Probability and Propositions

David J. Chalmers
Philosophy Program
Research School of Social Sciences
Australian National University

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. But the nature of propositions is itself contested. What is a proposition, such that it can serve as an object of belief?

In philosophy in recent years, tools from the philosophy of language have often been brought to bear on this problem. Analyses of the objects of linguistic assertion and about the reference of 'that'-clauses have been used to support various views of propositions, and these views have then been taken as suggestions about the objects of belief.

A popular class of views, strongly influenced by direct reference theories in the philosophy of language, involves *referentialism* about the objects of belief. Referentialist views say that insofar as beliefs attribute properties to individuals (e.g. the belief that Nietzsche is dead), the objects of these beliefs are determined by those individuals (e.g. Nietzsche) and those properties (e.g. the property of being dead). On one such view, the objects of belief are Russellian propositions composed from the individuals and properties that one's belief is about. On another such view, the objects of belief are sets of possible worlds in which the individuals in question have the relevant properties.

In this paper I will approach the objects of belief from a different starting point. The starting point will be 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. I will approach the question of what objects of belief are by asking: what sort of objects can play the role that Bayesian confirmation theory needs them to play?

I will argue that if the objects of belief are understood this way, referentialism 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.

The arguments involve cases in which a subject has two different ways of thinking about the same object, with apparent differences in associated credences. As such, the argument can be considered a probabilistic version of Frege's puzzle, with credence playing the role of cognitive value. The discussion of Frege's puzzle in the philosophy of language has reached something of a stalemate in recent years, so 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. So it is at least enlightening to work through these issues as they appear in a probabilistic context.

Attention to these issues is also of independent interest in understanding the interpretation of probability theory as it is used in the analysis of belief. 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 may be 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. Toward the end of the paper I will sketch an alternative nonreferential account of the objects of credence, and will discuss their relationship with the objects of chance.

I take referentialism to be committed to at least the following claims. If A and B are two names for the same object, then in believing that A has ϕ, one believes the same proposition that one believes in believing that B has ϕ. Likewise, the belief that one expresses when one asserts 'A has ϕ' has the same propositional object as the belief that one expresses when one asserts 'B has ϕ'.

I take Bayesianism to be committed to the following claims. Individuals have unconditional and conditional credences in propositions. When a hypothesis H and and an evidential proposition E are such that one's prior unconditional credences in H and E are p(H) and p(E) and one's prior conditional credence in E given H is p(E|H), and if one's 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).

2 Two cases

The Jekyll case

Utterson knows that Dr. Jekyll has P, but he has no information about whether he has Q. He knows that 1 in 10 people with P have schizophrenia, that 1 in 5 people with property Q have schizophrenia, and that 9 in 10 people with P and Q have schizophrenia. Let JS be the proposition that Jekyll has schizophrenia, and let JQ be the proposition that Jekyll has Q. Then Utterson's current credences are as follows:

(J1) $p(JS) = 0.1$

(J2) $p(JS \mid JQ) = 0.9$

Shortly thereafter, Utterson sees Mr. Hyde and observes that he has Q. If HQ is the proposition that Mr. Hyde has Q, then

(J3) $p(HQ) = 1$.

Now Dr. Jekyll is in fact Mr. Hyde, although Utterson has no evidence for this and has no idea that this is so. If referentialism is true, it follows that JQ = HQ. But now, it follows from Bayesian principles that

(J4) $p'(JS) = p(JS \mid JQ) = 0.9$.

That is, Utterson should update his beliefs so that he has credence 0.9 in the hypothesis that Jekyll has schizophrenia. But that seems wrong. He is in no position to raise his credence, since he has no evidence that Jekyll is Hyde. So referentialism combined with Bayesian principles yields a false prediction here. So either referentialism or Bayesianism is false.

The Tinasky case

Here is the first paragraph of the Wikipedia entry on Wanda Tinasky:

> Wanda Tinasky, ostensibly a bag lady living under a bridge in the Mendocino County area of Northern California, was the pseudonymous author of a series of playful, comic and erudite letters sent to the *Mendocino Commentary* and *Anderson Valley Advertiser* between 1983 and 1988. These letters were later collected and published as *The Letters of Wanda Tinasky*. In them, Tinasky weighs in on a variety of topics – most notably local artists, writers, poets and politicians – with an irreverent wit and literate polish at odds with her apparently

straitened circumstances. ... At the time, the identity of Tinasky was completely unknown, and subject to much local speculation. Tinasky was thought by many to be novelist Thomas Pynchon, but is now widely believed to be an obscure Beat Generation poet named Tom Hawkins.

The crucial step was taken by English professor Don Foster in 1998. He discovered striking similarities between the texts of Tinasky and Hawkins, while many fewer similarities between the texts of Tinasky and Pynchon. Let us suppose, for simplicity, that before he examined the texts, Foster had ruled out all hypotheses identifying Tinasky with other candidates other than Pynchon and Hawkins, and that he was indifferent between those two. Let us also suppose that the crucial evidence was new evidence that Foster had never seen before.

A Bayesian might represent all this as follows. Let H be the proposition that Tinasky is Hawkins, and let E be a proposition characterizing the evidence Foster acquired from the texts. Then Foster's prior credences, before examining the texts, could be represented as follows:

(T1) $p(H) = 0.5$

(T2) $p(\sim H) = 0.5$

(T3) $p(E|H) = 0.99$

(T4) $p(E|\sim H) = 0.01$

Bayesianism entails that after examining the texts, Foster should update his credences as follows.

(T5) $p'(H) = p(E|H)p(H)/p(E) = 0.99$

(T6) $p'(\sim H) = p(E|\sim H)p(\sim H)/p(E) = 0.01$

This appears to be a perfectly standard application of Bayesian reasoning. It appears to yield both a reasonable normative account of a process that Foster should have gone through, and a reasonable descriptive account of a process that Foster might in fact have gone through, assuming that he 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 H_2O, that Twain is Clemens, and so on.

If this application of Bayesian principles is correct, however, then referentialism is false. Suppose for the purposes of *reductio* that referentialism is true. Then the proposition that

Tinasky is Hawkins is identical to the proposition that Tinasky is Tinasky. (Here and in what follows, I will assume that Tinasky really is Hawkins.) This is presumably a proposition in which Foster has credence 1, as he is certain that Hawkins is Hawkins. (The alternative of suggesting that Foster has credence 0.5 in the proposition expressed by 'Hawkins is Hawkins' leads to problems that are analogous but even worse.) Likewise, the proposition ~H is identical to the proposition expressed by 'it is not the case that Hawkins is Hawkins', in which Foster presumably had credence 0. So $p(\sim H) = 0$, both before and after acquiring the crucial evidence. So (T1) and (T3) above are incorrect. Contrapositively: if this application of Bayesian principles is correct, then referentialism is false.

3 Analysis

The Tinasky and Jekyll cases can be seen as complementary. In effect, the Tinasky case brings out the *success* of Bayesian theory when it is employed in a nonreferential way, while the Jekyll case brings out the *failure* of Bayesian theory when it is employed in a referential way.

The dialectical force of the two examples is correspondingly somewhat different. In the Tinasky case, a referentialist might respond flat-footedly by saying that the description of the case begs the question against them, as their position is committed to the falsity of (T1). The residual burden here for the referentialist is that of explaining the success of the Bayesian reasoning that Foster employs. In the Jekyll case, an analogous flat-footed response requires the claim that Utterson really should update his beliefs so that his credence in the hypothesis that Jekyll has schizophrenia is 0.9. But that seems clearly incorrect. The residual burden here for the referentialist is that of either justifying this claim, or explaining why referentialism is not committed to it.

To help organize the 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 'Tinasky is Tinasky' and 'Tinasky is Hawkins', we can avoid begging any questions about whether these assertions express different propositions.

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

(1) "Tinasky is Hawkins."

"Wanna bet?".

Furthermore, any *pair* of assertions A and B 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 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. (Such an association does not require the framework of betting odds, but that framework provides useful motivation.) 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). (So "pr" will be used for credences associated with assertions, and "p" for credences associated with propositions.)

We can now consider versions of the cases above that involve assertions. For example, in the Tinasky case, corresponding to claims (T1)-(T6) about credences of assertions, we can construct (T1')-(T6') about credences of assertions. Here H' is a prior assertion by Foster of 'Tinasky is Hawkins' and E' is a prior assertion of 'the evidence will be such-and-such'. In addition, for later purposes we can let T' be a prior assertion by Foster of 'Tinasky is Tinasky'. For an assertion A, pr(A) is Foster's prior credence in an assertion and pr'(A) is what his later rational credence should be.

(T0') pr(T') = 1

(T1') pr(H') = 0.5

(T2') pr(~H') = 0.5

(T3') pr(E'|H') = 0.99

(T4') pr(E'|~H') = 0.01

(T5') pr'(H') = pr(E'|H')pr(H')/pr(E') = 0.99

(T6') pr'(~H') = pr(E'|~H')pr(~H')/pr(E') = 0.01

Now it seems very plausible that whatever the status of (T1)-(T6), something like (T1')-(T6') could be correct. For example, it seems plausible that the rational betting odds that Foster associates with an *assertion* of 'Tinasky is Hawkins' may well be 0.5. It is plausible that the rational bettings odds associated with assertions H' and E' might yield pr'(E'|H')=0.99. And if so, it seems plausible, according to Bayesian reasoning, that Foster

should update his credence in the original assertion H' so that pr'(H')=0.99. Note that these claims about assertions do not immediately commit us to claims about propositions. But this sets things up for different sorts of referentialist response, depending on their attitude to (T1')-(T6').

One sort of referentialist might *reject* the application of Bayesian reasoning to this case, in effect by rejecting (T1')-(T6') as well as (T1)-(T6). They might do this by repudiating Bayesianism wholesale, perhaps rejecting the whole idea of credences, or perhaps rejecting the claim that they can be rationally updated in a Bayesian way. Or they might do this by invoking the need for idealized conditions in the application of Bayesian reasoning, holding that these conditions fail in the cases above.

Alternatively, a referentialist might *reconstruct* the application of Bayesian reasoning to this case in such a way that it supports the intuitively correct results (T1')-(T6') without contradicting referentialism by endorsing (T1)-(T6). They might do this by relativizing the credences of propositions in some fashion, so that relativized versions of (T1)-(T6) 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 (T1)-(T6) invoking these entities are true.

In what follows I will discuss all of these strategies in turn.[1]

Although referentialism about thought 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 a sentence. Here one would need simply claims (T0') and (T1') above, and the additional premise that a subject's credence in an assertion is identical to the subject's credence in the proposition expressed by assertion. From these, it follows that assertions of

---

[1] There are other, less important replies that I will not discuss at length. One reply holds that in the Jekyll case, Utterson really does have credence 0.9 in the proposition that Jekyll has schizophrenia, although he does not realize it. It might be thought that this reply is analogous to the referentialist view that Lois Lane really does believe that Superman is Clark Kent, although she does not realize it. But there is no analogy here: the claim about Lois is analogous at best to the referentialist claim that Utterson has credence 0.2 in the proposition that Jekyll has schizophrenia, without realizing it, in virtue of the associated "Hyde has schizophrenia" mode of presentation. There is no candidate mode of presentation under which Utterson associates the proposition with credence 0.9.

Alternatively, a referentialist may appeal to restrictions on evidence, perhaps holding that evidential propositions are qualitative rather than object-involving, 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 P' and 'Hyde has Q' have credence 1, we can generate the problematic results.

'Tinasky is Tinasky' and 'Tinasky and Hawkins' express different propositions, which entails that referentialism about language is false.

A related problem for referentialism about language follows from the analysis of probabilistic claims themselves. For example, we might have described the Tinasky case as follows.

(L1) p(Tinasky is Tinasky) = 1

(L2) p(Tinasky is Hawkins) = 0.5

If (L1) and (L2) are correct, and if referentialism about language is construed so that it applies to such sentences, then referentialism about language is false. Of course these sentences are at best sentences of a technical version of English, so that it is not obvious that referentialism must apply to them, but they also have nontechnical counterparts. In these cases, moves analogous to those discussed below will once again be options for the referentialist., Many of the same issues will arise, although there will be differences in some cases. I will focus mainly on referentialism about thought, but it is worth keeping in mind these related problem for referentialism about language.[2]

Strategy 1: Reject Bayesianism

Most straightforwardly, a referentialist could reject Bayesianism entirely. I think that this is the worst option. First, while Bayesianism has controversial aspects (tied to the origin of prior probabilities, for example), 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 (T1), or from the obvious fact that Foster has credence between 0 and 1 in the hypothesis that Tinasky is Hawkins. Taken alone, however, this would lead to a familiar dialectic in which the referentialist denies (T1) while giving a pragmatic explanation of its intuitive plausibility.

---

[2] The most likely strategies for the referentialist about language are to say that (L2) above is false (although pragmatically assertible), or to say that a true reading for (L2) requires that a mode of presentation be tacitly specified. The second option is most naturally combined with strategy 3 or 4 below. The first option may also be combined with strategy 1 or 2.

Once (T1) is embedded in the machinery of a successful theory, then this denial has a much greater cost. However, many different theories of confirmation will suffice to generate an argument analogous to the argument here. All we need is a theory that allows that confirmation involves increase in credences, and that identity 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 H_2O', 'electricity is magnetism'). Likewise, it is a datum that highly rational agents can associate partial 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.

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, it is well known that probability theory requires an idealization according to which all logical truths are assigned probability 1. This idealization leads to the following problem cases, which are structurally reminiscent of the Tinasky and Jekyll cases:

> **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 not logically omniscient. E is a piece of evidence that Jones has just acquired, and F is a claim that follows from E only when conjoined with a highly 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 might say the analogy between the Smith and Jones cases and the Tinasky and Jekyll cases 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, Foster and Utterson are in non-ideal circumstances. All of them lack knowledge of certain necessary truths: logical truths in one pair of cases, and truths about identity in the other. Foster is like Smith: in these cases, intuitive assignments of credence contradict Bayesian verdicts, but this is to be expected in non-ideal circumstances. Utterson 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. 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 judgments about non-ideal creatures.

Here, the most obvious reply is that unlike Smith and Jones, neither Foster nor Utterson appear to be reasoning non-ideally at all. We can assume that both of them are Vulcans, with extraordinary reasoning acumen, and nothing about the cases will change. As Bayesian principles are usually understood, the idealization required to apply them involves this sort of acumen, but nothing more. So there should be no problem with the application of these principles to the Tinasky and Jekyll cases.

In fact, nothing in standard probability theory requires that all necessary truths be assigned probability 1. The axioms of probability theory are often formulated to require that all logical tautologies 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 (omega) is assigned probability 1. But this does not entail that necessary truths are assigned probability 1, unless we make the additional assumption that the probability of a statement is the probability of a corresponding set, and that all necessary truths correspond to the universal set. Nothing in probability theory or in Bayesianism requires this assumption. Later in the paper, I will outline a framework on which identity statements (such as 'Tinasky is Hawkins') correspond to non-universal sets,

for the purposes of Bayesian analysis. So an idealization requiring knowledge of all true identity statements is certainly not required by Bayesianism.[3]

Apart from being unnecessary, this idealization is unreasonable. Knowledge of all true identity statements requires something roughly equivalent to total omniscience about the state of the universe. For example, for any statement 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.

Strategy 3: Relativize credences

Henceforth I will set aside the rejection and restriction strategies. It is near-obvious that something like (T0')-(T6'), at least, are correct, and that Bayesian principles can be applied to them. The best option for a referentialist is to show how the truth of these claims is

---

[3] An added discomfort for the referentialist: even if a Bayesian idealization requires that subjects know all necessary *propositions*, then referentialism entails that Foster and Utterson know the relevant identity propositions, in virtue of knowing that Tinasky is Tinasky or that Jekyll is Jekyll. So the referentialist needs the stronger requirement that a subject knows these propositions under every mode of presentation, or that the subject has credence 1 associated with any sentence expressing such a proposition. But the basis for strengthening the requirement in this way is unclear.

compatible with referentialism. This requires giving an account of the relation between assertions, propositions, and credences, so that reconstructed versions of (T1)-(T6) might be true even though (T1)-(T6) themselves are defective.

One option is to *relativize credences*. This sort of view takes its cue from referentialist view on which, whenever a subject believes a proposition, they do so under some guise.[4] On this view, if H is the proposition that Tinasky is Hawkins, then even before examining the text he believed H under one guise (G1), but after examining the text he believed it under another (G2). One might then suggest that one can talk of *credence under a guise*. Under G1, $p(H) = 1$ and $p'(h) = 1$. Under G2, $p(H) = 0.5$ and $p'(H) = 0.99$. Or one might suggest that probabilities must be dyadic functions of guises and propositions, so that $p(G1,H) = 1$, $p(G2, H) = 0.5$, and so on.

This is a radical revision of Bayesianism, however. 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, in effect, that the objects of credence (and the entities characterized as propositions by Bayesian theory) are such ordered pairs. If so, the objects of credence behave in a non-referential way, and insofar as we understand objects of credence as objects of belief, then referentialism will be false.

---

[4] See 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.

Strategy 4: Invoke surrogate objects of credence

Perhaps the most important referentialist strategy reconstructs the Bayesian analysis of the crucial cases by reassigning the credences involved to different entities. According to this strategy, while (T1)-(T6) as they stand are not correct, they may be correct if we replace H in these sentences with something other than the proposition that Tinasky is Hawkins. That is, there is some *surrogate* entity H* such that p(H*)=0.5, p'(H*)=0.99, and so on, where H* is distinct from H. On one version of the view, H* will be a proposition distinct from H. On another version of the view, H* 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 Foster's credences in assertions of 'Tinasky and Hawkins', at least, fit claims (T0')-(T6'). 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 Tinasky 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 will 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 (T1)-(T6) are correct is a reading involving sentences, not propositions.

A problem immediately arises. We often want to apply Bayesian principles to subjects who speak 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 in effect reduces to the second option.

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

The mirror case can be used to generate the same sort of reasoning as the Tinasky case. But in this case, there are no surrogate sentences available. 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.

A worse problem for all of these views is that they rule out the use of reasoning about *content* in analyzing the Bayesian updating of belief, at least in these cases. But the apparatus of content is often given a central explanatory role in Bayesian analysis. For example, Bayesian theorists typically take it that the objects of credences either are sets of possibilities or at least determine sets of possibilities, and reason about credences using these sets (so that when one set contains another, it will be associated with greater or equal credence, and so on). But if the objects of credence are merely sentences, we cannot do this. If we cannot, then we lose a major source of explanatory power.

A referentialist might respond by saying that sentences themselves are associated with sets of possibilities, which we can then use in this sort of analysis. But if these sets behave in a referential way (so that 'Tinasky is Tinasky' and 'Tinasky is Hawkins' are associated with the same set, then we cannot apply reasoning about these sets in the usual manner to cases such as (T1)-(T6). And if these sets 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 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 H* that renders the Bayesian reasoning above correct, in that Foster's initial credence p(H*) was 0.5

and his later credence p(H*) was 0.99, updated in the Bayesian way. It is just that this proposition differs from H, the proposition that Tinsley is Hawkins. The most obvious version of this strategy takes H* to be some descriptive proposition, such as the proposition that the author of such-and-such letters is the author of such-and-such poems.

There is something awkward about referentialists suggesting that we can invoke descriptive surrogates for names. After all, 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 them 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('Gödel is the D' | E') is low. But clearly pr('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('Gödel is Smiffy' | E') is low while pr('the D is the F' | 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('Gödel is the D' | 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 'Tinasky is Hawkins' is something like "The person called 'Tinasky' is the individual called 'Hawkins'", or perhaps "'Tinasky is Hawkins' is true", or something like that.[5] It could be argued that "The person called 'Tinasky'" is guaranteed to corefer with 'Tinasky', 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, 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. Here we can adapt a familiar case due to David Lewis (1983).

---

[5] This 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, and is given a display screen giving him access to 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 symmetrical, so that Zaphod has a qualitatively indistinguishable twin Zaphod', who is president of a qualitatively indistinguishable galaxy. Zaphod sees both twins on the screen, and points to them. He names the one he points to first 'Alpha' and the one he points to second 'Beta'. He wonders which of the two twins he is, saying "Am I Alpha? Or am I Beta?". (Of course he sees both twins asking the question, but this does not help.) He rationally has credence 0.5 associated with his assertion 'I am Alpha', while he has credence 1 associated with his assertion 'Alpha is Alpha'. In fact, Zaphod is Alpha, the twin he pointed to first.

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 Alpha' 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 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 we can take these objects to be nonreferential objects of credence, and therefore nonreferential objects of belief.

The second horn suggests that an appeal to nonreferential surrogates undermines the force of referentialissm. This horn deserves a little more attention, as this issue has arisen repeatedly. One can spell out the argument in more detail as follows:

(1) Nonreferential surrogates are the entities to which credences are assigned in Bayesian updating.

(2) The entities to which credences are assigned in Bayesian updating are objects of credences.

(3) Objects of credence are objects of belief

_____

(4) Nonreferential surrogates are objects of belief.

Premise (1) is built into the strategy that we are considering. Premise (2) is at least a very natural claim. Premise (3) seems hard to question, because of the internal connections between the notion of credence and the notion of belief. If one were to question any premise of this argument, it would presumably be premise (2). Or perhaps one might hold that the notion of an "object of credence" is ambiguous between a notion that satisfies (2) and a notion that satisfies (3).

Around here, the issues are partly terminological. The notions of an "object of belief" and an "object of credence" are somewhat technical. It seems reasonable that there is at least one reading of "object of credence" that satisfies (2). And it seems reasonable to say that if "object of credence" is ambiguous, then "object of belief" is ambiguous in a corresponding way, because of the internal connection between the two. If so, there will be at least one natural reading of "object of belief" that satisfies (4). Of course this is compatible with there being other readings of "object of belief" that do not satisfy (4).

For example, a referentialist may suggest that the entities that credences are assigned to are nonreferential, while the referents of 'that'-clauses are referential. This is a somewhat uncomfortable combination, as one can express credences by saying 'My confidence that Tinasky is Hawkins' 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 $p(H)$ is low, while conveying the true proposition that $p(H^*)$ is low. Presumably they will say the same for sentences (T1')-(T6') earlier, and most other subjective probability claims involving names. Of course this requires

denying the truth of claims that play a highly successful and systematic explanatory role. But such denials are not foreign to the referentialist.

The broader point is that the arguments so far have not ruled out the possibility of *pluralism* about the objects of belief. Perhaps beliefs have multiple objects, or perhaps some entities are objects of belief in one sense and other entities are objects of belief in another. Still, even if some sort of pluralism is true, we have seen enough to reject *strong referentialism*, the view that all objects of beliefs about individuals are referential. And more importantly, we are faced with the challenge of giving an account of the nonreferential objects of credence and of belief.

8 What are the objects of credence?

I will now set aside the issue of referentialism, and ask: what are the objects of credences? We can bypass the terminological questions above by stipulating that the objects of credences are the entities to which credences are assigned in analyses of Bayesian confirmation. I will 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 will assume that there is an association relation between assertions and objects of credence, so that in 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 the previous section 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 will take it as an initial requirement that objects of credence can be associated with sets of possibilities of some sort. This is required in order to use the full explanatory apparatus of the probabilist in the analysis of credences. This does not require that the objects *be* sets of possibilities. They may be more fine-grained abstract objects, 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 the familiar metaphysically possible worlds, as probability theory does not obviously require this.

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 in 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 Alpha', 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 Alpha' and also with the disjoint set associated with 'I am Beta', 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 Alpha in one world and by Beta 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 'Alpha' and 'Beta'. For example, these might be qualitatively identical worlds such that Zaphod is the D in one world and the E in the other world.

Of course the claim that pairs of worlds that merely differ in some haecceitistic respect exist 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 nonhumanness. 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 experience'. The strategy above will then requires times, places, and experiences are associated with haecceitistic properties than can be instantiated in different places in qualitatively identical worlds, in a fairly unrestricted way.[6] This requires a highly baroque conception of the space of possible worlds.

9 Objects of credence as primary intensions

Fortunately there is a more straightforward, and relatively familiar way to handle these cases. We can invoke sets of *centered worlds* instead.[7] A centered world is usually conceived as an ordered triple of a world, an individual who is present in that world, and a time in that world. For present purposes, we can also conceive of it as involving an ordered set of experiences had by the subject at the center of the world, although this point will usually not be central. Now we can simply say that an assertion of the form 'I am ϕ' will be true at all worlds in which the individual at the center has the property associated with 'ϕ'. Something analogous goes for claims involving 'now' (which invoke the time at the center of the world)

---

[6] The strategy suggested here corresponds in an interesting way to a Bertrand 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 associate 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.

[7] The 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" paradox (Elga 2000). If I am right, then this role goes beyond cases involve simple indexicals to cases involving expressions of all sorts.

here' (which invokes the place where the individual at the center is located at the time at the center), and experiential demonstratives (which invoke corresponding experiences at the center).

Now, in the Beeblebrox case, we can say that Zaphod's complete objective knowledge is associated with just one possible world, but with more than one centered world. In particular, his objective knowledge is compatible with a world centered on him and a world centered on his twin. His assertions 'I am Alpha' and 'I am Beta' are associated with sets of worlds where the individual at the center has associated properties ϕ_1 and ϕ_2 respectively.[8] Each of these two sets of centered worlds may then have credence 0.5.

The set of centered worlds associated with assertions like this is highly reminiscent of what I have called the *primary intension* of the assertion. Primary intensions (which are closely related to entities discussed by Lewis 1983 and Jackson 1998) are 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 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.

Of course, showing how indexical assertions can be associated with sets of centered world does not show that every assertion can be associated with such a set. But in fact I think that there is good reason to think that every assertion, and every thought, whether indexical or nonindexical, can be associated with such a set. I have made that case at length elsewhere and will not repeat it here. Instead, I will sketch how one might use the apparatus of conditional

---

[8] Here ϕ_1 and ϕ_2 might involve, among other things, the properties of being pointed at first or second by the individual at the center of the world. Of course this relies on there being certain asymmetries in Zaphod's environment, which is compatible with the original case as described. If we move to an extreme version of the case in which Zaphod's environment and experience are themselves symmetrical, then Zaphod will have to perceive them simultaneously, entertaining the questions 'Am I that one?' and 'Am I that one?', with two symmetrical demonstratives. Here distinguishing the two hypotheses will require appeal to two experiences marked at the center of the worlds, where each "slot" for an experience is aligned with a demonstrative. His first question will be satisfied at all centered worlds where the individual at the center is the cause of the first marked experience; his second question will be satisfied at centered worlds where the individual at the center causes the second marked experience. One can give a closely related treatment to the hall-of-mirrors case, or to extreme versions of it akin to the "Two Tubes" case of Austin 1990.

probabilities to characterize primary intensions,[9] and then will examine how they might play the relevant explanatory roles.

How might an arbitrary assertion — say, Foster's assertion of 'Tinasky is Hawkins' — be associated with a primary intension? One way would be to first associate the names with (possibly indexical) descriptions, such as 'the person who wrote such-and-such and who bears such-and-such relation to me'. The assertion could then be associated with the set of centered worlds that satisfy the descriptive claim 'the D_1 is the D_2', where any indexicals involved are assigned to the entities at the center in the natural way. However, strategy presupposes a highly controversial equivalence between names and descriptions.

In order to avoid such a presupposition, one can instead appeal to conditional probabilities to characterize the relevant sets of centered worlds. Here, one first associates every centered world w with a corresponding world-sentence S(w). This sentence will say, in effect: my world is that world, and I am the person at the center. It will need to be an extraordinarily long sentence, characterizing the qualitative nature of the world in a neutral vocabulary, and using indexical claims such as 'I am ...' and 'now is ...' to specify one's location at the center of the world. Here the notion of a neutral vocabulary has to be characterized with care (see Chalmers 2004), but at a first pass we can think of this as a vocabulary excluding ordinary names and natural kind terms that involve hidden essences, as well as most context-dependent terms.

These world-sentences will themselves enter into conditional probabilities. For any assertion A by a subject and any world-sentence S, we can ask what is the rational conditional credence pr(A|S) for the subject. Of course, this requires a major idealization, as in practice S will be too long for any subject to grasp in full, but idealizations are familiar terrain for a Bayesian approach. If such a subject *did* accept S, then this detailed information would give them a sort of complete objective and indexical knowledge of the world. And this knowledge would in turn put them in a position to ascertain the truth or falsity of all sorts of assertions A. That is, for many A, this knowledge would put them in a position to assign credence 1, or close to 1, to A. And even hypothetically accepting S would put a subject in a position to make corresponding hypothetical conclusions involving A.

For convenience, let us say that pr(A|w)= pr(A|S(w)). We could then define the primary intension of A as the set of centered worlds w such that pr(A|w)=1.[10] Or better, to capture the

[9] This might be considered a project in "probabilistic semantics": the use of probabilistic notions to characterize semantic values.

[10] This characterization of primary intensions is closely related to the definition that I have given elsewhere, e.g. in Chalmers (2004), but it is not quite the same. According to that definition, the primary intension of A is true at

structure of cases where there are probabilities between 0 and 1 (see below), one can define the primary intension of A as a function from centered worlds to real numbers mapping w to pr(A|w). From here, the set above can be straightforwardly recovered, as can other sets corresponding to thresholds less than 1. We can then still say that the primary intension of A *contains* w when its value at w, i.e. pr(A|w), is 1 (or alternatively, when pr(A|w)is high). When A is such that knowledge of its truth can in principle be recovered from complete neutral and indexical information about w, then pr(A|w) will be 1 (or high), so that the primary intension of A will contain w. When knowledge of A's falsity can be recovered from this information, then the primary intension of ~A will contain w.

Of course there may be some A such that pr(A|w) falls in the midrange between 0 and 1, so that the primary intension of neither A nor ~A contains w. Cases of vagueness and other sorts of indeterminacy may provide relatively innocent instances of this. A more serious issue is the possibility of cases where A is determinately true in the actual world @ (for example), but the information in S(@) does not suffice to determine that A is true. In this case, one might say that the truth of A is *inscrutable* relative to S(@). This would be undesirable, so things work best if one requires that enough is built into world-sentences that the following scrutability thesis obtains:

Scrutability: For all truths T, pr(T|@) = 1.

Of course if the vocabulary involved in S(w) were unrestricted, then scrutability would hold automatically. But if one restricts the vocabulary, then whether scrutability holds depends on the precise restriction. In particular, the claim that scrutability holds for a neutral and indexical vocabulary is a substantive claim that I will not argue for here. Still, I think there are good reasons to accept this claim. (For arguments that can be used to support this thesis, see Chalmers and Jackson 2001, though it should be noted that the present framework does not require that world-sentences be restricted to microphysical and phenomenal terms.)

---

w iff a material conditional 'S(w)⊃A' is a priori. The present characterization does not invoke the notion of the a priori. Rather, it is a version of what I call there a "Ramsey intension", defined in terms of a sort of conditional inferential role (akin to the one in Ramsey's test for evaluating indicative conditionals). The two definitions give the same results insofar as p(A|W) is 1 iff 'S(w)⊃A' is a priori. The left-to-right direction here is not obvious, as the probability on the left might in principle be justified by empirical background beliefs. However, the "antecedent" S(w) is sufficiently complete that it will usually render the justificatory role of background beliefs superfluous (for an argument to this effect, see Chalmers and Jackson 2001, part 5(6)). The right-to-left direction has some plausibility, at least given a sufficiently demanding notion of apriority. Perhaps the most likely potential counterexamples arise from cases involving cognitively impairing drugs, discussed below.

In any case, whether scrutability holds or not, primary intensions are well-defined and will generally determine nontrivial classes of worlds.[11]

We can apply this machinery to Foster's assertion of 'Tinasky is Hawkins'. If Foster were given a complete qualitative description of the actual world @ and his place in it, he would plausibly (at least under idealization) be in a position to trace the causal source of relevant letters to a newspaper he has read, determine that the source is someone named 'Hawkins' who wrote certain poems, and so on. This would put him in a position to have very high credence in 'Tinasky is Hawkins'. This suggests that pr('Tinasky is Hawkins' | @) is high. Given a description of a quite different centered world w, he could trace the causal source back to a different individual named 'Pynchon' who wrote books with titles such as *Gravity's Rainbow*, which would put him in a position to have high credence in 'Tinasky is not Hawkins'. So pr('Tinasky is not Hawkins' | w) is low. So the primary intension of 'Tinasky is Hawkins' will include @ and will exclude w. This is quite different from the primary intension of 'Tinasky is Tinasky', which will include both @ and w.

Something similar applies to all the other cases discussed in this paper. In general, when a claim such is 'A is A' is trivial, its conditional probability given any world-sentence will be high, so its primary intension contains all worlds. But when a claim such as 'A is B' is such that one has non-negligible credence in its negation, there will be typically be a centered world w such that pr('A is B'| w) is low, so its primary intension will be false at w. So these claims, although referentially equivalent, will have different primary intensions. And these primary intensions will be associated with different credences, corresponding to the different credences of the sentences.

For this framework to satisfy the axioms of probability, we need to ensure that the total credence over the space of centered worlds is 1. This can be ensured with the following claim:

> **Plenitude**: For all assertions A, if pr(A | w) = 1 for all centered worlds w, then pr(A) = 1.

---

[11] A worry in the vicinity of scrutability arises from centered worlds involving cognitive impairments to the subject at the center. For example, if w is a world where the subject at the center has been given cognitively impairing drugs, and if A is an assertion of '57+65=122', then arguably pr(A|W) is midrange between 0 and 1 (given information that one has taken such drugs, one should rationally suspend judgment about the mathematical claim). If so, then the primary intension of A as defined here will be false at W. More generally, if one has nonzero credence that such a world is actual, then no nontrivial statement will have credence 1. For this reason I think that for present purposes the notion of conditional probability has to be understood in such a way that pr(A|W) is 1. Alternatively, one can define primary intensions in terms of a priori conditionals, as in the previous footnote.

Plenitude ensures that the space of centered worlds is large enough to exhaust the hypothesis space, so that we can be certain that we are in one such centered world. This thesis is plausible, but it is not obvious. On some philosophical views it will be false. Suppose that it is necessary that there is an omniscient being, but nevertheless one can have rational credence 0.5 in the nonexistence of an omniscient being. Let B be 'There is an omniscient being'. Then for all w, $pr(B|w) = 1$, but $pr(B)=0.5$. In effect, the space of centered worlds does not exhaust the hypothesis space, so that there are not enough centered worlds to contain all the credence associated with B.

Any such counterexample to plenitude requires something closely akin to what I have elsewhere called "strong necessities": necessities that are not knowable a priori but that are nevertheless true relative to all centered worlds. I have argued (Chalmers 2002) that there are no strong necessities. If so, then the framework of centered worlds is adequate. If not, then one has to expand the framework slightly to include a broader class of epistemically possible scenarios (which need not be metaphysically possible worlds). For example, to handle the case above, one would need there to be scenarios w such that pr('There is no omniscient being'| w) is high.

If this is needed, we might simply posit a space of scenarios that works this way, or one might construct the space in other terms, for example by appealing to maximal a priori consistent sets of sentences (see Chalmers 2004). If one does this, then one can use scenarios to play the role that centered worlds play above. Scenarios will be associated with scenario-sentences, which one can use to define conditional credences, and thereby associate arbitrary assertions and thoughts with primary intensions, now regarded as sets of scenarios. In this case an analog of plenitude will be guaranteed to hold. If one takes this line, one might identify the objects of credence as sets of scenarios rather than sets of centered worlds. But if the original plenitude thesis is correct, as I believe it is, then sets of centered worlds will suffice.

If plenitude holds, then we are guaranteed to avoid problems akin to those that plague a referentialist view of the objects of credence. These problems arose because there are pairs of assertions A and B with the same object of credence (construed referentially) but such that $pr(A) \neq pr(B)$. If A and B have the same primary intension, however, then $pr(A\equiv B|w) = 1$ for all w, from which it follows from plenitude that $pr(A\equiv B)=1$, so $pr(A)=pr(B)$.

All this gives us a natural candidate for what the objects of credence are, at least given the constraint that these objects determine sets of some sort: objects of credence are primary intensions. These may be taken to be sets of centered worlds, or possibly as sets of epistemically possible scenarios (or functions defined over these sets), depending on the

status of the plenitude thesis. These intensions are associated nonreferentially with assertions and beliefs, so they will avoid all the problems for referential objects discussed earlier. And they have a deep tie to conditional probabilities, so that they are able to play the role that is required of objects of credence in Bayesian theory.

10 What are the objects of belief?

Finally, we can ask: what are the objects of belief? We have seen that primary intensions can play the role of objects of credence for Bayesian theory, but this does not entail that they play all roles for objects of belief, though. They have obvious deficiencies for some aspects of that role. Primary intensions lack structure, so that the obvious difference between the quantified thought *all elephants are elephants* and the identity thought *3+7=10* is not reflected in their primary intensions, which are identical. Likewise, primary intensions lack referential information, so that they do not carry information about the individuals that thoughts are about. So your thought *I am hungry* and my thought *I am hungry* have the same primary intension despite being about the different things, while the reverse goes for your thought *I am hungry* and my thought *You are hungry*.

Another problem concerns the connection between credence and objective chance. The "Principal Principle" put forward by Lewis (1980) says that under certain conditions, when one knows that the chance of a proposition P is ϕ, one's credence in P should be ϕ. This statement assumes that the objects of chance and credence are the same. But it is not plausible that the objects of chance are primary intensions, or sets of centered worlds: the objective chance that an object has a property (e.g. the chance that a certain radium atom decays in a certain time) is not sensitive to a mode of presentation, and indexical phenomena seem to be irrelevant here. On the face of it, it is more plausible that the objects of chance are something like Russellian propositions, or perhaps sets of objective possible worlds. If we are not to give up the connection between chance and credence entirely, it is desirable to have an understanding of the objects of credence that makes their connection to the objects of chance transparent.

To address these issues, we need a finer-grained conception of the objects of credence and belief. The considerations in previous sections require that objects of credence *determine* sets of centered worlds, but they do not require that objects of credence *be* sets of centered worlds. So it is open to us to invoke finer-grained objects of credence that build in information about structure and about reference.

There is a natural approach here. To capture structure, we can invoke *structured primary intensions*. Here we assume that beliefs and assertions are structured entities, with a logical form involving simple concepts or simple expressions. It is not hard to extend the earlier framework so that simple concepts and simple expressions are associated with their own primary intensions, now considered as functions from centered worlds to objects, properties, or other extensions in those worlds. We can then associate an assertion with a structured entity involving the primary intensions of its simple expressions (along with any unarticulated constituents), structured according to its logical form. We can do the same for a belief, where the structured primary intension will involve the primary intensions of its simple concepts, structured according to its logical form. Now, a belief *3+7=10* and *all elephants are elephants* will have entirely different structured primary intensions.

We can go further by building in referential information, as follows. Let us say that the *enriched intension* of a simple expression or a simple concept is an ordered pair of its primary intension and its extension, where its extension is an object, property, or relation in the world. We can then say that the *enriched proposition* of an assertion or a belief is a structured entity made up of the enriched intensions of its simple parts, structured according to its logical form.

Enriched propositions are finer-grained than either structured primary intensions or Russellian propositions, as they encapsulate the information in both. 'Tinasky is Tinasky' and 'Tinasky is Hawkins' will have distinct enriched propositions, in virtue of the fact that 'Hawkins' and 'Tinasky' have distinct primary intensions. But at the same time, the Russellian part of the two enriched propositions will be the same, reflecting what they have in common. The same goes for my utterance of 'you are hungry' and my utterance of 'I am hungry'. The associated enriched propositions will be similar in their referential aspect (the Russellian part), and will be distinct in their nonreferential aspect (the primary intension part). This reflects an intuitively reasonable taxonomy of the objects of belief and assertion.

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. The enriched proposition associated with a belief will determine a set of centered worlds, so all the previous apparatus for the analysis of credence can be brought to bear. At the same time, it determines a Russellian proposition, and we can identify the objective chance of an enriched proposition with that of the associated Russellian proposition. This allows us to preserve

Lewis's Principal Principle: for an enriched proposition P, under appropriate conditions, if one believes that the chance of P is ϕ, then one's credence in P should be ϕ.[12]

It is not unreasonable to speculate that enriched propositions can play many of the roles of propositions. We seen that they determine sets of centered worlds and Russellian propositions. They can also be seen as structures involving the presentation of Russellian propositions under guises, as discussed earlier, if we understand guises in terms of primary intensions. Finally, enriched intensions and enriched propositions behave in a manner quite reminiscent of Fregean senses and Fregean thoughts, so that cognitively distinct beliefs will be associated with distinct enriched propositions.[13]

A theme of this paper 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. Here I have concentrated largely on the roles associated with probability theory. These roles have led us to primary intensions and to enriched propositions. But there are other roles for objects of belief: roles associated with cognitive significance, with truth-conditions, with attitude ascriptions, and so on. I have not argued here that enriched propositions can play all these roles, and 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. Still, the analysis of probabilistic reasoning suggests that attention to a multiplicity of roles may yield significant rewards in the analysis of objects of belief.[14]

---

[12] An alternative approach would be to hold that objects of credence are structured primary intensions, that objects of chance are Russellian propositions, and that a Russellian proposition can be entertained under the guise of a structured primary intension. A modified Principal Principle would then say that under appropriate conditions, if one believes 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. The approach in the text allows a closer connection between the objects of credence and of chance.

[13] For much more on enriched propositions and their relationship to Fregean thoughts, see "Propositions and Attitude Ascriptions: A Two-Dimensional Account".

Bibliography

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

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

Chalmers, D.J. 2002. Does conceivability entail possibility? In (T. Gendler & J. Hawthorne, eds) *Conceivability and Possibility*. Oxford University Press.

Chalmers, D.J. 2004. Epistemic two-dimensional semantics.

Chalmers, D.J. & Jackson, F. 2001. Conceptual analysis and reductive explanation. *Philosophical Review* 110:315-61.

Elga, A. 2000. Self-locating belief and the Sleeping Beauty problem. *Analysis* 60:143-47.

Lewis, D. 1980. A subjectivist's guide to objective chance. In (R.C. Jeffrey, ed) *Studies in Inductive Logic and Probability, Vol. II*. University of California Press.

Lewis, D. 1983. Attitudes de dicto and de se. *Philosophical Review* 88:513-43.

Jackson, F. 1998. *From Metaphysics to Ethics*. Oxford University Press.

Stalnaker, R. 1984. *Inquiry*. MIT Press.