

# Epistemic Modals

SETH YALCIN

Epistemic modal operators give rise to something very like, but also very unlike, Moore's paradox. I set out the puzzling phenomena, explain why a standard relational semantics for these operators cannot handle them, and recommend an alternative semantics. A pragmatics appropriate to the semantics is developed and interactions between the semantics, the pragmatics, and the definition of consequence are investigated. The semantics is then extended to probability operators. Some problems and prospects for probabilistic representations of content and context are explored.

## 1. A problem

I want to make some observations about the language of epistemic modality and then draw some consequences.

The first observation is that these sentences sound terrible.

- (1) # It is raining and it might not be raining
- (2) # It is raining and possibly it is not raining
- (3) # It is not raining and it might be raining
- (4) # It is not raining and possibly it is raining

All of these sentences are odd, contradictory-sounding, and generally unassertable at a context. They all contain modal operators which, in these sentential contexts, are default interpreted *epistemically*. (Just what the epistemic reading of modal operators is remains to be made precise—getting clearer on that is the point of this paper—but the motivation for calling the reading 'epistemic' is the intuitive idea that epistemically modalized clauses convey information about some epistemic state or a state of evidence.) I will take it that at the relevant level of abstraction, the logical form of the first two sentences is this:

$$(\phi \ \& \ \diamond \neg \phi)$$

and the logical form of the next two is this:

$$(\neg\phi \ \& \ \diamond\phi)$$

using ‘ $\diamond$ ’ schematically for natural language epistemic possibility operators.<sup>1</sup> We will have a need to refer back to conjunctions of these forms often, so let me call an instance of one of these two schemata an *epistemic contradiction*.

Epistemic contradictions are defective. Why?

It is tempting to try to connect the defect to Moore’s paradox, as follows. As Moore and others have noted, sentences like these:

- (5) It is raining and I do not know that it is raining
- (6) It is not raining and for all I know, it is raining

are odd, contradictory-sounding, and unassertable, just like (1)–(4) above. Now plausibly, we have a grip on *why* Moore-paradoxical sentences are defective: they involve the speaker in some kind of pragmatic conflict. For instance, if it is conventionally understood that, in making an assertion in a normal discourse context, one usually represents oneself as knowing what one says, then in uttering (5) or (6), one will end up representing oneself as both knowing something and also as knowing that one does not know it. It is not coherent to intend to represent oneself in this way, and so one therefore expects (5) and (6) to strike us as defective. (The appeal to some pragmatic tension like this one is the usual response to Moore’s paradox, though the details vary across theorists.<sup>2</sup>) Note that this line of explanation does not appeal to any semantic defect in these sentences. In particular, it does not appeal to the idea that (5) or (6) are contradictory in the sense that their conjuncts have incompatible truth-conditions, or in the sense that they mutually entail each other’s falsity.

Now we could take this sort of pragmatic account of Moore’s paradox on board, and then try extending it to our epistemic contradictions. The simplest way to do that would be to conjecture that each epistemic contradiction entails, in a way obvious to any competent speaker, a Moore-paradoxical sentence. For instance, we could try saying that, holding context and speaker fixed, (1) and (2) each entail (5), and that (3) and (4) each entail (6). Since it is plausible that anything that obviously entails a Moore-paradoxical sentence will itself sound

<sup>1</sup> I take it that in English these operators include, on the relevant readings, the pure modals ‘might’, ‘may’, and ‘could’, sentential operators constructible via expletives from these (‘it might be that’ etc.), and the sentential operators ‘possibly’ and ‘it is possible that’. I will abstract from any tense information contributed by the pure modals. Let me stress that by ‘ $\diamond$ ’ I do *not* have in mind complex operators containing overt attitude verbs, such as ‘for all I know it might be that’.

<sup>2</sup> See Hintikka 1962 and Unger 1975 for classic statements of the pragmatic approach, and Williamson 2000 and Stalnaker 2000 for more recent discussions.

paradoxical, this would give us an explanation for why (1)–(4) sound defective. Note that this explanation would assume that epistemic possibility clauses licence the following entailments:

$\diamond\neg\phi \models$  I do not know  $\phi$

$\diamond\phi \models$  For all I know,  $\phi$

—relative, again, to a fixed context and speaker.

It is at least *prima facie* plausible that epistemic possibility sentences in context do licence these entailments, so perhaps something like this line of explanation for the infelicity of our epistemic contradictions will ultimately prove correct. But I am not actually interested in pursuing this issue now. Rather, my aim in this section to highlight a way in which epistemic modals give rise to their own sort of ‘paradox’, one that differs from Moore’s paradox in significant respects. The puzzle I want to focus on emerges when we attempt to embed our epistemic contradictions. It turns out these conjunctions are much more difficult to felicitously embed than Moore-paradoxical sentences, and careful attention to this fact points to some interesting constraints on any theory of the meaning of epistemic modal operators.

Consider the following sentences.

(7) # Suppose it is raining and it might not be raining

(8) # Suppose it is not raining and it might be raining

Here we have (1) and (3) embedded under the attitude verb ‘suppose’. The resulting imperatival sentences are not acceptable. Indeed they are not even obviously intelligible. Substituting other natural language epistemic possibility operators yields equally defective sentences. Take ‘possibly’, for instance:

(9) # Suppose it is raining and possibly it is not raining

(10) # Suppose it is not raining and possibly it is raining

The fact is a general one about epistemic possibility modals. Intuitively, there is some element of inconsistency or self-defeat in what these sentences invite one to suppose.

We get similar results when we attempt to embed our epistemic contradictions in the antecedent position of an indicative conditional. For instance:

(11) # If it is raining and it might not be raining, then ...

(12) # If it is not raining and it might be raining, then ...

An indicative conditional that begins in one of these ways will strike any competent speaker as unintelligible, regardless of the consequent chosen to finish off the conditional. Even a conditional which merely repeats one of the conjuncts in the antecedent—say,

- (13) # If it is raining and it might not be raining, then (still) it is raining

—strikes us as unintelligible rather than trivially true, the usual judgement for such conditionals. Again, as the reader may confirm for herself, this is a general fact about epistemic possibility modals, not an idiosyncratic feature of ‘might’. The intuitive judgements about these conditionals are not surprising, given the intuitive judgements about the ‘suppose’ sentences just described. For the interpretation of an indicative conditional plausibly involves something like temporary supposition of the antecedent, and again, we see there is some element of inconsistency or self-defeat in what these antecedents invite one to entertain.

Here are the facts in schematic form.

- # Suppose  $(\phi \ \& \ \diamond \neg\phi)$
- # Suppose  $(\neg\phi \ \& \ \diamond \phi)$
- # If  $(\phi \ \& \ \diamond \neg\phi)$ , then  $\psi$
- # If  $(\neg\phi \ \& \ \diamond \phi)$ , then  $\psi$

Our first observation was that epistemic contradictions are not acceptable as unembedded, stand-alone sentences. Our second observation is that epistemic contradictions are also not acceptable in the embedded contexts described above.<sup>3</sup> We need an explanation for this second set of facts.

Finding an explanation proves not to be trivial. For starters, note that we will have no luck trying to explain this second set of facts by piggy-backing somehow on a pragmatic explanation of Moore’s paradox. Although our Moore-paradoxical sentences (5) and (6) are not felicitous unembedded, they are perfectly acceptable in the embedded contexts just described:

- (14) Suppose it is raining and I do not know that it is raining
- (15) Suppose it is not raining and for all I know, it is raining<sup>4</sup>

<sup>3</sup> Plausibly they are not acceptable in *any* embedded context, but it will be useful to focus on the two contexts just described.

<sup>4</sup> Feel free to replace the indexical ‘I’ in the imperatival sentences (14) and (15) with ‘you’, if you think that better makes the point.

- (16) If it is raining and I do not know it, then there is something I do not know
- (17) If it is not raining but for all I know, it is, then there is something I do not know

(Indeed, a reason often cited in favour of the view that Moore-paradoxical sentences are not, semantically, contradictions is the very fact that sentences like (14) and (15) strike us as coherent requests.) Moore-paradoxical sentences serve to describe totally clear possibilities, possibilities we can readily imagine obtaining. The same apparently does not apply to epistemic contradictions. These sentences do not seem to describe coherent possibilities, as witness the fact that an invitation to suppose such a conjunction strikes us as unintelligible. The upshot here is that, unlike the unembedded case, there is no obvious way to explain the unacceptability of our epistemic contradictions in embedded contexts by appeal to Moore's paradox. Moore-paradoxical sentences are quite acceptable in these contexts. We might describe the situation roughly as follows. Like Moore-paradoxical sentences, epistemic contradictions are not assertable; but unlike Moore-paradoxical sentences, they are also not *supposable*, not *entertainable as true*.

How are we to explain this novel feature of our epistemic contradictions? Let me put the question in a somewhat more theoretically-loaded way. What truth-conditions for epistemic contradictions could suffice to explain why they do not embed intelligibly under 'suppose' and in indicative conditional antecedents? To answer this question, we need to know the truth-conditions of epistemic possibility clauses. But when we look closely at the facts, it turns out that we face a certain dilemma concerning the logical relationship between epistemic possibility clauses ( $\Diamond\phi$ ) and their nonepistemic complements ( $\phi$ ), one which makes it hard to say what exactly the truth-conditions of epistemic possibility clauses, and hence our epistemic contradictions, could be. Let me explain.

To fix ideas, focus on epistemic contradictions of the form ( $\neg\phi \ \& \ \Diamond\phi$ ), and hold context fixed.<sup>5</sup> Now either  $\neg\phi$  is truth-conditionally compatible with  $\Diamond\phi$ , or it is not. Suppose first that the two are truth-conditionally compatible. Then their conjunction is, under some conditions or other, true; the truth-conditions of the conjunction ( $\neg\phi \ \& \ \Diamond\phi$ ) are non-empty. If the truth-conditions of the conjunction are non-empty, it seems there should be nothing at all preventing us from hypothetically entertaining the obtaining of these conditions. We ought to be able to do this simply as a matter of semantic competence.

<sup>5</sup>Where it creates no confusion, I will be loose about use and mention.

But we cannot. Evidently there is no coherent way to entertain the thought that it is not raining and it might be raining.

That suggests that we should drop the supposition that the two conjuncts actually are compatible. If we take it instead that  $\neg\phi$  is truth-conditionally incompatible with  $\diamond\phi$ , then we will have a ready explanation for our inability to entertain their conjunction. If there simply is no possible situation with respect to which  $(\neg\phi \ \& \ \diamond\phi)$  is true, then that explains why it is so hard to envisage such a situation. The conjunction is just semantically a contradiction. But although this line of explanation covers our intuitions about epistemic contradictions in embedded contexts, it comes at an unacceptably high price. If  $\neg\phi$  and  $\diamond\phi$  are contradictory, then the truth of one entails the negation of the other. On ordinary classical assumptions, this means that  $\diamond\phi$  entails the negation of  $\neg\phi$  — that is, it means  $\diamond\phi$  entails  $\phi$ . But that result is totally absurd. It would imply that the epistemic possibility operator  $\diamond$  is a factive operator, something it very clearly is not. (It might be raining, and it might not be raining; from this we obviously cannot conclude that it both is and is not raining.)

So it appears we face a dilemma.

- $\neg\phi$  and  $\diamond\phi$  should be modelled as having incompatible truth-conditions, in order to explain why it is not coherent to entertain or embed their conjunction; but
- $\neg\phi$  and  $\diamond\phi$  should be modelled as having compatible truth-conditions, in order to block the entailment from  $\diamond\phi$  to  $\phi$ .

A semantics for epistemic possibility modals should resolve this apparent tension. Note all of the preceding can be repeated *mutatis mutandis* for  $(\phi \ \& \ \neg\phi)$ , our second kind of epistemic contradiction.

It will be helpful to give the problem an alternative formulation, in terms of consequence. This will let us state the problem at a somewhat higher level of generality. (It will also let us sidestep the intuitive, but at this point imprecise, notion of truth-conditions.) We can think of the problem as a tension between the following three constraints on the notion of consequence appropriate to the semantics of natural language.

**Consequence is classical:**  $\models$  respects classical entailment patterns.

**Nonfactivity of epistemic possibility:**  $\diamond\phi \not\models \phi$

**Epistemic contradiction:**  $(\neg\phi \ \& \ \diamond\phi) \models \perp$ <sup>6</sup>

<sup>6</sup> We would also want the principle that  $(\phi \ \& \ \neg\phi) \models \perp$ . If certain classical principles were assumed, we would get this second principle from the first for free. It will be convenient to just focus on the first principle for now.

The principle of the nonfactivity of epistemic possibility is obvious. The principle of epistemic contradiction is much less obvious, but it is motivated by sentences like (8), (10), (12), and ordinary reflection on our inability to simultaneously coherently entertain instances of  $\neg\phi$  and  $\diamond\phi$ . Despite motivation for both principles, however, it is clear that the principles are not jointly compatible, if the consequence relation is assumed to be classical.<sup>7</sup>

The nonfactivity of epistemic possibility is surely nonnegotiable. Given that we keep it, we seem to face a choice between the principle of epistemic contradiction and the thesis that the consequence relation is classical. If we reject epistemic contradiction, we need to explain what it is about our epistemic contradictions that makes them semantically defective in embedded contexts. This does not look easy to do. Again, if epistemic contradiction is false and  $\neg\phi$  and  $\diamond\phi$  really are consistent in the sense appropriate to the correct semantics of the language, it is not clear why they should not be simultaneously entertainable as true, or why their conjunction does not embed intelligibly. On the other hand, if we keep epistemic contradiction, we need to clarify the nonclassical alternative notion of consequence in play.

That sets the stage. The task now is to spell out a logic and semantics for epistemic modals which makes sense of the facts, which resolves the tension just described. Here is the plan. I give a semantics which explains the phenomena in section 3. I consider the question of what notion of consequence is appropriate to that semantics in section 4. The discussion of consequence will set us up for a discussion, in section 5, of the pragmatics appropriate to the semantics. Equipped with a reasonable grip on the semantics and pragmatics of epistemic possibility operators, I turn in section 6 to the semantics of epistemic necessity operators. I then consider, in section 7, prospects for the extension of the semantics to probability operators. Probability operators, we will see, give rise to the same kind of problem epistemic possibility operators do, but also introduce their own challenges for analysis. In a closing discussion of outstanding issues, I attempt to catalogue some of the new questions raised by the semantics I give for these operators.

Before introducing the positive proposal for the semantics of epistemic possibility modals, I want to begin by explaining why the problem

<sup>7</sup>If this is not obvious, remember that classically,  $(\neg\phi \ \& \ \psi) \models \perp$  iff  $\psi \models \phi$ . Substituting  $\diamond\phi$  for  $\psi$  in this schema, we have the principle of epistemic contradiction on the left:  $(\neg\phi \ \& \ \diamond\phi) \models \perp$  iff  $\diamond\phi \models \phi$ . Epistemic contradiction therefore classically entails factivity. (Note I use ‘factivity’ to describe an entailment property, not a presuppositional property.)

I have set out in this section cannot be plausibly handled by a routine accessibility relation semantics for epistemic modals, since a semantics along those lines is perhaps the most familiar approach to the modals of natural language. This will help to clarify and motivate the need for the alternative semantics I describe.

## 2. Relational semantics for epistemic possibility

The idea for the semantics I want to consider and reject in this section is rooted in the classic work of Hintikka (1962), though to my knowledge Hintikka himself did not suggest it. The idea is to treat an epistemic modal clause effectively as a kind of covert attitude ascription, and to assume that attitude ascriptions are to be given the kind of semantics we find in epistemic logics of the sort inspired by Hintikka—logics conventionally interpreted on accessibility relation-based models (so-called *relational* or *Kripke* models). To make the semantics a little more realistic with respect to context-sensitivity, let me spell out the idea within a Kaplan-style two-dimensional semantics (see Kaplan 1989, Lewis 1980).

Sentences in context are true (false) relative to possibilities. We may take possibilities to be possible worlds, or world-time pairs, or centered worlds, etc.; I will talk in terms of worlds, but nothing hangs on this. Natural language modals are treated as analogous to the modal operators of ordinary normal modal logic, with truth-conditions for modal clauses stated via quantification, in the metalanguage, over a domain of possibilities. Possibility modals—‘may’, ‘might’, ‘could’, ‘possibly’, etc.—require existential quantification. (Necessity modals—‘must’, ‘has to’, ‘necessarily’, etc.—require universal quantification.) The basic structure of the semantics of a possibility clause is this:

$$\llbracket \Diamond \phi \rrbracket^{c,w} \text{ is true iff } \exists w' (wRw' \ \& \ \llbracket \phi \rrbracket^{c,w'} \text{ is true})^8$$

We assume that the accessibility relation  $R$  is, in any given case, provided by context.<sup>9</sup> On the approach to epistemic modals I now want to con-

<sup>8</sup> ‘ $\llbracket \cdot \rrbracket$ ’ denotes the interpretation function of the model of the language, which maps well-formed expressions to their extensions relative to choice of context  $c$  and possible world  $w$ . By ‘is true’, I mean ‘= Truth’.

<sup>9</sup> How is  $R$  provided by context? It could be the semantic value of a covert element in the underlying syntax of modal clauses; it could be specified as part of the definition of a model for the language; it could be an ‘unarticulated constituent’; or we could enrich the points of evaluation in our model, relativizing the truth of a sentence, not only to contexts and possibilities, but also to accessibility relations. Or perhaps something else. The choice does not really matter for our purposes. I am only interested in a general idea right now, namely that, relative to a fixed context, modals, in particular epistemic modals, express quantification over a domain of worlds which is determined as a function of the world at which the modal clause is evaluated.



sider, what makes a modal *epistemic* is the kind of accessibility relation used in the truth-conditions for the clause. (Cf. Kratzer 1977, 1981, Lewis 1979b.) The accessibility relation  $R$  associated with an epistemic modal clause is one which relates the world  $w$  at which the clause is evaluated to a set of worlds not excluded by some body of *knowledge* or *evidence* in  $w$ . Let us think of a body of knowledge or evidence  $S$  in a possible world as determining a set of possibilities, the possibilities still left open by that knowledge or evidence in that world. Then the accessibility relation  $R$  associated with an epistemic modal is a relation of the form

$$wRw' \text{ iff } w' \text{ is compatible with evidential state } S \text{ in } w$$

where world  $w'$  is compatible with  $S$  just in case  $w'$  is left open by  $S$  in  $w$ . Think of 'S' as standing in for a description of an evidential state—'what  $x$  knows', 'what  $x$  has evidence for', and so on—for some contextually specified  $x$ . It determines a function from worlds to sets of worlds.

Put simply, then, the idea is that  $\Diamond\phi$  is a sort of description of an evidential state. Its truth turns on whether  $\phi$  is left open by that evidential state in the world at which the clause is evaluated.

There has been much discussion of what exactly the rules are for determining  $S$  (and therefore the epistemic accessibility relation  $R$ ) precisely—for determining the state of knowledge or evidence relevant to evaluating the truth of an epistemically modalized sentence in any given context. When we ask whether 'It might be raining' is true as tokened in a given context, whose state of knowledge do we look to in order to settle the question? Should  $S$  be understood as the epistemic state of the speaker of the context? Is it something broader—say, the group knowledge of the discourse participants? Does  $S$  include the knowledge possessed by nearby agents not party to the conversation? Does it include evidence readily available, but not yet known, to the interlocutors? And so on. (For relevant discussion, see Hacking 1967, DeRose 1991, Egan et al. 2005, MacFarlane 2006.) It is a striking fact that these questions do not have obvious answers.

Let us set aside these questions for now. For even bracketing the question of whether it is actually possible to sort out what the right  $S$  is in any given case, we can see that there is a more basic problem with this semantics. It is the problem this paper we began with. On a relational semantics of the sort just described, epistemic contradictions are mistakenly predicted to be entertainable as true, and mistakenly predicted to be felicitous in embedded contexts. Consider again (3):

- (3) It is not raining and it might be raining

According to the basic structure of the account on the table, this has non-empty truth-conditions. It is just the conjunction of a meteorological claim with (roughly) a claim about a contextually determined agent or group's ignorance of this meteorological claim. More precisely, the sentence in context is true at a world  $w$  just in case, first, it is not raining at  $w$ , and second, there is some world  $w'$  compatible with what some specific contextually determined agent or group in  $w$  knows (or has evidence for, etc.) in  $w$  such that it is raining in  $w'$ . Who exactly the agent or group is, and what exactly their epistemic or evidential relation is to the body of information said to be compatible with rain is, we assume, settled in some more detailed way by  $R$ . The point is just that however these details are cashed out, we will have a totally clear, entertainable possibility in (3). We have the sort of thing that is completely coherent to hypothetically suppose. The semantics of this clause will interact in a perfectly nice way with attitude contexts such as 'suppose' and with indicative conditional antecedents, at least on conventional assumptions about the semantics of these environments. (Indeed, the sentence should be exactly as embeddable as a Moore-paradoxical sentence, for the underlying idea of the semantics is that sentences like (3) just *are* Moore-paradoxical sentences.)

We can illustrate the point with an example. Consider the defective indicative conditional:

- (18) # If it is not raining and it might be raining, then for all I know, it is raining

Now if the accessibility relation  $R$  for the epistemic modal in the antecedent is cashed out so that, whatever it is, it guarantees

$$\Diamond \phi \models \text{For all I know, } \phi$$

is valid given a fixed context—a weak assumption, and a standard one in the current literature—then we should expect (18) to strike us as sounding true. But clearly, the conditional is not true. It does not even make sense. The conditional is semantically defective, but this semantics does not capture the defect. This approach therefore misses the facts.

Why does it miss the facts? The problem, I suggest, is the idea, practically built into a relational semantics for modals, that the evidential state relevant to the truth of an epistemic modal clause is ultimately determined as a function of the evaluation world—the world coordinate of the point at which the modal clause is evaluated. If we model epistemic modals as if they behaved that way, epistemic modal clauses

end up acting like (covert) descriptions of epistemic states. And as a result, sentences like (1)–(4) are incorrectly predicted to be as embeddable as the overtly epistemic-state-describing counterparts of these sentences—that is, Moore-paradoxical sentences.

### 3. Domain semantics for epistemic possibility

If we want to keep the intuitively reasonable idea that epistemic possibility clauses indicate, in some sense, that their complements are compatible with some evidential state or state of information, we need a better way of representing informational states in the semantics than via accessibility relations. Here is a fix.

Start again with a two-dimensional semantics in the style of Kaplan. Let me be a little more precise now about what the two dimensions are. The points of evaluation relative to which extensions are defined have two coordinates: a *context* coordinate and an *index* coordinate. Contexts are locations where speech acts take place. Following Lewis (1980), we may think of them as centered worlds, determining both a possible world and a spatiotemporal location within that world. Contexts have indefinitely many features—speakers, audiences, indicated objects, standing presuppositions, etc.—and these features may figure in the truth of sentences said in that context in indefinitely many ways. Indices are  $n$ -tuples of specific features of context, those features which are independently shiftable by operators in the language. Which features of the context are shiftable depends on what operators the language contains. Our indices include at least a world parameter, since the fragment of English we consider has operators which shift the world at which a clause is evaluated.

Above our tacit assumption was that the index consisted *only* of a world parameter. Consequently there was no need to introduce the more general notion of an index. This notion only comes in handy when one posits an index with more than one parameter. That is what we do now. In addition to a world parameter, let our index include also an *information* parameter  $s$ . This coordinate will range over bodies of information, where a body of information is modelled as a set of worlds. Indices are therefore now pairs,  $\langle s, w \rangle$ ; and the intension of a sentence relative to a fixed context is now a function from such pairs into truth values, rather than simply a function from worlds to truth-values. Our plan is to use this new  $s$  parameter to supply the domain of quantification for epistemic modal clauses. I will call this a *domain* semantics. Rather than quantifying over a set of worlds that stand in

some  $R$  relation to the world of evaluation, as in a relational semantics, epistemic modals will be treated as quantifying over a domain of worlds provided directly by the index.<sup>10</sup> Here are the truth-conditions:

$$\llbracket \Diamond \phi \rrbracket^{c,s,w} \text{ is true iff } \exists w' \in s : \llbracket \phi \rrbracket^{c,s,w'} \text{ is true}$$

Epistemic possibility modals simply effect existential quantification over the set of worlds provided by the information parameter. No covert material is assumed, and no accessibility relation is appealed to.

We can observe immediately that iterating epistemic possibility operators adds no value on this semantics:  $\Diamond \Diamond \phi$  is semantically equivalent to  $\Diamond \phi$ . The outer modal in  $\Diamond \Diamond \phi$  serves only to introduce vacuous quantification over worlds. (This may explain why iterating epistemic possibility modals generally does not sound right, and why, when it does, the truth-conditions of the result typically seem equivalent to  $\Diamond \phi$ . I will generally ignore iterated epistemic modalities below.)

We can take it that the semantic role of  $s$  will be relatively minimal. Although denotations are now technically all relativized to a value for  $s$ , in most cases extensions will not be sensitive to it. Predicates will be assigned extensions relative only to worlds, as usual; logical connectives will be defined as usual;<sup>11</sup> and nothing new need be assumed about the semantics of names, generalized quantifiers, etc. Most clauses will continue to place conditions only on the world coordinate of the index, and will therefore retain their ordinary possible worlds truth-conditions. In such cases the information parameter  $s$  will be idle. We exploit  $s$  mainly in the definition of truth for epistemic modal talk (as above), and for certain constructions embedding such talk—in particular, attitude verbs and indicative conditionals. Let me now describe a domain semantics for these latter two constructions which will give us the desired predictions for our epistemic contradictions in embedded contexts.

Start with our troublemaking attitude verb ‘suppose’. For this verb, let us assume essentially an off-the-shelf possible worlds semantics, with one adjustment: the attitude verb will be taken to shift the value of

<sup>10</sup> In adding a parameter to represent a set of worlds to the index and using it to give semantics for epistemic modals, I follow MacFarlane (2006). MacFarlane’s work helped me to see a cleaner formalization of the ideas in a previous draft of this paper. MacFarlane does not motivate (what I am calling) a domain semantics as over a relational semantics in the way I do here. He also does not enrich the information parameter probabilistically in the way described later (Sect. 7), and he has a quite different conception of the pragmatics of epistemic modal claims and of their informational content. I hope to discuss these differences elsewhere.

<sup>11</sup> In particular, since negation and conjunction will occur often:  $\llbracket \neg \phi \rrbracket^{c,s,w}$  is true iff  $\llbracket \phi \rrbracket^{c,s,w}$  is false, and  $\llbracket \phi \ \& \ \psi \rrbracket^{c,s,w}$  is true iff  $\llbracket \phi \rrbracket^{c,s,w}$  is true and  $\llbracket \psi \rrbracket^{c,s,w}$  is true.

$s$  for its complement, replacing it with the set of worlds compatible with the agent's suppositions. The truth-conditions of 'x supposes  $\phi$ ' are as follows:

$$\llbracket x \text{ supposes } \phi \rrbracket^{c,s,w} \text{ is true iff } \forall w' \in S_x^w : \llbracket \phi \rrbracket^{c,S_x^w,w'} \text{ is true}$$

where

$$S_x^w =_{\text{def}} \text{the set of worlds not excluded by what } x \text{ supposes in } w$$

Roughly: when you suppose what  $\phi$  says, your state of supposition, abstractly represented by a set of worlds, includes the information that  $\phi$ . What is supposed is what is true at every world compatible with what is supposed.

Semantically, the attitude verb does two things. First, it quantifies over the set of possibilities compatible with the attitude state. Second, it shifts the value of  $s$  to that set of possibilities. The second effect is what is unique to a domain semantics. This effect matters only when we come to evaluating the complement of the clause. Most complements of 'suppose' ascriptions will not have truth-conditions which consult the  $s$  parameter in determining truth, and therefore this shiftiness will have no overall effect on truth-conditions. In such cases, the above semantics will yield the same predictions as a conventional accessibility relation semantics for attitude verbs. One type of complement which *will* consult the  $s$  parameter, however, is a complement containing an epistemic modal clause. As per the semantics just given above, epistemic possibility modals quantify over the set of worlds provided by the information parameter. Hence such a modal, when embedded under 'suppose', will quantify over supposition-worlds. We can see the interaction of the attitude verb and the modal by stating the truth-conditions for 'x supposes  $\Diamond\phi$ ' at the relevant level of abstraction:

$$\forall w' \in S_x^w : \exists w'' \in S_x^w : \llbracket \phi \rrbracket^{c,S_x^w,w''} \text{ is true}$$

We have two quantifiers here, one (universal) introduced by the attitude verb, the other (existential) introduced by the modal. They quantify over the same domain, since the quantificational domain of the epistemic modal is parasitic on that of the attitude verb. The modal picks up its domain from the information parameter, which has been shifted by 'suppose'. Notice that the universal quantifier introduced by the attitude verb is *vacuous*. It has been 'trumped', as it were, by the epistemic modal. The attitude verb influences truth in this case only because it has provided the domain over which the embedded epis-

temic possibility modal quantifies. So the truth-conditions are really just this:

$$\exists w' \in S_x^w : \llbracket \phi \rrbracket^{c, S_x^w, w'} \text{ is true}$$

where here we have simply removed the vacuous universal quantifier.

The nonstandard way in which the modal and the attitude verb interact here is precisely what we want, for it lets us explain what is wrong with embedding epistemic contradictions. Take, for instance, a sentence of the form ‘*x* supposes that ( $\neg\phi$  &  $\Diamond\phi$ )’. It is straightforward to verify that, on the semantics just given, this sentence will be true just in case, first, in all the worlds compatible with what *x* supposes,  $\neg\phi$  is true, and second, there is some world compatible with what *x* supposes where  $\phi$  is true. That is, the truth-conditions are, at the relevant level of abstraction:

$$(\forall w' \in S_x^w : \llbracket \neg\phi \rrbracket^{c, S_x^w, w'} \text{ is true}) \ \& \ (\exists w' \in S_x^w : \llbracket \phi \rrbracket^{c, S_x^w, w'} \text{ is true})$$

Obviously, there is no state of supposition *S* that could make this condition true, for the condition imposes contradictory demands on the state. (The same is true for ‘*x* supposes ( $\phi$  &  $\Diamond\neg\phi$ )’, since its truth-condition is the same, save for a switch in the location of the negation.) And this explains what is wrong with asking someone to suppose an epistemic contradiction. It is a request to enter into an impossible state of supposition, a request that cannot be satisfied.

We can motivate a domain semantics of the sort I have been describing from a second direction, separate from the whole issue of epistemic contradictions. I have discussed only ‘suppose’ so far, but it is very natural to extend a domain semantics of this type to other attitude verbs, such as ‘believe’, ‘suspect’, ‘think’, and ‘know’. Take a sentence like:

(19) Vann believes that Bob might be in his office

On the natural reading of this sentence, it is intuitively plausible that the epistemic modal in the complement of this sentence is understood as directly quantifying over Vann’s belief worlds.<sup>12</sup> If we gave ‘believe’ a domain semantics structurally analogous to ‘suppose’ above, we could capture this easily. Again, the verb would shift the information parameter (this time to the set of worlds not excluded by Vann’s beliefs in the world of evaluation), and the modal would existentially quantify over that parameter. The sentence would be true just in case Bob’s being in his office is compatible with what Vann believes. That is the intuitively correct result.

<sup>12</sup> Here I am indebted to work by Tamina Stephenson; see Stephenson 2007.

By contrast, the story would have to be more complicated in a relational semantics. On the usual formulation of that semantics, (19) would be treated as a *second-order* attitude ascription. It would be understood as saying, roughly, that Vann believes that it is compatible with what Vann believes that Bob is in his office. This second-order ascription would entail the first-order ascription (i.e. that it is compatible with what Vann believes that Bob is in his office) in a relational semantics *only if* we made an assumption about the modal logic of belief—namely, the assumption that whatever you believe to be compatible with what you believe actually *is* compatible with what you believe. We can avoid the need to make such assumptions in a domain semantics.

Second, the second-order truth-conditions of relational semantics, whether or not they entail the truth-conditions supplied by the domain semantics, are plausibly just too strong to be right. Suppose my guard dog Fido hears a noise downstairs and goes to check it out. You ask me why Fido suddenly left the room. I say:

(20) Fido thinks there might be an intruder downstairs

That is good English. What does it mean? Does it mean, as a relational semantics requires, that Fido believes that it is compatible with what Fido believes that there is an intruder downstairs? That is not plausible. Surely the truth of (20) does not turn on *recherché* facts about canine self-awareness. Surely (20) may be true even if Fido is incapable of such second-order beliefs.

Let me close this digression on attitudes by stating a certain apparently true generalization about the logical relation between (some) attitude verbs and epistemic possibility modals. Following in the tradition of standard logics of knowledge and belief, we have treated attitude verbs as modal operators—specifically, as boxes, to be interpreted in terms of universal quantification over possibilities. What we have been observing is that, a least for many attitude verbs  $\blacksquare$ , it appears that  $\blacksquare$  interacts with the epistemic possibility operator as follows:

$$\blacksquare \diamond \phi \leftrightarrow \blacklozenge \phi$$

That is: attitude verb + epistemic possibility modal = dual of the attitude verb.<sup>13</sup> What is nice about a domain semantics is that it underwrites this generalization easily, and without the need to make extra assumptions about the logics of the relevant attitude verbs.

<sup>13</sup>Note that the principle admits of certain exceptions, some of which are discussed below (Sect. 5).

Turn now to our other problematic embedded context, indicative conditional antecedents. Recall once more what needs to be explained:

# If  $(\phi \ \& \ \diamond \neg\phi)$ , then  $\psi$

# If  $(\neg\phi \ \& \ \diamond \phi)$ , then  $\psi$

The explanation to be offered will mimic the explanation just given for attitude contexts. Again, we want to understand our epistemic contradictions as serving to place incompatible demands on the information parameter. We therefore need our semantics for indicative conditionals to interact in the right way with this parameter.

Here is what I suggest. Let us think of indicative conditionals as behaving semantically like epistemic modals. They place conditions, not on the world parameter of the index, but on the information parameter. The truth-conditions are as follows:

$\llbracket \alpha \rightarrow \psi \rrbracket^{c,s,w}$  is true iff  $\forall w' \in s_\alpha: \llbracket \psi \rrbracket^{c,s_\alpha,w'}$  is true

with  $s_\alpha$  being a certain non-empty subset of  $s$ . This semantics likens indicative conditionals to epistemic necessity claims.<sup>14</sup> The only difference is that, rather than quantifying over all of  $s$ , the quantification is restricted to a certain subset of  $s$ . Which subset? What we want, intuitively, is simply the largest subset of  $s$  such that the information in the antecedent is included in that subset. Define  $s_\alpha$  as follows:

$s_\alpha =_{\text{def}} \text{MAX } s' \subseteq s : (s' \neq \emptyset \ \& \ \forall w' \in s': \llbracket \alpha \rrbracket^{c,s',w'} \text{ is true})$

The MAX term here supplies the largest nonempty subset  $s'$  of  $s$  satisfying the property specified (where  $s$  the value of the information parameter for the conditional). A maximizing operation is needed because  $s_\alpha$  is meant to be the 'minimal change' to  $s$  needed to add to it the information contained in the antecedent  $\alpha$ .

If we like, we can think of the semantics as proceeding in two steps. First, the antecedent of a conditional shifts the information parameter, 'updating' it with the information the antecedent contains. Second, universal quantification occurs over that updated parameter. The whole conditional is true just in case the information in the consequent is 'already included' in the updated parameter.<sup>15</sup>

<sup>14</sup> Assuming, that is, that epistemic necessity modals are the semantic duals of epistemic possibility modals, hence that they universally quantify where possibility modals existentially quantify. See section 6 for further discussion of epistemic necessity clauses.

<sup>15</sup> It may be that the two steps are the result of distinct compositional ingredients (Kratzer 1986). Perhaps 'if'-clauses serve to shift the information parameter only, with the universal quantification introduced separately by a (usually covert) epistemic necessity modal. We need not take a stand on the issue here.



Of course, it would take much more space than I have to defend a semantics of this form for indicatives adequately. I will just settle for pointing out that it gets the right result for our problem conditionals. The reason is that by the semantics, a conditional  $\alpha \rightarrow \psi$  is true only if there is exists a nonempty set  $s_\alpha$  such that

$$\forall w \in s_\alpha : \llbracket \alpha \rrbracket^{c, s_\alpha, w} \text{ is true}$$

Now if  $\alpha$  is an epistemic contradiction, there will be no such set. This is for just the same reason as in the attitude case discussed above. An antecedent which is an epistemic contradiction will impose incompatible demands on the information parameter. If the antecedent is  $(\neg\phi \ \& \ \diamond\phi)$ , the semantics will require that the information parameter be shifted to a set of worlds  $s_\alpha$  satisfying the following conjunctive condition:

$$(\forall w \in s_\alpha : \llbracket \neg\phi \rrbracket^{c, s_\alpha, w} \text{ is true}) \ \& \ (\exists w \in s_\alpha : \llbracket \phi \rrbracket^{c, s_\alpha, w} \text{ is true})$$

Again, there is no state of information  $s_\alpha$  that could make this condition true; the condition imposes contradictory demands. (The same remarks go, *mutatis mutandis*, for  $(\phi \ \& \ \diamond\neg\phi)$  in antecedent position.) This predicts that conditionals with epistemic contradiction antecedents are never true, hence that they should sound semantically defective. We have the desired result.<sup>16</sup>

There is a clear sense in which our puzzle about epistemic possibility modals is now dissolved. Consider again our first formulation of the puzzle, as a dilemma about truth-conditions.

- $\neg\phi$  and  $\diamond\phi$  should be modelled as having incompatible truth-conditions, in order to explain why it is not coherent to entertain or embed their conjunction; but
- $\neg\phi$  and  $\diamond\phi$  should be modelled as having compatible truth-conditions, in order to block the entailment from  $\diamond\phi$  to  $\phi$

We see that we have taken the second path, but avoided the associated horn, essentially by working with an enlarged conception of truth-conditions. Rather than modelling epistemic modal clauses as placing con-

<sup>16</sup> Let me note that the semantics I have given for indicative conditionals is essentially a restricted strict conditional analysis. One may prefer a variably strict analysis, along the familiar lines of Stalnaker 1968 or Lewis 1973. This could be done by imposing further constraints on  $s_\alpha$ . Such an analysis would be compatible with explanation just offered of the defect in embedding our epistemic contradictions in indicative antecedents. So long as it is necessary condition on the truth of an indicative that the relevant  $s_\alpha$  be such that for all  $w$  in it, the antecedent is true, the explanation will go through.

ditions on possible worlds relative to context (as would be typical on a relational semantics), we construed them as placing conditions on sets of worlds.  $\neg\phi$  and  $\diamond\phi$  have compatible truth-conditions on our semantics because, relative to context, they place conditions on different index coordinates:  $\neg\phi$  places a condition on the world parameter of the index, and  $\diamond\phi$  a condition on the information parameter. The incoherence of their conjunction in the various embedding environments discussed is explained, not by their joint truth at a point of evaluation being impossible, but by their failing to be jointly ‘acceptable’ by a single state of information in the way that those environments require.

In the next section this notion of acceptance is more precisely defined, and its relevance to the appropriate definition of consequence for the semantics is considered.

#### 4. Consequence

We were able to dissolve our puzzle without defining any notion of consequence. Our problem was solvable without any explicit commitment on that issue. Nevertheless, it is of interest to ask what notion of consequence is most appropriate to the semantics just provided—especially given our second setup of the puzzle, as a tension between the principle of epistemic contradiction and classical consequence. In this section, I will describe three notions of consequence, suggest that two are of primary interest, and ask where each of the two stand with respect to epistemic contradiction.

First, consequence might preserve *truth at a point of evaluation*, the notion recursively defined by our intensional semantics. We could call this *standard consequence*.

$\phi$  is a *standard consequence* of a set of sentences  $\Gamma$ ,  $\Gamma \vDash_s \phi$ , just in case for every point of evaluation  $p$ , if every member of  $\Gamma$  is true at  $p$ , then  $\phi$  is true at  $p$

I mention standard consequence only to set it aside. It is arguably not the notion we want if we are looking for a notion which tracks the intuitive notion of a conclusion following from a collection of premisses. The trouble is that the notion of truth that standard consequence preserves is, in an important sense, too general as applied to the *unembedded* sentences which constitute a set of premisses and a conclusion. To give a simple illustration, take the unembedded sentence ‘Jones has red hair’. Suppose we consider an occurrence of this sentence with respect

to a context in which Jones has black hair (that is, a context which is such that in the world of the context, Jones has black hair). Is the sentence, as it occurs in this context, true or false? False, intuitively. But given only our definition of truth at a point of evaluation, the question does not really make sense. According to that definition, sentences have truth values only with respect to a whole point of evaluation (a context and an index), and in stating the question, we have only specified the context coordinate of the point. But evidently we do have an intuitive notion of the truth or falsity of a sentence in context *simpliciter*. Given that we do, it would seem natural to define consequence so that it preserves this intuitive notion of truth.

Following Kaplan (1989), we can do that by first defining truth at a context in terms of truth at a point of evaluation. Let us write ‘ $\phi_c$ ’ for an occurrence of a sentence  $\phi$  in a given context  $c$ . Then we can say that:

$\phi_c$  is true iff  $\llbracket \phi \rrbracket^{c, s_c, w_c}$  is true

where  $w_c$  is the world of the context  $c$ , and  $s_c$  is the state of information determined by  $c$ . (More on  $s_c$  shortly.) A sentence in a context is true just in case it is true with respect to the point consisting of the context and the index determined by that context. Reflection on cases suggests that this definition does track the intuitive notion we intended to capture.<sup>17</sup>

With this notion of truth in hand, we can define our second notion of consequence. Call it *diagonal consequence*.

$\phi$  is a *diagonal consequence* of a set of sentences  $\Gamma$ ,  $\Gamma \vDash_d \phi$ , just in case for any context  $c$ , if every member of  $\Gamma_c$  is true, then  $\phi_c$  is true.

Diagonal consequence preserves truth at context. It is perhaps the most intuitively natural definition of consequence available in a Kaplan-style two-dimensional semantics—given, at least, that consequence is to be understood in terms of some form of truth-preservation. Note that the only points of evaluation that matter in evaluating an argument for diagonal consequence are those points which are pairs of a context and the index determined by that context. We can call such points *diagonal points*, since these are the points that would constitute the diagonal of

<sup>17</sup> e.g. ‘Jones has red hair’ is correctly predicted to be false with respect to the context described above, because it is false with respect to world coordinate of the index determined by the context. See Kaplan 1989 for further discussion.

the two-dimensional matrix associated with any given sentence. (Diagonal points are also sometimes called *proper* points.)<sup>18</sup>

Now let us raise the question of epistemic contradiction with respect to diagonal consequence. Is a contradiction a diagonal consequence of an epistemic contradiction such as  $(\neg\phi \ \& \ \diamond\phi)$ ? Or equivalently: is this sentence true at any diagonal points? Or equivalently again: are  $\neg\phi$  and  $\diamond\phi$  diagonally consistent? To answer, we need to know when  $\diamond\phi$  is true at a context. To know that, we need a grip on what  $s_c$ , the state of information determined by a given context  $c$ , is.

But, as already alluded to above (Sect. 2), that last issue is a difficult one, and it is one I have avoided addressing. When is  $\diamond\phi$  true at a context? What body of information is relevant to determining whether a simple unembedded epistemic possibility claim is true or false? The answer is not clear. Obvious choices—such as the knowledge state of the speaker of the context, or the distributed knowledge of the discourse participants—appear to be subject to counterexamples, as noted already by Hacking (1967); and recent work (Egan et al. 2005, MacFarlane 2006, Egan 2007) suggests that the fix, if there is one, is not going to be straightforward.

Again, I want to sidestep this issue for now. Fortunately, we can answer our question about epistemic contradiction under diagonal consequence without a full theory of how the information parameter is ‘initialized’ by context. We need only capture some of the basic structural features the information parameter must have at diagonal points of evaluation. Two in particular are plausible. First:

**Reflexivity:** For every diagonal point of evaluation  $\langle c, s, w \rangle$ ,  $w \in s$

Roughly: what is true at a context is epistemically possible at that context. This is uncontroversial. Second,

**Non-collapse:** For some diagonal point of evaluation  $\langle c, s, w \rangle$ ,  $\{w\} \neq s$

Roughly: with respect to some contexts, what is possible is not, or not merely, what is actual. This, too, is uncontroversial. (And indeed presumably it is true for practically *all* diagonal points.) Given Reflexiv-

<sup>18</sup>Note that we could also define diagonal consequence in terms of truth at diagonal points of evaluation, as follows:

$\Gamma \vDash_d \phi$  just in case for every diagonal point of evaluation  $p$ , if every member of  $\Gamma$  is true at  $p$ , then  $\phi$  is true at  $p$ .

This makes it obvious that diagonal consequence is a restricted version of standard consequence. (Standard consequence implies diagonal consequence, but not vice versa.)

ity, it merely states that epistemic possibility does not collapse into truth. More than one world may be epistemically possible with respect to a context.

Now it should be obvious, given these properties, that  $\neg\phi$  and  $\diamond\phi$  are diagonally consistent, hence that  $(\neg\phi \ \& \ \diamond\phi) \not\#_d \perp$ . For if more than one world may be epistemically possible with respect to a context, then for some  $\phi$  false at that context,  $\diamond\phi$  is true. Hence  $(\neg\phi \ \& \ \diamond\phi)$  is true at the diagonal point determined by that context.<sup>19</sup> So if the principle of epistemic contradiction is understood in terms of diagonal consequence, it is false.

What does this show? It shows that a *prima facie* natural, classical notion of consequence—diagonal consequence—is in fact compatible with our semantic explanation of the problematic embedding behaviour of our epistemic contradictions.<sup>20</sup> What it shows is that strictly speaking, diagonal consequence is not under direct threat by our puzzle about epistemic modals.

Diagonal consequence is under threat, however, from two other directions. First, as pointed out above, this notion of consequence requires the notion of truth at a context to be well-defined for epistemic modal claims. The current lack of consensus about how that definition is supposed to go—in our terms, about how the information parameter is to be ‘initialized’ by context—calls this assumption into question. Second, it may be argued that diagonal consequence misses quite elementary patterns of inference. I have in mind especially the following line of objection:

Surely, any formal regimentation of the intuitive notion of consequence should substantially track our intuitions concerning what follows *on the supposition* of what. Now suppose that it is not raining. Given that supposition, might it be raining? Obviously not! Hence  $\neg\phi$  and  $\diamond\phi$  are incompatible. Diagonal consequence misses this.

<sup>19</sup> Same goes, *mutatis mutandis*, for  $(\phi \ \& \ \diamond\neg\phi)$ , as usual.

<sup>20</sup> A proponent of diagonal consequence still needs to explain what is wrong with epistemic contradictions in the *unembedded* case, given he cannot appeal to epistemic contradiction. But this could be done by piggybacking on Moore’s paradox. If we made the following popular assumption:

**Speaker inclusion:** For every diagonal point of evaluation  $\langle c, s, w \rangle$ ,  $s \subseteq S$ , where  $S$  is the set of worlds not excluded by the knowledge of the speaker at  $c$ .

i.e. if we assumed that the information state determined by the context includes at least the knowledge of the speaker, then, unembedded, our epistemic contradictions would diagonally entail Moore-paradoxical sentences. Their badness could then be explained by whatever pragmatic explanation we give for the badness of Moore-paradoxical sentences generally.

(A line of thought rather like this one was voiced by Łukasiewicz, who proposed the following as an intuitive ‘general theorem’: ‘*If it is supposed that not- $\phi$ , then it is (on this supposition) not possible that  $\phi$ .*’ Łukasiewicz 1930, p. 156, his italics.)<sup>21</sup>

It would be of interest to find an intuitive notion of consequence for our semantics which did not face these two threats. In fact it is not difficult, given the semantics already in place, to define a such a notion of consequence. The notion of consequence I have in mind preserves, not truth, but a different property of sentences in context—one they have in relation to a state of information. We might call this property *acceptance*:

$\phi_c$  is *accepted* in information state  $s$  iff for all worlds  $w$  in  $s$ ,  $\llbracket \phi \rrbracket^{c,s,w}$  is true.

The definition of acceptance mimics the domain semantics for attitudes given above. Intuitively, think of a sentence in context as determining a constraint on a state of information. A state of information accepts a sentence in context just when it satisfies the constraint determined by that sentence. If  $\phi_c$  is nonepistemic, it places a condition on worlds, and the constraint it determines on a state of information is that each world compatible with the information satisfy that condition. If  $\phi_c$  is epistemically modalized, then it places a global condition on a state of information (set of worlds), and the constraint on a state of information it determines is just that the state itself satisfy this global condition.

Now we can define a notion of consequence according to which consequence preserves acceptance. Call it *informational consequence*:

$\phi$  is an *informational consequence* of a set of sentences  $\Gamma$ ,  $\Gamma \vDash_i \phi$ , just in case for every context  $c$  and body of information  $s$ , if every member of  $\Gamma_c$  is accepted in  $s$ , then  $\phi$  is accepted in  $s$ .

If  $\phi$  is an informational consequence of a set of sentences  $\Gamma$ , then any state of information which satisfies all the informational constraints imposed by the sentences of  $\Gamma$  (all evaluated with respect to a given context) already satisfies the informational constraint imposed by  $\phi$  (at that context).

Informational consequence avoids the two threats described above. Unlike diagonal consequence, informational consequence does not require the idea of a diagonal point to be well-defined. It requires only

<sup>21</sup> Together with some other ‘general theorems’, Łukasiewicz used this principle to motivate his trivalent logic, which he interpreted as a modal logic.

the notion of truth at a point of evaluation, not the Kaplanian notion of truth at a context. Hence it avoids the assumption that this notion is in fact definable for epistemic modal claims. Second, informational consequence respects the intuitive pattern of inference from  $\neg\phi$  to  $\neg\Diamond\phi$ . Indeed, it is worth noting that informational consequence validates the following three principles.

**Lukasiewicz's principle:**  $\neg\phi \vDash \neg\Diamond\phi$

**Epistemic contradiction:**  $(\neg\phi \ \& \ \Diamond\phi) \vDash \perp$

**Nonfactivity of epistemic possibility:**  $\Diamond\phi \not\vDash \phi$

The first principle, which I have called *Lukasiewicz's principle*, expresses the intuition about consequence our objector had in mind above. The intuitive reason for its truth is that  $\neg\phi$  and  $\neg\Diamond\phi$  impose the same informational constraint. Given an arbitrary context, if a state of information accepts  $\neg\phi$ , then the state excludes all  $\phi$ -possibilities; hence  $\phi$  is not a possibility according to the state, hence  $\neg\Diamond\phi$  is accepted with respect to that state. Epistemic contradiction is correct for a similar reason:  $\neg\phi$  and  $\Diamond\phi$  are associated with incompatible informational constraints. Given a fixed context, they cannot be both accepted by a single state of information.

Most important, these two principles are correct for the semantics under informational consequence *together* with the (non-negotiable) nonfactivity of epistemic possibility. The truth of nonfactivity is also easy to see. Relative to context,  $\Diamond\phi$  merely asks for a state to contain at least one  $\phi$ -world, whereas  $\phi$  requires a state to be such that every world in the state is a  $\phi$ -world. Hence  $\Diamond\phi$  does not suffice for acceptance of  $\phi$ .

Informational consequence is a nonclassical notion of consequence. This is because, as pointed out above, epistemic contradiction and nonfactivity are classically incompatible. (Nonfactivity is also classically incompatible with Lukasiewicz's principle.) A more detailed study of the logic that results from the combination of domain semantics for epistemic possibility modals plus informational consequence is better reserved for elsewhere, but let me just make an informal remark about the nature of the non-classicality. Informational consequence is built around the notion of acceptance. Acceptance is a gappy notion. Fixing context, there is a gap between (nonepistemic)  $\phi$  being accepted with respect to some  $s$  and  $\neg\phi$  being accepted ( $\phi$  being *rejected*) with respect to that  $s$ . It may be that  $\phi$  is neither accepted nor rejected. Nonepistemic sentences are therefore what we might call *acceptance-trivalent*.

The epistemic possibility operator  $\diamond$  exploits this trivalence: semantically it maps acceptance-trivalent sentences onto acceptance-*bivalent* ones. (Along with negation, it can be construed as an *acceptance-functional* operator.) It is the existence of a third acceptance value which introduces the nonclassical behaviour, and which lets us have both epistemic contradiction and nonfactivity.<sup>22</sup>

Our semantics lets us define two notions of consequence, diagonal consequence and informational consequence. The former rejects epistemic contradiction, the latter accepts it. Informational consequence seems to have two theoretical advantages: it avoids the apparently troubled notion of truth at a context for epistemic modal claims, and it validates some natural forms of inference invalidated by diagonal consequence. In the next section I will discuss a further consideration relevant to the question of which of these notions of consequence has greater theoretical interest.

## 5. Content and communication

Distinguish two questions.

1. What is the compositional semantics of an epistemic modal clause?
2. What informational content do utterances of epistemic modal sentences communicate?<sup>23</sup>

The questions are obviously related, but they should not be conflated. Very roughly, the first is a question of semantics, the second of prag-

<sup>22</sup> Informational consequence is similar to a notion of validity Frank Veltman defines over his update semantics (what he calls ‘validity<sub>3</sub>’; see Veltman 1996, p. 224). Though I lack the space to adequately discuss Veltman’s important work here, it should be noted that his semantics is, from an abstract point of view, very similar in its treatment of epistemic possibility modals to the domain semantics given above. On both approaches, the basic idea is to think of epistemic possibility clauses as expressing conditions on sets of worlds. Like domain semantics, Veltman’s semantics has no difficulty with epistemic contradictions. (Indeed, the facts discussed above concerning these conjunctions in embedded contexts provide strong evidence in favour Veltman’s semantics as over a relational semantics for epistemic modals.) Whether a static domain semantics for epistemic modals is preferable to a dynamic semantics along Veltman’s lines is not a question I consider here. In focusing only on a static domain semantics, I have two simple motivations: first, to contribute to understanding what a static alternative to Veltman’s proposal might look like; and second, to make for an easier approach into the analysis of probability operators, by separating out questions of dynamics. The second motivation—only a methodological one—will become clearer later (Sects. 7–8).

<sup>23</sup> Cf. Dummett’s distinction between ‘ingredient sense’ and ‘assertoric content’ in Dummett 1973 and elsewhere.



matics (or of the semantics-pragmatics interface). It is important to be clear that we have said a lot about the first question, and almost nothing about the second.

The second question has intrinsic interest. It might also be considered relevant to the question of what notion of consequence is of the most general theoretical relevance. For it might be held that a reasonable notion of consequence should be such as to preserve, in some relevant sense, the content communicated by sentences in context. (That is, it might be held that if  $\Gamma \models \phi$ , then for all  $c$ , the informational content communicated by  $\phi_c$  is included already in the informational content communicated by the sentences in  $\Gamma_c$ .) That is not exactly a radical view, so it is worthwhile, in comparing diagonal and informational consequence, to ask whether one of them dovetails better with the actual informational or communicative content of epistemic possibility claims.

That requires asking what the communicative content of epistemic possibility claims is. In this section I will discuss just two of the myriad possible answers to this question. Then I will say how each answer connects to the issue of consequence. The two views about communicative content I want to describe both assume the same abstract picture of linguistic communication, so let me start by spending three paragraphs sketching that picture.<sup>24</sup>

Think of linguistic communication as foremost a matter of coordination on a body of information. Participants in conversation begin with certain information presumed to be in common or mutually taken for granted, and the speech acts they perform in context are directed, and mutually understood to be directed, at variously influencing that common body of information. The attitude that communicating agents take towards the body of information they share is the attitude of *presupposition*. Presupposition is, in the intended sense, a public attitude: one presupposes propositional content  $p$  only if one presupposes that one's interlocutors also presuppose that  $p$ . When things are going as they should, the interlocutors of a discourse all make the same presuppositions, and we can say that everything that any participant presupposes is *common ground*, in the following sense:

It is *common ground* that  $p$  in a group just in case all members of the group presuppose that  $p$ , and all know that all presuppose that  $p$ , and all know that all know that all presuppose that  $p$ , etc.

<sup>24</sup>The picture is due in essentials to Stalnaker (see e.g. Stalnaker 1975), but there are some non-trivial differences in formulation.

What is common ground is what is common knowledge about what is presupposed. (When the agents in a discourse context are not all making the same presuppositions, something has gone wrong—the agents are misled about what is common ground—and the discourse context is defective, although the defect may never reveal itself.)

Given only what is common ground among a group of agents, one does not yet know how the agents of the context mutually regard the propositions in the common ground with respect to their other cognitive attitudes. To be given the common ground is only to be given a set of propositions mutually understood to be presupposed; it is not yet to be given that the agents also regard those presuppositions as knowledge, or as warranted belief, or conjecture, or fiction, or whatever. Using the notion of common ground, we can define a second notion which will let us articulate the status that the agents of a given context attach to the propositions they presuppose. Call this notion *conversational tone*:

An attitude is the *conversational tone* of a group of interlocutors just in case it is common knowledge in the group that everyone is to strike this attitude towards the propositions which are common ground.

(It may be that a conversation is plausibly understood as having more than one conversational tone, but let me focus on the case where there is just one. And let me stipulatively exclude presupposition itself from the class of possible conversational tones.) When interlocutors coordinate on a conversational tone, they come into agreement about what counts as the correct non-public attitude to take towards what is common ground. This will be a reflection, *inter alia*, of the purpose of the discourse. If the conversational tone of our discourse is knowledge, then we regard our common ground as common knowledge, and we take our discourse to be trafficking, and aiming to traffic, in factual information. Similarly with belief. If the conversational tone is pretense, then we are not attempting to keep the common ground compatible with the truth, and we take ourselves to be trafficking in fiction. And so on, for all the various attitudes around and in between: the conversational tone may be belief, or suspicion, or supposition, or high-credence-that, or ironic non-belief, etc., depending on the interests and purposes of the interlocutors. It may also be a conditional attitude: the conversational tone may, for instance, be belief (in each  $q$  in the common ground) *conditional on* some specified  $p$ .

Now speech acts on this picture are understood as influencing, and intended to influence, the information that is common ground. (Their appropriateness is therefore partly dependent on the conversational tone.) Of central interest to us is assertion. We will take assertion to be a speech act whose conventionally understood effect is to update the common ground of the conversation by adding the informational content of the speech act to the common ground. To assert informational content  $p$  is just to propose to change the common ground in a certain way, viz., by adding to it  $p$ . The assertion is successful when the proposal is accepted.

That is the picture. The view of assertion it comes with carves out a certain theoretical role, viz., that of the *informational content* of an assertion. (We might also like to call it the *proposition asserted*, or *what is said* by the sentence in context.) What we need to do now is to say what occupies this role in the case of epistemic possibility claims. What content do such claims serve to assert?

According to the first of the two views I want to consider, the informational content of an unembedded epistemic modal claim is the *diagonal proposition* determined by the two-dimensional matrix provided by the semantics of the clause. We can write it as follows:

$$\lambda c. \llbracket \Diamond \rrbracket \phi^{c, s_c, w_c}$$

The diagonal is a function from centered worlds (contexts) to truth values, or equivalently, a set of centered worlds.<sup>25</sup> The diagonal of  $\Diamond \phi$  is true with respect to a centered world just in case  $\phi$  is compatible with the state of information determined by that centered world. There is a certain obvious theoretical attraction in taking the diagonal as the communicative content of (not just epistemic modal claims but) assertions in general. Since most ignorance can be understood as ignorance of features of context—ignorance of features of the world of the context, or of the location of the context within the world—we never know what context we occupy. So it is a natural idea to represent interlocutors as communicating information by uttering sentences which determine conditions on the context. Obviously, the idea of a diagonal proposi-

<sup>25</sup> This proposal about the informational content of unembedded epistemic modal claims is akin to that of Egan 2007, though Egan does not arrive at his proposal via diagonalization.

I should stress that here I want to remain agnostic, in so far as I can, on the question of the metaphysical nature of the epistemic possibilities the diagonal carves up. My semantics makes it technically convenient to take the diagonal to divide the space of centered worlds, and that view of epistemic possibilities could be buttressed by Lewis 1979a; nevertheless, it would be acceptable for my purposes to take diagonals to divide the coarser space of possible worlds (Lewis 1980 defines a diagonal along such lines), or perhaps something else. I abstract also from the pragmatic complexities introduced by the assumption of centered worlds; see Egan 2007 for discussion.

tion requires the notion of a diagonal point to be well-defined: it requires a position on how context supplies a value for the information parameter.

As you might predict, diagonal content goes naturally with diagonal consequence. Diagonal consequence preserves diagonal content in the sense that  $\Gamma \vDash_d \phi$  iff every centered world  $c$  where the diagonals of the premisses in  $\Gamma$  are true at  $c$  is such that the diagonal of  $\phi$  is true at  $c$ . Informational consequence does not preserve diagonal content in this sense, since (for example) there are centered worlds where the diagonals of  $\Diamond \phi$  and  $\neg \phi$  are both true.

Now let me consider a second, very different response to the question of what content epistemic possibility claims serve to assert. According to this response, the question is actually confused. It just mistakes the speech act force of epistemic possibility claims. To say  $\Diamond \phi$  is not to propose to add some informational content, some proposition, to the common ground, as with assertions. Rather, it is to make explicit that  $\phi$ -possibilities are compatible with the common ground—to make ‘explicit that the negation of  $\phi$  is not presupposed in the context’ (to quote a passing suggestion of Stalnaker, 1970, p. 45). Suppose we followed Stalnaker in representing the information that is common ground by a *context set*, the set of possibilities where the propositions presupposed are all true. Then we could formalize this idea about the pragmatic effect of an epistemic possibility claim with the notion of acceptance defined above. To make an epistemic possibility claim in some context, on the present idea, is to propose to make it accepted with respect to the context set. What this speech act move exploits is not the diagonal of the epistemic modal sentence *per se*, but rather its horizontal at the context of utterance. The horizontal of an epistemic possibility claim determines a global condition on states of information (sets of possibilities), and the idea here is that in making such a claim, the speaker is proposing to make (or make explicit that) the context set satisfies this condition. (The horizontal associated with  $\Diamond \phi_c$  might be expressed as  $\lambda s. [[\Diamond \phi]]^{c,s,w_c}$ .) To agree, in context, on  $\Diamond \phi$  is to explicitly coordinate on a body of presuppositions compatible with  $\phi$ . A speaker who says  $\Diamond \phi$  is not expressing a proposition believed (known, etc.), but rather is expressing the compatibility of her state of mind with  $\phi$ .

This account of the pragmatics of epistemic possibility claims goes naturally with informational consequence, because according to it, the communicative impact of such claims is understood in fundamentally in terms of acceptance, and informational consequence is what preserves acceptance. There is no ‘proposition expressed’ by an epistemic

possibility claim on this picture, so there is no question of whether the proposition expressed is true or false. At most we can ask whether the claim is appropriate to accept or not, given the conversational tone(s) of the conversation.

Helping ourselves to the idea of a context set, we might summarize the two views just described about the pragmatics of epistemic possibility claims as follows.

- **Diagonal view:** To say  $\diamond\phi$  in a context  $c$  is to propose to make  $\dagger\diamond\phi_c$  accepted with respect to the context set of  $c$ <sup>26</sup>
- **Informational view.** To say  $\diamond\phi$  in a context  $c$  is to propose to make  $\diamond\phi_c$  accepted with respect to the context set of  $c$

How to choose?

The informational view has the same advantages over the diagonal view that informational consequence has over diagonal consequence: it avoids the need to define diagonal points of evaluation, and it gels better with intuition when it comes to inferences involving epistemic claims. The first point is obvious; let me give an illustration of the second.

Suppose the following. (1) Nobody—including ourselves—knows whether or not there is lead on Pluto, and indeed nobody is even close to having any evidence on the question of whether there is lead on Pluto. (2) As a matter of fact, there is no lead on Pluto. Now, on the basis of the information provided by these two premisses, is the following sentence true or false?

There might be lead on Pluto

There is strong pull to answer ‘false’. What that suggests is that the unembedded sentence ‘There might be lead on Pluto’ is not really understood as literally describing the condition of some agent’s evidential state, as on the diagonal view. (If it were, you would presumably be inclined to say ‘true’, since we have stipulated that, in the envisaged scenario, there is lead on Pluto for all anyone knows.) Rather, the behaviour of the sentence is akin to its behaviour in embedded contexts. The epistemic possibility operator is sensitive, not to the *possession* of some body of information by some agents, but rather only, as it were, to what is possessed: to the information itself. Its role is to place a condition on a possible body of information. In the sentence above, the modal is

<sup>26</sup>The dagger ‘†’ is a two-dimensional modal operator which takes the diagonal of the sentence it embeds and projects it onto to the horizontal. See Lewis 1973, p. 63–4 for discussion.

understood relative to the information conveyed by the premisses I asked you to suppose. When you evaluated ‘There might be lead on Pluto’ for truth, plausibly what you considered was whether lead’s being on Pluto would be compatible with the information you were asked to take for granted. We could say that you assessed whether the sentence was acceptable (in the technical sense) with respect to a certain temporary or ‘derived’ context set, one which included the information provided by premisses I asked you to take as given. Your judgement of falsity, on this interpretation of the facts, was really a (correct) judgement that the sentence could not be accepted with respect to that body of information.

I am about to conclude that intuition favours the informational view and informational consequence. Before that, let me consider a worry about that conclusion. The worry is that epistemic modal claims sometimes seem to communicate some kind of objective information, and it is not obvious how the informational view explains this. Take for instance:

- (21) Cheerios may reduce the risk of heart disease
- (22) Late Antarctic spring might be caused by ozone depletion<sup>27</sup>

We tend to hear these sentences as (not just making certain possibilities explicit but) serving to communicate real information. Indeed, they strike us as the result of some actual research. This is intelligible on the diagonal view, according to which a proposition is expressed by epistemic modal claims. But how can we understand it on the informational view?

As follows. Sometimes when we converse we do so with the tacit aim of keeping our presuppositions compatible with (as it might be) the knowledge of the relevant experts. We try to get our presuppositions to relevantly overlap with expert knowledge. We try to obey a rule like:

Presuppose  $\phi$  iff  $\phi$  is known by the relevant experts

Our conversational tone is something like: treat as known by the relevant experts. In such cases epistemic modal claims, which on the informational view are pragmatically understood as imposing a condition on the information presupposed, will be assessed for correctness according to whether the informational condition they express is actually satisfied by the ‘target’ information—in this case, the relevant expert knowledge. They will therefore be ‘heard’ as communicating information about the knowledge of the relevant experts—concerning, as it might be, the

<sup>27</sup>The title of a 1987 article in *Nature* (Anderson 1987).

health effects of Cheerios, or the causes of the late Antarctic spring. In such contexts, epistemic possibility claims will be harder to make appropriately. These are contexts where it may be quite natural to say something like

- (23) I do not know whether the late Antarctic spring might be caused by ozone depletion

On the natural reading of (23), what one grants is that one does not know whether something is an open possibility according to the target state of information that our presuppositional context aspires to. In these cases, we need to allow that interpretation may involve a tacit shift in the information parameter under the scope of 'knows', a shift to the target state of information for the context. Aside from Gricean considerations of charitable interpretation, it is not obvious whether general principles are involved in the interpretation of such tacit shifts.

(Of course, expert knowledge need not be the only sort of target information we attempt to keep our presuppositions in line with. We may have some specific body of evidence in mind, or we may be interested in what *could* be known about a topic if the investigative circumstances were ideal etc.)

I conclude that intuition favours the informational view and informational consequence. This conclusion suggests that we should not—or at least, we need not—actually think of the semantics proposed as a two-dimensional semantics. Were the semantics two-dimensional, the existence of diagonal points would be guaranteed. But if our conclusion is right, we need not assume the existence of diagonal points at all; hence we need not assume a purely two-dimensional semantics. The information parameter is perhaps better treated as semantically *sui generis*, not parasitic on Kaplanian contexts in the way that indices by definition are.

(If a purely two-dimensional semantics were found to be desirable on independent grounds, however, perhaps the best way to preserve the diagonal view in the face of the threats described above would be to effectively collapse it into the informational view, by letting the *s* parameter of a diagonal point be the context set of the context of that point. If diagonal points are defined this way, the two pragmatic moves technically come to the same thing. Note that this move would require abandoning Reflexivity, since the context set of a conversation need not include the actual world. For this reason, the resulting definition of truth at a context would perhaps not be intuitive.)

## 6. Epistemic necessity operators

I have focused on the attractions of a domain semantics for epistemic possibility modals. A domain semantics for epistemic necessity modals has similar attractions. Here is the appropriate ‘dual’ semantics for epistemic necessity operators:

$$\llbracket \Box \phi \rrbracket^{c,s,w} \text{ is true iff } \forall w' \in s : \llbracket \phi \rrbracket^{c,s,w'} \text{ is true}$$

This semantics has three nice features. First, it explains what is wrong with

(24) # Suppose it is not raining and it must be raining.

and its ilk. The explanation is along precisely the same lines as the domain semantics for epistemic possibility modals:  $(\neg\phi \ \& \ \Box\phi)$  is an unacceptable sentence.

Second, it captures, to some degree, the sense in which epistemic necessity modals serve to indicate that a conclusion is being drawn from some (perhaps tacit) premisses. The reason is simple. On a domain semantics,  $\Box\phi$  expresses a condition, not on possible worlds, but on bodies of information (sets of worlds). A body of information satisfies the condition expressed just in case  $\phi$  follows from that information.

Third, in conjunction with our semantics for indicative conditionals, it explains the following observation. Observation:  $(\alpha \rightarrow \psi)$  and  $(\alpha \rightarrow \Box\psi)$  usually sound equivalent. Illustration:

Either the butler or the gardener did it. Therefore:

(C1) If the butler did not do it, the gardener did.

(C2) If the butler did not do it, the gardener must have.

(C1) and (C2) sound semantically equivalent. The explanation for why these two sentences sound equivalent on a domain semantics is that the sentences *are* equivalent. (C2) merely involves some additional vacuous quantification: the universal quantification introduced by the indicative conditional connective in (C2) is trumped by the quantification introduced by the embedded epistemic necessity modal. The conditional connective only influences interpretation in (C2) by shifting the information parameter over which the modal quantifies.<sup>28</sup>

<sup>28</sup> If the connective  $\rightarrow$  is decomposed into two semantic ingredients along the Kratzerian lines of n. 15 above, a second interpretation of the facts emerges. On the second interpretation, the only difference between (C1) and (C2) is in surface syntax: the epistemic modal explicit in (C2) is covert in (C1). Whether this interpretation is preferable depends on whether there is independent evidence for the presence of a covert modal in (C1)—not a question I will look into here. Suffice to say that both interpretations can be expressed in a domain semantics.



## 7. Toward probability operators

More trouble ...

- (25) # Suppose it is not raining and it is likely that it is raining
- (26) # Suppose it is raining and it probably is not raining.
- (27) # If it is not raining and it is probably raining, then ...
- (28) # If it is raining and it is likely that is not raining, then ...

Certain probability operators—‘it is likely that’, ‘probably’, etc.—give rise to epistemic contradictions in the same way that epistemic possibility and necessity modals do. In the remaining pages I will sketch, in broad strokes, an approach to these operators, one developed in more detail elsewhere (Yalcin forthcoming; Yalcin in preparation).

Abbreviate ‘it is probable that  $\phi$ ’ and its kin (‘probably  $\phi$ ’, ‘it is likely that  $\phi$ ’) as  $\Delta\phi$ . The project is to state truth-conditions for  $\Delta\phi$ . The basic idea of the approach I want to recommend is simple: just upgrade the kind of object the information parameter can take as a value, from a set of worlds to a *probability space*. The intension of a sentence, relative to context, will be a function from world-probability space pairs to truth values. We will take it that a probability space  $P$  determines a probability measure  $Pr_P$  over sets of possible worlds, and this measure is exploited in the semantics of  $\Delta$  as follows:

$$\llbracket \Delta\phi \rrbracket^{c, P, w} \text{ is true iff } Pr_P(\{w : \llbracket \phi \rrbracket^{c, P, w} \text{ is true}\}) > \frac{1}{2}$$

Relative to context,  $\Delta\phi$  determines a condition on probability spaces. The condition is satisfied just in case (roughly)  $\phi$  is more likely than not according to the probability measure of the space.

There are some subtleties concerning what definition of ‘probability space’ is best suited to natural language probability operators. What follows is just one path through the decision tree; certainly, others are possible and worth exploring.

Think of a probability space (a state of information) as a certain triple  $\langle \Pi, \pi, Pr \rangle$ . Let me describe each member of the triple in turn. First,  $\Pi$  is partition over the space all possible worlds. The cells of this partition will represent the space of possible alternatives that are ‘recognized’ by the probability space (information state), in the sense that the grain of this partition will determine the possible worlds propositions that the probability measure of the space is defined over. A given  $\Pi$  may be said to recognize a possible worlds proposition  $p$  as an alternative just in case every cell in  $\Pi$  *classifies* with respect to  $p$ : just in case every

cell  $\iota$  in  $\Pi$  is such that, either every world in  $\iota$  is a  $p$ -world, or every world in  $\iota$  is  $\neg p$ -world. To use a visual metaphor,  $\Pi$  provides a kind of ‘resolution’ over logical space: propositions not classified by  $\Pi$  are not ‘seen’ by the information state. (Cf. Lewis 1988 on subject matters.)

Second,  $\pi$  is a subset of  $\Pi$ . (It is therefore also a partition.) The cells of  $\pi$  are to be the *live* possible alternatives: they reflect what is *really* epistemically possible according to the probability space (information state). All of the probability mass of the probability measure will be located on the  $\pi$ -region of logical space.

Last, define  $Pr$  so that:

- (i)  $Pr$  assigns each cell  $\iota$  in  $\pi$  a real value in the closed interval from zero to one, such that these values all sum to one
- (ii) For all propositions  $p$  that  $\Pi$  classifies,  $Pr(p) =_{\text{def}} \sum_{\iota \subseteq p} Pr(\iota)$ ; otherwise  $Pr(p)$  is undefined

My distinguishing  $\pi$  and  $\Pi$  may seem unnecessary. Why distinguish a special set  $\pi$  of epistemic possibilities? Why not leave  $\pi$  out of the formalism and let the epistemic possibilities just be those cells with nonzero probability? Because, at least not without further assumptions, it would be mistake to collapse epistemic possibility with nonzero probability. Continuous sample spaces in which probability zero events may nevertheless happen provide the usual counterexamples. See McGee 1994 and Hájek 2003 for further discussion of this issue.<sup>29</sup>

Since we have changed the formal representation of information associated with the information parameter, and since epistemic possibility and necessity modals access this parameter, we need to update our semantics for these operators. The obvious idea would be to understand them as determining conditions on  $\pi$ , the epistemic possibilities associated with the relevant  $\mathbf{P}$ . To state the new semantics, a space-saving definition comes in handy: define truth and falsity with respect to a context, a probability space, and a *cell*  $\iota$  as follows:

$\llbracket \phi \rrbracket^{c, \mathbf{P}, \iota}$  is true (false) iff  $\forall w \in \iota : \llbracket \phi \rrbracket^{c, \mathbf{P}, w}$  is true (false)

Then the semantics for epistemic possibility and necessity modals is this:

$\llbracket \diamond \phi \rrbracket^{c, \mathbf{P}, w}$  is true iff  $\exists \iota \in \pi_{\mathbf{P}} : \llbracket \phi \rrbracket^{c, \mathbf{P}, \iota}$  is true

$\llbracket \square \phi \rrbracket^{c, \mathbf{P}, w}$  is true iff  $\forall \iota \in \pi_{\mathbf{P}} : \llbracket \phi \rrbracket^{c, \mathbf{P}, \iota}$  is true

<sup>29</sup> Thanks here to Alan Hájek, Kenny Easwaran, and an anonymous reviewer. I have benefitted also from Aidan Lyons’s unpublished work on this topic.

The probabilistic semantics for  $\Delta\phi$ ,  $\Diamond\phi$ , and  $\Box\phi$  just given calls for one further assumption. Let us say that epistemic modal clauses carry a *classification presupposition*, to the effect that the partition  $\Pi$  of the  $\mathbf{P}$  they are evaluated with respect to *classifies* the possible worlds proposition expressed by their complements  $\phi$ . A probability space speaks to the question of whether a proposition is possible or probable only if the proposition is classified according to the space.

There are a lot of questions to be raised about this semantics, both technical and philosophical. Reserving extended discussion for elsewhere, let me devote the remainder of this section to a straight technical question: how do we use this semantics to explain what is going on with our probabilistic epistemic contradictions in (25)–(28)? The obvious thing to do would be to follow the same strategy used earlier: first, define a notion of acceptance according to which the conjunctions are unacceptable, and second, give a semantics for the relevant embedding environments according to which these environments require acceptable complements.

The first step of this strategy is simple enough. We can update our definition of acceptance as follows:

$\phi_c$  is *accepted* with respect to  $\mathbf{P}$  iff  $\forall u \in \pi_{\mathbf{P}} : \llbracket \phi \rrbracket^{c, \mathbf{P}, u}$  is true

On this definition of acceptance, it is trivial to verify that all the relevant epistemic contradictions ( $(\neg\phi \ \& \ \Delta\phi)$ ,  $(\phi \ \& \ \Diamond\neg\phi)$ , etc.) are unacceptable.

The second step, however, is not as simple. Defining a semantics for attitude verbs and for indicative conditionals in the current probabilistic setting is a subtle matter, one raising considerations beyond the scope of this paper. I will have to settle for some sketchy and preliminary remarks on these constructions, the aim being only to give a sense of the prospects for probabilistic analyses and of the decision points that arise.

First, it is natural to conjecture that the semantics for acceptance attitude verbs ('believes', 'knows', 'accepts', 'supposes', etc.) can straightforwardly mirror our earlier domain semantics (Sect. 3). Let these verbs shift the value of the information parameter to the information state corresponding to the attitude state of the subject, and let the whole ascription require, for truth, that the complement of the verb be accepted with respect to that information state. The information parameter ranges over probability spaces, so the semantics assumes that these attitude states can be modelled by such spaces. The question

arises how exactly to interpret the probabilities that go into modelling these attitude states. This issue is discussed, inconclusively, in the final section (Sect. 8).

Second, indicative conditionals. As with the attitudes, it would be natural to expect their analysis to be a probabilistic analogue of the domain semantics presented above. Here is a first pass. Conditionals express properties of probability spaces: an indicative conditional ( $\phi \rightarrow \psi$ ) in context is true with respect to a probability space  $P$  just in case a certain other probability space (determined as a function of  $P$ ) which accepts the antecedent also accepts the consequent. So the interpretation of an indicative will again involve an information parameter shift, a shift to a probability space accepting the antecedent. Which space do we shift to? The one that involves the ‘minimal change’ to  $P$  needed to make the antecedent accepted with respect to that space. Here one can expect various theories of minimal change, which will need testing on specific examples. In the case where the antecedent is nonepistemic, a natural idea would be to shift to the probability space whose measure is just the conditionalization of the antecedent on the measure associated with  $P$ . In the less common case where the antecedent is epistemic, it is less obvious what to say. (We could try shifting to the space  $P'$  whose measure satisfies the condition and which is such that the relative entropy between the measures of  $P'$  and  $P$  is minimized; but care would need to be taken to avoid certain pitfalls for relative entropy minimization. See Grünwald and Halpern 2003 for a sense of the issues and references.)

It should be noted that a probabilistic semantics for indicative conditionals along the lines just described has a familiar independent motivation. Thanks to the triviality results of Lewis (1976) and others, it is well-known that if indicative conditionals express possible worlds propositions, the probabilities of the propositions they express could not in general be identical to the probabilities of their consequents conditional on the corresponding antecedents. It is also widely thought (thanks especially to Adams 1975) that our tendency to accept an indicative conditional correlates closely with our intuitions about the corresponding conditional probability. Impressed by Adams’s thesis of acceptability and by the triviality results, many theorists have been tempted to conclude that indicative conditionals do not express possible worlds propositions; and from this they are tempted to effectively abandon semantics for indicatives altogether. I sympathize with the first temptation, but not the second. We can deny that indicative conditionals have possible worlds truth-conditions without denying that

they have compositional semantic values. We can do it by saying that their compositional semantic values relative to context are effectively conditions on probability spaces—specifically, conditions on the relevant conditional probabilities. This would let us keep a tight semantic connection between indicative conditionals and the corresponding conditional probabilities without having to maintain, implausibly, that compositional semantics stops at ‘if’.

## 8. Outstanding issues

A lot of questions remain open. Here are some of them.

**Outstanding semantic issues.** There is still plenty of formal semantics left to do. The above semantics for epistemic modals should be connected in a natural way with the semantics of epistemic adjectives (as in ‘This a possible design for the new museum’). The work on probability operators should be connected with work on gradable adjectives generally, since these operators take all the same morphology and occur in comparative form (‘as probable as’, ‘more likely than’). The interaction of tense with this semantics for epistemic modals needs investigation. The attitude semantics given above should be shown to interact with plausible story about hyperintensionality. Finally, a detailed comparison of the static semantics I have sketched and a dynamic semantics for epistemic modals (along the lines of Veltman 1996 and Beaver 2001) is in order.<sup>30</sup>

**The representation of uncertainty.** In the previous section, I assumed without question that the representation of uncertainty appropriate to what I am calling probability operators is the probability space. But as it well known, there are numerous ways to represent uncertainty formally (see Halpern 2003), and it may be questioned whether probability spaces really are appropriate to the semantics of (what superficially appears to be) natural language probability talk. Hamblin 1959, an impressive early investigation into this question, seems to favour a plausibility measure approach; and Kratzer 1991 gives a semantics for probability operators in terms of nonnumerical qualitative orderings of possibilities. It would be desirable to demonstrate, in so far as possible, that the resources of probability theory are in fact needed.

<sup>30</sup> For recent work on epistemic adjectives, see Swanson 2006. For recent work on the interaction of tense and epistemic modality, see Condoravdi 2002 and MacFarlane 2006.

**The interpretation of probabilities and probabilities in interpretation.**

Having assumed that probability spaces are in fact appropriate to the modelling of probability operators and related constructions, the question arises how best to understand the notion of probability at work in the semantics. There are a number of options here. Let me just mention two of interest, again reserving extended discussion for elsewhere.

First, we can try interpreting the probabilities along Bayesian lines, thinking of them as measuring degrees of confidence. On this interpretation, we can use the semantics to formalize the idea that, in saying that it is probably raining, one thereby expresses one's credence in the proposition that it is raining—where this is not the same as saying that one's credence in the proposition is thus and so. What is the distinction between expressing one's credence and saying something about one's credence?

The contrast is the same here as it is with the expression of straightforwardly factual beliefs. Let Cleopatra say

Antony's fleet outnumbers the enemy's

She thereby expresses her belief that Antony's fleet outnumbers the enemy's, but she does not say that she has this belief. She is talking about the opposing fleets, not about her beliefs. (Gibbard 1990, p. 84)

Exactly right. (I quote Gibbard out of context—the contrast he refers to is not actually my distinction between expressing one's credence and saying something about one's credence, but rather his distinction between expressing one's acceptance of a system of norms and saying that one accepts the system—but his analogy is perfect for the contrast I want to draw.) Now suppose Cleopatra says

Antony's fleet probably outnumbers the enemy's

On the Bayesianism-inspired interpretation of the semantics I want to consider, Cleopatra here expresses her state of high credence, or her confidence, in the proposition that Antony's fleet outnumbers the enemy's. She does not say that she is in this state of confidence. The only proposition in the vicinity is one about the opposing fleets.

How can our probabilistic semantics help to formalize this idea? Relative to context, the intension of 'Antony's fleet probably outnumbers the enemy's' determines a set of probability spaces—namely, the set of spaces in which the possible worlds proposition that Anthony's fleet outnumbers the enemy's receives a probability greater than one-half. If the probability spaces of the semantics are interpreted as idealized representations of credal states, then we can think of the sentence, relative to con-

text, as expressing a property of credal states. Cleopatra expresses an aspect of her state of credence by uttering a sentence which, relative to context, expresses a property her credal state has, namely, the property of giving greater than one-half credence to the proposition that Anthony's fleet outnumbers the enemy's. Thereby she 'gives voice' to that aspect of her credal state, but without uttering a possible worlds proposition about her credence. In a nod to the structurally similar view in metaethics, we could call this view about probability talk *credal expressivism*.

Credal expressivism is, I think, already tacit in the way that many Bayesians tend to informally describe epistemic modal beliefs. The usual way of modelling, within a Bayesian framework, someone describable as believing (for example) that it is probably raining would be to let the credence function characterizing their credal state map the proposition that it is raining to some highish value. Whether someone accepts what an epistemic modal clause says is thus generally taken to be a matter of their credence in the proposition expressed by the sentence embedded under the modal—not a matter of their credence in a proposition about their credence. In the attitude report, the modal tends to be treated, as it were, adverbially: the object of the agent's attitude is the proposition that it is raining, and the modal tells how strongly the proposition is believed. Attitude semantics for 'believes' along the probabilistic lines briefly sketched above can make semantic sense of this Bayesian tendency.

A second way of interpreting the probabilities is as measuring 'how far evidence supports or counts against various hypotheses about the world' (Mellor, 2005, p. 79), or 'something like the intrinsic plausibility of hypotheses prior to investigation' (Williamson, 2000, p. 211). This is sometimes called the *epistemic* or *evidential* interpretation of probabilities. While on the Bayesian interpretation probabilities measure the strength of an agent's confidence, on the evidential interpretation probabilities measure something like the objective degree of confirmation a body of propositions confers on a given proposition. We can still use the probability spaces of our semantics to characterize attitude states on this interpretation of the probabilities. But the import of the representation is quite different. The probabilities are now to be understood as a part of the informational content of the attitude state—not as measures of the strength of the attitude towards content. We could call this view *content probabilism*. It is the view that informational content itself is probabilistically articulated.

A challenge for the credal expressivist about epistemic modal talk is to make sense of the felicitous occurrence of epistemic modals in attitude contexts for which the corresponding attitude does not, intuitively, come in degrees. A mixed strategy may therefore be in order as far as the interpretation of the probabilities goes: perhaps credal expressivism and content probabilism are each appropriate to different fragments of our folk probabilistic talk. The question of how the two might be integrated deserves investigation.

**Probability in context.** It is a short step from a probabilistic semantics and probabilistic representations of attitude states to a probabilistic pragmatics. Above (Sect. 5) I followed Stalnaker in treating the attitude of presupposition as central to characterizing the informational context (common ground) of a conversation, and in treating the informational context as central to the characterization of speech acts like assertion. Suppose now that, along content probabilist lines, we took it that the informational content of a state of presupposition could be characterized by a probability space, or by a set of such spaces. This would lead us naturally to a view we could call *context probabilism*.

*Context probabilism:* the common ground of a conversation is characterizable as a probability space, or as a set of such spaces.

Rather than representing the common ground by a context set, a set of possible worlds, we would represent it as a probability space—call it a *context probability space*—or as a set of probability spaces—what we might call a *context representor*. We could then think of sentences uttered in context as serving to determine constraints on probability spaces, and thereby on the common ground. To utter  $\phi$  in  $c$ , we could try saying, is to propose to make it accepted by the context probability space (or by all the probability spaces in the context representor). This would let us model the communication of information, not only in terms of the elimination of possibilities, but also in terms of the elimination or evolution of the possible *probabilities* over possibilities. It would let us represent the transfer of purely probabilistic information.

What is at issue in the choice between representing the common ground as a context representor or as a single context probability space? If each thing presupposed in context determines a constraint on a probability space, then the context representor can be understood as just the set of probability spaces satisfying all those constraints. We can think of these as the probability spaces that are admissible given what is presupposed. This representation of the common ground seems to have an



advantage over the representation in terms of a single context probability space: it allows us to avoid the idealization that interlocutors in context coordinate on precise probabilities for the propositions they are concerned about. Nevertheless, we might try constructing a single context probability space from the context representor, by finding the probability space that satisfies all the constraints associated with representor and which otherwise maximizes entropy. Both representations deserve investigation.

The probabilistic representation of context we settle on—a context probability space, a context representor, or something else—will determine our options for modelling the dynamics of context change. In this paper I have mainly focused on static acceptability conditions: I have asked what properties an agent's presuppositional state must satisfy in order to count as accepting the various epistemic modal claims I have discussed. What I have left out is an account of the dynamics of presupposition. If you are in some presuppositional state, and your interlocutor proposes that you move to a presuppositional state satisfying property *F*, exactly how should you shift your state in order to satisfy *F*? Any complete version of context probabilism will have to address this question. Individual probability spaces can be understood to evolve by various forms of conditionalization, and by shifts in what counts as epistemically possible according to the space; context representors can be understood to evolve by changing their members, permitting or eliminating new probability spaces. If something like context probabilism is on track, the proper treatment of dynamics will be among the leading questions.<sup>31</sup>

*Department of Linguistics and Philosophy*  
*Massachusetts Institute of Technology*  
 77 Massachusetts Avenue, 32-D940  
 Cambridge, MA 02139  
 USA  
 yalcin@mit.edu

SETH YALCIN

<sup>31</sup> I have benefitted from conversations on the topics of this paper with many people. In addition to those mentioned already above, let me thank David Chalmers, Kai von Fintel, Ephraim Glick, Valentine Hacquard, Irene Heim, Angelika Kratzer, Vann McGee, Sarah Moss, Agustin Rayo, Tamina Stephenson, and Eric Swanson. For extended discussions, special thanks to Andy Egan, John MacFarlane, Dilip Ninan, Bob Stalnaker, and Stephen Yablo. Support from the American Council of Learned Societies and the Andrew W. Mellon Foundation is gratefully acknowledged.

## References

- Adams, Ernest 1975: *The Logic of Conditionals*. Dordrecht: D. Reidel.
- Almog, J., J. Perry, and H. Wettstein (eds) 1989: *Themes from Kaplan*. Oxford: Oxford University Press.
- Anderson, Alun 1987: 'Late Antarctic spring might be caused by ozone depletion'. *Nature*, 330, p. 686.
- Beaver, David 2001: *Presupposition and Assertion in Dynamic Semantics*. Stanford, CA: CSLI Publications.
- Beaver, David, L. Casillas, B. Clark, and S. Kaufmann (eds) 2002: *The Construction of Meaning*. Stanford, CA: CSLI Publications.
- Condoravdi, Cleo 2002: 'Temporal Interpretation of Modals: Modals for the Present and for the Past,' in Beaver, Casillas, Clark, and Kaufmann 2002, pp. 59–88.
- DeRose, Keith 1991: 'Epistemic Possibilities'. *Philosophical Review*, 100, 4, pp. 581–605.
- Dummett, Michael 1973: *Frege: Philosophy of Language*. Cambridge, MA: Harvard University Press.
- Eells, E. and B. Skyrms (eds) 1994: *Probability and Conditionals*. Cambridge: Cambridge University Press.
- Egan, Andy 2007: 'Epistemic Modals, Relativism and Assertion'. *Philosophical Studies*, 133, 1, pp. 1–22.
- Egan, Andy, John Hawthorne, and Brian Weatherson 2005: 'Epistemic Modals in Context'. In Preyer and Peter 2005, pp. 131–68.
- Egan, Andy and Brian Weatherson (eds) forthcoming: *Epistemic Modality*. Oxford: Oxford University Press.
- Eikmeyer, H.-J. and H. Rieser (eds) 1981: *Words, Worlds, and Contexts, New Approaches to Word Semantics*. Berlin: Walter de Gruyter.
- Engel, P. (ed.) 2000: *Believing and Accepting*. Dordrecht: Kluwer.
- Farley, A., P. Farley, and K.-E. McCullough (eds) 1986: *Papers from the Parasession on Pragmatics and Grammatical Theory*. Chicago: Chicago Linguistics Society.
- Gibbard, Allan 1990: *Wise Choices, Apt Feelings*. Cambridge, MA: Harvard University Press.
- Grünwald, Peter and Joseph Halpern 2003: 'Updating Probabilities'. *Journal of Artificial Intelligence Research*, pp. 243–78.
- Hacking, Ian 1967: 'Possibility'. *Philosophical Review*, 76(2), pp. 143–68.
- Hájek, Alan 2003: 'What Conditional Probability Could Not Be'. *Synthese*, 137, pp. 273–323.
- Halpern, Joseph 2003: *Reasoning about Uncertainty*. Cambridge, MA: MIT Press.
- Hamblin, Charles L. 1959: 'The Modal "Probably"'. *Mind*, 68, pp. 234–40.

- Hintikka, Jikko 1962: *Knowledge and Belief: An Introduction to the Logic of Two Notions*. Ithaca: Cornell University Press.
- Kanger, S. and S. Ohman (eds) 1980: *Philosophy and Grammar*. Holland: D. Reidel.
- Kaplan, David 1989: 'Demonstratives'. In Almog, Perry, and Wettstein 1989, pp. 481–563.
- Kratzer, Angelika 1977: 'What "Must" and "Can" Must and Can Mean'. *Linguistics and Philosophy*, 1, pp. 337–55.
- Kratzer, Angelika 1981: 'The Notional Category of Modality'. In Eikmeyer and Rieser 1981, pp. 38–74.
- Kratzer, Angelika 1986: 'Conditionals'. In Farley, Farley, and McCullough 1986, pp. 1–15.
- Kratzer, Angelika 1991: 'Modality'. In von Stechow and Wunderlich 1991, pp. 639–50.
- Lewis, David 1973: *Counterfactuals*. Malden: Blackwell.
- Lewis, David 1976: 'Probabilities of Conditionals and Conditional Probabilities'. *Philosophical Review*, 85, pp. 297–315.
- Lewis, David 1979a: 'Attitudes *De Dicto* and *De Se*'. *Philosophical Review*, 88, 4, pp. 513–43.
- Lewis, David 1979b: 'Scorekeeping in a Language Game'. *Journal of Philosophical Logic*, 8, pp. 339–59.
- Lewis, David 1980: 'Index, Context, and Content'. In Kanger and Ohman 1980, pp. 79–100.
- Lewis, David 1988: 'Relevant Implication'. *Theoria*, 54, pp. 161–74.
- Łukasiewicz, Jan 1930: 'Philosophische Bemerkungen zu mehrwertigen Systemen des Aussagenkalküls.' *Comptes rendus des seances de la Societe des Sciences et des Lettre de Varsovie*, 23. Translated as 'Philosophical Remarks on Many-Valued Systems of Propositional Logic' in Łukasiewicz 1970, pp. 153–78.
- Łukasiewicz, Jan 1970: *Selected Works*. Amsterdam: North Holland.
- MacFarlane, John 2006: 'Epistemic Modals are Assessment-Sensitive'. Unpublished manuscript.
- McGee, Vann 1994: 'Learning the Impossible'. In Eells and Skyrms 1994, pp. 179–200.
- Mellor, D. H. 2005: *Probability: A Philosophical Introduction*. New York: Routledge.
- Preyer, Gerhard and G. Peter (eds) 2005: *Contextualism in Philosophy*. Oxford: Oxford University Press.
- Puig-Waldmüller, E. (ed.) 2007: *Proceedings of Sinn und Bedeutung*, 11. Barcelona: Universitat Pompeu Fabra.
- Rescher, N. (ed.) 1968: *Studies in Logical Theory*. Oxford: Basil Blackwell.

- Stalnaker, Robert 1968: 'A Theory of Conditionals'. In Rescher 1968, pp. 98–112.
- Stalnaker, Robert 1970: 'Pragmatics'. *Synthese*, 22, pp. 272–89.
- Stalnaker, Robert 1975: 'Indicative Conditionals'. *Philosophia*, 5, pp. 269–86.
- Stalnaker, Robert 2000: 'On "Moore's Paradox"'. In Engel 2000, pp. 93–100.
- Stephenson, Tamina 2007: 'A Parallel Account of Epistemic Modals and Predicates of Personal Taste'. In Puig-Waldmüller 2007, pp. 583–97.
- Swanson, Eric 2006: *Interactions with Context*. Ph.D. thesis, MIT.
- Unger, Peter 1975: *Ignorance: a Case for Scepticism*. Oxford: Oxford University Press.
- Veltman, Frank 1996: 'Defaults in Update Semantics'. *Journal of Philosophical Logic*, 25, 3, pp. 221–61.
- von Stechow, A. and D. Wunderlich (eds) 1991: *Semantics: An International Handbook of Contemporary Research*. Berlin: de Gruyter.
- Williamson, Timothy, 2000: *Knowledge and its Limits*. Oxford: Oxford University Press.
- Yalcin, Seth forthcoming: 'Nonfactualism about Epistemic Modality'. Forthcoming in Egan and Weatherson forthcoming.
- Yalcin, Seth in preparation: 'Probability Operators'.