vec{g}} A \textit{ or } w \Vdash_{\vec{g}} B \\
w \Vdash_{\vec{g}} A \rightarrow B & \textit{iff} & w \nVdash_{\vec{g}} A \textit{ or } w \Vdash_{\vec{g}} B \\
w \Vdash_{\vec{g}} K_i A & \textit{iff} & u, \vec{g} \Vdash_{\vec{g}} A \textit{ for all } u \in \vec{g}^i w \\
w \Vdash_{\vec{g}} \triangle A & \textit{iff} & w \Vdash_{\vec{h}} A \textit{ for all } \vec{h} \in \alpha_M \\
w \dashv\Vert_{\vec{g}} A & \textit{iff} & w \nVdash_{\vec{g}} A
\end{array}
$$

For impossible worlds $w \in \mathbb{W}^I$:

$$
\begin{array}{rcl}
w \Vdash_{\vec{g}} A & \textit{iff} & A \in V^+ w \\
w \dashv\Vert_{\vec{g}} A & \textit{iff} & A \in V^- w
\end{array}
$$

Definition 6 (**n-entailment**) *A pointed model is a pair $M' = \langle M, w \rangle$ where M is as above and $w \in \mathbb{W}^P$ in M. We define truth relative to M′ as:*

$$M' \Vdash A \quad \textit{iff} \quad w, \langle f_1, \ldots, f_n \rangle \Vdash A$$

$M' \Vdash \Gamma$ iff $M' \Vdash A$ for each $A \in \Gamma$. For any $n \in \mathbb{N} \cup \{\omega\}$, logical n-entailment is then defined as:

$$\Gamma \vDash_n A \textit{ iff, for every pointed model } M \textit{ of rank} \geq n,\ M \Vdash \Gamma \textit{ only if } M \Vdash A$$

It is then easy to see that $\vDash_n$ extends classical entailment, and that our previous definition of epistemic entailment can be put equivalently in terms of epistemic models:

Theorem 2 *For any $n \in \mathbb{N} \cup \{\omega\}$:*

(a) If Γ classically entails A, then $\Gamma \vDash_n A$.

(b) $\Gamma \vDash^e_n A$ iff, for all epistemic models M of rank n, all worlds $w \in \mathbb{W}^\cup$ in M and all $\vec{g} \in \alpha_M$, if $M, w \Vdash_{\vec{g}} B$ for each $B \in \Gamma$ then $M, w \not\dashv\Vert_{\vec{g}} A$.

Proof: For (a), suppose $\Gamma \nvDash_n A$. Then there is an epistemic model M, world $w \in W^P$ in M and alternative $\vec{g} \in \alpha_M$ such that $w \Vdash_{\vec{g}} B$ for all $B \in \Gamma$ but $w \nVdash_{\vec{g}} A$. Let $v : \mathcal{P} \longrightarrow \{\textit{true}, \textit{false}\}$ be a classical valuation such that $vp = \textit{true}$ iff $p \in V^+ w$ in M. Now extend v for '¬', '∧', '∨' and '→' in the usual way. We show that vC satisfies each $C \in \Gamma \cup \{\neg A\}$ by induction on the complexity of C. If $C := p$, then $p \in V^+ w$ hence $vp = \textit{true}$. Now assume that, for all $C' \in \Gamma \cup \{\neg A\}$ of lower complexity than C, $vC' = \textit{true}$ iff $w \Vdash_{\vec{g}} C'$. If $C := \neg C_1$ then $w \Vdash_{\vec{g}} \neg C_1$, hence

$w \not\Vdash_{\vec{g}} C_1$ and, by hypothesis, $vC_1 = \textit{false}$. Then $vC = \textit{true}$. If $C := C_1 \wedge C_2$, then $w \Vdash_{\vec{g}} C_1$ and $w \Vdash_{\vec{g}} C_2$ and, by hypothesis, $vC_1 = vC_2 = \textit{true}$, hence $v(C_1 \wedge C_2) = \textit{true}$. The '$\vee$' and '$\rightarrow$' cases are similar. Hence $\Gamma \not\vDash_n A$ only if $\Gamma \cup \{\neg A\}$ is classically satisfiable, and (a) follows by contraposition.

For (b), 'only if' direction: assume that $\Gamma \vDash_n^e A$ and $M, w \Vdash_{\vec{g}} B$ for each $B \in \Gamma$, where M has rank n. If $w \in W^P$ in M then, by theorem 1(c), Γ classically entails A, hence by (a), $\Gamma \vDash_n A$ and so, by definition, $M, w \Vdash_{\vec{g}} A$. Suppose instead that $w \in W^I$ and let $E = \langle W^I, V^+, V^- \rangle$, where W^I, V^+, V^- are taken from M. Then $\langle E, w \rangle \Vdash \Gamma$ and so $\langle E, w \rangle \Vdash A$. This gives us $A \notin V^- w$ and hence $M, w \not\dashv\!\Vdash_{\vec{g}} A$. 'If' direction: assume the r.h.s. of the theorem and that $\langle \langle W_1, V^+, V^- \rangle, w \rangle \Vdash \Gamma$ where $w \in W_1$ and $\langle W_1, V^+, V^- \rangle$ is an epistemic space of rank n. Let $M = \langle W^P, W_1, V^+, V^-, \vec{g} \rangle$ for some W^P and $\vec{g}$. Then by definition, $w \Vdash_{\vec{g}} C$ iff $C \in V^+ w$, hence $M, w \Vdash_{\vec{g}} B$ for each $B \in \Gamma$ and so $M, w \not\dashv\!\Vdash_{\vec{g}} A$. By definition, $A \notin V^- A$ and so $\langle \langle W_1, V^+, V^- \rangle, w \rangle \not\dashv\!\Vdash A$. It follows that $\Gamma \vDash_n^e A$. ■

This result tells us that we the extra structure added to epistemic spaces to obtain epistemic models does not interfere with epistemic entailments. We could have defined epistemic entailment in terms of epistemic models, rather than in terms of epistemic spaces, and the definitions would have been equivalent. It is important to note, however, that epistemic entailments are sensitive to *all* pairs $\langle M, w \rangle$, and not just those for which $w \in W^P$ in M. This contrasts with the *logical* entailments, which are sensitive only to possible worlds $w \in W^P$.

I now turn to the main result of the paper, concerning trivial inference, knowledge and (in)determinacy:

Theorem 3 *For any $n \in \mathbb{N} \cup \{\omega\}$, if $\Gamma \vDash_n^e A$ then $\{\triangle K_i B \mid B \in \Gamma\} \vDash_n \neg\triangle\neg K_i A$.*

Proof: Assume that $\Gamma \vDash_n^e A$ and, for all $B \in \Gamma$, $\langle M, w \rangle \Vdash \triangle K_i B$, where $M = \langle W^P, W^I, V^+, V^-, \vec{f} \rangle$ and $w \in W^P$. Then for all $B \in \Gamma$ and all $\vec{g} \in \alpha_M$, $M, w \Vdash_{\vec{g}} K_i B$. Hence $M, u \Vdash_{\vec{g}} B$, for all for all $u \in \vec{g}_i w$ and all $B \in \Gamma$. Given $\Gamma \vDash_n^e A$ and theorem 2(b), it follows that $M, u \not\dashv\!\Vdash_{\vec{g}} A$, for all $u \in \vec{g}_i w$ and all $\vec{g} \in \alpha_M$. This guarantees that $f_i w \subseteq \{v \mid A \notin V^- v\}$ and so, by definition 3:

$$f_i^A w = f_i w \cap \{v \mid A \in V^+ v\} \cup (f_i w \cap W^P)$$

Then, for every world $u \in f_i^A w$ and alternative $\vec{h} \in \alpha_M$ such that $\vec{h}_i = f_i^A$, $M, u \Vdash_{\vec{h}} A$. Since $w \in W^P$, this gives us $M, w \Vdash_{\vec{h}} K_i A$, hence $M, w \not\Vdash_{\vec{f}} \triangle\neg K_i A$ and so $M, w \Vdash_{\vec{f}} \neg\triangle\neg K_i A$. This gives us $\langle M, w \rangle \Vdash \neg\triangle\neg K_i A$ and so $\{\triangle K_i B \mid B \in \Gamma\} \vDash_n \neg\triangle\neg K_i A$. ■

Corollary 1 *For any $n \in \mathbb{N} \cup \{\omega\}$:*

(a) *If $\Gamma \vDash_n^e A$ then $\{\triangle K_i B \mid B \in \Gamma\} \cup \{\neg K_i A\} \vDash_n \triangledown K_i A$.*

(b) *If $n \geq 2$, then $\vDash_n \neg\triangle\neg K_i(A \vee \neg A)$ and $\neg K_i(A \vee \neg A) \vDash_n \triangledown K_i(A \vee \neg A)$.*

Proof: For (a), suppose $\Gamma \vDash_n^e A$. Then by theorem 3, $\{\triangle K_i B \mid B \in \Gamma\} \vDash_n \neg\triangle\neg K_i A$ and, since $\triangle B \vDash_n B$, $\neg K_i A \vDash_n \neg\triangle K_i A$. Hence $\{\triangle K_i B \mid B \in \Gamma\} \cup \{\neg K_i A\} \vDash_n \triangle\neg K_i A \wedge \neg\triangle\neg K_i A$ and, by definition of '$\triangledown$', $\{\triangle K_i B \mid B \in \Gamma\} \cup \{\neg K_i A\} \vDash_n \triangledown K_i A$. For (b), suppose $(A \vee \neg A) \in V^- w$ in any space E. Then $\#w \leq 2$, hence E has a rank ≤ 2. Hence for any pointed space E' of rank > 2, $E' \Vdash A \vee \neg A$ and so, for all $n \geq 2$, $\vDash_n^e A \vee \neg A$. By theorem 3, $\vDash_n \neg\triangle\neg K_i(A \vee \neg A)$ and by (a), $\neg K_i(A \vee \neg A) \vDash_n \triangledown K_i(A \vee \neg A)$. ■

These results are a very pleasing feature of the theory. Recall that we think of each $n \in \mathbb{N} \cup \{\omega\}$ as a sharpening of epistemic notions such as 'trivial inference', 'epistemic space' and 'epistemic accessibility'. Some (informally specified) range of n gives the admissible sharpenings, and what is (actually) determinately true is what is true on all of those sharpenings. Theorem 3 then tells us that, however we sharpen these notions, if the inference from Γ to A is trivial, then determinate knowledge of Γ entails the agent does not determinately lack knowledge of A. So if A is a trivial consequence of what agent i knows, then it is never determinate that i fails to know that A. Equivalently (as the corollary says), if agent i does not know some trivial consequence A of what she knows, then it is indeterminate whether she knows that A. So, on the account proposed, there are no *determinate* epistemic blindspots. Equivalently, if an agent is in fact in an epistemic blindspot, then it is indeterminate whether she is in that blindspot.

Since what is indeterminate is not rationally assertible, it is then never rational to assert that agent i is in a particular epistemic blindspot. If an agent is not logically omniscient, then we can be sure that she suffers from epistemic blindspots. Indeed, it is determinate that real-world agents are not logically omniscient, and hence determinate that real-world agents suffer from epistemic blindspots. But we can never say what they are. Whenever we focus on a particular trivial consequence A of agent i's knowledge, it is never rational to assert that she does not know that A. Epistemic blindspots are elusive, just as counterexamples to tolerance principles for vague predicates are.

7 Conclusion

Epistemic closure principles are very much like tolerance principles for vague predicates such as 'greedy'. In any sorites case, we must deny the tolerance principle, but we must also explain why the principle seems so reasonable. When a case a is a counterexample to a tolerance principle for a predicate 'F', it is unknowable and hence not assertible that a is F. We can never rationally assert any counterexample to the tolerance principle, and this is how the principle acquires its rational appeal. It is false that, if taking n sweets isn't greedy, then taking $n+1$ sweets isn't greedy. What is true is that one extra sweet never takes us from a determinate case of non-greediness to a determinate case of greediness. If it is determinate that taking n sweets isn't greedy, then it is not determinate that taking $n+1$ sweets is greedy.

In just the same way, to avoid treating real-world agents as being logically omniscient, we must deny the epistemic closure principles. Yet must also explain why they seem so reasonable. My answer is that no counterexample to any *trivial* closure principle is ever rationally assertible. It is false that agents know all trivial consequences of what they know. What is true is that agents never determinately fail to know any trivial consequence of their determinate knowledge. Trivial inferences never take us from clear cases of knowledge to clear cases of knowledge failure. So we can never rationally assert any counterexample to the knowledge closure principles.

I developed epistemic models which support this idea. These models allow us to draw a distinction at the level of content between trivial and non-trivial deductions (§4). The former are the deductions that correspond to epistemic entailments. What is known by an agent may be an indeterminate matter, according to these models (§6). Using this approach, we can show that, however we sharpen the relevant epistemic notions, there are no determinate instances of epistemic blind spots (theorem 3). In this way, the formal approach supports the philosophical contention that epistemic blindspots are always elusive.

References

Belnap, N. (1977). A useful four-valued logic, in J. Dunn and G. Epstein (eds), *Modern Use of Multiple-valued Logic*, D.Reidel, Dordrecht.

Bjerring, J. (2012). Impossible worlds and logical omniscience: an impossibility result, *Synthese* .

Buss, S. (1998). *Handbook of proof theory*, Vol. 137, Elsevier Science, Amsterdam.

Dummett, M. (1978). The justification of deduction, *Truth and other enigmas*, Harvard University Press, Cambridge, MA, pp. 166–185.

Fagin, R., Halpern, J. and Vardi, M. (1990). A nonstandard approach to the logical omniscience problem, in R. Parikh (ed.), *Proceedings of the Third Conference on Theoretical Aspects of Reasoning about Knowledge*, Morgan Kaufmann, pp. 41–55.

Hintikka, J. (1962). *Knowledge and belief: an introduction to the logic of the two notions*, Cornell University Press, Ithaca, N.Y.

Hintikka, J. (1975). Impossible possible worlds vindicated, *Journal of Philisophical Logic* 4: 475–484.

Ho, D. (1995). Logical omniscience vs. logical ignorance, in C. Pereira and N. Mamede (eds), *Proceedings of EPIA'95*, Vol. 990 of *LNAI*, Springer, pp. 237–248.

Ho, D. (1997). Reasoning about rational, but not logically omniscient, agents, *Journal of Logic and Computation* 5: 633–648.

Lakemeyer, G. (1986). Steps towards a first-order logic of explicit and implict belief, in J. Y. Halpern (ed.), *Proceedings of the First Conference on Theoretical Aspects of Reasoning About Knowledge*, Morgan Kaufmann, San Francisco, California, pp. 325–340.

Lakemeyer, G. (1987). Tractable metareasoning in propositional logic of belief, *Proceedings of the Tenth International Joint Conference on Artificial Intelligence*, pp. 401–408.

Levesque, H. J. (1984). A logic of implicit and explicit belief, *Proceedings of the Fourth National Conference on Artificial Intelligence*, pp. 198–202.

Lewis, D. (1986). *On the Plurality of Worlds*, Blackwell, Oxford.

Lewis, D. (1996). Elusive knowledge, *Australasian Journal of Philosophy* 74(4): 549–567.

Lewis, D. (2004). Letters to priest and beall, in B. Armour-Garb, J. Beall and G. Priest (eds), *The Law of Non-Contradiction—New Philosophical Essays*, Oxford University Press, Oxford, pp. 176–177.

Rantala, V. (1975). Urn models, *Journal of Philosophical Logic* 4: 455–474.

Rantala, V. (1982). Impossible worlds semantics and logical omniscience, *Acta Philosophica Fennica* 35: 18–24.

Sorensen, R. (1984a). Conditional blindspots and the knowledge squeeze: a solution to the prediction paradox, *Australasian Journal of Philosophy* 62(2): 126–135.

Sorensen, R. (1984b). Uncaused decisions and pre-decisional blindspots, *Philosophical Studies* 45(1): 51–56.

Stalnaker, R. (1976). Propositions, in A. MacKay and D. Merrill (eds), *Issues in the Philosophy of Language*, Yale University Press, New Haven, pp. 79–91.

Stalnaker, R. (1984). *Inquiry*, MIT Press, Cambridge, MA.

Williamson, T. (1992). Inexact knowledge, *Mind* 101(402): 217–242.