The Aristotelian Society

133rd session – 2011/2012

9 January 2012 | 16.15 – 18.00

Bayesian Expressivism

Seth Yalcin
University of California, Berkeley

The Woburn Suite
Senate House (South Block)
University of London
Malet Street
London WC1E 7HU
United Kingdom

BIOGRAPHY

Seth Yalcin is an Assistant Professor of Philosophy at the University of California, Berkeley, where he is also a member of the Group in Logic and the Methodology of Science. Prior to that he was an Assistant Professor of Philosophy at New York University. He holds a PhD in philosophy from the Massachusetts Institute of Technology. He works mostly in the philosophy of language, on descriptive and foundational issues in natural language semantics. Lately his work has borrowed ideas from formal epistemology and from metaethical expressivism to develop accounts of the meaning of epistemic and deontic modals, probability operators, conditionals, attitude verbs, and the language of spatial orientation. He also has research interests in metaphysics, on questions about the nature of modality, information, and randomness. Seth Yalcin is an Assistant Professor of Philosophy at the University of California, Berkeley, where he is also a member of the Group in Logic and the Methodology of Science. Prior to that he was an Assistant Professor of Philosophy at New York University. He holds a PhD in philosophy from the Massachusetts Institute of Technology. He works mostly in the philosophy of language, on descriptive and foundational issues in natural language semantics. Lately his work has borrowed ideas from formal epistemology and from metaethical expressivism to develop accounts of the meaning of epistemic and deontic modals, probability operators, conditionals, attitude verbs, and the language of spatial orientation. He also has research interests in metaphysics, on questions about the nature of modality, information, and randomness.

Bayesian Expressivism*

Seth Yalcin

Draft of Dec 29 2011

yalcin@berkeley.edu

1 A Bayesian bottleneck

There is a tension between the broadly decision-theoretic view of agents and a certain standard picture of linguistic communication. If we adopt both, we are driven to the view that linguistic communication involves a curious bottleneck. We do better, ultimately, to change the standard picture of communication. Let me explain.

Communicating agents aim, generally speaking, to achieve some coordination on aspects of their states of mind by exploiting common knowledge of relevant dimensions of the syntax, semantics, and pragmatics of their language. Now suppose, in the spirit of the decision-theoretic perspective, that agents engaged in conversation are equipped with doxastic states representable by probability spaces and preference states representable by utility functions. Then if agents wish to coordinate on aspects of their credences *per se*, or on aspects of their utilities *per se*, by using language, their route must be indirect, given usual ways of modeling linguistic communication. The reason is that standard approaches in semantics and pragmatics have generally framed the enterprise of communication in a way which takes the transmission of nongraded, 'binary' belief as the paradigm. What one generally 'puts into play', in making an assertion, is a potential object of (nongraded, binary) belief—a proposition. What the compositional semantics of the language is charged with doing, fundamentally, is determining propositions relative to context. The attitude of mutual presupposition which supports the informational *common ground* of a conversation—the body of information we aim to influence and update via speech acts like assertion—is assumed to be analogous to the state of nongraded, binary belief. From a decision-theoretic perspective, this picture of communication is fixated on the objects of credence and preference, and lacks explicit room for the structure of credence and preference *per se*: it is designed around coordination on propositional objects for attitudes, and not for coordination on the probabilities or utilities to be assigned to those objects. And this makes for a bottleneck when our Bayesian agents wish to coordinate, not just on certain objects of credence or preference, but also on how likely to regard an outcome, or how much to prefer it.

Here is a simple example. Blaise doesn't know what to think about whether it is raining in Paris, and it matters for his plans. Thomas happens to have highish credence in the proposition that it is raining in Paris. As I would like to put it, Thomas is in a certain

*Various references still to be added. My aims here overlap in important ways with those of Rothschild [2012 forthcoming].

informational state assigning that proposition a highish probability. Thomas and Blaise are nearby, friendly, speak a common language, and other things are equal. Thomas wants to impart the relevant information his state of mind carries to Blaise. To accomplish this on the usual picture, Thomas effectively needs to find a way to pack his graded, non-binary state of mind into a nongraded, binary one. He must find a proposition to add to the black-and-white common ground he shares with Blaise, such that when Thomas proposes to add it to the common ground (asserts it), Blaise will get the idea that the thing to do, if Thomas is right, is to enter a certain gray doxastic state of mind and assign highish credence to the proposition it is raining in Paris. He must resort to this indirect path because on the usual picture, while one can express the *objects* of credence directly, one cannot directly give voice to the particular ways one apportions probabilities to those objects.[1] The standard framework is not designed to accommodate the possibility of "expressing" aspects of one's credal state besides their propositional content.

Instead of expressing his credal state directly, Thomas must therefore take other measures. For instance, he might elect to *describe* the relevant features of his state of mind, by finding and expressing a proposition about his credence. He might express the proposition that his confidence in a rainy Paris is within such-and-such bounds, or say something like 'I am quite confident it is raining in Paris'. Thereby a proposition about Thomas and his confidence level would enter the common ground. This, presumably, would pragmatically cue Blaise to adjust his credence in the relevant proposition (at least given appropriate background assumptions, such as that Blaise ought to take Thomas's confidence level on this kind of matter as a guide to his own). But the simple point here is that this is a quite indirect path. The attitude state Thomas meant to give voice to here—his highish credence in a rainy Paris—was one about the rain in Paris, not about himself. And ultimately his aim is to get Blaise to apportion probability in a certain way, not accept something about his confidence level. (Compare: if Thomas wanted to express his ordinary belief that Paris is the capital of France, he could simply say, in an appropriate context: 'Paris is the capital of France'. He would not have to resort to expressing a belief about his belief state, by saying 'I believe Paris is the capital of France'.)

Or again, suppose the question at issue concerns which of a class of actions it would be best for us to realize. We are deliberating how to spend the day, and the question arises whether to take off for Paris. It is common knowledge among us, suppose, that we slightly prefer spending the day in a sunny Paris to staying home, and would greatly prefer both to spending the day in a rainy Paris. Armed as I am with high confidence in rain, I want to make it clear to you that heading to Paris is not the thing to do—in Bayesian terms, not a way of maximizing expected utility. What I want to coordinate us on is a certain complex feature of a rational state of mind, a feature which is a function both of the probabilities I take to be assigned to outcomes as well as the utilities I take to be associated with them.

[1] A threshold view of assertion (according to which assertion, as a matter of convention, paradigmatically calls for, say, .9 credence, or some number fixed by context) would enable some probabilistic information to get through pragmatically, but it would still leave a substantial bottleneck. We can, anyway, stipulate that Thomas's credence in a rainy Paris is below the putative threshold.

What I want to get across is the idea that on all of the admissible ways of apportioning the probabilities and utilities, heading to Paris is not a relevantly utility-maximizing option. To accomplish this on usual picture, I again must take an indirect route. I need to find a way to pack this complex property of my state of mind into some kind of propositional package, so that it can be transmitted across our black-and-white common ground. I cannot give direct voice to my utilities, or to the structure of my preferences, in the way that I can give direct voice to the objects of my credence. To get aspects of my preference or utility state across, I am forced to resort to indirect measures, such as the strategy of asserting propositions about myself.

The resulting situation looks strange. To communicate with language, our Bayesian agents have adopted a practice which compels them to squeeze their decision-theoretically-structured states of mind into a binary, ungraded propositional medium for transmission. This is logically possible practice, but it is a curiously restrictive one.

Certainly it is possible for things to have been set up differently. A more direct practice might have been in place. Rather than a black-and-white common ground, our agents might have had a common ground allowing for probabilistic shades of gray. Or to switch metaphors: rather than a perfectly flat common ground, with all propositions on a par, our agents might have enjoyed a common ground admitting of probabilistic landscaping. That is to say that our agents might have had states of presupposition admitting of a representation akin to that of graded belief, or credence. This richer common ground would enable a wider array of possible rational conventional practices for changing or updating it. Among other things, it would make room for the idea that agents can coordinate on, or express, or assert (not just possible objects of credence, but) aspects of their credal states. Thomas, for instance, could express that aspect corresponding to his assigning highish credence to rain in Paris. His speech act move could be to propose to add this property to the probabilistically landscaped common ground.

The idea can be developed formally. If credal states are modeled probabilistically, a property of a credal state corresponds to a set of probability spaces: the set of probability spaces having the property. We can model a common ground as such a set: this would reflect the set of probability distributions compatible with what is mutually presupposed. Thomas now has a direct route for expressing the relevant property of his credal state of mind (assuming, that is, that his language takes advantage of the extra structure). He can perform a speech act tantamount to proposing to eliminate from the common ground those probability spaces lacking the relevant property—as it might be, again, the property of assigning highish probability to rain in Paris. If Blaise cooperates, their states of mutual presupposition would then become directly coordinated with respect to that feature of Thomas's doxastic state. No indirect measures required. Thomas could directly express his confidence, and without describing it.

We can see the possibility of saying something much the same for utilities. One can imagine a semantics and pragmatics for a language which allowed for direct expressions of aspects of one's utility distribution. We follow the same basic recipe. First, add further

structure to our representation of the conversational common ground—this time, by equipping the probability spaces left open in conversation with utility functions, the utilities compatible with what we are mutually presupposing. Then allow for speech acts whose characteristic conversational impact is to change the common ground by eliminating from it those ways of apportioning utility lacking the relevant expressed property.

I hope it seems uncontroversial that a language whose semantics and pragmatics worked in this way is at least logically possible. If not, what follows is meant to remove any doubt. Further, I want to suggest in this paper that we can be profitably modeled as actually speaking just such a language. When it comes to certain fragments of ordinary natural language, that is, we do best to adjust the received view of linguistic communication. We do best to operate with a richer, Bayesian conception of the conversational common ground, along the lines just described; and we do best to operate with a richer conception of how speech acts can serve to change the common ground. When we so operate, what we end up with is a view analogous in important ways to a certain recent strand of the expressivist tradition in metaethics—the strand largely due to Gibbard [1990, 2003]. What we end up with, I explain, is a kind Bayesian expressivism.

My focus will chiefly be on a fragment of modal language: the fragment containing probability operators (*probably*, *likely*, and the like), and containing the deontic modal *ought*. I will describe a semantics and pragmatics for a fragment containing these modals as well as epistemic modals and indicative conditionals, attitude verbs, negation and conjunction. The story about probability operators and indicatives is one I have told in fragments elsewhere (Yalcin [2007, 2010, 2011a]). The story about *ought* I sketch follows important work by Levinson [2003] and Lassiter [2011], among others. While I will eventually be interested here in the semantic analysis of some specific constructions and their empirical motivations, my larger concern is to engage with some of the central conceptual questions which arise within this sort of project, and with clarifying the sense in which the view can be called 'expressivist'.

2 Semantic value, content, context-change potential

Before setting out the semantics and pragmatics which will be our main concern, we do well to clarify at the outset some terminology and separate some issues. Distinguish three questions one might ask about any given declarative sentence of natural language:

(I) What is the **semantic value** of the sentence?

(II) What is the **content** of the sentence relative to a given context?

(III) What is the **context-change potential** of the sentence?

The questions each involve a different technical term. The first two terms will be familiar to philosophers; the third will be more familiar to linguists working in semantics. I will understand all of these terms as corresponding to functional role notions, so that it will be intelligible to debate for each one just what the realizer of the relevant role is. Let me say a bit about each of the three functional roles.

2.1 Semantic value

Start with 'semantic value'. Here is the kind of explanatory work I assume the realizers of the semantic value role are to perform:

(a) PRODUCTIVITY. The semantic values of expressions should play a central role in explaining *productivity* of language use: the empirical fact that competent speakers can understand and produce complex expressions that they have never before encountered.

(b) COMMUNICATION. The semantic values of sentences should play a central role in explaining how speakers *communicate*, and in particular, *transfer information*, by using the sentence.

(c) CONSEQUENCE. The semantic values of sentences should play a central role in explaining which sentences *follow* from which, and which sentences are *incompatible* with which.

Like most theorists, I mean to adopt the hypothesis that understanding a language requires knowing, in some theoretical sense, what the semantic values of the expressions of the language are; and I mean to adopt the hypothesis that in order to satisfy PRODUCTIVITY, we must assume that semantic values are compositional.[2] I take it this way of using the term overlaps substantially with what has typically been intended by 'semantic value' in the literature. This specification is of course just a rough beginning; in the context of further assumptions, the theoretical role may be further refined.

2.2 Content

The notion of content is more fraught and contested. Among those prepared to theorize with a notion of content, it is easy to find disagreement, not merely over what sort of entities play the content role, but about how best to carve the role out.

I elect to start with an understanding of the content role one can find in a certain familiar representationalist tradition within the philosophy of mind—partly because it is relatively congenial to a Bayesian prerogative. On the target conception, content (or 'propositional content') has two interrelated jobs to perform. First, content participates in an account of the representational character (intentionality, aboutness) of mental states—particularly, the states of belief and desire. Representation is, we take it, a matter of bearing content. The realizers of the content role are whatever they need to be in order to tell a satisfying story about mental representation. Second, the contents of mental states are assumed to be causal-explanatory properties of them vis-a-vis action. Particular hypotheses about the belief and desire contents of an agent generate *ceteris paribus* predictions about how the agent will be disposed to act in various circumstances. The realizers of the content role must serve to ground these explanations in an intelligible way. In particular, I assume the following as a rough necessary condition for being in a belief state with the content P:

[2] That is, we must assume that the semantic value of a complex expression is generally a function of the semantic values of its parts and its syntactic structure.

RATIONAL ACTION. To be in a belief state with content P is, at least, to be disposed to act in ways that would tend to satisfy one's desires in possible worlds where P, together with one's other beliefs, are true.[3]

This condition assumes that contents are the sorts of things which determine truth-conditions—some way the world might have gone. It also assumes a kind of holism about the explanatory or predictive import of content possession by belief and desire states. We presume further that a notion of content so grounded can be extended, somehow, to other apparently representational states of mind—states like imagination, wondering, and presupposition, for instance. We might hope to do this in part by elaborating the interconnections between the functional roles of belief and desire and these other states.

That supplies a very crude, but for our purposes serviceable enough, sense of what kind of work we expect the realizer of the content role to perform.[4] Adopting this conception of the content role is, of course, a nontrivial move. Note that this way of specifying the content role makes no explicit reference to language, or to syntactic structure, or to concepts, or to "what is said," and it makes no explicit appeal to any technical notions of semantics. Our notion of content is initially grounded outside of the formal semantics of natural language. This point is important. If some *sentences*, or sentences in context, are supposed to have content—if there is some content which is normally the 'object of assertion' when a sentence is tokened—we must extend the present notion, or give it some derivative sense.

Here is a simple, and I take it relatively traditional (if not uncontroversial) way of doing that. Assume as above that normal linguistic communication centrally involves coordinated states of presupposition. Like states of belief and desire, states of presupposition are states of mind bearing content. The coordinated states of presupposition among a group of interlocutors jointly determine a *common ground* for their conversation, as follows:

A content P is *common ground* among a group of interlocutors just in case everyone in the group presupposes that P, and it is common knowledge within the group that everyone in the group presupposes that P.

Following Stalnaker [1978], we take it that the mutually understood proximal rational aim of the speech act of assertion is to change the common ground of the conversation, by adding some content to the common ground. This idea requires the assumption that a declarative sentence tokened in a discourse context is mutually understood by competent speakers of the language to determine some item of content—if you like, the object of assertion. Supposing that assumption is generally correct, we then have a useful sense in which sentences in context can be said to have content. Briefly, the *content of a sentence in context* is that which is mutually understood by the speakers of the context to be the content which is to be added to the common ground, if an assertion of the sentence in the context is performed and accepted.

[3] Note this condition assumes a binary, rather than graded, conception of belief. The Bayesian analogue of this condition would advert to the idea that agents act so as to maximize utility. We will return to the matter below.

[4] Broadly in the spirit of, e.g., Ramsey [1931], Stalnaker [1984], Dretske [1988], Lewis [1994], Braddon-Mitchell and Jackson [2007], among others.

Suppose declarative sentences generally do have content in this sense. How do speakers coordinate on the content of a sentence? The natural idea is that interlocutors leverage their common knowledge of the compositional semantics of the language, together with any conventional pragmatic rules. One logically possible way this could work is as follows. Suppose semantic values are partly a function of context. Then the following might be the rule common knowledge of which enables competent speakers to recover the content expressed by a sentence ϕ in context:

Kaplan's Rule. The content of ϕ in $c = [\![\phi]\!]^c$.

That is, the content of a sentence ϕ in context is just identical to the semantic value of ϕ (i.e., $[\![\phi]\!]$), when the latter is evaluated relative to the context. We could conceive of this rule as a component of the pragmatics of the language, or as a rule at the semantics-pragmatics interface.

The relationship between semantic value and content articulated by Kaplan's Rule is often assumed without serious question in the philosophy of language.[5] Or so anyway it appears; it is not obvious how much of the appearance is due to the terminological morass surrounding 'content' (or 'proposition'). Some theorists, notably Kaplan [1977/1989], seem close to just using the word 'content' as just synonymous with 'semantic value in context'.[6] On such a use, the above rule would be true by stipulation. Be that as it may, it is important to emphasize for present purposes that given the way we are mooring the notion of content, the rule reflects a substantive empirical hypothesis about how the content of the sentence is fixed in context. Certainly, there is nothing logically inevitable about it.

Indeed, it is worth reviewing two ways it is not inevitable. First, consider for example a fictional tribe, the Negators (cf. the 'tribe of Liars' discussed by Lewis [1980]). The Negators speak a language equipped with the same interpretation function as English, but their rule for determining content in context looks roughly like this:

Negators' Rule. The content of ϕ in $c = [\![\neg\phi]\!]^c$.

The Negators are an odd bunch, no doubt; but the point is that their linguistic practice is perfectly possible, and it is one where Kaplan's Rule would not apply.

Second and perhaps more interestingly, Kaplan's Rule will be incorrect if the semantics of the language is such that the kind of object you get, when you evaluate the semantic value of a sentence relative to context, is just *not the kind of object which plays the content role*.

[5] Cappelen and Hawthorne [2009], MacFarlane [2009]... see Lewis [1980], Ninan [2010].

[6] Cappelen and Lepore [1997] read Kaplan as grounding the content role in ordinary intuitions about *what is said*: content is that which tracks intuitions about when what is said is the same or different. On this reading, Kaplan's Rule would be a hypothesis about how what is said relates to semantic value. This reading of Kaplan, though it has nontrivial textual support, conflicts with Kaplan's (correct) remarks to the effect that what operators a language contains (and hence what kind of object $[\![\phi]\!]^c$ is) is "largely a matter of language engineering" (504). It seems to me plausible that Kaplan simply did not have a very clear functional role for the notion of content in mind.

Let me illustrate. Suppose one comes to the view that the best realizers for the content role (as described above) are sets of metaphysically possible worlds (as in, say, Stalnaker [1984]). And suppose, after developing one's formal semantics for natural language, one comes to the view that the best realizers of the semantic value role, for sentences, are a certain kind of intension: namely, functions from contexts, worlds, and times to truth-values (as in, say, the formal system of Kaplan [1977/1989], *modulo* relativization to variable assignments).[7] Thus the interpretation function for the language would be relativized to contexts, worlds, and times. (Maybe this amount of structure is required in order to model the semantics of indexicals, modal operators, and temporal operators in the language; we needn't get into the details.) In this kind of case, Kaplan's Rule would not apply. For in such a case, $[\![\phi]\!]$, evaluated at a particular context, would be a function from a pair of a possible world and a time to a truth value. And this would just not be the same kind of thing as the kind of thing which (our theory of content told us) plays the content role: sets of possible worlds.

In such a case, we would have to posit a different rule for determining content in context. There are many logical possibilities. We could, for instance, have this rule:

> **Eternalizing Rule.** The content of ϕ in $c = [\![\phi]\!]^{c,t_c}$.

(Where 't_c' is the time of the context c.) Thereby we would evaluate the semantic value of the sentence with respect to a particular context and time, and what we would be left with is something only variable (or potentially variable) with respect to possible world. This, in any case, would at least leave us with something of the appropriate type to play the content role.

The larger point I wish to stress here is simply that the question of how the semantic value of a sentence in context relates to its content is a nontrivial, substantive question, at least if we see the notion of content as moored in the way I have described. Accordingly I do not wish to assume, at this point, a particular rule for recovering the content of a sentence in context from its semantic value. By now it should be clear that to assume a particular rule, I would have to take a stand on a variety of controversial issues in semantics, syntax, pragmatics, and the philosophy of mind. For now, I only wish to consider the hypothesis that for natural languages like English there is such a rule, that the rule is finitely specifiable, and that it forms a part of the presumed common knowledge among speakers of the language. We can think of the rule as of this form:

> The content of ϕ in $c = f([\![\phi]\!], c)$.

where f is assumed to be finitely specifiable.

2.3 Context-change potential

We come finally to what is, in certain ways, the most straightforward of the three notions. The context change potential (CCP) of a declarative sentence describes its characteristic impact

[7] I do not mean for the preceding to suggest that two theoretical projects need to occur in isolation; that is not the case.

upon any given common ground—its characteristic way of changing or updating what is being mutually presupposed. It is given formally as a function from (prior) common grounds to (posterior) common grounds.[8]

The preceding discussion of content, and in particular of the 'object of assertion'-role, already incorporates a simple idea about how any given sentence in context might be associated with a CCP. To give an illustration, suppose we model a common ground simply as a set of propositional contents. Write the CCP of ϕ in c as $[\phi_c]$, and (as is traditional) assume postfix notation for context-change potentials, so that arguments of functions are written to their left.[9] Then for any declarative sentence ϕ and common ground s in context c, the following would be a way of specifying the function $[\phi_c]$, given our recent discussion:

Stalnaker Update. $s[\phi_c] = s \cup \{\text{the content of } \phi \text{ in } c\}$

This says, along Stalnakerian lines, that the characteristic impact of assertion is to add the content of the sentence in context to the common ground.[10] This rule can be refined, obviously, once a particular rule for recovering content from semantic value is specified. If we assumed Kaplan's Rule, for example, we could restate the above as:

$$s[\phi] = s \cup \{[\![\phi]\!]^c\}$$

Thereby we would associate each sentence in context with its context-change potential.

The term 'context-change potential' originated within the dynamic tradition in semantics (Heim [1982, 1983]). In the dynamic tradition (or the relevant part thereof), the compositional semantic value of a sentence is *identified* with a context-change potential. Given the genesis of the term, one might think that in order to speak of the CCP of a sentence (in context), one must assume a dynamic semantics. But not so. One can speak with propriety about the CCP of a sentence without assuming that the compositional semantics for the language must take a dynamic form *per se*. The possibility of using a rule like Stalnaker Update to associate sentences with CCPs illustrates this simple point.

Having distinguished these three notions, we can go on to ask what sort of semantic values, contents, and context-change potentials are associated with the target modal fragment we are concerned with: talk of what is likely and of what ought to be the case. As my introduction hinted, I will work backwards. I will start by specifying the context-change potentials of these sentences, beginning with talk about what is likely: I will address the question how exactly such sentences characteristically change the common ground. In order to tell the story, we first need to upgrade our conception of the common ground in the manner informally described at the outset.

[8] The 'context' in 'context-change potential' adverts to the common ground—the informational context—of the conversation, not to the concrete discourse location.

[9] This enables us to use the left-right order to reflect the order of update. Thus $s[\phi][\psi]$ denotes the result of updating s with the CCP of ϕ, followed by the CCP of ψ.

[10] Stalnaker Update is a generalization of Stalnaker's actual technical proposal.

3 Context probabilism

We want to move beyond the idea of modeling the common ground of a conversation as a body of propositional content. We want to add Bayesian structure: probabilities and utilities. And we want to add this additional structure in as conservative a manner as possible, so that we needn't give some radically new story about language which doesn't obviously call for Bayesian resources.

Start with the probabilities. We need some notion of a probability space. The following will suit our purposes:

Def. Given a domain of worlds W, an **information state** i is a pair $\langle s, Pr\rangle$ of

1. a set s, $s \subseteq W$ (call it the **domain** of i) and
2. a probability function Pr on the elements of some Boolean algebra of subsets of W which is normalized on s (so that $Pr(s) = 1$.)

An information state is a probability space conditionalized on a primitively given domain of possibilities. (The term 'information state' reflects our intended modeling application.) Below this domain will be used to represent the possibilities regarded as open according to a conversation.

Now at this stage we have at least the following two options for going 'probabilist' about presupposition, and hence about the common ground. We can

(i) represent states of presupposition as information states; or

(ii) represent states of presupposition as sets of information states.

There is good reason to prefer the latter, especially if we retain the assumption that in the optimal case, interlocutors are coordinated with respect to their presuppositions. Suppose you ask me whether it is raining in Paris, and I say that it might be. It is now unambiguous that the possibility of rain in Paris is open for all we mutually presuppose. Exactly how likely is this possibility, for all we each presuppose? If our states of presupposition are probabilistically precise, we can't help but each associate specific numbers with this outcome. But that is excessive. Even if we each had precise credences on the matter, why suppose that our states of *presupposition* must be precise on the matter, especially given nothing explicit has been said about its exact likelihood?

Moreover, even if each of our presuppositional states associated specific numbers with the possibility of rain in Paris, then absent incredible luck, the numbers would differ. If they differed, then *a fortiori* we would seem not to be in coordinated states of presupposition—we seem not to be in the same presuppositional state. Our conversation would then be in violation of what is usually taken to be a standing norm of presupposition. It would be defective. Since cases like this are ubiquitous—normal communication rarely involves any

attempt to coordinate on specific probabilities for every mutually recognized live possibility—the model would thus seem to imply extensive failures of coordination in presuppositional states. But that is quite implausible.

Thus we have some initial strikes against the idea of representing a state of presupposition with a single probability space. (We will encounter a further difficulty in the next section.) If it is open whether it is raining in Paris and nothing has been presumed about just how likely that is, we should like to say that essentially any way of assigning this outcome a probability is compatible with what is mutually presupposed. This calls for sets of information states. Let a state of presupposition, and likewise a common ground, be representable as a set of information states. To have a name for it, call this thesis *context probabilism*:

> **Context probabilism.** The common ground of a conversation is representable as a set of probability spaces (information states).

(I have discussed this thesis in a few other places; see Yalcin [2005, 2007, 2011a].) We endorse context probabilism while still maintaining a relatively standard schematic characterization of when something is common ground, as follows (cf. [Stalnaker, 2002, 716]):

> It is common ground that ϕ in a group if all members presuppose that ϕ, and it is common knowledge in the group that all members presuppose that ϕ.

Note this schema is neutral on the 'objects' of the state of presupposition. In particular, it does not assume that the truth-conditions of sentences of the form 'A presupposes that ϕ' require that A stand in some binary presupposition relation to a propositional content. That is, of course, one standard direction of development, but we are interested in another.

4 Probabilistic context-change potentials

In particular, we are interested in capturing the following idea: to presuppose PROBABLY ϕ is to be in a state of presupposition associating the ϕ-outcome with a probability greater than .5. For example, to presuppose *that it is probably raining in Paris* is to be in a state of presupposition ruling out the option of associating the rain-in-Paris outcome with even-or-worse odds.[11] That is, one presupposes *that it is probably raining in Paris* if and only if the information states i left open by one's presuppositional state are all such that

$$Pr_i(\texttt{rainy Paris}) > .5$$

If this state of presupposition is construed as a relation to an 'object', the best candidate for that object would simply be the set of information states satisfying the above constraint—what we might call a *probability condition*.[12] Talk of 'objects of attitudes' should be used

[11] The assumption of a .5 cutoff point here is an idealization; the correct account of 'probably' is more complicated and context-sensitive. See Yalcin [2010] for discussion. The required subtleties are orthogonal to the present discussion.

[12] This object is much like what Jeffrey [1983b] calls a *probasition*, and what van Fraassen [1990] calls a *representor*.

with caution here, however, for we are not assuming that presupposition is a *sui generis* relation to such objects.[13]

We may now observe a further advantage gained by theorizing about presupposition with sets of information states. We should like for there to be a difference between (a) presupposing that it is not the case that an outcome is likely, and (b) not presupposing that the outcome is likely. Absent some incoherence, being in the former state entails being in the latter state, but not so the reverse; the former state is in some sense a more opinionated one. It is difficult to see how to mark this difference with just a single probability space. If your state of presupposition is given by a single information state, then if it fails to be a state according to which a given outcome is likely—say, by assigning the outcome .374 probability—*ipso facto* it is a state according to which it is not the case that the outcome is likely. By contrast, the additional structure afforded by sets of information states allows us to mark the distinction. If it is not the case that for all information states i that one's state of presupposition leaves open,

$$Pr_i(\texttt{rainy Paris}) > .5$$

then one counts as failing to presuppose that rain in Paris is likely. But this alone doesn't suffice to count as presupposing that it is not the case that rain in Paris is likely. For that, we require that for all information states i that one's state of presupposition leaves open,

$$Pr_i(\texttt{rainy Paris}) \leq .5$$

Thus by resort to sets of probability spaces, we can mark the apparently undeniable difference between 'not presupposing likely' and 'presupposing not likely.'

This conception of what it is for the sentence 'It is probably raining in Paris' to be common ground in context gives us a clear idea about what the context-change potential of 'It is probably raining in Paris' needs to do: we want it to take a common ground and output a new common ground which is such as to leave open only information states satisfying the relevant probability condition, and which otherwise leaves the context minimally changed. We can achieve this result very simply as follows: given a common ground I (a set of information states i), the CCP of 'It is probably raining in Paris' should map I to:

$$\{i \in I : Pr_i(\texttt{rainy Paris}) > .5\}$$

[13] An analogy: suppose we describe my height in centimeters as falling within some interval of numbers. Then this interval would constitute an 'object of my height in centimeters,' and my height could be glossed as a relation to this object. We can talk this way if we like, but it has little *prima facie* appeal, and we'd have to remember that these 'objects' are only being used to isolate a possible property of a state of height. (Compare Stalnaker [1984] describing the classic possible worlds model for belief: "Attitudes are primarily attitudes to possible states of the world and not to the propositions that distinguish between those states. A belief state can be represented as a set of possible worlds. Individual beliefs are properties of such a belief state: to believe that P is for the proposition that P to be true in all the possible worlds in the belief state" (69).

Thus the sentence takes a common ground and eliminates from it those information states not satisfying the probability condition determined by the sentence. Saying that it is probably raining in Paris is a way of winnowing down the possible ways of apportioning probability over the outcomes mutually recognized to be open in context. It adds a probability condition to the common ground.

I suggest that doing this—changing, or proposing to change, the probabilities open according to the common ground in this way—can be a way of expressing one's credence. It is a way of expressing one's credence where the mutually recognized aim of the conversation is one of belief or knowledge transference.

We can make this more precise. Say that 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. If the (or a) conversational tone of a conversational is belief, then performing a speech act tantamount to proposing to add a probability condition to the common ground is, we can say, a direct way of expressing one's credence. (If the conversational tone is different—if one is just telling a fairy tale, for instance—then performing a speech act tantamount to proposing to add a probability condition to the common ground is not plausibly construable as a case of expressing credence. The context-change potential we have associated with claims about what is likely is thus not constitutively tied to belief or credence.[14])

We noted that we wanted our context probabilism to be conservative with respect to language not obviously calling for Bayesian resources. A simple example of such language would be the unmodalized sentence 'It is raining in Paris'. Let us ask: what is it to presuppose this ordinary, non-probabilistic item in a context probabilist setting?

On this we can maintain a version of the classic picture from Stalnaker [1978]: to presuppose what an ordinary factual assertion says is to be in a presuppositional state which rules out those possibilities with respect to which what is said is not the case. To presuppose that it is raining in Paris is to be in a presuppositional state excluding possible worlds wherein it is not raining in Paris. Here is where we help ourselves to the assumption that the common ground is not merely a set of probability *measures*, but rather a sets of probability *spaces* each of which is equipped with a primitively given domain of possibilities. This allows us to say that to presuppose that it is raining in Paris is for the following to be the case: for all information states i that one's state of presupposition leaves open, it is raining in Paris in all worlds $w \in s_i$. Ordinary factual information—the sort that characterizes ways the world might have been, rather than ways of apportioning probability over possibilities—is in effect reflected by the sample spaces of the probability spaces left open by the common ground.

Likewise we can maintain the Stalnakerian conception of assertion in connection with ordinary factual talk. Stalnaker represents a common ground as a set of possible worlds (the worlds left open by what is presupposed), and models assertion as a sort of proposal to eliminate possibilities from the common ground. We can say something exactly analogous.

[14] A good thing, since we can imagine, suppose, or presuppose things about what is likely compatible with possessing any credal distribution. For further discussion see Yalcin [2007].

Given a common ground I, the CCP of 'It is raining in Paris' should map I to:

$$\{i \in I : \forall w \in s_i : w \in \texttt{rainy Paris}\}$$

An assertion of this sentence thus eliminates from the common ground any information state allowing for the possibility that it is not raining in Paris.

This enables us to see that we can have our context probabilist story about talk of what is probable or likely, without needing an entirely new account of the pragmatic impact of ordinary non-probabilistic, factual talk. The version of context probabilism just sketched can be viewed as a conservative extension of a familiar non-probabilistic model of the common ground and of assertion—Stalnaker's. When the modeling of Bayesian language is not at issue, we can if we wish ignore the probabilistic structure assumed by context probabilism altogether, and just speak instead of the possibilities left open or eliminated by the common ground.

5 Bayesian expressivism and content

5.1 Pulling apart two explanatory roles

I have explained what the context-change potential of sentences of the form PROBABLY ϕ generally look like for our Bayesian expressivist. Turn to the next question: what is the content of such sentences on the target view?

Above I noted that if we wish to talk about the 'object of assertion' for such sentences, the best candidate for that role would be a probability condition. This reflects what is added, in the relevant sense, to the common ground. Should we therefore take probability conditions to be contents? Should we say that the content of 'It's probably raining in Paris' is given by the set of information states assigning the rain in Paris outcome better-than-even odds?

We could use the word 'content' in this way, but I fear it would blur over distinctions we should like to make. Let me explain.

Compare the state of believing that it is probably raining in Paris with the state of believing that it is not likely that it is raining in Paris. According to the sort of textbook Bayesian picture I want to assume here, to be in the former state is to lend high enough credence—better-than-even odds—to the possibility of rain in Paris. To be in the latter state is to be in a state lending even-or-worse odds to that possibility. Now we can ask, do these two doxastic states differ in content?

Plausibly the answer is: they do and they don't. They do in as much as they are states of mind intentionally directed at the same possible state of affairs. The states do not differ in what they are *about*. They are representationally alike. Insofar as we wanted the notion of content to track purely representational similarities and differences between states, there is a strong pull to say that these states do not differ in content.

Yet plainly the states are different. Most obviously they correspond to different causal-explanatory properties of agents vis-a-vis action. Other things being equal, we can expect different dispositions to act from an agent who believes that it is probably raining in Paris

as compared to an agent believing that this is not likely. Insofar as we wanted the notion of content to track causal-explanatory differences vis-a-vis action between doxastic states, then we might wish to mark the distinctions between these states as distinctions of content.[15]

The situation seems to be this: the two jobs we have asked the realizers of the content role to perform are, in a Bayesian setting, better understood as being performed by different objects. The what-is-represented role of content is performed, in the Bayesian setting, by the outcomes (options, possibilities)—the possessors of probability and utility. But when we wish to point to features of belief states which correspond to characteristic dispositions to act (given assumptions about the agent's desires), the features we want to point to are properties of the agent's credence distribution. After all, if we know only that an agent has some doxastic attitude or other towards a certain outcome, but have no indication of how much credence they have in the outcome, then while we know what the relevant state of mind represents, we can predict little if anything about how the agent is disposed to act.[16]

5.2 Force and content in a Bayesian setting

I can imagine a source of resistance to the view I have just described, rooted in a certain classic way of understanding credence. Here is the way the understanding works. Start first with the ancient distinction between *attitude type* (*force, mode*) and *content*, a distinction which has its natural home in a binary conception of the attitudes. We explain the distinction by example: *believing it is raining* and *imagining it is raining* are states with different forces, but the same content; *believing it is raining* and *believing it is snowing* are states with the same force, but different contents. Now to extend this distinction to a Bayesian conception of belief, wherein binary belief gives way to varying degrees of credence, construe each possible degree of credence as corresponding to a new force or attitude type. Thus, for instance, having .55 confidence in some outcome and having .56 confidence in that outcome are states with different forces, but the same content. And now returning to our previous example, *believing that it is probably raining in Paris* and *believing that it is not likely that it is raining in Paris* would be understood as states with the same content, but differing forces. On this

[15] On this point, we should pause to note that the RATIONAL ACTION condition described above, which assumes a binary conception of belief, must be revised when we move to a Bayesian representation of agents. A standard way of capturing roughly the same idea in a Bayesian setting would be the following:

> BAYESIAN RATIONAL ACTION. Agents act so as to maximize expected utility, relative to their credences and utilities.

Observe that (unlike the earlier condition) there is no explicit mention of the notion of content *per se* here; there is only the presumption that agents are in states of mind that associate probabilities and utilities with outcomes or options. One could of course impose the notion of content onto the Bayesian model, using that word for some aspect of the structure; but the model, considered on its own, seems not to cry out for the additional vocabulary.

[16] This isn't to suggest that the answer to the question what outcome an agent's doxastic state is intentionally directed toward is just *irrelevant* to the agent's disposition to behave on a Bayesian picture; that is obviously false. All I am noting is that this piece of information *per se* is much less predictive vis-a-vis action than one might have hoped for realizer of the content role.

conception, to construe the probabilistic dimension of a belief state as an aspect of content is just to mix up the content of a state with its force.

It seems to me that this way of employing terminology is unhelpful at best. There is, after all, a completely straightforward sense in which *believing that it is probably raining in Paris* and *believing that it is not likely that it is raining in Paris* (or having .55 confidence in some outcome and having .56 confidence in that outcome, etc.) are states of exactly the same attitude type. They are doxastic states—as distinguished from states of preference, imagination, presupposition, and so on.

Stepping back, we can say that a Bayesian model of belief recognizes two joints in nature where the binary conception marks just one. On the binary conception, once you fix an attitude type of a state, there is just one further question: what is the content of the state? On a Bayesian model of belief, by contrast, there are two further questions: first, what possibility does the state concern? Second, what probability does the state associate with the outcome? If we shift to a Bayesian perspective, we should not continue assuming that there is just one joint here, a joint marked by the old force-content distinction. We should instead adapt our terminology to the nuance revealed by the model.

One might object that to describe a state of mind as associating a certain probability with an outcome *just is* to describe it as a belief state. Hence the questions 'What type of attitude is it?' and 'What probability does the state associate with the outcome?' do not come apart logically in the way that both come apart from the question 'What is the content of the state?' This brings us to a crucial point. Our Bayesian expressivist rejects the assumption that only states of belief may be probabilistically articulated. Indeed, it is constitutive of the view as we have described it that states of presupposition, for instance, admit of probabilistic structure. The expressivist approach I am in the process of recommending does *not* associate sentences of the form PROBABLY ϕ with a condition on credal states *per se*. Rather, the objective is to associate the sentence with a probability condition—a property of information states. Since various attitude types—belief, supposition, imagination, doubt, presupposition—may be modeled as information states (or sets thereof), the language of probability is not constitutively tied to the state of belief.

So let us adapt our terminology. I suggest we make the following threefold distinction among the characteristic properties of intentional mental states:

I. Attitude type

II. Structural character

III. Representational content

'Attitude type' corresponds most closely to the classical notion of force. 'Representational content' corresponds most closely to the classical notion of content. 'Structural character' adverts to certain properties states have in virtue of their particular formal structure *per se*. (For now we have in mind chiefly the probabilistic structure of information states, but this is meant to cover in a generic way any kind of structure one might posit in modeling broadly

representational states of mind. As advertised, eventually we will consider the addition of utility-theoretic structure.) We can illustrate the distinctions by explaining how the following three attitude states come apart in respect of these features:

(a) believing it is very unlikely to be raining

(b) believing it is probably snowing

(c) imagining that it is probably raining

The states (a) and (b) are the same in respect of attitude type, differing in that respect from (c). The states (a) and (c) are the same in respect of representational content, differing in this regard from (b). The states (b) and (c) are the same in respect of structural character—they are both states of associating high probability with an outcome—and they differ in this regard from (a), which is a state of associating a very low probability to an outcome.

The probability condition associated with an informational state of mind is a property the state has in virtue both of its structural character and its representational content. If we wish to speak more generally about differences between states owing *either* to structural character *or* representational content, we can speak generically of *informational* differences. Thus we can say that the state described by (a) is informationally different than the state described by (b). Or again, if one agent presupposes that it is probably raining and the other presupposes it is not raining, we can describe this as an informational difference between the states.[17]

5.3 Representational content and informational content

At the beginning of this section we suggested the two explanatory roles associated with our notion of content begin to come apart in the Bayesian setting: the roles seem realized, on the most straightforward understanding, by different objects. So be it: the underlying Bayesian reality is more complicated. Therefore we supplant our earlier jargon. Where we previously had a single notion of content, let us now have two notions: *representational content* and *informational content* (or just *information*). The former kind of thing determines a condition on possible worlds or situations, and reflects what the state being characterized is about. The latter kind of thing corresponds to a condition on information states—a probability condition. Varying assumptions about the informational content of a doxastic state, holding fixed desires, yields varying predictions for the agent's behavioral dispositions. At the most general, it is information which is the sort of thing that gets added to the common ground: if there are 'objects of assertion', these objects are informational contents.[18]

[17] This tracks, incidentally, the way of talking about information familiar from information theory. In that theory the amount of information possessed by a state is a function, not just of what possibilities one can eliminate, but also of what probabilities the state assigns to the open possibilities. Two states leaving open the same possibilities but distributing probabilities over them very differently are described as informationally different.

[18] As in effect noted above (p. 16), the notion of informational content can be seen as a generalization of the notion of representational content, inasmuch as any distinction between possible worlds can be mimicked by

6 Probabilistic compositional semantics

It remains to show how the context change potentials we have hypothesized for sentences about what is probable or likely can be compatible with compositional semantics for the probability operators *probably* and *likely*; and it remains to show that some such semantics is not merely possible, but a contender as an analysis of the relevant fragment of natural language. I have addressed these issues in some detail elsewhere (Yalcin [2007, 2010, 2011a,b]); what follows largely reviews relevant aspects of this work.

I will describe two possible semantic analyses of probability operators. One takes the form of a traditional intensional semantics. The other takes the form of a dynamic semantics. I will be neutral on which to prefer. My main motivation in describing two semantic accounts, rather than just one, is to clarify the relationship between accepting Bayesian expressivism and accepting a particular formal semantics. It is sometimes thought that expressivism about a fragment of discourse is a particular thesis about the formal semantics of the discourse. The expressivist is described as proposing an "expressivist semantics." As I eventually argue below, this conception of expressivism is mistaken. Expressivism, although it imposes demands on compositional semantics, is not itself a thesis of semantics. This point will be easier to see once we have two different expressivist-friendly semantic accounts on the table.

6.1 An intensional semantics for probability operators

Aspects of the following account are discussed in Yalcin [2007, 2010, 2011a,b] (see also Kolodny and MacFarlane [2010], MacFarlane [2011], Rothschild [2011a,b], Klinedinst and Rothschild [2011], Gillies [2010], Khoo [2011]).

Suppose it is correct that for assertions of sentences like 'It is probably raining in Paris', the relevant object of assertion—item added to the common ground—is a certain probability condition, as described above. This imposes a demand on the compositional semantics of the sentence. The demand is that the semantic value of the sentence be able to determine such a condition.

To repeat a point emphasized above: the demand is *not* that the compositional semantic value of the sentence be *identical* to this probability condition (informational content). That requirement would be needlessly strong. It would suffice for the semantic value of a sentence to merely *determine* its informational content. As long there is there a finitely specifiable function from the semantic value of a sentence to its informational content, we will be able to describe the update impact of the sentence (its CCP) in terms of the idea of adding that information to the common ground.

To get to the formal proposal, assume a simplified formal language:

> **Syntax of $\mathcal{L}$.** The primitive expressions include sentence letters $p_1, p_2, ...$; the one-place operators $\neg$, $\Diamond$ (epistemic *might*), $\triangle$ (*probably*); the two-place operator

a distinction between information states. So in making this claim we do not deny that there is an interesting sense in which ordinary factual assertions can be said to add representational content to the common ground.

$\wedge$; and parentheses. Sentence letters are well-formed. If ϕ and ψ are well-formed, so too are $\neg\phi$, $\Diamond\phi$, $\triangle\phi$, $(\phi \wedge \psi)$.

We take $\mathcal{L}$ to reflect the logical forms of the relevant fragment of English at the suitable level of abstraction. Our interest is in giving a model-theoretic intensional semantics for this fragment. Models will include a space of possible worlds, and a valuation function:

Defs. A **model** $\mathcal{M}$ for $\mathcal{L}$ is a pair $\langle \mathcal{W}, \mathcal{I} \rangle$ where $\mathcal{W}$ is a set of possible worlds, $\mathcal{I}$ is an **valuation function** mapping the propositional letters of $\mathcal{L}$ to subsets of $\mathcal{W}$.

With the valuation function, we abstract from subsentential compositional structure.

Any given model is assumed to determine a space of possible information states in the obvious way. We use the notions of a possible world and of an information state to define the notion of a *point of evaluation* in a model:

Def. An **index** in $\mathcal{M}$ is any world-information state pair $w_{\mathcal{M}}, i_{\mathcal{M}}$.

The semantics takes the form of a recursive characterization of truth at an index for sentences of $\mathcal{L}$, as follows:

Def. For any $\mathcal{M}$, an **interpretation** $[\![\ ]\!]^{w,i}$ for $\mathcal{M}$ is a function assigning either 0 or 1 to each wff relative to each index w, i in $\mathcal{M}$ subject to the following constraints, where α is any propositional letter, ϕ and ψ are any wffs (I surpress relativity to models):

$$
\begin{array}{rll}
[\![\alpha]\!]^{w,i} = 1 & \text{iff} & w \in \mathcal{I}(\alpha) \\
[\![\neg\phi]\!]^{w,i} = 1 & \text{iff} & [\![\phi]\!]^{w,i} \neq 1 \\
[\![\phi \wedge \psi]\!]^{w,i} = 1 & \text{iff} & [\![\phi]\!]^{w,i} = 1 \text{ and } [\![\psi]\!]^{w,i} = 1 \\
[\![\Diamond\phi]\!]^{w,i} = 1 & \text{iff} & s_i \cap [\![\phi]\!]^{i} \neq \emptyset \\
[\![\triangle\phi]\!]^{w,i} = 1 & \text{iff} & Pr_i([\![\phi]\!]^{i}) > .5
\end{array}
$$

(Where $[\![\phi]\!]^{i}$ is an abbreviation for $\{w : [\![\phi]\!]^{w,i} = 1\}$.) Note that negation and conjunction have their classical semantics.[19] Observe also we set aside context-relativity for the interpretation function here, as the fragment of language we are concerned with is not context-sensitive in the relevant sense.

To capture the probability condition (informational content) associated with any given sentence, we define the notion of acceptance:

Def. An information state i **accepts** ϕ iff $\forall w \in s_i : [\![\phi]\!]^{w,i} = 1$.

The informational content associated with a sentence is then just a matter of the states which accept it:

Def. The **informational content** of $\phi = \{i : i \text{ accepts } \phi\}$

[19] For discussion of disjunction, see Rothschild [2012 forthcoming].

Thus the semantic value of a sentence—given, in effect, by a set of world-information state pairs—is just a different kind of thing than its informational content—a set of information states.

Finally, we can define consequence in terms of preservation of informational content, as follows:

> Def. $\phi_1, ..., \phi_n \vDash \psi$ iff no information state which accepts $\phi_1, ..., \phi_n$ fails to accept ψ.

This suffices to illustrate the possibility of a semantics which determines the sort of informational contents needed for our Bayesian expressivism about talks of what is likely or probable.

Why think a semantics along such lines is attractive *qua* analysis of natural language? For some general motivation for the use of probability theory *per se* in the semantics of natural language probability operators, see Yalcin [2010]. The motivations I would to highlight here, however, emerge when our target fragment is extended to incorporate attitude verbs and conditionals.

6.2 Extension to informational attitudes

In the extension of the semantics to 'informational' or 'acceptance' or 'informational' attitudes such as supposition, imagination, belief, and presupposition, we face a choice point: we can model such states using a single information state, or we can model such states via sets of information states. We have already put our cards on the table as concerns presupposition: there we saw good reason to opt for sets. A uniform model of these attitudes would therefore call for each of them to be modelled via sets of information states. Fortunately, the motivatations we offered in the case of presupposition apply *mutatis mutandis* to these other states.[20] The generic semantics for these attitudes would then look as follows. Let there

[20] A number of theorists have investigated the use of sets of probabilities ("representors," "vague probabilities," "mushy credence," "probabisitions" etc.) to model rational belief states: see for instance Levi [1980], Jeffrey [1983a], van Fraassen [1990], Kaplan [1996], Joyce [2005], van Fraassen [2006], Sturgeon [2008]. (Elga [2010] raises an interesting worry for the idea that a rational belief state can fail to assign sharp probabilities. See Joyce [2010] for one reply. See also White [2010].)

From the point of view of our semantics, it is hard to see how to get by with less than something like sets of probability spaces. The challenge is exactly parallel to the one we faced with presupposition. We should like to be able to distinguish the truth-conditions of pairs like these:

(1) It is not the case that John believes it is probably raining.
$\neg B_x \triangle \phi$

(2) John believes that it is not the case that it is probably raining.
$B_x \neg \triangle \phi$

If, as seems *prima facie* possible, John's state of belief can be agnostic on the matter of the exactly likelihood of rain (something most, if not all, of the above cited theorists take as a datum), then he can be such as to make (1), but not (2), true. But if his belief state is given by a single (precise) probability space, then (unless we modified other aspects of the semantics) (1) and (2) would mutually entail each other, the incorrect result.

For further relevant discussion and motivation see Rothschild [2012 forthcoming].

correspond to each agent x in informational attitude A the set of information states left open by that attitude $\mathcal{A}_x$. Then the schematic form of the semantics for arbitrary informational attitude verbs can be given as follows:

$$[\![A_x\phi]\!]^{w,i} = 1 \quad \text{iff} \quad \forall i \in \mathcal{A}_x : i \text{ accepts } \phi$$

I briefly mention three empirical advantages of the foregoing. First, this semantics has the nice result that an ascription like:

(3) John believes that it is probably raining.

is a way of saying that John assigns high credence to the rain outcome—rather than a way of saying that John has a certain second-order belief, a belief about his credence.

Second, it explains the defect in sentences like:

(4) # John is imagining it isn't raining and that it probably is.

The reason is that given our semantics, sentences of the form $\neg\phi$ and $\triangle\phi$ are generally not jointly acceptable. (See Yalcin [2007] for detailed discussion.)

A third advantage of the apparatus is that it enables an attractive Bayesian semantics for 'doubt,' one according to which it means in essence the same as 'believes it is unlikely that':

$$[\![D_x\phi]\!]^{w,i} = 1 \quad \text{iff} \quad \forall i \in \mathcal{B}_x : Pr_i([\![\phi]\!]^i) < .5$$

Plausibly, to doubt that P is just to believe that P is unlikely; and very plausibly, one can doubt that P without having a belief concerning one's credence. We can have these results with the semantics above. It is not easy to see how to get them otherwise.

6.3 Extension to indicative conditionals

As with the epistemic modals and probability operators, many theorists have found it more fruitful to theorize about indicative conditionals by asking, not about the conditions under which they are true or false, but by asking what it is to be in a state of mind which accepts them.[21] And in addressing that question, it has become *de rigueur* to start with the 'Ramsey Test' (Ramsey [1931]):

> If two people are arguing 'If p will q?' and are both in doubt as to p, they are adding p hypothetically to their stock of knowledge and arguing on that basis about q; ... they are fixing their degrees of belief in q given p. (247)

[21] The affinity between epistemic modals and indicative conditionals is not surprising if Kratzer [1986] is right that if-clauses are devices for restricting (perhaps tacit) modal operators. See Rothschild [2011b] for an excellent recent discussion, and for a Kratzerian adaptation of a version of the conditional semantics to follow.

Ramsey's idea is usually associated with approaches to indicative conditionals which emphasize some very tight connection between indicative conditionals and conditional probabilities, usually at the expense of truth-conditions and of systematic compositional semantics (see for instance Adams [1975, 1965], Gibbard [1981], Edgington [1986, 1995], and more recently Bennett [2003]). But it is interesting to note that the classic truth-conditional modal account of Stalnaker [1968] was also motivated by appeal to Ramsey's thought. After agreeing that Ramsey had answered the question how to decide whether to *believe* a conditional statement, Stalnaker writes:

> ... the problem is to make the transition from belief conditions to truth conditions; that is, to find a set of truth conditions for statements having conditional form which explains why we use the method we do use to evaluate them. The concept of a *possible world* is just what we need to make this transition, since a possible world is the ontological analogue of a stock of hypothetical beliefs. (102)

Stalnaker then goes on to give his justly famous variably strict conditional account, using the notion of nearness between possible worlds.

The trouble with the above passage, however, is that the final sentence is not plausible: a possible world is not "the ontological analogue of a stock of hypothetical beliefs." At best a possible world might serve to model the belief state of an omniscient, or completely decided, agent. For such agents, however, a substantive question about whether to accept any given indicative conditional cannot really arise.[22] The doxastic states characteristically expressed by indicative conditionals are states of uncertainty—specifically, uncertainty about whether the antecedent is true.

Fortunately, we do not need to associate indicative conditionals with truth-conditions in Stalnaker's sense—with representation contents, to use our earlier jargon—in order to capture the Ramseyian idea Stalnaker is after within a compositional semantics. In place of Stalnaker's possible worlds, we put information states. The idea will be that $\phi \rightarrow \psi$ is accepted with respect to an information state i just in case, when you go to the nearest information state i where ϕ is accepted, ψ is accepted. And we define nearness using the probabilistic idea of conditionalization, as follows:

> Def. The **nearest information state to i accepting ϕ**, $i + \phi$, is defined as follows:
>
> $$i + \phi =_{\text{DEF}} \langle s_i \cap [\![\phi]\!]^i, Pr_i^{\phi} \rangle, \text{ where}$$
> $$Pr_i^{\phi}(x) =_{\text{DEF}} Pr_i(x|[\![\phi]\!]^i)$$

We then add the following clause to our recursive characterization of truth with respect to an index:

[22] Inasmuch as such agents will take the question as based on a mistake (if they believe the antecedent is false), or as just equivalent to the question whether the consequent is true.

$[\![\phi \rightarrow \psi]\!]^{w,i} = 1$ iff $i + \phi$ accepts ψ

Thereby we in effect compositionally associate indicative conditionals with probability conditions (informational content), conditions which do no reduce to possible worlds truth-conditions (representational content). And we do so in manner which explains the idea that one's confidence in what an indicative conditional says goes by one's credence in the consequent conditional on the antecedent. Given the appropriate sort of context, we can say that on this story, indicative conditionals are devices for expressing one's conditional credences.[23]

I briefly mention three empirical advantages of the foregoing account.

First, as already noted, it makes for a straightforward connection between indicative conditionals and conditional probabilities. That is an advantage, and not merely because Ramsey declared there was such a connection: there is not a small amount of psychological data which supports the thesis that there is some tight connection between indicative conditionals and conditional probabilities. (See among others Hadjichristidis et al. [2001], Oberauer and Wilhelm [2003], Evans et al. [2003], Over and Evans [2003], Over et al. [2007], Douven and Verbrugge [2010].)

Second, the semantics explains the defects in indicative conditionals like:

(5) If it is not raining and it is probably raining, ...

The reason is as before: sentences of the form $\neg\phi$ and $\triangle\phi$ are not jointly acceptable. (See again Yalcin [2007] for detailed discussion.)

Third, the semantics explains failures of Modus Tollens in connection with probability operators. I illustrate with an example discussed in detail in Yalcin [2011a]. Suppose an urn contains 100 marbles: a mix of blue and red, big and small. The distribution is as follows:

	blue	red
big	10	30
small	50	10

A marble is selected at random and placed under a cup. This is all the information given about the situation. Against this background, the following claims about the marble under the cup are licensed:

(P1) If the marble is big, then it's likely red.

(P2) The marble is not likely red.

However, from these, the following conclusion does not intuitively follow:

[23] It is sometimes thought, due in no small part to the triviality results of Lewis [1976], that one must choose between giving a compositional semantics for indicative conditionals and allowing for some systematic connection between indicative conditionals and the corresponding conditional probabilities. The thought is mistaken. I suspect it rests on a conflation of the notion of the semantic value of a sentence with the notion of representational content.

(C1) The marble is not big.

But this conclusion would follow, were Modus Tollens (MT) valid. So MT is not generally valid. So we require a theory of conditionals according to which MT is not generally valid. The semantics and definition of consequence already given is such an account: according to it, (P1) and (P2) do not entail (C1). The reason, simply put, is that an information state can accept (P1) and (P2)—it can satisfy their probability conditions, or incorporate their information content—without accepting (satisfying the probability conditions of, incorporating their information content of) the conclusion. Any information state i with the follow three properties will do the job:

$$Pr_i(\texttt{red}|\texttt{big}) > .5$$

$$Pr_i(\texttt{red}) \leq .5$$

$$s_i \cap \texttt{big} \neq \emptyset$$

We can add that our semantics will validate MT in the relevant modal-free cases. Thus if we replace (P1) with:

(P1′) If the marble is big, then it's red.

The resulting argument would indeed be valid according to our semantics, the intuitively correct result.

6.4 A dynamic implementation

Before I make some general remarks about how to interpret the preceding semantic apparatus, I want to give a second, dynamic implementation of the same basic ideas (one first given in appendix of Yalcin [2011a]). My point in doing this, again, is to make it clear that our Bayesian expressivism, while it imposes some abstract demands on semantics, is not wedded to a particular semantic framework.

In our intensional semantics, sentential semantic values are in effect sets of indices. The bridge from these semantic values to the context-change potentials of sentences was supplied as follows: first, we gave a rule for recovering a probability condition from a sentential semantic value; second, we gave a rule explaining what it is to add a probability condition to the common ground. These rules specified, if we like, core aspects of the semantics-pragmatics interface. By comparison, a typical dynamic semantics takes a more streamlined approach: the compositional semantic contribution of a sentence is literally identified with its context-change potential. As a result we require no additional pragmatic, or semantic-pragmatic, rules to enable us to understand what the characteristic impact of a sentence on the common ground is.

Strictly speaking, the semantics I wish to sketch now is an *atypical* dynamic semantics; it might be called 'quasi-dynamic'. Like a static semantics, it involves two stages *en route* to CCP; but unlike a static semantics, the first stage recursively associates sentences (not with

sets of indices, but) with functions from information states to information states. Once every sentence is associated with such a function, we then give a general rule for inducing from this an update function on sets of information states—that is, contexts in our sense. Thereby we associate all sentences with context change potentials.

To state the semantics, assume the same models as did our intensional semantics above. Then recursively associate sentences with functions from information states to information states, as follows:

> Def. A **dynamic valuation function** $[\cdot]$ is a function from wffs of $\mathcal{L}$ to functions from information states to information states subject to the following constraints, where α is any propositional letter, ϕ and ψ are any wffs:
>
> $i[\alpha] = \langle s_i \cap \mathcal{I}(\alpha), Pr_i(x|s_i \cap \mathcal{I}(\alpha))\rangle$
> $i[\neg\phi] = \langle s_i - s_{i[\phi]}, Pr_i(x|s_i - s_{i[\phi]})\rangle$
> $i[\phi \wedge \psi] = i[\phi][\psi]$
> $i[\Diamond\phi] = i$ iff $s_{i[\phi]} \neq 0$, else $\langle \emptyset, Pr(x|\emptyset)\rangle$
> $i[\triangle\phi] = i$ iff $Pr(s_{i[\phi]}) > .5$, else $\langle \emptyset, Pr(x|\emptyset)\rangle$
> $i[\phi \rightarrow \psi] = i$ iff $i[\phi] = i[\phi][\psi]$, else $\langle \emptyset, Pr(x|\emptyset)\rangle$
> $i[A_x\phi] = \langle \{w \in s_i : \forall i \in \mathcal{A}_x^w : i[\phi] = i\}, Pr_i(x|\{w \in s_i : \forall i \in \mathcal{A}_x^w : i[\phi] = i\})\rangle$
>
> Def. A **dynamic interpretation function** $[[\cdot]]$ is a function from wffs of $\mathcal{L}$ to functions from sets of information states to sets of information states defined in terms of $[\cdot]$, as follows: for any wff ϕ,
>
> $$I[[\phi]] = \{i \in I : i[\phi] = i\}$$

I leave it to the reader to verify that this semantics has the same empirical advantages adduced for the intensional semantics described above.

7 Why expressivism is not a semantic thesis

We have implemented the target expressivist idea that one can express aspects of ones credal state directly, without merely describing one's state of mind; we have supplied two distinct compositional semantic systems which can serve the purpose; and we have marshaled some empirical evidence for these systems taken as analyses of the relevant fragments of English.

The second point highlights the fact that our expressivism is not identical to a particular semantic theory. It is best understood as a claim about the characteristic context-change potentials of the target epistemic and probabilistic language. It is the view that the primary role of these sentences in discourse is not, or not merely, that of adding some representational content to the common ground—not a matter of characterizing the world as being some way rather than another. The impact on the common ground of a sentence like 'It is probably

raining in Paris' is better articulated by reference to the characteristic properties had by the state of mind which accepts the sentence.

This point bears repeating: expressivism about some fragment of language is best taken as a thesis about the context-change potentials of the relevant fragment. It is not, in any direct sense, a thesis about the semantic values of the sentences of the fragment. (This is, of course, compatible with a point we have already noted, namely that a view about the context-change potential of a given sentence constrains the space of options for the semantics of the sentence.)

Perhaps the only case where one can in principle 'read-off' of a semantics in isolation whether it is 'expressivist' is the case where one adopts a further substantive thesis, namely the dynamic thesis that the semantic value of a sentence just is its context-change potential. For then there is no gap between the CCP of a sentence and its compositional semantic value. But absent the dynamic thesis, it is not in general true that one can read off of a formal semantics whether it is expressivist in character.

Notably, the intensional compositional semantics described above could be accepted without modification by non-expressivists of various stripes. For instance, an off-the-shelf contextualist might fully accept our intensional compositional semantics, but reject our rules for moving from sentential semantic values to CCPs. She can give different rules instead. She can declare, for instance, that the context of utterance fixes some value for the information state parameter, call it i_c; and she can then say that all sentences ϕ determine ordinary representational contents, as follows:

$$\{w : [\![\phi]\!]^{w,i_c}\}$$

Sentences which express (nontrivial) probability conditions for the expressivist are, for this contextualist, sentences with ordinary truth-conditions (representational content)—in particular, they are sentences which serve to describe some body of information determined by context. This contextualist can say that it is always representational contents that are added to the common ground. Credence would then only be describable, not expressible. The contextualist can say this, again, despite agreement with our expressivist on all matters purely internal to compositional semantics.[24]

One might also accept the static compositional semantics we have given, but go for a radical form of relativism. See MacFarlane [2011] for such a view.[25]

For these reasons I have avoided describing the current project as one about giving an 'expressivist semantics'. Expressivism is not helpfully construed as a thesis in formal semantics. The same goes for contextualism and relativism, for same reasons. If the expressivist, contextualist, and relativist can in principle all *agree entirely* on the relevant compositional semantic values, their disagreement is not properly semantic.

The preceding may make it clear why I would not characterize expressivism as follows:

[24]The basic point here is discussed further in Yalcin [2007], Ninan [2010].

[25] The above semantics could also be accepted by the very different relativism of Egan [2007].

> The basic idea of expressivism is that it is the job of a semantic theory is to explain what a sentence, 'P', means, by saying what it is to think that P. [Schroeder, 2008, (704)]

The intensional semantic theory given above does not give the meaning of 'It's raining in Paris'—or even, for that matter, 'It's probably raining in Paris'—by "saying what it is think" these things. The situation is just more subtle. The job of a semantic theory, narrowly construed, is to supply semantic values which perform the semantic value role, as described above. Within our intensional semantics, these semantic values were certain sets of indices. These sets give the meaning, in the technical sense, of the relevant expressions. To effect such an association between sentences and sets of indices compositionally suffices to give a semantic theory. But it is not yet to explain what it is to be in the various states of mind the sentences can be used to express, and it is certainly not yet to endorse expressivism.

Expressivism might seem to offer a rival conception of semantic theory if one started with a conception of semantics roughly like this:

> The job of a semantic theory to explain what a sentence means by saying under what conditions it is true—that is, by associating the sentence with some representational content.

But this conception of the task of natural language semantics is simplistic, question-begging, and confused in the first place. It is simplistic in that it leaves unmentioned most of what theorists traditionally expect semantic theory to assist in *explaining*—for instance, lingusitic productivity, relations of entailment, the transfer of information, and so on. It is question-begging and confused because it conflates the *job* of a semantic theory with a substantive *proposal about how to carry out that job* (namely, by associating sentences with truth-conditions). And it is confused inasmuch as it conflates, at the outset of theorizing, the notions of semantic value and content.

8 Truth and objectivity

Expressivism about a fragment of discourse is sometimes described as committed to the view that the relevant sentences are not truth-apt, or that they lack truth-conditions. For example, Jackson and Pettit [1998] write:[26]

> Expressivism... holds, first, that ethical sentences lack truth conditions—they are not truth apt, truth-assessable etc.—and do not serve to report anything that the speaker believes to be so. (239)

Modulo the restriction to 'ethical' sentences, is our Bayesian expressivism correctly described in this way?

[26] Compare, among many others, Stoljar [1993], Jackson et al. [1994], Harman and Thomson [1996]...

Here we encounter another place where it pays dividends to distinguish clearly between semantic value and content. There is a notion of truth that can arise in connection with the realizers of the semantic value role, and there is a notion of truth that can arise in connection with the realizers of the content role. When we encounter statements like the quote above, we must step back and ask which sense of 'true' (if either) is at issue.

To expand: our intensional semantics availed itself of a technical notion, the notion of truth at an index. We used this notion to recursively associate sentences with sets of indices, as part of the larger project of capturing the structure of our semantic competence with the target fragment of language. The notion, again, is technical; is one cannot read off of the machinery set out above what the semantics of the ordinary word 'true' is (though some choices will of course seem more natural than others). Separate from this technical notion of truth at an index for sentences, we have the notion of the truth as it might apply to content—in particular, given the apparatus constructed so far, to representational contents, or to propositions as traditionally conceived. Such contents can be true or false. Representational contents are, or determine, a way the world might have gone; they are, or determine, a condition on possible worlds or situations. This notion of truth, too, is technical, at least insofar as our theory of content is stated formally.

This gives us at least two senses of 'true', and thus two senses of 'truth-conditions'. Our Bayesian expressivist, if he signs up for our intensional semantics, thereby signs up for the semantic notion of truth in connection with all declarative sentences: sentences have truth-conditions in the sense of determining a condition on indices in the relevant sort of model.[27] But as we have reviewed, our Bayesian expressivist denies that the sentences of the target fragment—talk of what is likely and unlikely, and affiliated epistemic modal and conditional talk—have representational contents. These sentences express probability conditions instead. A probability condition is not a way the world could have gone. The sentences of the target Bayesian fragment are not truth-apt in this sense.[28]

One might think this means our Bayesian expressivism will have trouble dealing with sentences such as:

(6) It is true that it is probably raining in Paris.

Not so. 'True' is a word in the object language, calling for a semantics like any other. So far we have said nothing on the matter, but if we like, we can take a stand. Here is a natural candidate semantics:

$$[\![\text{IT IS TRUE THAT } \phi]\!]^{w,i} = 1 \text{ iff } [\![\phi]\!]^{w,i} = 1$$

[27] Alternatively, of course, she might instead adopt the dynamic implementation, in which case she is not thereby committed to any such a notion of truth.

[28] Are probability conditions—informational contents—truth-apt? Probability conditions are a technical construct, so this seems to me not obviously a substantive question, but rather a question how to talk. I myself see little advantage to talking this way, and so pass. (Perhaps, however, if one assumed a notion of epistemic or evidential probability akin to that of Williamson [2000]—crudely, the idea that there is a particular conditional probability space reflecting the true evidential relations between propositions—one would have motivation for recognizing a substantive kind of truth-aptness for probability conditions.)

Business may now carry on as usual: 'It is true that ϕ'-ascriptions will inherit whatever nonfactualist, expressivist character infects ϕ. The sentence (6) will express a probability condition when tokened unembedded, just like the sentence it embeds.

The above semantics for 'It is true that' might appear in the spirit of a deflationary or minimalist conception of truth. Is it the case, according to such a semantic analysis, that

> There are sentences ... in which the word 'truth' seems to stand for something real ... [but] our analysis has shown that the word 'truth' does not stand for anything. [Ayer, 1946, 89]

The answer is: not really. In particular it should be clear that we are not in the deflationary territory of, say, Horwich [1998] or Field [2001]. We are entirely within the classic model-theoretic tradition in semantics; we are assuming an 'inflationary' notion of compositionality; we are allowing for a thick, representational notion of truth (though we deny it is straightfowardly expressed by the ordinary language truth predicate); we are assuming a fairly traditional representationalist position about mental states such as belief; and we are maintaining that ever so many sentences serve in discourse to represent the world. Our expressivism is thus not helpfully described as wedded to a deflationary conception of truth.

Finally, one might worry that if irreducibly probabilistic states of belief are not truth-apt, then we face some challenge explaining how they are 'objectively disciplined', or subject to some standard of correctness—a challenge of explaining why it is not the case that anything goes when it comes to credence. But the Bayesian has always had that challenge; indeed, *anyone* who thinks that belief may come in degrees has this challenge. The challenge does not arise in connection with Bayesian expressivism *per se*; so it is not particularly incumbent on the Bayesian expressivist to meet the challenge. For a leading attempt to meet the challenge in a Bayesian setting, see Joyce [1998, 2009].

9 Bayesian expressivism for deontic modality

Having come this far in the Bayesian direction, it is hard to resist going all the way, by adding (not just probabilistic structure to the common ground, but) also utility-theoretic structure. In this final section I briefly sketch a path here and some motivation, reserving extended discussion for elsewhere.

We may add utilities by adding structure to information states. A common ground can remain a set of information states, but information states now include further structure: they are *utility-laden*, supplying utility functions, and (hence) expected utility functions, defined over the relevant outcomes. Thereby we can give a semantics and pragmatics for (*inter alia*) deontic uses of modals such as *ought* in terms of the following idea: x ought to ϕ just in case x's ϕ-ing has an appreciably greater expected utility than the alternatives. Roughly as follows (I build here on Goble [1996], Levinson [2003], Kolodny and MacFarlane [2010], Lassiter [2011]):

$$[\![O\phi]\!]^{w,i} = 1 \text{ iff } e_i([\![\phi]\!]^i) \succ e_i([\![\neg\phi]\!]^i)$$

(Where e_i is the expectation function determined by i, and letting $\succ$ mean "appreciably greater than," allowing this notion to be context-sensitive.) As talk of what is probable changes the context by eliminating information states as a function of their probability measures, talk of what ought to be the case changes the context by eliminating information states as a function of their expected utility functions.

As noted by Lassiter [2011], a semantics for deontic modals employing expected utility along the lines above has at least the following three empirical advantages:

1. **Information-sensitivity.** It allows for the information-sensitivity of deontic modals, particularly as they occur in indicative conditionals. This feature of deontic modals is nicely illustrated by the miners case discussed by Kolodny and MacFarlane [2010]:

   > Ten miners are trapped either in shaft A or in shaft B, but we do not know which. Flood waters threaten to flood the shafts. We have enough sandbags to block one shaft, but not both. If we block one shaft, all the water will go into the other shaft, killing any miners inside it. If we block neither shaft, both shafts will fill halfway with water, and just one miner, the lowest in the shaft, will be killed.

   Against this background information, most are disposed to judge the following true:

   (7) We ought to block neither shaft.

   (8) If the miners are in shaft A, we ought to block shaft A.

   We can vindicate these judgments given the semantics so far described: blocking neither shaft is what maximizes expected utility under the appropriate uncertainty about the location of the miners; given certainty about miners location, however, it does not. On our semantics, an indicative conditional consequent is semantically evaluated relative to a shifted information state, one whose probability measure is conditionalized on the antecedent. When probability shifts, expected utility too shifts, as expected utility is a function of probability; and since what ought to be the case is a matter of expected utility, what ought to be the case shifts.

2. **Ross's paradox.** The following inference is obviously invalid:

   > You ought to mail the letter.
   > ∴ You ought to mail the letter or burn it.

   The invalidity of this pattern illustrates the fact that deontic *ought* is not upward monotonic—a traditional source of difficulty for deontic logics, particularly those treating deontic modals as quantifiers over some class of ideal worlds. But the concept of expected utility gets at what we want: if the expected utility of your mailing the letter is appreciably greater than your not mailing it, it does not follow that the expected utility of your mailing the letter or burning it is appreciably greater than your doing neither.

3. **Failure of distribution over conjunction.** Jackson supplied a counterexample to the idea that deontic *ought* distributes over conjunction, roughly as follows:

> Dr. Procrastinate has been asked to review a book, on a subject for which he is by far the most qualified to judge. If he accepts, he very probably will not write it, being the procrastinator he is. If he declines, someone less qualified will do it, and that person will write a worse review than the review Dr. Procrastinate would have written, had he both accepted and followed through.

In this case, it is not hard to hear the following as both true:

(9) Dr. Procrastinate ought not accept the invitation to write the review.

(10) Dr. Procrastinate ought to accept the invitation to write the review and follow through.

We can again accommodate this result using the notion of expected utility: the expected utility of Dr. Procrastinate's both accepting the invitation *and* writing the review can appreciably exceed that of outcome in which neither obtains, even if the expected utility of his not accepting, considered in isolation, does appreciably exceed his accepting.

With a semantics and pragmatics for deontic modals in place, we might turn to the question whether other fragments of normative or evaluative language merit treatment in Bayesian expressivist terms—for instance, talk of what is good, or right, or beautiful, or delicious. Can we construe any of this kind of talk as expressing, perhaps *inter alia*, constraints on preference, utility, or expected utility? The question is obviously beyond our narrow focus here, but my hope is that the preceding has rendered a positive answer to this question newly intelligible.

References

Ernest Adams. The logic of conditionals. *Inquiry*, 8(1):166–97, 1965. 22

Ernest Adams. *The Logic of Conditionals*. D. Reidel, 1975. 22

A.J. Ayer. *Language, Truth and Logic*. Dover, New York, NY, 2nd edition, 1946. 29

Jonathan Bennett. *A Philosophical Guide to Conditionals*. Oxford University Press, Oxford, 2003. 22

David Braddon-Mitchell and Frank Jackson. *The Philosophy of Mind and Cognition: An Introduction*. Blackwell Publishers, 2007. 2nd edition. 6

Herman Cappelen and John Hawthorne. *Relativism and Monadic Truth*. Oxford University Press, 2009. 7

Herman Cappelen and Ernest Lepore. On an alleged connection between indirect quotation and semantic theory. *Mind and Language*, 12(278-296):76, 1997. 7

Igor Douven and Sara Verbrugge. The Adams family. *Cognition*, 117(3):302–318, 2010. 23

Fred Dretske. *Explaining Behavior: Reasons in a World of Causes*. MIT Press, Cambridge, MA, 1988. 6

Dorothy Edgington. Do conditionals have truth-conditions? *Critica*, XVIII(52):3–30, 1986. 22

Dorothy Edgington. On conditionals. *Mind*, 104(414):235–329, 1995. 22

Andy Egan. Epistemic modals, relativism and assertion. *Philosophical Studies*, 133(1):1–22, 2007. 26

Adam Elga. Subjective probabilities should be sharp. *Philosophers Imprint*, 10, 2010. 20

J.S.B.T. Evans, S.J. Handley, and D.E. Over. Conditionals and conditional probability. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 29(2):321, 2003. 23

Hartry Field. *Truth and the Absence of Fact*. Oxford University Press, USA, 2001. 29

Allan Gibbard. *Thinking How to Live*. Harvard University Press, Cambridge, 2003. 4

Allan Gibbard. *Wise Choices, Apt Feelings*. Harvard University Press, Cambridge, 1990. 4

Allan Gibbard. Two recent theories of conditionals. In W. Harper, R. Stalnaker, and G. Pearce, editors, *Ifs: Conditionals, belief, decision, chance, and time*, pages 211–47. D. Reidel, 1981. 22

Anthony Gillies. Iffiness. *Semantics and Pragmatics*, 3:1–42, 2010. ISSN 1937-8912. 18

Lou Goble. Utilitarian deontic logic. *Philosophical Studies*, 82(3):317–357, 1996. 29

C. Hadjichristidis, R.J. Stevenson, D.E. Over, S.A. Sloman, J. Evans, and A. Feeney. On the evaluation of if p then q conditionals. In *Proceedings of the Twenty-Third Annual Conference of the Cognitive Science Society*, pages 409–414, 2001. 23

Gilbert Harman and Judith J. Thomson. *Moral Relativism and Moral Objectivity*. Wiley-Blackwell, 1996. 27

Irene Heim. *The Semantics of Definite and Indefinite Noun Phrases*. PhD thesis, University of Massachusetts, 1982. 9

Irene Heim. On the projection problem for presuppositions. In M. Barlow, D. Flickinger, and M. Wescoat, editors, *WCCFL 2: Second Annual West Coat Conference on Formal Linguistics*, pages 114–25. Stanford University, 1983. 9

Paul Horwich. *Truth*. Clarendon Press, 2nd edition, 1998. 29

Frank Jackson and Philip Pettit. A problem for expressivism. *Analysis*, 58(4):239–251, 1998. 27

Frank Jackson, Graham Oppy, and Michael Smith. Minimalism and truth aptness. *Mind*, 103(411):287–302, 1994. 27

Richard Jeffrey. Bayesianism with a human face. In John Earman, editor, *Testing Scientific Theories*, volume 10 of *Minnesota Studies in the Philosophy of Science*, pages 133–156. 1983a. 20

Richard Jeffrey. Bayesianism with a human face. In J. Earman, editor, *Testing Scientific Theories*, volume X, pages 133–56. Minnesota Studies in the Philosophy of Science, 1983b. 11

James Joyce. A nonpragmatic vindication of probabilism. *Philosophy of Science*, pages 575–603, 1998. 29

James Joyce. How probabilities reflect evidence. *Philosophical Perspectives*, 19(1):153–178, 2005. 20

James Joyce. Accuracy and coherence: Prospects for an alethic epistemology of partial belief. *Degrees of belief*, pages 263–297, 2009. 29

James Joyce. A defense of imprecise credences in inference and decision making. *Philosophical Perspectives*, 24(1):281–323, 2010. 20

David Kaplan. Demonstratives. In J. Almog, J. Perry, and H. Wettstein, editors, *Themes from Kaplan*, pages 481–563. Oxford University Press, 1977/1989. 7, 8

Mark Kaplan. *Decision Theory as Philosophy*. Cambridge University Press, 1996. 20

J. Khoo. Operators or restrictors? a reply to gillies. *Semantics and Pragmatics*, 4(4):1–43, 2011. 18

Nathan Klinedinst and Daniel Rothschild. Connectives without truth tables. *Natural Language Semantics*, forthcoming, 2011. 18

Niko Kolodny and John MacFarlane. Ifs and oughts. *Journal of Philosophy*, CVII(3):115–43, March 2010. 18, 29, 30

Angelika Kratzer. Conditionals. In A. Farley, P. Farley, and K.-E. McCullough, editors, *Papers from the Parasession on Pragmatics and Grammatical Theory*. Chicago Linguistics Society, 1986. 21

Daniel Lassiter. *Measurement and Modality: The Scalar Basis of Modal Semantics*. PhD thesis, PhD dissertation, New York University, 2011. 4, 29, 30

Isaac Levi. *The Enterprise of Knowledge*. The MIT Press, 1980. 20

D. Levinson. Probabilistic model-theoretic semantics for'want'. In R. Young and Y. Zhou, editors, *Proceedings of SALT XIII*, volume 13, pages 222–239, 2003. 4, 29

David K. Lewis. Probabilities of conditionals and conditional probabilities. *Philosophical Review*, 85:297–315, 1976. 23

David K. Lewis. Index, context, and content. In S. Kanger and S. Ohman, editors, *Philosophy and Grammar*, pages 79–100. D. Reidel, Holland, 1980. 7

David K. Lewis. Reduction of mind. In S. Guttenplan, editor, *A Companion to the Philosophy of Mind*. Blackwell, 1994. 6

John MacFarlane. Epistemic modals are assessment-sensitive. In Andy Egan and Brian Weatherson, editors, *Epistemic Modality*, pages 144–78. Oxford University Press, 2011. 18, 26

John MacFarlane. Nonindexical contextualism. *Synthese*, 166(2):231–250, 2009. 7

Dilip Ninan. Semantics and the objects of assertion. *Linguistics and Philosophy*, pages 1–26, 2010. 7, 26

K. Oberauer and O. Wilhelm. The meaning(s) of conditionals: Conditional probabilities, mental models, and personal utilities. *Journal of Experimental Psychology*, 29(4):680–693, 2003. 23

D.E. Over and J.S.B.T. Evans. The probability of conditionals: The psychological evidence. *Mind & language*, 18(4):340–358, 2003. 23

D.E. Over, C. Hadjichristidis, J.S.B.T. Evans, S.J. Handley, and S.A. Sloman. The probability of causal conditionals. *Cognitive Psychology*, 54(1):62–97, 2007. 23

Frank Ramsey. General propositions and causality. In *The Foundations of Mathematics and Other Logical Essays*. Routledge and Kegan Paul, London, 1931. 6, 21

Daniel Rothschild. Do indicative conditionals express propositions? *Nous*, forthcoming, 2011a. 18

Daniel Rothschild. A note on conditionals and restrictors. 2011b. 18, 21

Daniel Rothschild. Expressing credences. In *Proceedings of the Aristotelian Society*. 2012 forthcoming. 1, 19, 20

Mark Schroeder. What is the frege-geach problem? *Philosophy Compass*, 3(4):703–720, 2008. 27

Robert Stalnaker. A theory of conditionals. In N. Rescher, editor, *Studies in Logical Theory*, pages 98–112. Basil Blackwell, Oxford, 1968. 22

Robert Stalnaker. Assertion. In P. Cole, editor, *Syntax and Semantics 9: Pragmatics*. Academic Press, 1978. 6, 13

Robert Stalnaker. Common ground. *Linguistics and Philosophy*, 25:701–21, 2002. 11

Robert Stalnaker. *Inquiry*. MIT Press, Cambridge, 1984. 6, 8, 12

D. Stoljar. Emotivism and truth conditions. *Philosophical Studies*, 70(1):81–101, 1993. 27

S. Sturgeon. Reason and the grain of belief. *Nous*, 42(1):139–165, 2008. 20

Bas van Fraassen. Vague expectation value loss. *Philosophical studies*, 127(3):483–491, 2006. 20

Bas van Fraassen. Figures in a probability landscape. In M. Dunn and A. Gupta, editors, *Truth or Consequences*, pages 345–56. Kluwer, Dordrecht, 1990. 11, 20

Roger White. Evidential symmetry and mushy credence. *Oxford Studies in Epistemology*, 3:161, 2010. 20

Timothy Williamson. *Knowledge and its Limits*. Oxford University Press, Oxford, 2000. 28

Seth Yalcin. Epistemic modals. In Bernard Nickel Seth Yalcin Jon Gajewski, Valentine Hacquard, editor, *New Work on Modality*, volume 51, pages 231–72. MIT Working Papers in Linguistics, Cambridge, MA, 2005. 11

Seth Yalcin. Epistemic modals. *Mind*, 116(464):983–1026, Nov. 2007. 4, 11, 13, 18, 21, 23, 26

Seth Yalcin. Probability operators. *Philosophy Compass*, 5(11):916–37, 2010. 4, 11, 18, 20

Seth Yalcin. A counterexample to Modus Tollens. *Journal of Philosophical Logic*, forthcoming, 2011a. 4, 11, 18, 23, 24

Seth Yalcin. Nonfactualism about epistemic modality. In Andy Egan and Brian Weatherson, editors, *Epistemic Modality*, pages 295–332. Oxford University Press, 2011b. 18