

Frege's Puzzle and the Objects of Credence

David J. Chalmers

Abstract

The objects of credence are the entities to which credences are assigned for the purposes of a successful theory of credence. I use cases akin to Frege's puzzle to argue against referentialism about credence: the view that objects of credence are determined by the objects and properties that one's credence is directed at. I go on to develop a nonreferential account of the objects of credence in terms of sets of epistemically possible scenarios.

1 Introduction

What are the objects of belief? That is, what are the things we believe, when we believe that it is sunny outside and that Nietzsche is dead? Usually these things are taken to be propositions: the proposition that Nietzsche is dead, for example. But the nature of propositions is itself contested. What is a proposition, such that it can serve as an object of belief?

A popular class of views, strongly influenced by direct reference theories in the philosophy of language, involves *referentialism* about belief. Loosely speaking, referentialism about belief says that in so far as beliefs attribute properties to individuals (e.g., the belief that Nietzsche is dead), the objects of these beliefs are wholly determined by those individuals (e.g., Nietzsche) and those properties (e.g., the property of being dead). On one version of referentialism, the objects of

belief are Russellian propositions composed from the individuals and properties that one's belief is about. On another version, the objects of belief are sets of possible worlds in which the individuals in question have the relevant properties.

Referentialism about belief can be understood to have at least the following more specific commitments. If A and B are two names for the same object, then in believing that A has ϕ and in believing that B has ϕ (e.g., in believing that Hesperus is a planet and in believing that Phosphorus is a planet), one believes the same proposition. Likewise, in sincerely asserting ' A has ϕ ' ('Hesperus is a planet') and in sincerely asserting ' B has ϕ ' ('Phosphorus is a planet'), one expresses belief in the same proposition.

The debate over referentialism has approached something of a stalemate in recent years. Referentialism has simplicity and power on its side, weighed against intuitive costs. Perhaps the most important cost stems from versions of Frege's puzzle. For example, many find it hard to accept that a belief that Mark Twain is a writer and a belief that Samuel Clemens is a writer are beliefs in the same proposition. This has led to an extensive debate over the semantics of belief ascriptions, centering on the analysis of ascriptions such as 'John believes that Samuel Clemens is a writer'. Referentialist analyses of these ascriptions are much simpler than nonreferentialist analyses, but the most popular referentialist analyses still have intuitive costs that many see as unacceptable, and progress has begun to stall.

In this paper I approach the objects of belief from a different starting point. I start not from the philosophy of language but from the philosophy of probability. In particular, I start with one of our most successful theories in which objects of belief play a central role: namely, Bayesian confirmation theory. Bayesian theory is cast in terms of degrees of belief, or credences. And

credences are always credences in certain objects. Since credences are degrees of belief, it is natural to hold that the objects of credence are the same as the objects of belief. So one can approach the question of what the objects of belief are by asking: what sort of objects can play the role that Bayesian confirmation theory needs objects of credence to play?

These issues are of independent importance in understanding the interpretation of probability theory. Although much attention has been given to the interpretation of probabilities, much less attention has been given to the interpretation of the *objects* of probability: that is, the entities to which probabilities are assigned. It is now widely accepted that different sorts of probability are at play in the domains of objective chance and of subjective credence, but the status of the objects of probability in these domains is unclear. In this paper I focus especially on the objects of credence, and toward the end I discuss their relation to the objects of chance.

In the first half of what follows (sections 2 through 8), I argue that if the objects of belief and credence are understood as above, referentialism about belief is false. In order for objects of belief to do the work that Bayesianism needs them to do, a non-referential view of objects of belief is required. This is to say that in a certain sense, Bayesianism is inconsistent with referentialism. As Bayesianism is an extremely successful theory, this can be considered significant evidence against referentialism.

In the second half of what follows (sections 9 through 14), I develop a nonreferential account of the objects of credence that avoids the problems for referential accounts, using tools from two-dimensional semantics. Here the key idea is that objects of credence correspond to sets of epistemically possible scenarios, rather than to sets of metaphysically possible worlds.

My arguments in the first half involve probabilistic analogs of Frege's puzzle, centering on

cases in which a subject has two different ways of thinking about the same object, with apparent differences in associated credences. Given the lengthy discussion of Frege's puzzle in the philosophy of language, one might wonder whether the probabilistic case has anything new to add. I think that although the underlying issues are closely related, Bayesianism provides an additional ingredient: a formal normative theory in which credences play a central role. In the context of this sort of theory, old issues appear in a new light, and old moves on behalf of the referentialist are much less comfortable in the new setting. Correspondingly, the opponent of referentialism need not rely simply on an appeal to intuition, but instead can appeal to theoretical explanatory power. An additional benefit is that the probabilistic approach allows us to largely set issues about the semantics of 'that'-clauses to one side, and thereby allows us to approach referentialism about mental states in a way that is not as deeply entangled with referentialism about language.

I take Bayesianism to be committed to the following claims. Individuals have unconditional and conditional credences in propositions: that is, for every individual at a given time, there is a credence function p that maps (at least some) propositions H to an unconditional credence $p(H)$, and that maps (at least some) pairs of propositions (H, E) to conditional credence $p(H|E)$. When a hypothesis H and an evidential proposition E are such that one has prior conditional credence $p(H|E)$ in H given E , and if one's total relevant new evidence is specified by E , one should revise one's credence in H so that one's posterior credence $p'(H) = p(H|E)$. Bayesianism is also usually taken to involve the claim that one's credences should conform to the standard axioms of probability theory, but for my purposes this claim plays a less essential role.

I take it as a stipulation about 'object of credence' that objects of credence are those objects in which individuals have credences: that is, they are the objects that an individual's credence

function at a time maps to credences, for the purposes of a successful theory of credence. I also call these entities *propositions*, following standard practice, but nothing of substance rests on this usage.

We can then define *referentialism about credence* in parallel to referentialism about belief. Referentialism about credence is the view that when one has a certain credence concerning some objects and/or properties (say, a high credence that Nietzsche is dead), the corresponding object of credence is determined by those objects and properties. I take referentialism about credence to be committed to at least the following claims, which parallel the commitments of referentialism about belief outlined earlier. If *A* and *B* are two names for the same object, then in having a certain credence that *A* has ϕ and in having a certain credence that *B* has ϕ , the corresponding objects of credence are the same. Likewise, when one sincerely asserts '*A* has ϕ ' and '*B* has ϕ ', one expresses high credence in the same object of credence.

In the first instance, I am arguing against referentialism about credence (which I just call 'referentialism' in what follows). This way, the issues can be cast wholly in terms of probability theory, allowing us to set to aside other issues that arise with referentialism about belief and language (I return to some of these issues later in the paper). For example, I do not take it as a stipulation that objects of credence are referents of 'that'-clauses, so my arguments do not disprove the claim that the referents of 'that'-clauses are referential. Still, I take it that if referentialism about credence is false, this is a significant strike against the referentialist program in the philosophy of mind.

2 The Jekyll/Hyde case

My central argument uses a Frege-style case involving Jekyll and Hyde, with two stages.

Stage One: Olivia is conducting clinical research on the genetic basis of a hereditary disease. She is especially interested in two genetic factors, which we can call the alpha gene and the beta gene. It is known that 10% of people with the alpha gene have the disease, and that 20% of the people with the beta gene have the disease, that the two genes are independent, that each is had by 1% of the population, and that 90% of people with both the alpha and beta gene have the disease. People who have the disease show no outward sign of it until death, which always occurs after age 60. She is surveying fifty individuals under 60 (two per day) randomly selected from the telephone book for their occurrence of both genes.

In the morning, she interviews Dr. Jekyll, and takes a swab for his genetic material. After he leaves, she examines the alpha gene locus, and finds that that Jekyll has the alpha gene. In the evening, she interviews Mr. Hyde, and takes another swab. After he leaves, she examines the beta gene locus, and finds that Hyde has the beta gene. Although she has no evidence for this, Jekyll and Hyde are the same person, a master of disguise leading a double life.

Let JA , JB , and JD be the propositions that Jekyll has alpha, beta, and the disease respectively, and let HA , HB , and HD be the corresponding propositions about Hyde. During the morning examination, Olivia acquires evidence JA . Afterward, her unconditional credence $p(JD)$ is 0.1, and her conditional credence $p(JD|JB)$ is 0.9. During the evening examination, Olivia acquires total relevant evidence HB . What should Olivia's credence $p'(HD)$ be now?

If referentialism is true, HD and JD are the same proposition, and so are HB and JB . It follows from Bayesian principles that $p'(HD) = p'(JD) = p(JD|HB) = p(JD|JB) = 0.9$. That

is, Olivia should update her credences so that she has credence 0.9 in the hypothesis that Hyde has the disease. But this seems incorrect. Olivia is in no position to raise her credences, despite being fully rational, as she has no evidence that Hyde and Jekyll are the same person. Instead, Olivia's credence $p'(HD)$ should be 0.2. So referentialism combined with Bayesianism yields a false prediction here. So either referentialism or Bayesianism is false.

Stage Two: Months later, Olivia is told by a reliable source that on exactly one of the 25 days of her study, the subject in the morning and evening were the same person. Each pair of subjects looked and behaved about as similarly to each other as the other pairs. (One could also run a version of the case where each pair is perceptually indistinguishable to Olivia.) So for any given pair, Olivia has credence $1/25$ (0.04) that they are the identical pair. Later, she examines the unexamined loci for Jekyll and Hyde, and discovers that Jekyll has the beta gene. Only 1% of subjects with the alpha gene have the beta gene, so the chance of this discovery if Jekyll and Hyde were distinct would be one in a hundred. So Olivia is now much more confident that Jekyll is Hyde.

Given Bayesianism, it is natural to analyse this situation as follows. Let JH be the proposition that Jekyll is Hyde, and let E be the proposition JB . Just before examining the unexamined loci, Olivia's credences are $p(JH) = 0.04$, $p(E|JH) = 1$, $p(E|\neg JH) = 0.01$, $p(E) = p(E|JH)p(JH) + p(E|\neg JH)p(\neg JH) = 0.04 + 0.01 * 0.99 =$ (approximately) 0.05. When examining the unexamined loci, Olivia acquires total relevant evidence E . It follows from Bayesian principles that the resulting credence $p'(JH) = p(JH|E) = p(E|JH)p(JH)/p(E) =$ (approximately) $0.04/0.05 = 0.8$. So Olivia's new credence that Jekyll is Hyde should be around 0.8.

This appears to be a perfectly standard application of Bayesian reasoning. It appears to yield

both a reasonable normative account of a process that Olivia should have gone through, and a reasonable descriptive account of a process that Olivia might in fact have gone through, assuming that she was rational. One could apply broadly similar treatment to many other hypotheses about the identity or distinctness of individuals or of kinds, in science and elsewhere: the hypotheses that Hesperus is Phosphorus, that water is H₂O, that Twain is Clemens, and so on.

If this application of Bayesian principles is correct, however, then referentialism is incorrect. Assume for *reductio* that referentialism is true. Then JH , the proposition that Jekyll is Hyde, is identical to JJ , the proposition that Jekyll is Jekyll. It is natural to hold that $p(JJ) = 1$. (The alternative of holding that $p(JJ) = 0.04$ leads to problems that are similar but worse. I discuss the alternative of holding that there can be multiple credences in the same proposition under option 3 below.) If so, $p(JH) = 1$. It follows that $p'(JH) = 1$, so that Olivia's new credence in JH should be 1. So the analysis in the previous paragraph is incorrect. Contrapositively: if this application of Bayesian principles is correct, then referentialism is false.

3 Analysis

Stage one and stage two can be seen as complementary. Both center on the status of certain normative claims about credences, and in particular on claims about the value of $p_1(HD)$ and $p_2(JH)$, the credences that Olivia should have in the proposition that Hyde has the disease (after stage one) and the proposition that Jekyll is Hyde (after stage two).¹ I have suggested that the referentialist Bayesian is committed to:

¹Here $p_1(HD)$ and $p_2(JH)$ correspond to $p'(HD)$ and $p'(JH)$ in the cases above. The subscripts serve to distinguish the two different posterior probability functions after the two stages.

$$(R1) p_1(HD) = 0.9$$

$$(R2) p_2(JH) = 1$$

while the more reasonable Bayesian verdicts are

$$(B1) p_1(HD) = 0.2$$

$$(B2) p_2(JH) = 0.8.$$

I think that stage one has somewhat more force than stage two, and the comparison brings out its distinctive force. In stage two, the prediction (R2) is counterintuitive but perhaps defensible. A referentialist might respond in a familiar way, saying that the analysis yielding (B2) begs the question against referentialism, by effectively assuming from the start that *JH* and *JJ* are distinct. If the referentialist view is accepted, then these are a single proposition, and it is clear that the subject has credence 1 in that proposition all along, at least under its ‘Jekyll is Jekyll’ mode of presentation. Nevertheless, stage two leaves the referentialist with the significant burden of explaining the apparent success of Bayesian reasoning that leads to (B2), and the apparent reasonableness of Olivia having credence 0.8 in the proposition that Jekyll is Hyde.

In stage one, by contrast, the prediction (R1) is not just counterintuitive but near-indefensible. Even on a referentialist view, there is not plausibly any mode of presentation under which Olivia (if rational) has credence 0.9 in *HD*. The only relevant modes of presentation are the ‘Hyde has the disease’ and ‘Jekyll has the disease’ modes of presentation, and Olivia has credence 0.2 in *HD* under the former, and 0.1 in *HD* under the latter. (The stage-one analog of (R2) is not (R1) but the somewhat less bizarre claim that $p_1(HD) = 0.1$.) So a response in this familiar style appears to be unavailable to the referentialist here.

To help organize the further discussion of referentialist responses, it is helpful to introduce a framework on which credences can be associated not just with propositions but with *assertions*, construed as utterances of assertive sentences. By initially associating credences with assertions of sentences such as ‘Hyde is Hyde’ and ‘Hyde is Jekyll’, we can avoid begging any questions about whether these assertions express different propositions. One could also try to associate credences with sentences themselves, but doing things in terms of assertions avoids worries tied to context-dependent sentences such as ‘You are hungry’.²

To associate assertions with credences, we can note that on the standard understanding of credences, they can be characterized in terms of rational betting odds. Any assertion can yield a corresponding bet, for example as follows:

‘Hyde is Jekyll.’

‘Wanna bet?’

Furthermore, any *pair* of assertions *A* and *B* by the same speaker can be associated with a conditional bet, such that the speaker will gain a certain amount if the first is true and the second true, will lose a certain amount if the first is true and the second false, and will not lose or gain anything otherwise. Some such bets will be such that the speaker rationally accepts them, and some will be

²Does the fact that assertions can be associated with credences entail that assertions are the objects of credence?

No: a successful general theory of credence is more likely to associate credence with contents rather than assertions, not least because there are many cases in which subjects have credences without making corresponding assertions. It is natural to hold that the association between assertions and credences itself derives from a more basic association between contents and credences, for example because credence in an assertion always coincides with credence in the content of the assertion. But for present purposes, the characterization of credence in an assertion in terms of betting odds gives us a grip on the notion of credence in an assertion that does not require first establishing a theory of content.

such that the speaker rationally rejects them. So this understanding yields a straightforward way to associate many assertions with credences and many pairs of assertions with conditional credences. The betting-odds understanding is only a first approximation at defining this association (there are obvious problems for assertions such as ‘Bets never pay off’, for example), but it is good enough for our purposes.

When A is an assertion with a credence, I will call this credence $pr(A)$. When A and B are a pair of assertions with an associated conditional credence, I will call this credence $pr(B|A)$. Henceforth ‘ pr ’ will be used for rational credences associated with assertions, and ‘ p ’ for rational credences associated with propositions.

We can now consider versions of the cases above that involve assertions. Say that HD' is an assertion by Olivia of ‘Hyde has the disease’, that JH' is an assertion by Olivia of ‘Jekyll is Hyde’, and so on. Then the analogs of (B1) and (B2) for assertions are

$$(B1') \quad pr_1(HD') = 0.2$$

$$(B2') \quad pr_2(JH') = 0.8$$

Whatever we say about (B1) and (B2), it seems hard to deny (B1') and (B2'). That is, it is hard to deny that the betting odds that Olivia should associate with an assertion of ‘Hyde has the disease’ after stage one are around 0.2, and that the betting odds that she should associate with an assertion of ‘Jekyll is Hyde’ after stage two should be around 0.8. A natural form of Bayesian reasoning gets us to both claims, and they seem plausible in any case. These claims about assertions do not immediately commit us to claims about propositions. But they set things up for different referentialist responses, depending on their attitude to (B1') and (B2').

A radical response rejects (B1') and (B2') as well as (B1) and (B2). One rejection strategy might repudiate Bayesianism wholesale, perhaps rejecting the whole idea of credences, or perhaps rejecting the claim that credences should be updated in a Bayesian way. Another rejection strategy accepts Bayesianism, but rejects the Bayesian reasoning that leads to the rejected claims, perhaps on the grounds that Bayesian analysis requires idealized conditions that fail in these cases.

A moderate response accepts (B1') and (B2') while rejecting (B1) and (B2), by distinguishing the credences associated with assertions from the credences in the propositions expressed by those assertions. One strategy relativizes the credences of propositions in some fashion, so that relativized versions of (B1) and (B2) are true even if the original versions are defective. Or they may do this by reassigning the credences to something other than propositions, so that modified versions of (B1) and (B2) invoking these entities are true.

A referentialist might also try to respond by appealing to claims about evidence, perhaps holding that evidential propositions are qualitative rather than object-involving—specifications of phenomenal states, say, or existential rather than singular claims about individuals—or perhaps holding that in these cases the total evidence goes beyond the object-involving evidence invoked above. These appeals do not help, however. As long as the referentialist allows that credences associated with assertions are credences in the associated object-involving propositions, and that assertions such as 'Jekyll has alpha' and 'Hyde has beta' have high credence, we can generate the problematic results.

In what follows I discuss two radical responses and two moderate responses in turn.³

³Although referentialism about mental states is my focus in this article, it is worth noting briefly that one can use these cases to raise related problems for referentialism about language: the thesis that a singular term contributes only its referent to the proposition expressed by an assertion of a sentence containing that term. For example, the falsity of

4 Strategy 1: Reject Bayesianism

Most straightforwardly, a referentialist can reject Bayesianism entirely. I think that this is the worst option. First, while Bayesianism has controversial aspects (tied to idealization, the problem of old evidence, and the status of prior probabilities, for example; see Glymour 1980 and Earman 1992), these are largely irrelevant to the cases above. The only Bayesian claim we need is the weak claim that when one has a rational credence $p(H|E)$ and then acquires evidence E , one should revise one's credences so that $p'(H) = p(H|E)$. For present purposes, it would even suffice to appeal to the principle that when $p(H|E) > p(H)$ and one acquires evidence E , one should revise one's credences so that $p'(H) > p(H)$. Few theorists reject this principle.

Strictly speaking a *theory* of confirmation such as Bayesianism is not needed to generate a problem for referentialism. One can generate a problem simply from claim (B2), or from the apparently obvious fact that Olivia has credence between 0 and 1 in the hypothesis that Jekyll is Hyde. Taken alone, however, this would lead to the familiar dialectic in which the referentialist denies (B2) while giving a pragmatic explanation of its intuitive plausibility. Once (B2) is embedded in the machinery of a successful theory, then this denial has a much greater cost. For this purpose, many different theories of confirmation will suffice to generate an appropriate argument. All we need is a theory that allows that confirmation involves increase in credences, and that identity referentialism about language follows from the plausible claims that $p_2(JH') = 0.8$ and $p_2(JJ') = 1$, and the additional premiss that a subject's credence in an assertion is identical to the subject's credence in the proposition expressed by that assertion. As I discuss in section 8, one can also raise problems for referentialism about language by focusing on the analysis of probabilistic claims themselves, such as the claim ' $p_2(\text{Jekyll is Hyde})=0.8$ '. But the issues here are close to familiar issues concerning referentialism about language, and the referentialist about language will have familiar replies.

statements can be confirmed.

There does not seem to be any room to deny that identity statements can be confirmed. It is a datum that many successful episodes in the history of science involve the confirmation of such statements ('Hesperus is Phosphorus', 'water is H₂O', 'electricity is magnetism'). Likewise, it is a datum that highly rational agents can associate intermediate credences with identity statements (whether or not these are associated with identity propositions), and that these credences can change in a rational way. If we do not give an account of this, the theory of confirmation is incomplete. But if we do give an account of this, then the same problems will rearise, and the referentialist will have to answer the challenge again, presumably by embracing one of the other three options below.

In any case, Bayesianism is a simple and highly successful theory, and it gives a straightforward and appealing explanation of what is going on in the relevant cases. If one has to choose between rejecting referentialism and rejecting the simple core of Bayesianism, the former would seem to be the obvious choice.

5 Strategy 2: Invoke an idealization

Alternatively, a referentialist could accept the core Bayesian principles, while invoking restrictions on the way that they can be applied in the cases above. Here, the obvious restriction is that Bayesian principles require *idealization*, and cannot be expected to function well with non-ideal subjects. In particular, standard Bayesianism requires credences to satisfy the axioms of probability theory, which require all logical truths to be assigned probability 1. This idealization leads to the familiar problem of logical omniscience for Bayesian theories (Garber 1983). The problem

can be brought out using the following two cases, which are structurally reminiscent of stage two and stage one of the Jekyll/Hyde case respectively.

The Smith case. Smith is rational but not logically omniscient. For a certain highly complicated logical truth L , his credence $p(L) = 0.5$. At a later time, he gains strong but not conclusive evidence for the truth of L , allowing him to update his credences so that $p'(L) = 0.9$. However, these intuitive results contradict standard Bayesianism, as standard Bayesianism satisfies the axioms of probability, which require that all logical truths have probability 1.

The Jones case. Jones is rational but not logically omniscient. E is a piece of evidence that Jones has just acquired, and F is a nonobvious logical consequence of E : Jones is not in a position to derive F from E , but it could be derived in principle using a complex logical truth L . By standard Bayesian principles, Jones' prior credence $p(F|E)$ was identical to $p(F|E \& L)$, which is 1. This entails that when Jones acquires evidence E , he should update his credences so that $p'(F) = 1$. But Jones is in no position to do this, as Jones has no idea that L is true.

The referentialist may suggest that the analogy between the Smith and Jones cases and the Jekyll/Hyde case runs deep. In particular, referentialists may suggest that Bayesian theory requires that all *necessary* truths be assigned probability 1. If so, then like Smith and Jones, Olivia is in non-ideal circumstances. All of them lack knowledge of certain necessary truths (or at least of certain necessarily true sentences): logical truths in one pair of cases, and truths about identity in the other.⁴ In stage one, Olivia is like Jones: in these cases, the dictates of Bayesian updating appear unreasonable, but this is because these are idealized dictates that cannot reasonably be expected to

bind non-ideal creatures. In stage two, Olivia is like Smith: in these cases, intuitive assignments of credence contradict Bayesian verdicts, but this is to be expected in non-ideal circumstances. Once this is recognized, we can see that the apparent tension between referentialism and Bayesianism is simply a new instance of an old tension between Bayesianism and our intuitive judgements about non-ideal creatures.

Here, the most obvious reply is that unlike Smith and Jones, Olivia does not appear to be reasoning non-ideally at all. We could assume that she is a Vulcan, with extraordinary reasoning acumen. As long as we spell out the cases the right way, so that there is no evidence that the same person has been seen twice, nothing relevant about the cases will change. Bayesian principles are usually understood in such a way that the idealization required to apply them involves reasoning acumen but nothing more. So there should be no problem with the application of these principles to the Jekyll/Hyde case.

A referentialist might still appeal to the idealization according to which subjects must know all necessary truths, on the grounds that probability theory requires that all necessary truths be assigned probability 1. This is not a straightforward line for the referentialist to run, as on their own account, Olivia does not lack knowledge of relevant necessarily true propositions: on their account JH and JJ are the same proposition, one known to her in virtue of knowing that Jekyll is Jekyll. But perhaps the referentialist might cast Olivia's deficit in terms of her lack of knowledge of necessarily true sentences such as 'Jekyll is Hyde', or something along those lines.

In any case, it is far from clear that standard probability theory requires that all necessarily true sentences be assigned probability 1. The axioms of probability theory are often formulated to require that all logical truths are assigned probability 1, but this requirement does not cover

all necessary truths. Perhaps more promisingly for the referentialist, the axioms are sometimes formulated in terms of sets, requiring that the universal set (Ω) is assigned probability 1. But this does not entail that necessarily true sentences are assigned probability 1, unless we make the additional assumption that the probability of a sentence is the probability of a corresponding set, and that all necessarily true sentences correspond to the universal set. Nothing in probability theory or in Bayesianism requires this assumption. Later in the paper, I outline a Bayesian framework on which the sets in question are sets of epistemically possible scenarios rather than metaphysically possible worlds. On this framework, truths that are metaphysically but not epistemically necessary, such as 'Jekyll is Hyde', correspond to non-universal sets and need not be assigned probability 1. So an idealization requiring knowledge of all true identity sentences is certainly not required by Bayesianism.

Apart from being unnecessary, this idealization is unreasonable. Knowledge of all necessarily true sentences requires something roughly equivalent to total omniscience. For any truth S , 'Actually S ' will be necessary. But an idealized subject who knows the truth of 'Actually S ' for all truths S will also know the truth of S for all truths S . Even if the claim is restricted to identity sentences, something similar applies. For any sentence S , one might introduce a name 'Fred' rigidly designating the truth-value of S . The identity statement 'Fred = True' will then be necessary. The idealization in question will require knowledge of the truth-value of this statement, which requires knowledge of the truth-value of S , for all S . If a creature has such knowledge, there is nothing that they will be uncertain about, so there will be no need for Bayesian principles. That is, the current idealization undercuts the very rationale for Bayesianism: namely, modeling reasoning under conditions of uncertainty.

Perhaps one might take a pluralist view, allowing that there is a coherent sort of referentialist Bayesianism that might apply under limited ideal conditions, in which subjects are not ignorant of any *relevant* identity statements. Such a view, assigning credences to propositions understood referentially, might have some usefulness in some cases in which different routes of access to the same object are not playing a central role. But this view will have little use in the cases we have discussed, and in many related cases.

Here, it is useful to note again how central the confirmation of identity statements and related necessary truths have been in the history of science. In many instances, this sort of confirmation is straightforwardly amenable to a successful Bayesian analysis. But if the current idealization is invoked, the Bayesian analysis is incorrect. So this idealization cripples Bayesianism as a tool for analyzing these episodes in the history of science. Likewise, it cripples Bayesianism as a tool for the confirmation or disconfirmation of identity statements in the present day. So once again, a fully general Bayesianism requires going beyond referentialism.

6 Strategy 3: Relativize credences

Henceforth I set aside the strategies above, and accept the near-obvious claims that (B1') and (B2') are correct. The best option for a referentialist is to show how the truth of these claims is compatible with referentialism. This requires giving an account of the relation between assertions, propositions, and credences, so that (B1') and (B2') are correct even though (B1) and (B2) are defective.

One option is to *relativize credences*. This sort of view takes its cue from referentialist views on which, whenever a subject believes a proposition, they do so under some guise.⁵ On this view,

it is natural to say that before examining the unexamined loci, Olivia believed JH under one guise, G_1 , but not under another guise G_2 . One might then suggest that one can talk of *credence under a guise*. Under G_1 , $p(JH) = 1$ and $p'(JH) = 1$ (where these are the prior and posterior probabilities at stage two). Under G_2 , $p(JH) = 0.04$ and $p'(JH) = 0.8$. Or one might suggest that probabilities must be dyadic functions of guises and propositions, so that $p(G_1, JH) = 1$, $p(G_2, JH) = 0.04$, and so on.

This is a major revision of Bayesianism. Guises never appear in standard Bayesian theory. Rather, there is a space of propositions with a single probability function over that space. The new proposal requires many probability functions over that space, one for each guise. Worse, there will be many instances where a credence involving one guise interacts with a credence involving another (perhaps information about George Bush is used to update one's beliefs about Laura Bush), but here Bayesian theory will yield no guidance about how the two probability functions should interact. Perhaps one could handle the interaction by constructing an even more nonstandard Bayesian theory in which conditional credences are relativized to two guises (one for each proposition involved). But this would now be something very different from standard probability theory, in which conditional credences associated with a probability function are constitutively connected to unconditional credences associated with the same function.

A system like this would work much better if guises are built into the space of objects over which credences are defined, rather than built into the credence function itself. For example, there might be a single credence function defined over a space of (guise, proposition) ordered pairs. But this view now says that the objects of credence, as we are understanding them, are such ordered

⁵See e.g. Braun (2000), who appeals to guises to address a number of potential problems for the Russellian referentialist, for example in making sense of psychological generalizations.

pairs. If so, the objects of credence behave in a non-referential way, and referentialism will be false of the objects of credence.

7 Strategy 4: Invoke surrogate objects of credence

An important strategy reconstructs the Bayesian analysis of the crucial cases by reassigning the credences involved to different entities. According to this strategy, while (B1) and (B2) as they stand are not correct, they may be correct if we replace *HD* and *JH* in these sentences with something other than the proposition that Hyde has the disease and that Jekyll is Hyde. On this view, there is some *surrogate* entity JH^* such that at stage two, $p(JH^*)=0.04$, $p'(JH^*)=0.8$, and so on, where JH^* is distinct from *JH* (and likewise for HD^* and *HD*). On one version of the view, JH^* will be a proposition distinct from *JH*. On another version of the view, JH^* will be something other than a proposition.

The surrogate strategy requires some constraints. For a start, it requires that there be a way of associating problematic *assertions* with surrogates in a manner that preserves probabilities. Once we have set aside the rejection and restriction strategies, we can stipulate that Olivia's credences in assertions about Jekyll and Hyde satisfy (B1') and (B2'). The surrogate strategy requires that there be surrogates associated with all such assertions such that for each assertion *S*, there is a surrogate S^* such that $p(S^*) = pr(S)$. In order to be able to apply conditional probabilities, this association relation must also be such that $p(S^*|T^*) = pr(S|T)$.

Furthermore, in order to apply principles of updating, then different assertions of the same sentence by the same speaker should be associated with the same surrogate, at least when these assertions do not involve indexicality, context-dependence, or meaning change. This suggests that

in cases such as the Jekyll case, which do not involve indexicality and the like, we should be able to associate surrogates with sentences (at least relative to a speaker), and not just with assertions, so that unconditional and conditional probabilities of sentences at times (for a speaker) coincide with unconditional and conditional credences of surrogates at times (for that speaker).

The strategy must apply to any assertion of ‘*A is B*’ that is rationally associated with a credence less than 1 by the speaker. In fact, it must apply to any such assertion that is rationally associated with a *conditional* credence of less than 1 by the speaker, conditional on some possible evidence *E*. So a very wide range of such assertions must be associated with surrogates. It is natural to suppose that surrogates behave compositionally: that is, that for any name *A*, there is a surrogate entity associated with it (for a speaker), so that the surrogate for the sentence ‘*A is B*’ is determined in some compositional manner from the surrogates for *A* and *B*. And once we have gone that far, it is natural to suppose that surrogates are associated *universally*, so that any assertions involving names, and perhaps any assertion at all, can be associated with a surrogate. I do not assume compositionality and universality in what follows, but I will indicate where they become relevant.

(1) Sentential surrogates

A natural suggestion is that surrogates might themselves be sentences. After all, we have seen above that probabilities *can* be assigned to sentences. A referentialist might suggest that Bayesian theory should be understood as assigning credences to sentences (construed either as types or tokens), and not as assigning credences to propositions. Or they might suggest that credences can be assigned to both, but that the reading according to which (B1) and (B2) are correct is a reading involving sentences, not propositions.

A problem immediately arises. We often want to apply Bayesian principles to subjects who

speaking languages other than our own. The question then arises: are the surrogate sentences to be understood as sentences of the *theorist's* language, or sentences of the *subject's* language? Here is a case.

The German case. Max has heard of a person called 'N1' and a person called 'N2'.

He wonders aloud 'Ist N1 N2'? For him, 'N1 ist N2' has a low credence, but later he receives evidence that raises his credence significantly.

If we take the first option, we are faced with the problem that no surrogate sentence may be available. Names 'N1' and 'N2' are not expressions of our language, and they may not have specific counterparts in our language. For example, there might be only one name for the referent of these names in English. If so, then there will not be distinct sentences to associate with Max's sentences 'N1 ist N1' and 'N1 ist N2'. Even if there is more than one term in English, there will often be no canonical association between the two. Often the best we can do is to import the German term into our language. But that strategy will suffer from the same problems as the second option below.

If we take the second option, we are faced with a different version of the problem that no surrogate sentence may be available. We sometimes want to apply Bayesian principles to subjects who do not express their beliefs in language.

The hall-of-mirrors case. Fred is in a room full of mirrors, with three red balls in the room. He has a visual experience as of hundreds of red balls, but he knows that there are only three in the room. He attends to a specific ball that he sees in a mirror to his left, and to a specific ball that he sees to a mirror on his right. Without speaking, he

entertains the hypothesis *that ball is the same as that ball*, and rationally assigns it a credence of approximately one-third.

The hall-of-mirrors case can be used to generate the same sort of reasoning as the Jekyll case. But in this case, there are no surrogate sentences in the subject's language available: the subject does not assert any sentences, and unasserted sentences such as 'That ball is . . . ' do not distinguish the two demonstratives. Likewise, one can also generate a version of the problem for creatures without language, assuming that rational belief without language is possible. Here once again, no surrogate sentences are available.

The last resort is to suggest that surrogates may be sentences in a subject's language of thought, or perhaps mental representations of some other sort. Of course it is not uncontroversial that such objects exist, and it would be surprising if a commitment to their existence were required for the general application of Bayesian theory. But in any case, this view is now very much in the spirit of an existing *rival* of referentialism: a view on which the objects of belief are sentences in a language of thought. Perhaps a referentialist could say that beliefs have two sorts of objects, or perhaps they could say that sentences in a language of thought are objects of credence but not objects of belief, or perhaps they could say that they are the objects to which credences must be assigned in these cases without being objects of credence. But of course the last two views are at best extremely awkward, while the first view concedes that strong referentialism about the objects of belief is false.

Another problem for this view is that non-sentential objects of credence are often given a central explanatory role in Bayesian analysis. The standard Kolmogorov axioms assume that the objects of credences are sets of some sort. While it is also possible to formulate the axioms in

terms of sentences, and some Bayesian theorists do this, it is very useful to be able to associate credences with sets as well, so that for example when one set contains another, it will be associated with greater or equal credence, and so on. If we cannot take sets to be objects of credence, then we lose this important source of explanatory power.

A worse problem for sentential views is that they rule out the use of reasoning about *content* in analyzing the Bayesian updating of belief, at least in these cases. Intuitively, our credences concern the world, not just sentences. Just as we capture this world-directedness of belief by analyzing it in terms of content, we should expect the same for credences. Bayesian reasoning exploits rational relations between beliefs, and it is especially natural to hold that these rational relations are to be explained in terms of relations between contents. In addition, multiple different sentences (synonymous sentences within a language or translated sentences across languages) can be used to express the same patterns of probabilistic reasoning. It is natural to explain the equivalence by appealing to the fact that the same content is involved in each case; but this requires us to be able to assign credences to contents.

A referentialist might respond by saying that sentences themselves are associated with contents, which we can then use in this sort of analysis. But if these contents behave in a referential way (so that ‘Jekyll is Jekyll’ and ‘Jekyll is Hyde’ are associated with the same contents), then we cannot apply reasoning about these contents in the usual manner to cases such as (B1) and (B2). And if these contents behave in a nonreferential way, then we are very close to the view on which nonreferential abstract entities serve as objects of credences, to be discussed below. So while the view that takes sentences and not contents to be the objects of credence is coherent, I think that it is very much a last resort.

(2) Propositional surrogates

On the propositional strategy, a referentialist suggests that there is a proposition JH^* that renders the Bayesian reasoning above correct, in that Olivia's initial credence $p(JH^*)$ was 0.04 and her later credence $p(JH^*)$ was 0.8, updated in the Bayesian way. It is just that this proposition differs from JH , the proposition that Jekyll is Hyde. The most obvious version of this strategy takes JH^* to be some descriptive proposition, such as the proposition that the patient Olivia met in the morning is the patient she met in the evening.

There is something awkward about referentialists suggesting that we can invoke descriptive surrogates for names. Referentialists often argue for their view in part by arguing that for many names, no descriptive surrogates available. For example, Kripke argued that a speaker may be competent with the name 'Gödel' despite being able to associate no identifying description with the name. And even when names are associated with identifying descriptions, referentialists argue that the description and the name are not equivalent: for any D , we might discover that Gödel was not the D . This can be leveraged into a serious problem for the descriptive surrogate strategy.

The Gödel case. Saul has heard talk of Kurt Gödel, but knows almost nothing about him apart from the fact that he is a famous mathematician. He has also heard talk by students around Princeton of someone called Smiffy, but knows almost nothing about Smiffy apart from the fact that some students think he is eccentric. One day he overhears students arguing about whether Smiffy is Gödel or Einstein. He does not know whether either suggestion is correct (in fact, the first suggestion is correct), but he now rationally gives the hypothesis that Gödel is Smiffy a significantly higher credence than he would have earlier.

On the current referentialist strategy, an assertion by Saul of ‘Gödel is Smiffy’ must be associated with some surrogate descriptive proposition—say, *The D is the F*, such that conditional and unconditional credences involving that proposition match those associated with the sentence. But if Kripke’s epistemological argument is correct, this cannot be so. The argument suggests that for any description *D*, there is evidence *E* such that if Saul discovers that *E* is the case, he will rationally reject ‘Gödel is the *D*’. So if *E'* is a sentence expressing *E*, then $pr(\text{‘Gödel is the } D\text{’} | E')$ is low. But clearly $pr(\text{‘Gödel is Gödel’} | E')$ is high. If we assume that surrogates behave compositionally and that ‘the *D*’ and ‘Gödel’ are associated with the same surrogate (*the D*), this yields a contradiction. And even if we do not assume this, one can easily extend Kripke’s strategy to yield cases where $pr(\text{‘Gödel is Smiffy’} | E')$ is low while $pr(\text{‘The } D\text{ is the } F\text{’} | E')$ is high. If so, then assuming that *The D is the F* is the surrogate for ‘The *D* is the *F*’, then we still obtain a contradiction.

In order for the descriptive surrogate strategy to work, a referentialist must reject Kripke’s epistemological argument. In effect, this allows that Saul’s use of ‘Gödel’ can be associated with some descriptive content *the D*, expressed by ‘the *D*’, such that Saul’s conditional credence $pr(\text{‘Gödel is the } D\text{’} | E) = 1$ for all evidential propositions *E*. The referentialist can then allow that all conditional and unconditional probabilities associated with Saul’s assertion of ‘Gödel is ϕ ’ are the same as those associated with an assertion of ‘The *D* is ϕ ’.

Of course this is an extremely awkward strategy for a referentialist to embrace, as Kripke’s epistemological argument and related arguments against associated descriptive contents are often thought to be the most powerful arguments for referentialism. Furthermore, the position seems to lead inevitably to the view that *all* singular terms and all object-directed beliefs are associated with

this sort of descriptive content, as one can generate cases such as the above for almost any such terms. This leads one to the conclusion that all such beliefs are at least associated with descriptive propositions, such that it is guaranteed that the belief is correct if and only if the descriptive proposition is correct. In such a case it seems natural to say that the descriptive proposition is a content of the belief, and the resulting view is close to an across-the-board descriptivism.

At this point a referentialist might advocate a special sort of descriptive propositional surrogate, namely a *metalinguistic surrogate*. The idea is that the surrogate proposition associated with ‘Jekyll is Hyde’ is something like ‘The person called ‘Jekyll’ is the individual called ‘Hyde’’, or perhaps ‘‘Jekyll is Hyde’ is true’, or something like that.⁶ It could be argued that ‘The person called ‘Jekyll’ is guaranteed to corefer with ‘Jekyll’, given any evidence, but that this result does not threaten Kripke’s epistemological argument, as the coreference is relatively trivial. However, we here once again run into the problem that cases like this arise when there are no associated linguistic items. In the hall of mirrors case above, for example, no relevant linguistic item is available. One might suggest associating their thought with a metacognitive surrogate proposition, such as *this thought is correct*, or *the object of this mental representation is the same as the object of that mental representation*. But it seems clear that subjects need not be entertaining any such metacognitive surrogate in these cases.

Further, all instances of the descriptive proposition strategy (including the metalinguistic and metacognitive strategies as special cases) have even worse problems in cases involving indexicals. Cases like this are familiar from the work of Perry (1979) and Lewis (1983). The following variant on Lewis’s ‘two gods’ case makes the point clearest for our purposes.

⁶This sort of strategy is reminiscent of the appeal to metalinguistic propositions by Stalnaker (1984).

The Beeblebrox case. Zaphod Beeblebrox has just been elected president of the galaxy. He enters the Total Perspective Vortex, which has a display showing everything going on in the universe. He uses this to gain complete objective knowledge of all individuals in the world and their properties. But the world is near-symmetrical, so that Zaphod has a nearly indistinguishable twin Zaphod', who is president of a counterpart galaxy. The only difference is that on their scalps underneath their hair, Zaphod has a red star and Zaphod' has a blue star. They have never seen these stars and the difference has never affected their evidence. Zaphod sees both twins on the screen and notices the stars: he introduces the name 'Red' for the twin with the red star and 'Blue' for the twin with the blue star. He wonders which of the two twins he is, saying 'Am I Red? Or am I Blue?'. (Of course he sees both twins asking the question, but this does not help.⁷) He rationally has credence 0.5 associated with his assertion 'I am Red', while he has credence 1 associated with his assertion 'Red is Red'.

In this case, a referentialist needs a nonreferential account of the objects of the credences. The descriptive surrogate strategy suggests that assertions of the form 'I am Red' be associated with a descriptive proposition of the form *The D is the F*. (Alternatively, it might be associated with a partially descriptive proposition of the form *I am the F*, but the same problems arise.) But for any

⁷Thanks to Robert Adams for pointing out a flaw in an earlier version of this case involving a fully symmetrical scenario. In a fully symmetrical scenario, Zaphod and his twin's utterances 'I am such-and-such' would have the same truth-value. Zaphod could determine this truth-value, perhaps by determining whether both are pointing to images originating in themselves, and could therefore determine whether or not he is such-and-such. In the partly asymmetrical scenario here, the utterances have different truth-values and this method will not work.

such proposition, at least if it is construed as an objective proposition, then Zaphod knows its truth already. So Zaphod's credence in these propositions will be much higher than 0.5, contradicting the requirement of matching probabilities.

Perhaps the referentialist could invoke a special class of non-objective propositions, such as egocentric propositions, as surrogates here. But this is to move even further away from a standard referentialist view. In any case, this suggestion comes to much the same thing as the suggestion of non-propositional surrogates considered below.

(iii) Other abstract surrogates

A final version of the surrogate strategy holds that credences are assigned not to propositions, or at least not to propositions as the referentialist usually understands them, but instead are assigned to some other sort of abstract object, such as events, sets of centered worlds, or Fregean propositions.

The problem here is predictable. If these abstract surrogates work in a referentialist way—say, if they are events or sets of possible worlds associated referentially with beliefs—then the original problem will arise again. But if they do not work in a referentialist way—if they are ordered pairs of guises and propositions, Fregean propositions, or sets of centered worlds associated nonreferentially with beliefs, say—then then these objects will be nonreferential objects of credence.

Summary

We have seen that there are specific problems for many potential surrogates: sentences, propositions, and so on. But we have also seen versions of the same dilemma throughout. If the surrogates behave referentially, then they are inadequate to serve as objects of credence. If the surrogates do not behave referentially, then in so far as they are objects of credence, referentialism is false. It

follows that if the objects of credence are surrogates, referentialism is false.

8 Referentialism about credence and referentialism about belief

I have argued that referentialism about credence is false. At this point, some referentialists may respond by conceding that referentialism about credence as I have defined it is false, while holding that referentialism about credence as they understand it remains untouched. A related strategy concedes that referentialism about credence is false, while holding that referentialism about belief is unthreatened. On these views, a referentialist will allow that in so far as objects of credence are defined as the entities to which credences are assigned by credence functions, referentialism about these objects is false. But they may allow that on some other way of understanding objects of credence and/or objects of belief, the way that is important to them, referentialism about these objects is true.

For example, one common view takes the objects of belief to be the referents of ‘that’-clauses. This view can be supported by arguing that our best semantics for attitude ascriptions is relational: ‘believes’ in ‘S believes that p’ expresses a relation between the referent of ‘S’ and the referent of ‘that p’. If so, the latter referent is literally the thing that is believed, and is therefore the object of belief. The same goes for ‘S doubts that p’ and ‘S finds it probable that p’: if so, the referent of ‘that p’ is literally the thing that is doubted and found probable, and is therefore the object of credence.

Referentialists about attitude ascriptions hold that the referent of ‘that p’ is associated referentially with ‘p’: so ‘that Hesperus is a planet’ and ‘that Phosphorus is a planet’ have the same referent, for example. When combined with the considerations in the paragraph above, referen-

tialism about attitude ascriptions naturally leads to referentialism about the attitudes themselves, and in particular to referentialism about belief and credence. Such a referentialist might note that my arguments have not established that referentialism about attitude ascriptions is false or that a relational semantics for attitude ascriptions is false, and might use this to argue that my arguments have not established that referentialism about credence and belief are false.

The resulting position is somewhat uncomfortable, as one can apply the arguments given earlier to credence ascriptions such as ‘Olivia has low confidence that Jekyll is Hyde is low’ and so on. Perhaps the referentialist will give some special treatment of such sentences, as they do for belief sentences. For example, they might argue that such a sentence expresses the false proposition that $p(JH)$ is low, while conveying the true proposition that $p(JH^*)$ is low. Alternatively, they might hold that it makes tacit reference to a contextually specified guise, saying correctly that $p(G_1, JH)$ is low.⁸ Presumably they will say the same for many or most other subjective probability claims involving names. The first option requires denying the truth of claims that play a highly successful and systematic explanatory role, while the second option requires significant complexity. These outcomes are not foreign to the referentialist, but I think they give reason to take seriously a view on which the referents of ‘that’-clauses are nonreferential too.

⁸David Braun suggests a version of this view in the Online Philosophy Conference. In this section I am indebted to an exchange with Braun. I note that my own account of attitude ascriptions, developed in Chalmers (forthcoming a), is a relational account that is nonreferentialist about the referents of ‘that’-clauses (those clauses refer to the enriched propositions associated with embedded sentences, as discussed later in this paper). But even in that paper I note that there is a near-equivalent ‘hidden indexical’ view that is referentialist about the referent of ‘that’ clauses while also giving a role to contextually specified constraints on modes of presentation (primary intensions), roughly as on Braun’s view here. I argue in that paper that there are considerations favoring the former view, but I do not claim to have decisive arguments against the latter, and I think that views of the latter sort are compatible with the data of this paper.

But even if referentialism and relationism about attitude ascriptions are correct, these at best entail that there is one sense—a sense tied to the semantics of attitude ascriptions—in which the objects of credence are referential. The arguments given here take as their starting point a different sense of ‘objects of credence’, defined in terms of the entities to which credences are assigned by our best theory of credence. If a version of the ‘contextually specified guise’ view is correct, for example, then it may turn out that objects of credence in the first sense are referential while objects of credence in the second sense are not. An opponent might insist that terms such as ‘object of credence’ and ‘object of belief’ express the first sense rather than the second, but I think that both senses are natural, and in any case the point is verbal. One could simply give such an opponent the term ‘object of credence’ and reformulate the points here using another locution such as ‘contents of credence’, without much loss.

If we set aside the verbal issue of what counts as an ‘object of credence’ or an ‘object of belief’, the broader point is that the arguments so far have not ruled out *pluralism* about the objects of credence and belief. There are multiple notions in the vicinity of these expressions, corresponding to different explanatory roles that we might want objects of credence and belief to play. One role concerns a formal theory of confirmation, while another concerns attitude ascriptions. The objector could then put their point by saying that just because referentialism is false of the notion tied to the former role, it does not follow that referentialism is false of the notion tied to the latter role.

For present purposes, I am happy to concede this point, with three qualifications. First, the falsity of the first sort of referentialism at least puts some pressure on the second sort of referentialism, for the reasons given above. Second, this point is enough to reject *strong referentialism*,

the view that referentially associated objects can play all the central explanatory roles of objects of belief.

Third, and perhaps most importantly, the falsity of the first sort of referentialism is enough to suggest that the structure of the underlying mental states is not best understood in referential terms. To fully understand the dynamics and the normative role of beliefs and of states of credence, we have to understand them as relations to objects that are nonreferentially associated with the mental states. This does not imply that there is no role for reference in understanding the nature of these states. Toward the end of the paper, I sketch a view on which belief and credence have referential as well as nonreferential components. But it does suggest that in the metaphysics of belief and credence, we have to go beyond a referential view.⁹

9 What are the objects of credence?

I now set aside the issue of referentialism, and ask: what are the objects of credence? As before, I stipulate that the objects of credences are the entities to which credences are assigned by a probability function for the purposes of an adequate theory of credence, which I take to include Bayesian confirmation theory. I assume that objects of credence must be uniformly available, so that they can be found in all cases of Bayesian confirmation (with single objects for unconditional credences, and pairs of objects for conditional credences). And for reasons discussed above, I assume that there is an association relation between assertions and objects of credence, so that in

⁹It is worth noting that numerous referentialists about language and about attitude ascriptions are already nonreferentialists about the metaphysics of belief. For example, Salmon (1986) holds that the metaphysics of belief involves a three-way relation between a subject, a proposition, and a guise, where the belief states associated with typical utterances of 'Hesperus is a planet' and 'Phosphorus is a planet' will involve the same proposition but different guises.

any case where an assertion has a credence, this credence will coincide with the speaker's credence in the object.

The discussion of surrogates in section 7 can be read as an initial canvassing of candidates to be objects of credence. A few candidates were left standing: sentences in the language of thought, descriptive propositions (apart from the worry about indexicals), nonreferentially associated sets of possible worlds, and various abstract objects such as ordered pairs of guises and propositions, Fregean thoughts, and sets of centered worlds.

To make progress, I take it as an initial requirement that objects of credence can be associated with sets of possibilities of some sort. The requirement is grounded in reasons given in the discussion of sentential surrogates in section 7: for example, it is required in order to use the full explanatory apparatus of the probabilist in the analysis of credences. Of course it is also possible to associate credence with sentences (perhaps in the language of thought) and not with sets. But as discussed earlier, this strategy will not deliver a notion of content, and a notion of content tied to rationality is part of what we are trying to understand here. So it makes sense to at least see how far we can get under the constraint imposed by this requirement.

This requirement does not require that the objects *be* sets of possibilities. They may be more fine-grained abstract objects with internal structure, for example. But they must at least determine such a set, and we must be able to associate credences with such sets. It is also not required that the possibilities in question be metaphysically possible worlds: probability theory does not require this, and we will see that such a requirement is hard to satisfy.

Given this requirement, we can usefully narrow our initial investigation to sets of possibilities, determining just what sorts of sets might be able to play the relevant role. Once this is done, we

can then work backwards to the question of which associated entities are the best candidates to be objects of credences.

The association between assertions and objects of credence and that between objects of credence and sets of possibilities yields an association between assertions and sets of possibilities, such that the credence of an assertion matches the credence of the associated set. Of course, to avoid the earlier problems, this association relation must be nonreferential, so it cannot be the familiar referential assignment of sets of possible worlds to assertions. But this does not rule out the idea that some *other* assignment of (ordinary, uncentered) metaphysically possible worlds to assertions could still serve. For example, as discussed earlier, distinct names might be associated with distinct descriptive contents, so that ‘*A is A*’ and ‘*A is B*’ determine distinct sets of possible worlds.

A serious problem for this idea arises from the Beeblebrox case, however. If Zaphod has complete objective knowledge, it appears that all of his credence is concentrated on a single possible world w . All sets containing w will have associated credence 1, and sets not containing w will have associated credence 0. If so, there is no set available to be associated with his assertion ‘I am Red’, which has credence 0.5. If so, sets of ordinary possible worlds cannot serve as the objects of credence.

The only way to avoid this problem is to say that Zaphod’s ‘complete’ objective knowledge is in fact associated with a set of at least two possible worlds. This set must overlap both with the set associated with ‘I am Red’ and also with the disjoint set associated with ‘I am Blue’, so that Zaphod’s credence can be divided between these sets. As we can stipulate that Zaphod has complete qualitative (physical, phenomenal, and so on) knowledge of the world, this seems to

require that there be worlds that are qualitatively identical but distinct in some other way. The natural suggestion is that the worlds are qualitatively identical but haecceitistically distinct, so that some haecceitistic property is instantiated by Red in one world and by Blue in another, or perhaps better, by the *D* in one world and the *E* in the other, where these are descriptions that Zaphod associates with the names 'Red' and 'Blue'. For example, these might be qualitatively identical worlds such that Zaphod is the twin with the red star in one world and the twin with the blue star in the other world.¹⁰

Of course the claim that there are pairs of worlds that merely differ in some haecceitistic respect is highly controversial. And even if we accept it, it is not quite clear how the proposal will work. Presumably any claim of the form 'I am ϕ ' will be associated with a set of worlds in which one's haecceity is coinstantiated with ϕ (or with a descriptive property associated with ' ϕ '). But this may run into problems with essential properties. For example, if I am essentially human, I might nevertheless give nonzero rational conditional credences to the hypothesis that I am not human, which requires a set of worlds associated with 'I am not human'. But there will be no worlds in which my haecceity is coinstantiated with non-humanness. Perhaps one might move to a world where it is coinstantiated with an associated descriptive property, but it is not clear that this move will always work. If any claim of the form 'I am the *D*' is necessary in cases where 'the *D*' is a 'qualitative' description that is simply associated with the property of being the *D*, then the move will not work. So this framework requires very strong restrictions on essential properties, and may work best if individuals have no nontrivial essential properties at all.

Furthermore, one can generate problems analogous to those that arise for 'I' with other indexicals such as 'here' and 'now', and perhaps also with experiential demonstratives such as 'This

¹⁰A version of this strategy is suggested by Stalnaker (2008).

experience'. The strategy above will then requires times, places, and experiences are associated with haecceitistic properties that can be instantiated in different places in qualitatively identical worlds, in a fairly unrestricted way.¹¹ But presumably if one time, place, or experience can be instantiated in all these different ways, the same goes for every time, place, or experience, so this strategy will generate a huge apparatus of qualitatively identical worlds, far beyond what is needed for the purposes of dealing with the problem at hand. This requires a highly baroque conception of the space of possible worlds.

10 Objects of credence as sets of epistemically possible scenarios

There is a better way to handle these cases. We can still understand the objects of credence as sets of possibilities, but the possibilities need to be understood as *epistemic* possibilities rather than as metaphysical possibilities. For example, even if it is not metaphysically possible that Jekyll is not Hyde, it is epistemically possible for Olivia that Jekyll is not Hyde. Intuitively, her nonzero credence associated with 'Jekyll is not Hyde' reflects that it is epistemically possible for her that Jekyll is not Hyde. And this epistemic possibility itself reflects the fact that there

¹¹The strategy suggested here corresponds in an interesting way to Russell's own 'Russellian' view of the objects of belief. Russell thought that all propositional contents are constructed from entities with which we are acquainted: namely ourselves, our sense-data, and certain universals. Occurrence of the first-person indexicals contribute oneself to the associated proposition; occurrences of demonstratives for sense-data contribute those sense-data; while ordinary names or concepts of external objects contribute a descriptive content constructed from the entities above. The strategy in the text effectively invokes a possible-worlds version of these Russellian propositions. (Russell does not explicitly invoke direct reference to times and places, but one can easily build this into his framework.) If one were committed to the claim that the objects of credence are Russellian propositions, then I think Russell's own Russellian propositions would be the best candidates.

are numerous epistemically possible scenarios open to her that can be seen as instances of the epistemic possibility that Jekyll is not Hyde.

Something similar goes for the Beeblebrox case. Even though it is not metaphysically possible that Zaphod is not Red (since 'Red' rigidly designates Zaphod), his nonzero credence in 'I am not Red' reflects the fact that it is epistemically possible for him that he is not Red. And this epistemic possibility itself reflects the fact that there are numerous epistemically possible scenarios open to him that can be seen as instances of the epistemic possibility that he is not Red.

On this picture, then, we understand the objects of credence as sets of epistemically possible scenarios, rather than as sets of metaphysically possible worlds. Here epistemically possible scenarios stand to epistemic possibility (and in particular to idealized notions of epistemic possibility, such as the notion of what one cannot rule out a priori) roughly as metaphysically possible worlds stand to metaphysical possibility.

Of course this suggestion immediately raises the question of just how epistemically possible scenarios are to be understood. I have discussed that question at length elsewhere (Chalmers 2011), but fortunately for present purposes there is a straightforward and relatively familiar way to understand them, at least to a first approximation.

For present purposes, we can understand epistemically possible scenarios as *centered worlds*.¹² A centered world is usually conceived as an ordered triple of a metaphysically possible world, an individual who is present in that world, and a time in that world (I discuss a refinement to this conception later). For any given subject s , at time t in (uncentered) world w , we can say that the centered world that s inhabits (at t in w) is $\langle w, s, t \rangle$.

¹²The locus classicus for sets of centered worlds as objects of belief is Lewis 1983. The idea that credence is distributed over sets of centered worlds has played a central role in the recent literature on the 'Sleeping Beauty'

When a subject considers what is true, they do not only consider what uncentered world they inhabit, they also consider (in effect) what centered world they inhabit. The centering is necessary to accommodate hypotheses about one's location within a world. For example, Zaphod knows which uncentered world he inhabits, but intuitively he does not know whether he inhabits a centered world centered on the twin with the red star or one centered on the twin with the blue star. Both of these centered worlds represent epistemically possible scenarios for him. More generally, it is arguable that for any centered world w and any subject, that subject cannot rule out a priori the hypothesis that they inhabit w (just what is involved in such a hypothesis is discussed in more detail in the next section). If so, we can see the centered world as representing an epistemically possible scenario, in the sense where epistemic possibility goes with what cannot be ruled out a priori.

Thinking in these terms, it is natural to say that Zaphod's assertion 'I am Red' will be true only at scenarios where the individual at the center has a scalp with a red star, and that his assertion 'I am Blue' will be true only at scenarios where the individual at the center has a scalp with a blue star. Each assertion can be associated with the set of scenarios at which it is true: intuitively, these correspond to the epistemic possibilities compatible with the assertion. Zaphod's two assertions will be associated with two distinct nonempty sets, and he will have credence 0.5 in each of them.

This much is relatively familiar from the literature on self-location. In this literature, it is often supposed that the role of centers in centered worlds is limited to cases involving 'I', 'here', 'now', paradox (Elga 2000). On my view this role for centered worlds is grounded in the fact that centered worlds correspond to epistemically possible scenarios better than uncentered worlds do. The correspondence works best if we assume that uncentered worlds are individuated qualitatively, so that there are no pairs of qualitatively identical worlds containing different individuals, and if we assume the principle of plenitude discussed in the next section.

and closely related indexical expressions. But in fact it is easy to see that the considerations that favor centered worlds in these cases arise in many other cases, including cases involving names and natural kind terms. For most such terms there will be cases such as the following analog of the Beeblebrox case (adapted from Putnam 1975a):

The Oscar case. Oscar lives on Earth. His counterpart Twin Oscar inhabits Twin Earth, which is just like Earth except that the watery stuff there is XYZ. Oscar and Twin Oscar have led phenomenally identical lives, and in particular Oscar has no prior evidence as to whether the watery stuff on his own world is H₂O or XYZ. Upon entering the Total Perspective Vortex, Oscar learns all about the objective structure of the world, including the existence of the twin planets, but he still does not know which planet he is on. He wonders ‘Is water H₂O? Or is it XYZ?’. He rationally has credence 0.5 associated with ‘Water is H₂O’, whereas he has credence 1 associated with ‘Water is water’.

Just as the Beeblebrox case suggests that Zaphod’s credence in ‘I am Alpha’ cannot be understood as a credence in a set of objective worlds (as all such credences for Zaphod are 0 or 1), the Oscar case suggests that Oscar’s credence in ‘Water is H₂O’ cannot be understood as a credence in a set of objective worlds. Instead, it is natural to understand it as a credence in a set of centered worlds. In particular, it might roughly be understood as a credence in the set of centered worlds in which the watery stuff in the environment of the subject at the center is H₂O.¹³ It is natural to hold that Oscar has credence 0.5 in this set, while having credence 0.5 in the corresponding set involving XYZ.

¹³This does not require that ‘water’ is an indexical term that picks out different referents in different contexts of

Something similar goes for ‘Hesperus is Phosphorus’. We could imagine a Beeblebrox-style case in which one twin inhabits an environment in which the morning and evening stars are distinct, and one inhabits an environment in which the morning and evening stars are identical. On entering the Total Perspective Vortex each twin might rationally have credence 0.5 associated with ‘Hesperus is Phosphorus’ despite having credence 1 associated with all sets of objective worlds. The first credence is best modeled as a set of centered worlds: something like the set of worlds at which the evening star around the individual at the center is identical to the morning star around the individual at the center.

The uncentered worlds that serve as the basis for centered worlds are typically be taken to be metaphysically possible worlds. But importantly, the truth of an assertion or a belief at an epistemically possible scenario will often come apart from its truth at the corresponding metaphysically possible world. This much is already clear from the case of ‘I’. On the possible-worlds model familiar from Kripke, Kaplan, and others, my assertion of ‘I am hungry’ will be true at all metaphysically possible worlds at which David Chalmers is hungry. By contrast, on the current model, my assertion of ‘I am hungry’ will be true at all epistemically possible scenarios at which the individual at the center is hungry. Something similar applies to ‘water is H₂O’ and ‘Hesperus is Phosphorus’. On the familiar model, these are true at all metaphysically possible worlds. On the current model, these are false at some epistemically possible scenarios.

utterance. Epistemically possible scenarios are not the same as contexts of utterance: the claim that ‘Water is H₂O’ is false in some epistemically possible scenarios turns on the epistemic properties of the utterance rather than on what it would pick out in different contexts, and does not entail that it could ever be uttered falsely. See Chalmers 2006 for discussion of the differences between the epistemic understanding of two-dimensional semantics (used here) and the contextual understanding developed by Kaplan, Stalnaker, and others.

There should be no question of these models competing with each other. It is a commonplace that it may be epistemically possible that water is H₂O even though it is not metaphysically possible. Correspondingly, we should expect possible-worlds analyses of epistemic and metaphysical possibility to behave differently. We just need to distinguish the notion of truth at an epistemically possible scenario from that of truth at a metaphysically possible world.

All this leads to a familiar sort of two-dimensionalism about the evaluation of beliefs and utterances: there is one sort of evaluation appropriate for analyzing epistemic possibility, and another sort for metaphysical possibility. For present purposes, epistemic possibility is the relevant sort of possibility, so what matters is evaluation at epistemically possible scenarios.

In this framework, the set of epistemically possible scenarios associated with an assertion is a version of what I have called the *primary intension* of the assertion. Primary intensions (which are closely related to the sets of centered worlds discussed by Lewis 1983) can be seen as sets of centered worlds that are nonreferentially associated with assertions and beliefs. They contrast with secondary intensions, which are the more familiar sets of uncentered metaphysically possible worlds that are referentially associated with assertions and beliefs. We have seen that secondary intensions cannot be objects of credences. But the cases above make it very tempting to suggest that primary intensions are the objects of credences. At least, if one wants a set of possibilities of some sort to play the role of objects of credences, then primary intensions appear to be natural candidates.

11 The probabilistic characterization of primary intensions

The informal discussion in the previous section serves to motivate the idea that objects of credence are primary intensions. For a rigorous grounding for this thesis, one would need an argument that arbitrary assertions and thoughts are associated with primary intensions, and an argument that primary intensions can play the core roles of objects of credence. I have given arguments of the first sort in other work (e.g. Chalmers 2006 and 2011), characterizing primary intensions in terms of the a priori. I will not repeat those arguments here, partly to avoid repetition and partly to avoid issues about the a priori. Instead, I make a case for the association by characterizing primary intensions in probabilistic terms, and I use this characterization to argue that primary intensions can serve as objects of credence.

This might be considered a project in ‘probabilistic semantics’: the use of probabilistic notions, such as credences involving an assertion, to characterize semantic values, such as the content of an assertion. One might think that such a strategy is circular: credence in an assertion is plausibly grounded in credence in its content, so the content cannot itself be grounded in credence in an assertion. But as noted earlier, we have a pretheoretical grip on the conditional and unconditional credences associated with assertions, one that does not require prior commitment to a theory of content. So we can use this grip to help explicate a notion of content. This direction of explication is entirely compatible with the reverse direction of grounding, and with the claim that primary intensions rather than assertions are the fundamental objects of credence. Just as we can use our knowledge of microscopes to help characterize the molecules in which they are grounded, we can use our knowledge of credences of assertions to help characterize the contents in which they are grounded.

The basic idea is that the primary intension of an assertion A by a subject includes a scenario w if the subject's idealized conditional credence $pr(A|w) = 1$ (or: is high). This probability is to be understood as the credence that the subject should have in A conditional on the hypothesis that they inhabit w , given ideal rational reflection. So, for example, the primary intension of 'Jekyll is Hyde' is true at a scenario w if the subject's idealized credence in 'Jekyll is Hyde' conditional on the supposition that they inhabit w is 1 (or high).

The hypothesis that one inhabits w can be understood more precisely as the hypothesis that $S(w)$ is true, where $S(w)$ is a sentence that serves as a *canonical specification* of w . Corresponding, the probability $pr(A|w)$ can be understood more precisely as $pr(A|S(w))$. To a first approximation, one can think of a canonical specification of a scenario as a complete qualitative specification of the world and of one's place within it. It will be a conjunction of a sentence D (specifying the objective character of the world) with sentences of the form 'I am the D_1 ' and 'Now is the D_2 ' (specifying one's location within the world), where D_1 and D_2 are descriptions of a person and a time.

Here D , D_1 , and D_2 are all required to use only qualitative vocabulary: one that excludes proper names, natural kind terms, and other expressions that generate a posteriori identity statements, as well as excluding context-dependent expressions.¹⁴ Doing so avoids obvious worries that would arise if one had expressions such as 'Jekyll' and 'Hyde' in the canonical specification. The remaining vocabulary might include some or all of causal, mental, spatiotemporal, logical, and mathematical vocabulary, along with microphysical vocabulary (either explicitly or implicitly via Ramsey sentences in the other vocabulary). But it need not be restricted to an austere vocabulary such as this. For present purposes it can be any qualitative vocabulary that satisfies the

Plenitude and Scrutability constraints outlined below.

We can characterize the primary intensions of sentences (for subjects at times) in a similar manner, at least when those sentences use only qualitative expressions along with indexical expressions such as ‘I’ and ‘now’: the primary intension of a sentence S is true at w (for a subject at a time) iff the ideal conditional credence $pr(S|S(w))$ is 1 (or high) for that subject at that time. This definition invokes the credence of a sentence conditional on another sentence, and the previous definition invokes the credence of an assertion conditional on a sentence. These notions can be understood much as we understood credence in an assertion conditional on another assertion. For example, we might gloss the notions as before in terms of the rational betting odds for a bet on A or S that pays only if $S(w)$ is true in the context of the relevant subject and time. Subjects can have ideal credences involving sentences without ever uttering those sentences. They need not even possess the relevant vocabulary, as long as the idealization stipulates mastery of that vocabulary.

We can say that a centered world w *verifies* an assertion A when the subject’s ideal conditional credence $pr(A|w) = 1$. On an alternative definition, w verifies A when $pr(A|w)$ is high, where this involves some threshold less than 1. The latter approach requires weaker assumptions, a cost of somewhat messier behaviour (for example, this approach allows that a conjunctive assertion might

¹⁴For more on the relevant notion of qualitiveness, see the discussion of semantic neutrality in Chalmers 2006. A useful heuristic is that an expression is qualitative when it does not give rise to Putnam-style ‘Twin Earth’ cases, in which intrinsic twins use the expression and a counterpart expression nondeferentially with distinct referents. The restriction to nondeferential use excludes Burge-style cases in which expressions are used with deference to a linguistic community: it is arguable that any expression can give rise to a Burge-style case, so without this restriction no term would count as qualitative. To avoid Jekyll-Hyde worries arising from deferential usage of expressions in $S(w)$, we can stipulate that the idealization involved in defining $pr(A|S(w))$ requires nondeferential mastery of the expressions in $S(w)$.

fail to be true at w even when the conjuncts are all true at w). I will use the former definition, but it is worth noting that both are available. Either way, the primary intension of A will be the set of scenarios that verify A .

These credences are all ideal credences: that is, credences that the subject should have on ideal rational reflection. We can assume that the idealization allows abstraction away from contingent cognitive limitations on memory, processing, and reasoning acumen, and that it allows arbitrary amounts of a priori reasoning. I need not claim that the ideal credences are always well-defined, but I think they are at least sometimes well-defined. In particular, if a subject supposes that they inhabit a certain scenario, as specified by a highly detailed canonical specification, then there will very frequently be rational verdicts about certain assertions conditional on this supposition.

We can apply this machinery to Olivia's assertion of 'Jekyll is Hyde'. If Olivia were given a complete qualitative description of her actual centered world @, she would not be given the information 'Jekyll is Hyde' directly. But she would be given qualitative information (using existential claims, for example) that enable her to determine that there is one man with various qualitative properties: who goes by both 'Jekyll' and 'Hyde', who came in for testing on the morning and evening of the same day, and so on for other aspects of his life and of his interaction with her. This would plausibly put her in a position where she should high credence in 'Jekyll is Hyde', and arguably should be certain of this if the specification is rich enough.¹⁵ This suggests that $pr(\text{'Jekyll$

¹⁴On an alternative definition (Chalmers 2006), the primary intension of an assertion of a sentence A is true at w iff a material conditional ' $S(w) \rightarrow A$ ' is a priori (setting aside complications due to context-dependent sentences). This definition will give the same result as the current characterization to the extent that for all A and w , the ideal conditional credence $p(A|S(w))$ is 1 iff ' $S(w) \rightarrow A$ ' is a priori. This claim is not obvious, but I argue (Chalmers forthcoming b) that the notions of apriority and of ideal conditional credence can be understood in such a way to make the claim true.

is Hyde' |@) is high and arguably 1. Given a description of a quite different centered world w , she could determine that she interacted with two quite different men on the relevant occasions, and so on, putting her in a position to conclude (arguably with certainty) that Jekyll is not Hyde. So $pr(\text{'Jekyll is Hyde'}|w)$ is low and arguably 0. So the primary intension of 'Jekyll is Hyde' will include @ and will exclude w . This is quite different from the primary intension of 'Jekyll is Jekyll', which will include both @ and w .

The same goes for Zaphod's assertion of 'I am Red'. Given a specification of the scenario of assertion @—that is, the centered world centered on the speaker and the time of the assertion—he will be in a position to determine (arguably with certainty) that he is the twin with the red star, and so to have high credence in 'I am Red'. Given a specification of an alternative centered world w centered on the twin with the blue star, he will be in a position to determine that he is the twin with the blue star, and so to have low credence in 'I am Red'). So $pr(\text{'I am Red'} |@)$ is high or 1, while $pr(\text{'I am Red'} |w)$ is low or 0. So the primary intension of 'I am Red' includes @ and excludes w , unlike the primary intension of 'Red is Red' which includes both.

All this illustrates the fact that primary intensions behave nonreferentially and avoid the problems for referential objects of credence. It should be noted that we have not had to antecedently assume a thesis such as descriptivism: the nonreferential behaviour of primary intensions follows directly from the nonreferential behaviour of credences in assertions.

The general principles underlying the analyses of the Jekyll-Hyde and Beeblebrox cases above are as follows. First, for any true assertion A , $pr(A|@) = 1$, where @ is the scenario of assertion. Second, for any epistemically possible assertion A , there exists a scenario w such that $pr(A|w) = 1$. We might call these claims Scrutability (all truths are recoverable from the truths in a canonical

specification of the actual world) and Plenitude (there are enough scenarios to verify every epistemically possible assertion) respectively. I will not attempt to argue conclusively for these key claims here, but I will set them out and discuss some consequences.

Plenitude: For all assertions A , if $pr(A) > 0$, then there exists a scenario w such that $pr(A|w) = 1$.

Plenitude (which we might also call Probabilistic Plenitude, to distinguish it from other plenitude theses) tells us that there are enough scenarios that sets of scenarios can serve as objects of credence whenever a subject's credence is nonzero. In effect, this ensures that epistemically possible scenarios stand to epistemic possibility (construed here as nonzero credence) as we would expect: whenever an assertion is epistemically possible in this sense, there is an epistemically possible scenario that verifies it.¹⁶ I have argued for a version of this thesis (cast in terms of apriority rather than probability), when scenarios are construed as centered worlds, in Chalmers 2009. Still, the requirement that there is a verifying centered world whenever rational credence is nonzero is a strong requirement, and I discuss some objections to it and possible refinements at the end of this section.

Scrutability: For all true assertions A , then if $@$ is the scenario of assertion, $pr(A|@) = 1$.

¹⁶If epistemic possibility is understood instead as the dual of apriority (S is epistemically possible when $\neg S$ is not a priori), the thesis cast in terms of epistemic possibility yields the version of Plenitude that is central in Chalmers 2006 and 2011 (and when scenarios are construed at centered worlds, it yields the Metaphysical Plenitude thesis of those papers). Probabilistic Plenitude is probabilistic analog of that thesis: the connection is particularly strong if one holds that a priori truths have ideal credence 1, although the two notions may still come apart for a posteriori claims with credence 1 ('I am conscious', perhaps).

Scrutability (which we might also call Conditional Scrutability) says that all true assertions are verified by the scenario of assertion: that is, the scenario inhabited by the speaker at the time of assertion. This ensures that the primary intension of true assertions contains the speaker's actual scenario, as it surely should. The thesis requires that a canonical specification $S(@)$ be such that $pr(A|S(@)) = 1$. As with Plenitude, Scrutability imposes a constraint on the vocabulary used for canonical specifications: $S(@)$ needs to be rich enough that all truths A can be recovered from it.

Scrutability is not quite entailed by Plenitude, but the two theses are closely related. When scenarios are centered worlds and canonical specifications are restricted to qualitative and indexical vocabulary, Scrutability requires that for all true assertions A (including nonqualitative assertions such as 'Jekyll is Hyde'), there is a qualitative and indexical truth S such that $pr(A|S) = 1$. Plenitude requires a related thesis: for all epistemically possible assertions A , there is an epistemically possible qualitative and indexical sentence S such that $pr(A|S) = 1$. These are both substantive theses; arguments for them (or for nearby theses) are given in Chalmers and Jackson 2001 and Chalmers forthcoming b.

A useful consequence of Plenitude is:

Coverage: For all assertions A , if $pr(A|w) = 1$ for all scenarios w , then $pr(A) = 1$.

Coverage is weaker than Plenitude, as can be seen by noting that it can be reformulated as the claim that if $pr(A) > 0$, then $pr(A|w) > 0$ for some w . The respects in which Coverage is weaker than Plenitude can be brought out by noting that Coverage would be satisfied if there were just two coarse-grained scenarios with canonical specifications 'It is raining' and 'It is not raining'. Still, Coverage captures the core idea that scenarios need to cover all of epistemic space.

These theses ensure that primary intensions avoid problems akin to those that plague a referen-

tialist view of the objects of credence. These problems arose because there are pairs of sentences or assertions A and B with the same object of credence (construed referentially) but such that $pr(S) \neq pr(T)$, for a subject at a time. If sentences A and B have the same primary intension, however, then $pr(A \leftrightarrow B|w) = 1$ for all w , from which it follows from Coverage that $pr(A \leftrightarrow B) = 1$, so $pr(A) = pr(B)$. Something similar goes for assertions.

We can use this result to characterize a subject's credence in a given primary intension (or set of scenarios) in terms of credences in sentences. A subject's credence in a given primary intension I at a time as the subject's credence in any sentence S whose primary intension is I (for the subject at the time). There will always be some such sentence (if only an infinitary disjunction of scenario specifications), and all such sentences will have the same credence, so this defines primary intensions as objects of credence. We can define conditional credences involving primary intensions in an analogous way.

We can then make the case that credences in primary intensions satisfy all the key Bayesian principles. Coverage ensures that the credence of the universal set of scenarios always be 1. Additivity of credences for disjoint sets of scenarios is ensured by additivity of credences for incompatible sentences. Bayesian updating for these sets will be ensured to the extent that Bayesian updating applies to sentences corresponding to these sets.

Of course all this is just using our grip on credences of sentences to help characterize primary intensions and our credences in them, and to make the case that they are coherent and well-behaved. Once we have done this, we can kick away the stepladder, and consider credences in primary intensions to be prior to the credences in the sentences that express them. So we have a reasonable case for primary intensions as objects of credence.

12 Challenges to Plenitude

Plenitude requires that there is a verifying scenario whenever a subject's ideal rational credence in a sentence or assertion is nonzero. This is a strong requirement, especially when given the constraint that scenarios are identified with centered metaphysically possible worlds. We have seen that this model of scenarios can handle the cases involving names and indexicals that caused problems for referentialism, but other challenges remain. In what follows I discuss some potential counterexamples: sentences S that arguably may have nonzero ideal credence while being verified by no centered world. Most of these cases (except perhaps challenge 2) can also be used to pose challenges to Scrutability.

Challenge 1: *Cognitive impairment* (Christensen 2008): Let L be a complex logical truth and let H be the empirical hypothesis that one is unreliable at logic. It is arguable that the ideal credence $pr(H)$ is nonzero (one cannot be certain that one has not just been given an anti-logic drug) and that the ideal conditional credence $pr(L|H)$ is less than one (ideally one should suspend judgement in L conditional on H). If so, the unconditional ideal credence $pr(L) < 1$, even though L is true in all worlds.

Challenge 2: *Borderline cases*: Suppose that 17 is a borderline case of a small number. Then one might have nonzero credence in '17 is a small number', even though there is no centered world that verifies it.¹⁷

¹⁷Another way to bring out problems involving vagueness is to note that there is a prima facie argument from Plenitude to a Determinacy thesis saying that for all A and w , $pr(A|w)$ is either 0 or 1. If $pr(A|w) < 1$, then if the speaker became certain that $S(w)$, their new credence $pr(A) > 0$. Then $pr(A \& S(w)) > 0$, where $A \& S(w)$ is the obvious conjunctive assertion. By Plenitude and simple reasoning it follows that $pr(A|w) = 1$. It is worth noting that Determinacy plus Coverage entails Plenitude and Plenitude entails Coverage; so to the extent that this argument is valid

Challenge 3: *The two-tubes case* (Austin 1990): Fred is looking down two tubes, one attached to each eye, and has a symmetrical experience as of two red balls. Fred is objectively omniscient and knows that in fact that tubes are connected to one red ball and one orange ball. He also knows his own location in the world and the current time, and so knows which centered world he inhabits. When he entertains the hypothesis *That is red and that is orange* (using two simultaneous perceptual demonstratives), he is not in a position to determine that it is true and has rational credence 0.5 in it, but there is no verifying centered world.

Challenge 4: *Strong necessities*: Strong necessities (Chalmers 2009) are a posteriori necessities that are unlike standard Kripkean a posteriori necessities (such as 'Jekyll is Hyde') in that strong necessities are verified by all centered worlds. For present purposes, we can take strong necessities to include any truth that is verified by all centered worlds (such truths include any qualitative necessary truths) without being knowable a priori with certainty. For example, suppose that 'There is an omniscient being' is metaphysically necessary but not knowable a priori. Then 'There is no omniscient being' will plausibly have nonzero ideal credence for some subjects, but no centered world will verify it. One might think something similar applies to unprovable mathematical truths, moral truths, metaphysical truths, and so on.

I will not try to provide conclusive responses to these cases, but I briefly explore the options for responding. In general there are four options for responding to such counterexamples: (i) hold that credence is 0 or 1 given an appropriate idealization; (ii) hold that there is a verifying world; (iii) move to a more fine-grained model of scenarios that provides a verifying scenario, or (iv) reformulate Plenitude to be compatible with the case.

(there is a possible loophole for credences that are indeterminate between 0 and other values), Plenitude is equivalent to Determinacy plus Coverage.

Appealing to idealization is perhaps the most natural for cases involving logical and mathematical truths. It is arguable that our own inability to know certain unprovable mathematical truths is due to contingent cognitive limitations that disappear on idealization. The case of cognitive limitations is trickier, but it is arguable (Chalmers forthcoming b) that a Bayesian should appeal to a special sort of idealization on which $pr(L|H)$ and $pr(L)$ are 1: an ‘insulated idealization’ on which subjects’ first-order reasoning is insulated from certain effects of higher-order beliefs about their cognitive capacities.

The reformulation strategy is most natural in the case of vagueness. One might simply formulate Plenitude to exclude sentences involving vague predicates, and then later give such sentences special treatment, perhaps involving supervaluation. Another reformulation strategy modifies the thesis to say that if $pr(det(A)) > 0$, then there exists w such that $pr(A|w) = 1$, where here $det(A)$ will be ‘Determinately, 17 is a small number’. Another strategy holds that the sentence has a zero credence, or perhaps better, a credence that is indeterminate between zero and other values, corresponding to indeterminacy in whether there is a verifying scenario (one might also apply a version of this move on the previous strategy, holding that in cases of higher-order vagueness. I will not try to choose between these options here.

The fine-graining option is often attractive. There is nothing sacrosanct about centered worlds: just as we moved from metaphysically possible worlds to centered versions of such worlds, we might move if necessary from centered metaphysically possible worlds to more fine-grained objects still. I think that a limited sort of fine-graining is certainly required to handle the two-tubes case.¹⁸ Here the best option is to reconceive centered worlds so they have more information in the center. As well as a subject and a time, they will contain one or more marked experiences,

corresponding to the bases of perceptual demonstratives. *That is red* will be true at such a centered world roughly when the object that causes the corresponding marked experience is red. Then Fred's credence will be distributed over two centered worlds, in which the relevant marked experience is caused by the red and the orange ball respectively. This method also provides a natural treatment of the hall-of-mirrors case and other cases involving perceptual demonstratives.

Fine-graining is also available in the case of strong necessities. My own view (Chalmers 2009) is that there are no strong necessities: truths verified by all centered worlds are always knowable a priori (under an appropriate idealization), and in fact are knowable a priori with certainty. So the candidate necessary truths will either have ideal credence 1 (as with unprovable mathematical truths), or will be contingent or indeterminate rather than necessary (perhaps as with 'There is an omniscient being'), or will have centered worlds that verify their negations (as with Kripkean a posteriori necessities such as 'Jekyll is Hyde' and 'Water is H₂O'). So on my own view, we do not need to fine-grain scenarios beyond centered worlds here. But many philosophers will disagree on this issue, and if so, the fine-graining strategy is available to them.

The natural way to fine-grain scenarios here is to introduce epistemically possible scenarios not as a construction from metaphysically possible worlds but as entities in their own right. For example, one might construct them in terms of maximal a priori consistent sets of sentences in an ideal language (see Chalmers 2011). We can use those sentences to provide canonical specifi-

¹⁸Proponents of standard centered-worlds theories may try appealing to associated descriptions such as 'The object I am attending to' or 'The object on the left', but Austin (1990) makes a convincing case that these moves do not work in general. I think the two-tubes case decisively refutes standard centered-worlds theories of the objects of belief, such as Lewis's property theory: it stands to those theories much as Lewis's two-gods case and the Beeblebrox case stand to possible-worlds theories.

cations of scenarios, and define verification of sentences by scenarios as before. Then, if neither ‘There is an omniscient being’ or its negation is a priori, then there will be scenarios verifying both of these sentences, even if one of them is metaphysically necessary. More generally, assuming that we understand apriority and ideal credences so that a priori sentences always have ideal credence 1 for all subjects, then given reasonable assumptions, any sentence with nonzero ideal credence will be verified by some scenario.

These fine-grained epistemically possible scenarios will be available to handle many potential counterexamples to Plenitude: if one thinks moral or metaphysical claims are counterexamples, for example, then one can have epistemically possible scenarios in which the moral or metaphysical truths are different from those that are metaphysically necessary. If one thinks that phenomenal truths are necessitated but not a priori entailed by physical truths, one might have epistemically possible scenarios in which the physical truths are the same as out world but the phenomenal truths are different. On an epistemic view of vagueness, one might even have scenarios that verify ‘17 is a small number’ and scenarios that verify its negation.

On the fine-grained view, one can still identify the objects of credence with primary intensions, construed as sets of epistemically possible scenarios, as long as one does not identify scenarios with centered worlds. If one denies that there are strong necessities (as I do), then the entities that result will have a close correspondence to centered metaphysically possible worlds, but the construction may still be useful for some purposes: there is something to be said for understanding epistemic space in epistemic terms.

13 Other explanatory roles

So far I have argued that the objects of credence are primary intensions: either sets of centered worlds, or sets of more fine-grained epistemically scenarios. My arguments for this thesis have concentrated on the explanatory role of analyzing a subject's ideal credences. There are other explanatory role that we may want objects of credence and belief to play, however: to help understand nonideal reasoners, to analyse the sameness of belief across subjects and across times, to analyse attributions of propositional attitudes, and to connect credence with chance, for example. These roles give rise to some natural objections to the current proposal, and in some cases suggest finer-grained candidates for the objects of credence.

Objection 1: All a priori truths have the same primary intension: the set of all scenarios. But nonideal reasoners can have different credences in different a priori truths: for example, I might have credence 1 that $2+2=4$ and credence less than one in Fermat's Last Theorem. Understanding these credences requires finer-grained objects of credence.

Response: One can develop finer-grained objects of credence for nonideal reasoners by introducing *structured primary intensions*.¹⁹ To do this, we assume that thoughts are structured entities, with a logical form involving simple concepts. We can then extend the earlier framework so that concepts are associated with their own primary intensions: functions from scenarios to extensions. We can then associate a thought with a structured entity composed from the primary intensions of its simple components, structured according to its logical form. Then a thought that $2+2=4$ and a thought whose content is Fermat's Last Theorem will be associated with quite different structured primary intensions.

¹⁹One could also try understanding objects of credence as sets of very fine-grained scenarios at which logical and mathematical truths can be false. But as Bjerring (2010) demonstrates, there are significant problems for constructing

Objection 2: When I accept ‘You are hungry’ and you accept ‘I an hungry’, there is a sense in which we accept the same thing, but our primary intensions (and our structured primary intensions) are quite different.

Response: We can supplement a structured primary intension with information about extension. Let us say that the *enriched intension* of a concept is an ordered pair of its primary intension and its extension, and that the *enriched proposition* associated with a belief is a structure composed of the enriched intensions of its simple components, structured according to its logical form. Then a beliefs enriched proposition determines a structured primary intension, but it also determines a Russellian proposition: a structure composed of the extensions of the concepts. In cases such as the above, the enriched propositions associated with a thought will have the same Russellian component, capturing the sense in which the two subjects accept the same thing.

Objection 3: Chance-credence principle, such as Lewis’s Principal Principle, almost always require that the objects of credence are the objects of chance. But the objects of chance are certainly not primary intensions: the chance of an electron decaying is not sensitive to a mode of presentation of the electron, for example.

Response: The fine-grained structure of enriched propositions allows them to serve both as bearers of credence and as bearers of chance, in a way that serves to connect the two. We can say that a subject’s credence in an enriched proposition is their credence in the associated structured primary intension, while the chance of an enriched proposition is the chance of the associated Russellian proposition. For example, one version of the Principal Principle holds that for a propo-

scenarios of this sort. A residual option is to go all the way to understanding scenarios as arbitrary sets of sentences with no interesting constraints, but intensions defined over these scenarios have much less explanatory power than the structured entities above.

sition p , if one is certain (without inadmissible information) that the chance of p is ϕ , then one's credence in p should be ϕ . Here, the left-hand side should be understood (like the right-hand side) as requiring that p is conceived under the mode of presentation corresponding to its structured primary intension. For example, if p is the enriched proposition expressed by 'Hesperus is a planet', both sides will require conceiving p under the mode of presentation associated with 'Hesperus'.²⁰

Objection 4: Primary intensions give the wrong results in some cases of belief updating. Suppose that at 3pm, I have credence 1 that it is now 3pm, while at 4pm, I have credence 0 that it is now 3pm. These credences are fully rational. At both 3pm and 4pm, these credences are associated with the same set of centered worlds: worlds at which the time at the center is 3pm. But on the standard Bayesian model, the only way to rationally update one's credence is by conditionalization, and it is impossible for one's credence in the same object to change from 1 to 0 by conditionalization. So the objects of credence here are not centered worlds.²¹

Response: The claim that it is impossible for a credence to change from 1 to 0 by conditionalization is grounded in claim that when $p(P) = 1$, then $p(P|E) = 1$ for all E , which is itself grounded

²⁰One might object that in effect, enriched propositions yoke together objects of credence (structured primary intensions) and objects of chance (Russellian propositions) into a single entity. After all, credence does not really depend on the Russellian component, and chance does not really depend on the mode-of-presentation component. There is something to this, but at best, it suggests that enriched propositions are not the *fundamental* objects of credence or of chance, not that they are not objects of credence or of chance. One could also formulate modified chance-credence principles under the assumption that objects of credence are structured primary intensions and objects of chance are Russellian propositions. For example, a modified Principal Principle might say that if one is certain (without inadmissible information) that the chance of p is ϕ , where p (a Russellian proposition) is entertained under guise q (a structured primary intension), then one's credence in q should be ϕ .

²¹Issues in this vicinity are discussed by Arntzenius (2003), Hitchcock (2004), and Titelbaum (2008), among others.

in the ratio formula for conditional probability: $p(P|E) = p(P \& E)/p(E) = p(E)/p(E) = 1$. But it is well-known (Popper 1959; Hajek 2003) that there are good reasons to reject the ratio formula in cases where $p(E) = 0$.²² So one might hold that if P is the set of centered worlds centered on 4pm and E is the set of centered worlds constituting one's evidence at 4pm, then at 3pm $p(P) = p(E) = 0$, but $p(P|E) = 1$. If so, one can update by conditionalization upon getting evidence E at 4pm here, resulting in $p(P) = 1$ then.

Someone might deny that in this case $p(P|E) = 1$ at 3pm: one may have a high credence in P 's becoming the case given that E becomes the case, but not a high credence in P 's being the case given that E is the case.²³ After all, if we discovered E at 3pm (immediately after looking at clocks saying 3pm), we would probably conclude that something is very wrong and suspend judgement about the time.

However, if E is total evidence including the evidence of memory and introspection, then conditionalizing on E should trump any current conflicting evidence, and should lead one to accept P . So I think $p(P|E)$ is plausibly 0 at 3pm. But if one denies that, one could instead hold that one has a *pseudo-conditional credence* $pcc(P|E)=0$, where a pseudo-conditional credence is a state

²²One might object that one cannot have credences conditional on propositions that one is certain is false. But this seems wrong: even when one is certain that an empirical proposition E is false, it is plausible that $p(E|E) = 1$, and likewise that $P(E \vee F|E) = 1$, and so on.

²³Alternatively, an opponent might see the relation between P and E as the sort of subjunctive or 'imaging' credence discussed by Joyce (1999), very roughly reflecting a low credence in 'if E were the case, P would be the case'. But I think this is less clear. One could argue that the nearest possible world in which E is the case is one in which everything is otherwise as it currently is (so that P is true) except that one has a misleading experience (so that E is true). If so, then the subjunctive credence in P given E will be high. For this reason, I think the future-directed version is a more promising alternative here.

capturing the relationship between P and E in the case above (roughly, an expectation about what will be the case on obtaining certain evidence). One can then embrace a principle of *pseudo-conditionalization*, holding that for a rational subject, when $pcc(P|E) = \phi$ and the subject then obtains total evidence E , then $p'(P) = \phi$. This principle will then deliver the desired result in the case above. Of course the principle of pseudo-conditionalization goes beyond the principle of conditionalization. But it is widely believed that a correct account of updating temporal belief requires going beyond conditionalization in any case, so this is not a huge cost.

Objection 5: Primary intensions give a poor account of the objects of belief when these are construed as the referents of that-clauses. For example, to say that ‘John believes that I am hungry’ is not to say that John has high credence in the primary intension expressed by ‘I am hungry’: doing so would require John to believe that he is hungry, not that I am hungry.

Response: Primary intensions are not plausible candidates for the referents of ‘that’-clauses, but enriched propositions are. In ‘Propositions and Attitude Ascriptions: A Fregean Account’, I argue for a treatment of ‘S believes that P’ on which ‘that P’ refers to the associated enriched proposition p , and on which the ascription requires that S endorses an enriched proposition co-ordinate with p , where co-ordination requires the same Russellian component and an appropriately related structured primary intension component.²⁴ This framework handles cases such as the one above straightforwardly.

The co-ordination relation might also be invoked to help analyse issues concerning the interpersonal relations and the dynamics of temporal belief, as objections 2 and 4. If I believe ‘You are hungry’ and you believe ‘I am hungry’, we believe co-ordinate enriched propositions. And if I believe ‘I am hungry now’ at 3pm and ‘I was hungry then’ at 4pm (referring to 3pm), I believe

co-ordinate enriched propositions. Co-ordination can then be used to capture the sense in which we believe the same things in these cases.

14 Conclusion

For the purposes of Bayesian confirmation theory narrowly construed, it may suffice to take the objects of credence and of belief to be primary intensions. But for the purposes of a broader theory of belief, it makes sense to take the objects of credence and of belief to be enriched propositions, in order to best handle phenomena such as nonideal reasoning, interpersonal relations between beliefs, and attitude ascriptions. Not much turns on the question of which is truly the object of credence, though, once we are clear on the explanatory roles that are played by each.

A theme of this article has been that to understand objects of belief, one needs to attend to roles that these objects need to play, and to determine what objects can play these roles. Roles tied to confirmation theory are just as important as roles tied to attitude ascriptions, for example. It remains an open question whether there is any single object that can play all the roles that we would like objects of belief to play, and we should not be surprised if it turns out that different roles are best played by different objects. Still, the analysis of credence suggests that attention to a multiplicity of roles may yield significant rewards in the analysis of objects of belief.²⁵

School of Philosophy

DAVID J. CHALMERS

²⁵Thanks to Alan Hajek, Matt Kotzen, Bernhard Nickel, Susanna Siegel, Scott Sturgeon, and audiences at ANU, Beijing, Cologne, UC Davis, and the Carolina Metaphysics Workshop for comments. Thanks also to three referees and the editor of *Mind*, and to David Braun for his commentary on an earlier version of this paper (then called 'Probability and Propositions') in the 2006 Online Philosophy Conference.

Australian National University

Canberra ACT 0200, Australia

Department of Philosophy

New York University

New York NY 10003, USA.

chalmers@anu.edu.au

References

Arntzenius, F. 2003: Some problems for conditionalization and reflection. *Journal of Philosophy*, 100, pp. 356–70.

Austin, D.F. 1990: *What's the Meaning of 'This'?* Ithaca, NY: Cornell University Press.

Bjerring, J.C. 2010: *Non-Ideal Epistemic Spaces*. Ph.D. thesis, Australian National University.

Braun, D. 2000: Russellianism and psychological generalization. *Nous*, 34, pp. 203–236.

Chalmers, D.J. 2006: The foundations of two-dimensional semantics. In Garcia-Carpintero and Macia 2006.

— —2009: The two-dimensional argument against materialism. In McLaughlin and Walter 2009. Expanded version in Chalmers 2010.

— —2010: *The Character of Consciousness*. New York: Oxford University Press.

— —2011: The nature of epistemic space. In Egan and Weatherson 2011.

— —forthcoming a: Propositions and attitude ascriptions: A Fregean account. *Nous*. <<http://consc.net/papers/proposit>>

— —forthcoming b: *Constructing the World*. Oxford: Oxford University Press.

<<http://consc.net/constructing/>>.

- —and Jackson, F. 2001: Conceptual analysis and reductive explanation. *Philosophical Review*, 110, pp. 315–61.
- Christensen, D. 2008: Does Murphy’s law apply in epistemology? Self-doubt and rational ideals. *Oxford Studies in Epistemology*, 2, pp. 3–31.
- Earman, J. (ed.) 1983: *Testing Scientific Theories*. Minneapolis: University of Minnesota Press.
- —1992: *Bayes or Bust? A Critical Examination of Bayesian Confirmation Theory*. Cambridge, MA: MIT Press.
- Egan, A. and Weatherson, B. (eds.) 2011: *Epistemic Modality*. Oxford: Oxford University Press.
- Elga, A. 2000: Self-locating belief and the Sleeping Beauty problem. *Analysis*, 60, pp. 143–47.
- Garber, D. 1983: Old evidence and logical omniscience in Bayesian confirmation theory. In Earman 1983.
- Garcia-Carpintero, M. and Macia, J. (eds.) 2006: *Two-Dimensional Semantics*. Oxford: Oxford University Press.
- Glymour, C. 1980: *Theory and Evidence*. Princeton: Princeton University Press.
- Hajek, A. 2003: What conditional probability could not be. *Synthese*, 137, pp. 273–323.
- Hitchcock, C. 2004: Beauty and the bets. *Synthese*, 139, 405–420.
- Jeffrey, R.C. (ed.) 1980: *Studies in Inductive Logic and Probability, Vol. II*. Berkeley: University of California Press.
- Joyce, J. 1999: *The Foundations of Causal Decision Theory*. Cambridge: Cambridge University Press.
- Lewis, D. 1980: A subjectivist’s guide to objective chance. In Jeffrey 1980.
- —1983: Attitudes de dicto and de se. *Philosophical Review*, 88, pp. 513–43.

- McLaughlin, B., Beckermann, A., and Walter, S. (eds.) 2009: *The Oxford Handbook of Philosophy of Mind*. Oxford: Oxford University Press.
- Popper, K. 1959: *The Logic of Scientific Discovery*. New York: Basic Books.
- Putnam, H. 1975a: The meaning of ‘meaning’. In his 1975b.
- —1975b: *Mind, Language, and Reality*. Cambridge: Cambridge University Press.
- Renyi, A. 1955: On a new axiomatic theory of probability. *Acta Mathematica Academiae Scientiarum Hungaricae*, 6, pp. 285–335.
- Salmon, N. 1986: *Frege’s Puzzle*. Cambridge, MA: MIT Press.
- Stalnaker, R. 1984: *Inquiry*. Cambridge, MA: MIT Press.
- —2008: *Our Knowledge of the Internal World*. Oxford: Oxford University Press.
- Titelbaum, M. 2008: The relevance of self-locating beliefs. *Philosophical Review*, 117, pp. 555–606.