Expert Deference *De Se*

Abstract: Principles of expert deference say that you should align your credences with those of an expert. This expert could be your doctor, the objective chances, or your future self, after you've learnt something new. These kinds of principles face difficulties in cases in which you are uncertain of the truth-conditions of the thoughts in which you invest credence, as well as cases in which the thoughts have different truth-conditions for you and the expert. For instance, you shouldn't defer to your doctor by aligning your credence in the *de se* thought 'I am sick' with the doctor's credence in that same *de se* thought. Here, I generalise principles of expert deference to handle these kinds of problem cases.

1 | Introduction

I have opinions about how likely various things are. I think that it's unlikely to snow in Dubai, that a flipped coin is just as likely to land heads as tails, and that global leadership is unlikely to take serious steps to address climate change. Others have these kinds of opinions, too. For instance, weather reporters have opinions about how likely it is to snow in Dubai; the objective chances have opinions about how likely it is that a flipped coin will land heads; and my future, better informed, self has opinions about how likely global leadership is to take serious steps to address climate change. Principles of expert deference say that I should treat the opinions of these experts—the weather reporters, the objective chances, or my future, better informed self—as a particularly strong kind of evidence. Roughly, given that one of these experts is n% sure of p, I too should be n% sure of p—that is to say, my own subjective degree of confidence in p, or my *credence* in p, should be n%.[1]

Principles like these are plausible, but they seem to give bad advice when *de se* thoughts are substituted for p. It is plausible that, when it comes to matters of my health, I should defer to my doctor. Let p be the thought 'I am sick'. Presumably, p concerns my health. But, just because my doctor is n% confident

Word Count: 5,753 in the body, 4,905 in appendices

1. For a nice introduction to principles of expert deference, see Dorst *et al.* (ms).

in 'I am sick', this doesn't give me a reason to be n% confident in 'I am sick'. After all, when my doctor entertains the thought 'I am sick', that thought will be true just in case *she* is sick, and when I entertain the thought 'I am sick', that thought will be true just in case *I* am sick. Or let p be 'It is now Monday'. On Sunday, I can know that one of my future, better informed selves will be nearly certain of 'It is now Monday'. But it doesn't follow that, on Sunday, I should be nearly 100% sure of 'It is now Monday'. (Some readers will have questions about what I mean by 'a thought'. I will address those questions in §2 below.)

Here, I will introduce and explore an emendation of principles of expert deference which allows them to deal with *de se* thoughts like these. This emendation comes in two parts. Standard principles of expert deference say that you should align your opinions about p with the expert's opinions about p. However, when it comes to thoughts like 'I am sick', I should not align my opinion with my doctor's opinions about 'I am sick', but instead with her opinion about some other thought—a *surrogate* for p. The first part of the emendation concerns this surrogate.

The second part of the emendation will be required to deal with expert deference in cases in which you or the expert are uncertain about who, when, or where you are. Roughly, I'll deal with these cases by suggesting that you should only align your opinions with those of the expert *conditional* on who you both are, and when and where you both are located in space and time.

In companion papers, I apply this general principle of expert defence to the particular experts of the objective chances and your future, better-informed selves, and show how they address some problems for Lewis (1980)'s *Principal Principle* and van Fraassen (1984)'s principle of *Reflection*. Here, I will simply attempt to introduce and motivate a general principle of expert deference which subsumes the principles of chance deference and the principle of deference to your future selves which I defend in the companion papers. In §2, I will explain more carefully what I mean by 'a thought'. In §3, I will introduce some traditional principles of expert deference. In §4, I will explain what it takes for a thought to be what I'll call a *location*, and I will show how to construct the *locational surrogate* of a given thought. Then, in §5, I will introduce and motivate a generalised principle of expert deference.

2 | CREDENCES AND THOUGHTS

Here, I am interested in principles of expert deference which govern your *credences*—that is, your subjective opinions about how likely various things are. I'll represent your credences with a function, C, from thoughts to real numbers between 0 and 1. If $C(p)$ is close to 1, then you are very confident in p—close to

0, and you are not very confident in p. I'll call the real number $C(p)$ 'your credence in the thought p' or 'your credence that p'. Of course, strictly speaking your credence in a thought is a psychological state, and not the abstract number we use to represent it. Likewise, your credences are not strictly speaking a function; instead, they are represented with a function. But I'll adopt this way of speaking as a convenient shorthand.

I use 'thoughts' as a technical term for the arguments of your credence function, whatever they happen to be. Thoughts are those things to which you assign credences. I'll suppose that, in many cases, at least, one and the same thought may be entertained by different people and at different times, that thoughts can be true or false, and that it makes sense to talk about the *negation* of a thought as well as the *conjunction* or *disjunction* of thoughts. I will, moreover, suppose that your thoughts are *closed* under both negation and arbitrary conjunction and disjunction. That is: if you have credal opinions about p, then you also have opinions about $\sim p$. And, if Σ is any set of thoughts about which you are opinionated, then you also have opinions about a thought, $\bigwedge \Sigma$, which is true iff every thought in Σ is true, as well as a thought, $\bigvee \Sigma$, which is true iff *some* thought in Σ is true. (More on this in §B.1 below.)

I use English sentences to give names to thoughts. Intuitively, a sentence names the thought to which a thinker would express commitment by asserting that sentence. This works well enough for my purposes here; though it's worth noting that, in some cases, a single sentence will correspond to multiple thoughts. For instance, for someone who falsely thinks that there are two famous Polish people with the name 'Pederewski', there will be two thoughts corresponding to the one sentence 'Paderewski is a musician'.[2] We will have to reach for more detailed sentences to name these thoughts.

I do not hold that thoughts are individuated by their truth-conditions; that is to say, I do not suppose that thoughts correspond one-to-one with sets of metaphysically possible worlds.[3] Nor should you. Your credence in 'Mark Twain is a gifted humorist' can differ from your credence in 'Samuel Clemens is a gifted humorist', though both of these are true in exactly the same metaphysically possible worlds (because Twain is identical to Clemens). So thoughts should be individuated in some respects more finely than their truth-conditions. And in other respects, they should be individuated more coarsely. When Hume and Heimson both entertain the thought 'I am sick', they thereby become re-

2. Cf. Kripke (1979)

3. I use 'truth-conditions' to mean the same thing as 'intension'. A thought's truth-conditions is just the set of possible worlds in which the thought is true.

lated to propositions with different truth-conditions. Hume becomes related to a proposition which is true iff Hume is sick, whereas Heimson becomes related to a proposition which is true iff Heimson is sick. However, Heimson might be confused about whether he is Hume or Heimson. He shouldn't thereby be confused about which thought he's investing credence in when he is 50% sure of 'I am sick'. So we should say that Hume and Heimson entertain the same thought, even if, by so doing, they become related to propositions with two different truth-conditions. In this respect, thoughts should be individuated a bit more coarsely than their truth-conditions.

Beyond these assumptions, I hope to remain officially neutral on what a thought is. I believe that my use of the term 'thought' is broad and ecumenical enough that, no matter your views about the objects of belief, you will be able to find something to play the role of my thoughts, and which can serve as the arguments of your credence function. For everyone should accept that there's an important difference between the belief state that I would report by saying 'I think that Twain is clever, and Clemens is not' and the belief state I would report by saying 'I think that Twain is clever and that Twain is not'. The first belief state could be rational, while the second could not be. So there is something to distinguish those belief states. Whatever that something is, thoughts should be partly individuated with respect to it, so that 'Twain is clever' and 'Clemens is clever' can be distinct thoughts. And everyone should accept that, when Hume and Heimson each have the belief they would express with the sentence 'I am sick', there is something that they thereby have in common. Whatever that something is, thoughts should be partly individuated with respect to it, so that when Hume and Heimson entertain the thought they would each express with the sentence 'I am sick', they are entertaining one and the same thought. In appendix A, I will say more about how to individuate thoughts, given some popular views about the objects of belief.

3 | DEFERENCE

If you're going to attempt to align your credences with those of an expert, then you must have some views about what the expert's opinions are. So all of the principles of expert deference I'm going to discuss take it for granted that you have credences in thoughts of the form $\mathcal{E} = E$, where the script $\mathcal{E}$ stands for the definite description 'the expert's credence function', and E is a particular probability function. Given thoughts of the form $\mathcal{E} = E$, we can construct thoughts of the form $\mathcal{E}(p) = n\%$. To get this latter thought, we take every thought of the form $\mathcal{E} = E$ such that $E(p) = n\%$ and disjoin them all. That is: the thought $\mathcal{E}(p) = n\%$ is the disjunction $\bigvee\{\mathcal{E} = E \mid E(p) = n\%\}$.

Several authors suggest the following general recipe for deference: conditional on the expert thinking that p is $n\%$ likely, you too should think that p is $n\%$ likely. That is, for every p and every $n\%$,[4]

$$\text{(D1)} \qquad C(p \mid \mathcal{E}(p) = n\%) \overset{!}{=} n\%$$

(A word on notation: I place an exclamation mark above an equation to indicate that it is a normative claim. Thus, D1 doesn't say that your credence in p, given $\mathcal{E}(p) = n\%$ *is* equal to $n\%$. Instead, it says that it *ought* to be equal to $n\%$.) Other authors suggest the following: conditional on E being the expert's probability function, your credence in p should be $E(p)$. That is, for every p and every E,[5]

$$\text{(D2)} \qquad C(p \mid \mathcal{E} = E) \overset{!}{=} E(p)$$

D2 is *nearly* equivalent to D1. If you satisfy D2, then you will satisfy D1, as well; and, if you satisfy D1—and you are not in an *incredibly* singular and special case—then you will satisfy D2, as well. The two constraints are not *precisely* equivalent, but they are equivalent for all philosophical purposes.[6] So we should not worry ourselves too much about the differences between them. (See [redacted for blind review] for more discussion.)

Finally, some authors suggest that your credence in p should be your *expectation* of the expert's credence in p.[7] That is, your credences should satisfy D3, for every p,

$$\text{(D3)} \qquad C(p) \overset{!}{=} \sum_E E(p) \cdot C(\mathcal{E} = E)$$

D3 is strictly weaker than D1 and D2; both D1 and D2 entail D3, but D3 does not entail either D1 or D2. Unlike the differences between D1 and D2, the difference between D3 and the other two principles is not negligible.

There are other forms which principles of expert deference take, but those other forms are introduced to deal with complications orthogonal to our interests here. For instance, D2 entails that the expert knows for sure what its

4. Principles with this form appear in Skyrms (1980) and Gaifman (1988), for instance.

5. Principles with this form appear in Elga (2007) and Pettigrew & Titelbaum (2014), for instance.

6. Gaifman (1988) provides a case in which you satisfy D1 but not D2. In [redacted for blind review], I show that Gaifman's case is, in a good sense, the *only* kind of case in which D1 and D2 come apart.

7. See, for instance, Ismael (2008) and Ismael (2015).

own probabilities are. In some applications, this seems implausible, and several authors have suggested the remedy of replacing the right-hand-side of D2 with $E(p \mid \mathcal{E} = E)$.[8] For the substituends for p I'll be looking at here, it is safe to suppose that whatever uncertainty the expert may have about their own probabilities won't make a significant difference to their probability in p, so this additional complication won't be relevant to anything I have to say here. And precisely the same remedy is available for the principle I will introduce in §5 below: just condition the expert's probability function on $\mathcal{E} = E$ on the right-hand-side.[9]

A source of difficulty for the traditional principles is that *your* thoughts 'I am sick' and 'Today is Monday' could be true without *the expert's* thoughts 'I am sick' and 'Today is Monday' being true. More generally, the truth-conditions of *your* thought p could differ from the truth-conditions of *the expert's* thought p. So you shouldn't want to align your credence in p with the expert's credence in that same thought, as the traditional principles of expert deference assume. To deal with this problem, I will suggest that, instead of deferring to the expert by aligning your credence in p with the expert's credence in p, you should defer by aligning your credence in p with the expert's credence in an appropriate *surrogate* of p. We want to find a thought which, when entertained by the expert, will have the same truth-conditions that p does, when it is entertained by you.

In the following section, I will define a family of appropriate surrogates (one for each *location* you might occupy—that is to say, one for each person you may be, place you may occupy, and time it may be). Then, in §5, I will use these surrogates that to propose a modification for principles like D1, D2, and D3.

4 | LOCATIONAL SURROGATE THOUGHTS

Here's a rough, first-pass suggestion: the general surrogate for your thought p should be 'your thought p expresses a truth'. For instance: the surrogate for 'I am sick' is 'my thought 'I am sick' expresses a truth'. The principle of expert deference this surrogate gives us says: your credence in p should be equal to the expert's credence that your thought p expresses a truth. This first-pass suggestion runs into difficulties when either you or the expert are unsure of who you

8. See, *e.g.*, Hall (1994), Lewis (1994), and Elga (2013).

9. Dorst (2020) defends a different principle of expert deference called 'Trust'. And Dorst *et al.* (ms) defend a slightly different principle they call 'Total Trust'. I believe that these principles can also be generalised in the ways I'm proposing, though I won't explore this any further here.

are, where you are, or what time it is. (We'll see several cases like this in §5 below.) To deal with these kinds of cases, I will introduce the notion of a *location*. A location is a thought which serves to settle any uncertainty about who you are, where you are, and what time it is. I'll define a surrogate for your thought p which is *relative to* a particular location. This surrogate will say, roughly and metaphorically, that p expresses a truth when entertained by somebody at that location. In these terms, the principle of expert deference I'll defend in §5 says: given that you are at the location λ, your credence in p should be equal to the expert's credence that p expresses a truth when entertained by somebody at λ.

I should ward off a potential confusion up front: a location is just a particular kind of thought. When I say, roughly, that someone is *at* a location, I just mean that that thought expresses a truth for them. For instance, if we're dealing with a small collection of thoughts, then 'I am Hume' could be a location. If I say that Hume is *at* that location and that Heimson is not, I just mean that this thought expresses a truth for Hume, and it does not express a truth for Heimson.

In §4.1 below, I'll introduce a familiar framework for modelling *de se* beliefs. In this framework, your thoughts are associated with sets of centred possible worlds. Within this framework, I'll say what a location is and what the locational surrogate of a thought is.

The reader may have worries with using this framework for the kinds of problem cases I'm interested in here. In the first place, while a framework which utilises centred *metaphysically* possible worlds will allow us to model uncertainty about *your own* identity, it will not easily or straightforwardly allow us to model uncertainty about *others'* identities, given the metaphysical necessity of identity. As we will see in §5, some interesting cases of expert deference involve uncertainty about the expert's identity. Some may wish to solve this kind of problem by taking the worlds to be *epistemically* possible worlds. Others may think that thoughts are more fine-grained than sets of epistemically possible worlds. They may think that thoughts are sentences in a language of thought, or algorithms for computing truth-value, or something else altogether. It's natural for such a reader to wonder how to define locations and locational surrogates, given their preferred conception of thoughts. Therefore, in appendix B, I give a general procedure for saying which thoughts are locations and what a locational surrogate is, given any arbitrary set of thoughts, howsoever you want to understand them. For the purposes of understanding the proposal in §5, the discussion in §4.1 below should suffice.

4.1 *Locations and Locational Surrogates with Centred Possible Worlds*

We can take a centred possible world to be a pair, (w, c), of a possible world, w, and a *centre*, c, where a centre is something like a triple of a person, place, and time. The interpretation is that (w, c) stands for both a maximal way things might be and for a person you might be, a place you might be, and a time it may be, if things are that way. I say that a centred world is a pair of a world and a centre, but if not every centre is occupied at every world, some pairs of worlds and centres will not be centred worlds. For instance, let w be a world in which Beyoncé doesn't exist and let c be a centre which includes Beyoncé. Then, (w, c) is not a centred world, since it's not possible for you to be Beyoncé if Beyoncé doesn't exist. Within this framework, we take the arguments of your credence function to be sets of centred worlds. So, given my terminology: in this framework, thoughts are sets of centred worlds.

A thought is *de dicto* iff it doesn't say anything about who you are, where you are, or what time it is. Within this framework, a thought is *de dicto* iff it draws no distinctions between centres at a single world. That is: the set of centred worlds p is *de dicto* iff, for any world w and any two centres c, c^* such that both (w, c) and $(\mathrm{w}, \mathrm{c}^*)$ are centred worlds, either both (w, c) and $(\mathrm{w}, \mathrm{c}^*)$ are included in p or neither is. A thought is *de se* iff it says something about who you are, where you are, or what time it is. Within the framework of centred worlds, a thought is *de se* iff it draws distinctions between centres at a single world. That is: p is *de se* iff there's some world w and two centres c, c^* such that both (w, c) and $(\mathrm{w}, \mathrm{c}^*)$ are centred words, and (w, c) is included in p while $(\mathrm{w}, \mathrm{c}^*)$ is not.

With this setup, it makes sense to say that two centred worlds have a centre in common. So we can take the set of all centred worlds which share the centre c, and get a thought which says exactly who you are, where you are, and what time it is. Any thought like this is what I will call a *location*. (As a notational matter, I'll use lowercase Greek letters for thoughts which are locations.) Notice that a location could have non-trivial *de dicto* consequences. For instance, if Beyoncé only exists at some worlds, then a location which tells you that you are Beyoncé will have the non-trivial *de dicto* consequence that Beyoncé exists. Now, take any set of centred worlds, p, and any location, λ. Let c_λ be the centre associated with λ. Then, we can define the λ-surrogate of p—which I'll write 'p_λ'—as follows: for every world w and every centre c, (w, c) is included in p_λ iff (a) (w, c) is a centred world and (b) $(\mathrm{w}, \mathrm{c}_\lambda)$ is included in p. That is, intuitively: p_λ is a thought which is true *anywhere* in a world, so long as the location λ exists at that world, and p is true at that location in that world. Or: so long as there's *somewhere* in a world at which both λ and p are true, p_λ is true *everywhere* in that world.

Notice that, by construction, the λ-surrogate of p, p_λ, will be a *de dicto* thought, whether or not p is *de dicto*. Notice also that the λ-surrogate of λ *itself*, λ_λ, will be a *de dicto* thought which is true at any world such that λ is true *somewhere* in that world. So λ_λ is a thought which says that the location λ is *occupied*. Of course, talk about your being *at* a location and a location being *occupied* is purely metaphorical. To be clear: when I say that you are *at* a location, λ, I just mean that λ expresses a truth for you, here and now. And when I say that a location, λ, is *occupied*, I just mean that λ expresses a truth for *someone*, at some place and time.

5 | DEFERENCE *DE SE*

In this section, I will use the locational surrogates defined in the previous section to formulate a generalisation of principles of expert deference which is capable of handling *de se* thoughts like 'I am sick'. In rough outline, the principle says that you should defer to the expert by setting your credence in p to the expert's credence in the locational surrogate of p, once both your and the expert's credence functions have been conditioned on both your and the expert's locations. (And not just your actual locations, but any locations which you and the expert might occupy.)

More carefully, here's the modification of the principle D2 from §3 that I'll be proposing: for any thought, p, any atomic locations λ and ϵ, and any potential expert function E, your credence in p, given that the expert's credence function is E, you are located at λ, and ϵ is occupied, should be equal to E's credence in the λ-surrogate of p, p_λ, given that λ is occupied, and given that they are located at ϵ.

(DSD2) $$C(p \mid \mathcal{E} = E \wedge \lambda \wedge \epsilon_\epsilon) \overset{!}{=} E(p_\lambda \mid \lambda_\lambda \wedge \epsilon)$$

(I explain what it is for a location to be *atomic*, and why it is necessary to restrict a principle like DSD2 to atomic locations, in §B.2 below. If you're happy to work within a framework of centred worlds, then so long as you have a thought for to every set of centred worlds, every location will be atomic, and you needn't concern yourself with this complication.) In the remainder of this section, I'll explain the principle DSD2 by walking through some examples to motivate successive revisions to D2, until we eventually arrive at the principle DSD2. Finally, in §5.2, I will discuss the natural generalisations of D1 and D3 as well as some of the applications of these principles which I explore in companion papers.

5.1 The Principle Explained

You and your doctor may faultlessly disagree about the *de se*—for instance, you may truly think 'I am sick' while she truly thinks 'I am not sick'. But you and your doctor may not faultlessly disagree about the *de dicto*. For this reason, coordination of opinion between you and your doctor should go by way of the *de dicto*. If your doctor's opinions are going to constrain your credence in the thought 'I am sick', then we must find a *de dicto* surrogate for this thought.

The question to ask yourself is this: 'how confident is the doctor that *my thought* 'I am sick' is true?' That is: 'how confident is she that, when *I* entertain the thought 'I am sick', it expresses a truth?' However confident she is of that is how confident you should be in 'I am sick'. In §§4.1 and B.3, I define the thought p_λ, which says that there's someone for whom the thought λ expresses a truth, and, for that person, p also expresses a truth. It is a *de dicto* surrogate for your thought p, given that you are at location λ (that is to say: given that λ expresses a truth for you). For instance, let β be a location at which you are Beyoncé and let i be the *de se* thought 'I am sick'. Then, i_β is *a priori* equivalent to 'Beyoncé is sick'. Let '$\mathcal{D}$' be the definite description 'the doctor's credence function', and let 'D' be any probability function. If you and your doctor both know that you are Beyoncé, then you should defer to your doctor's opinion by setting your credence in i, given that the doctor's credence function is D, equal to D's credence in the surrogate i_β, which will just be D's credence that Beyoncé is sick.

This suggests emending D2 to say that your credence in any thought p, given that the expert $\mathcal{E}$'s probability function is E, should be E's credence in p_λ:

(D2*) $$C(p \mid \mathcal{E} = E) \overset{!}{=} E(p_\lambda)$$

Above, I took for granted that you knew that you were Beyoncé. But this is something about which you may uncertain. Suppose you don't know whether you're Beyoncé or Kelly, and you know that your doctor thinks Beyoncé is very likely sick and Kelly is very likely not sick. In that case, you shouldn't be very confident in 'I am sick'. Instead, you should proportion your confidence in 'I am sick' to your confidence that you are Beyoncé and that you are Kelly. More generally, for each potential probability function D, and each atomic location λ, you should defer to your doctor's opinion by setting your credence in 'I am sick', i, given that $\mathcal{D} = D$ *and given* λ, equal to D's credence in the surrogate i_λ—that is, $C(i \mid \mathcal{D} = D \wedge \lambda)$ should be $D(i_\lambda)$. If λ is a location at which you are Beyoncé, then i_λ will say that Beyoncé is sick. And if λ is a location at which you are Kelly, then i_λ will say that Kelly is sick. This suggests that we should revise D2* to say that your credence in p, given that the expert $\mathcal{E}$'s probability function is E *and given the atomic location* λ, should be E's credence in p_λ:

$$(\text{D2}^{**}) \qquad C(p \mid \mathcal{E} = E \wedge \lambda) \stackrel{!}{=} E(p_\lambda)$$

Just as you can be uncertain about which locations you occupy, your doctor could be uncertain about which locations *are* occupied. Suppose you know for sure that you are Jekyll, and you are 50% confident that you are *also* Hyde. However, your doctor is very confident that Jekyll and Hyde are two different people. She has one medical record filed under 'Jekyll' and another filed under 'Hyde'. There are positive test results in the 'Hyde' file, and no test results in the 'Jekyll' file. The disease is very rare, so she thinks 'Jekyll is sick' is very unlikely; but the test is reliable enough that she thinks 'Hyde is sick' is very likely. However, were she to learn that Jekyll and Hyde are one and the same person, she would think it's very likely that that person is sick.[10]

Let η be an atomic location at which you are Jekyll and you are Hyde (that is to say: it is *a priori* knowable that, if η is true, then so too are 'I am Jekyll' and 'I am Hyde').[11] Then, if we abide by D2**, we'll say that your credence in 'I am sick', i, given η, should be the doctor's credence in the *de dicto* thought i_η. But the thought i_η is only true in worlds at which the location η is occupied—that is to say: it's only true at world at which 'I am Jekyll and I am Hyde' expresses a truth for *someone*. At any world at which Jekyll and Hyde are different people, i_η will be false. Since your doctor is very confident that Jekyll and Hyde are different people, she will think that i_η is very *unlikely*. So the principle D2** would tell you to not be very confident that you are sick, given that you are both Jekyll and Hyde. This is the wrong verdict. Your doctor is very confident that Hyde is sick, given that Jekyll and Hyde are the same person. So you should be very confident that you are sick, given that you are both Jekyll and Hyde.

The trouble with the principle D2** is that, on the left-hand-side, you have conditioned on some information—the information that you occupy location λ—which the expert may lack. If the expert doesn't have this information, then before deferring to them, you should first bring them up to speed by conditioning the function E on this information. Of course, the location λ is a *de se* thought; so you don't want to bring the function E up to speed by conditioning it on *this* information. Instead, you want to bring it up to speed by conditioning

10. Cf. Chalmers (2011*b*)

11. Or perhaps, it is a location at which you are Jekyll and Hyde, so long as both Jekyll and Hyde exist (see §B.2). In the present example, we can take for granted that both you and the doctor know for sure that Jekyll and Hyde both exist, so I'll ignore this caveat.

it on an appropriate *de dicto* surrogate of this information. That is, you want to condition it on the *de dicto* information that location λ is *occupied*: λ_λ. This thought, λ_λ, is true in any world at which λ is true for someone, somewhere and somewhen. In the case of the location η, η_η is a *de dicto* thought such that it's *a priori* knowable that, if η_η is true, then Jekyll and Hyde are the same person. Thus, your credence in i, given that your doctor's credence function is D, and given that you are at the location η, should be equal to D's credence in the η-surrogate i_η, *given that η is occupied*:

$$C(i \mid \mathcal{D} = D \wedge \eta) \stackrel{!}{=} D(i_\eta \mid \eta_\eta)$$

Since you know that your doctor is very confident that Jekyll/Hyde is sick, i_η, conditional on the information that Jekyll and Hyde are the same person, this principle will tell you, correctly, to be very confident that you are sick, given that you are Jekyll and Hyde.

These considerations suggest that we should modify D2** even further, by requiring that your credence in a thought, p, given that the expert $\mathcal{E}$'s probability function is E and you are at the atomic location λ, should be E's credence in the λ-surrogate p_λ, given that λ is occupied, λ_λ:

(D2***) $$C(p \mid \mathcal{E} = E \wedge \lambda) \stackrel{!}{=} E(p_\lambda \mid \lambda_\lambda)$$

But wait—what if the expert is uncertain about *their* location? In some cases, this may not matter. If her location is irrelevant to the question of whether you are sick, then you may defer to your doctor about whether you are sick without either of you knowing her location. However, in some cases, her location *may* be relevant to whether you are sick. Suppose, for instance, that both you and your doctor suffer from amnesia. However, you both know the following four things for sure: 1) you are either Alfred or Cyril, and she is either Blanche or Dinah; 2) Alfred and Blanche are fraternal twins, as are Cyril and Dinah; 3) the disease is congenital, so if one fraternal twin is sick, then the other one is sick, too; and 4) the doctor is very confident that she is sick.

In this kind of situation, you should only defer to your doctor conditional on a hypothesis about where you and she are both located. Let α be an atomic location at which you are Alfred.[12] Let β be an atomic location at which the doctor is Blanche.[13] And let 'δ' be an atomic location at which the doctor is

12. More carefully: α is an atomic location such that it's *a priori* knowable—at least, given your and the doctor's background information—that 'I am Alfred and I am the patient' is true if α is.

13. More carefully: it is *a priori* knowable—together with your and the doctor's shared background

Dinah.[14] Then, conditional on you being Alfred and the doctor being Blanche, your credence in 'I am sick' should be the doctor's credence in the α-surrogate i_α ('Alfred is sick'), given that you are Alfred and she is Blanche:

$$C(i \mid \mathcal{D} = D \wedge \alpha \wedge \beta_\beta) \overset{!}{=} D(i_\alpha \mid \alpha_\alpha \wedge \beta)$$

Since the doctor's credence in i_α is high, given that you are Alfred and she is Blanche, you should have a high credence in i ('I am sick'), given that you are Alfred and she is Blanche. Likewise, conditional on you being Alfred and her being Dinah, your credence in 'I am sick' should be her credence in the surrogate i_α ('Alfred is sick'), given that you are Alfred and she is Dinah:

$$C(i \mid \mathcal{D} = D \wedge \alpha \wedge \delta_\delta) \overset{!}{=} D(i_\alpha \mid \alpha_\alpha \wedge \delta)$$

Since the doctor's credence in i_α is low, given that you are Alfred and she is Dinah, you should have a low credence that you are sick, given that you are Alfred and she is Dinah.

Cases like this suggest replacing D2*** with the following principle: for any thought, p, any atomic locations λ and ϵ, and any potential expert function E, your credence in p, given that the expert function is E, you are at λ and the expert is at ϵ, should be equal to E's credence in the λ-surrogate p_λ, given that you are at λ and they are at ϵ:

(DSD2) $$C(p \mid \mathcal{E} = E \wedge \lambda \wedge \epsilon_\epsilon) \overset{!}{=} E(p_\lambda \mid \lambda_\lambda \wedge \epsilon)$$

5.2 *Further Discussion*

The same considerations motivate replacing principles with the form of D1 with the following: for any thought p, any atomic locations λ and ϵ, and any number $n\%$,

(DSD1) $$C(p \mid \lambda \wedge \epsilon_\epsilon \wedge \mathcal{E}(p_\lambda \mid \lambda_\lambda \wedge \epsilon) = n\%) \overset{!}{=} n\%$$

Likewise, notice that, just as D1 and D2 imply D3, the principles DSD1 and DSD2 imply the following, which is the natural generalisation of D3: for any thought

information—that 'I am Blanche and I am the doctor' is true if β is.

14. That is: it is *a priori* knowable, given your and the doctor's background information, that 'I am Dinah and I am the doctor' is true if δ is.

p and any atomic locations λ and ϵ,

(DSD3) $$C(p \mid \lambda \wedge \epsilon_\epsilon) \stackrel{!}{=} \sum_E E(p_\lambda \mid \epsilon \wedge \lambda_\lambda) \cdot C(\mathcal{E} = E \mid \lambda \wedge \epsilon_\epsilon)$$

Returning to the principle DSD2: notice that, if p is a *de dicto* thought, and both your and $\mathcal{E}$'s credence in p is independent of your and the expert's location, then this proposed principle entails the more familiar principle of expert deference D2. For, if your credence in p is independent of your and the expert's location (conditional on $\mathcal{E} = E$) then the left-hand-side of DSD2 equals $C(p \mid \mathcal{E} = E)$. And, if p is a *de dicto* thought, then p_λ is *a priori* equivalent to p, and the right-hand-side of DSD2 will be equal to $E(p \mid \lambda_\lambda \wedge \epsilon)$. If E's credence in p is independent of λ_λ and ϵ, then this is equal to $E(p)$. So the principle reduces to $C(p \mid \mathcal{E} = E) \stackrel{!}{=} E(p)$.

In certain applications, we may be able to ignore some of λ, ϵ, λ_λ, and ϵ_ϵ in the principle DSD2, because either you know your or the expert's location for sure, or else the expert knows your or their location for sure. For instance, the objective chance function at t, Ch_t, does not have a spatial location. Its only location is temporal, and that location is *a priori* knowable. If the objective chances at t were at t^*, they wouldn't be the objective chances *at* t.[15] So, when we are considering a principle of *chance* deference, DSD2 reduces to the following principle, which I defend in a companion paper: if C_0 is your *ur-prior* credences (the credences you're disposed to hold in the absence of any evidence), then[16]

(CD) $$C_0(p \mid Ch_t = Ch \wedge \lambda) \stackrel{!}{=} Ch(p_\lambda)$$

Here, I've motivated the principle DSD2 by considering cases where you wish to defer to another human, but I believe similar reasoning applies when we consider an expert like the objective chances. In that application, by the way, we will have to understand my talk of 'the expert's thoughts' slightly differently. When I talk about 'the expert's thoughts', I am still using the term 'thought' stipulatively to refer to refer to whatever the arguments of the expert's probability function happen to be. So, if the expert is the objective chances,

15. Of course, you may wish to defer, not to the *time t* chances, but instead the *current* chances. Then, even though the current chances know their temporal location, you may not (if you've lost track of the time). In that application, the current chances' temporal location cannot be ignored.

16. CD is actually a bit weaker than the principle I defend in the companion paper, since it does not include a clause about admissible evidence. See [redacted] for more.

then their thoughts are just truth conditions, or sets of metaphysically possible worlds. In that case, the difficulty isn't that the truth-conditions of your *de se* thoughts differ from the truth-conditions of chance's *de se* thoughts. The difficulty is that chance does not have *de se* thoughts in the first place. Just as with human experts, before you defer to chance, you need to find some *de dicto* surrogate thought of chance's to coordinate with your *de se* thoughts. And in my view, just as with human experts, this *de dicto* surrogate is the truth-condition that your thought expresses a truth. So, when it appears inside of of the objective chance function on the right-hand-side of the principle CD, 'p_λ' is a set of metaphysically possible worlds. It is the set of worlds at which the thought '$p \wedge \lambda$' expresses a truth for someone.

In another companion paper, I argue that we should defer to our future, better-informed selves using a principle of the form DSD3. Specifically, if $\mathcal{D}$ are the credences you are *d*isposed to adopt after undergoing some learning experience, then I say that, for any thought 'p', and any atomic locations 'λ' and 'δ',

$$C(p \mid \lambda \wedge \delta_\delta) \overset{!}{=} \sum_D D(p_\lambda \mid \lambda_\lambda \wedge \delta) \cdot C(\mathcal{D} = D \mid \lambda \wedge \delta_\delta)$$

In the companion paper, I show that this principle escapes several counterexamples to van Fraassen (1984, 1995)'s principle of *Reflection*.

A | THOUGHTS

In this section, I will say in some more depth how to think about thoughts, given some popular views about the objects of belief.

Some terminology: as I'll use the term here, a *proposition* is the referent of a 'that'-clause in an attitude ascription. So the referent of 'that he left the oven on' in an attitude ascription like 'John fears that he left the oven on' is a proposition. Some hold that propositions are *fine-grained*, in the following sense: the referent of 'that Twain is gifted humorist' is a different proposition than the referent of 'that Clemens is a gifted humorist'. Others want to identify these two propositions. Call the first group *fine-grainers*, and the second, *coarse-grainers*. Coarse-grainers say that, if you believe that Twain is clever and yet disbelieve that Clemens is clever, then you both believe and disbelieve one and the same proposition. Nonetheless, coarse-grainers will want to allow that you may still be rational (after all, you may not know that Twain is Clemens). They will want to distinguish your rational belief state from an irrational state of believing that Twain both was and was not clever. To do this, they will appeal to the notion of a *guise*. A guise is a way of being acquainted with a proposition. For coarse-grainers, when you bear an attitude to a proposition, you do so under some guise or other. The reason you can rationally believe and disbelieve one and the same proposition is that the guise under which you believe it is distinct from the guise under which you disbelieve it.

A coarse-grainer should distinguish thoughts from propositions. For I have supposed that your credences are a function from thoughts to real numbers. It follows that you cannot give one and the same thought two different credences: if $C(p) \neq C(q)$, then $p \neq q$. But, according to the coarse-grainer, you *can* give one and the same *proposition* two different credences. For instance, you can be confident that Twain is clever but not very confident that Clemens is. So, if you are a coarse-grainer, you should individuate thoughts by something other—or something more—than their propositional content.

I see two natural suggestions for the coarse-grainer: thoughts could be guises, or they could be guise-proposition pairs (where the paired proposition is the one you entertain *via* that guise).[17] Braun (2016) opts for the second option, though from my perspective, the first is more attractive. On a coarse-grained view of propositions, rational credence has everything to do with the guises under which propositions are entertained, and nothing to do with the propositions thereby entertained. Just to illustrate the point, consider

17. See the proposals discussed in Chalmers (2011*b*), Braun (2016), and Fitts (2014).

a coarse-grained view on which propositions are individuated by their truth-conditions—that is, if it is necessary that the propositions P and Q have the same truth-value, then $P = Q$. Take any thought, p, and let P be the proposition which your thought p expresses. So long as p is true, there is a guise such that you may be rationally certain in P, under that guise. For, if p is true, then p has the very same truth-conditions as $p \leftrightarrow @p$. Here, $@p$ says that p is *actually* true. (That is: if p is actually true, then $@p$ is necessarily true; and if p is actually false, then $@p$ is necessarily false.) We've supposed that p is true; that means that $@p$ is necessarily true. So the biconditional $p \leftrightarrow @p$ will have a necessarily true right-hand-side, and so it will be true iff its left-hand-side, p, is true. So p and $p \leftrightarrow @p$ have the very same truth-conditions. On the coarse-grained view we are considering, then, they correspond to precisely the same propositions. Now, $p \leftrightarrow @p$ is *a priori* knowable (it is *a priori* knowable that any thought of your is true if and only if it is *actually* true.) If a thought is *a priori* knowable, then you may be rationality certain of it. So you can be rationally certain of a thought with the same truth-conditions as p, for any true p whatsoever.[18] So, from my perspective, when it comes to rational credence, it's most natural for a coarse-grainer to think that propositions are an idle wheel, and so to identify thoughts with guises, not guise-proposition pairs.

Whether a coarse-grainer identifies thoughts with guises or guise-proposition pairs may make a difference to the book-keeping in this paper. Consider the guise associated with my belief that I am sick now. Suppose, for the sake of illustration, that both my future self and my doctor are also capable of entertaining a proposition under this guise—though that guise will determine different propositions for me, my future self, and my doctor. For me, the guise determines the proposition that [Author] is sick on August 16, 2020; whereas, for my future self, it determines the proposition that [Author] is sick on August 17, 2020, and, for my doctor, the guise determines the proposition that *she* is sick on August 16, 2020. If the thought 'I am sick now' is just the guise, then both me, my future self, and my doctor can have a credence in this one thought. On the other hand, if the thought 'I am sick now' is a pair of a guise and a proposition, then neither my future self nor my doctor is capable of entertaining my thought 'I am sick now'. I doubt that there is any substantive issue here. Suppose that thoughts are identified with guise-proposition pairs. Then, we may say that two thoughts are *similar* iff they have a guise-component in common. Then, even if my doctor cannot entertain my thought 'I am sick now', she can

18. I owe this observation to the first appendix in Gibbard (2012) and Yli-Vakkuri & Hawthorne (forthcoming); Gibbard shows that similar results hold for other kinds of coarse-grainers.

entertain a thought which is similar to it. And this will be enough for my purposes. In what follows, I'll opt for the first form of book-keeping, supposing that me and my doctor can both have credences in the thought I'd express with 'I am sick', but the reader should feel free to keep their own books differently, supposing instead that my and my doctor's thoughts are merely similar, and not identical. If you keep your books this way, then whenever I say "the thought *p*", you should read this as "a thought similar to *p*". So far as I can see, nothing substantive will change.

Turning now to fine-grainers: Insofar as they are happy to say that the proposition I express when I say 'I am sick' is the same as the proposition you expresses when you say 'I am sick', fine-grainers may identify thoughts with propositions. If they distinguish between these propositions, then their propositions are finer than my thoughts. Just as with the coarse-grainer, I don't think there is any substantive issue here. Even if you and I believe different propositions when we each believe the propositions we'd express with 'I am sick', there is nonetheless something that our belief states have in common. For instance, perhaps both of our beliefs are mediated by the same sentence in the language of thought. If you are this kind of fine-grainer, you may understand my talk about thoughts as talk about equivalence classes of belief states which have that feature in common. When I take two thoughts to be the same, you can understand me as talking about thoughts which are *similar*—being expressed by the same language in the sentence of thought, or what-have-you—though strictly speaking distinct. Again, so far as I can see, nothing substantive will change.

B | LOCATIONAL SURROGATES IN GENERAL

My goal in this appendix is to say what it is for a thought to be a *location*, λ, and to then define the λ-surrogate of an arbitrary thought. To do this, I don't believe that I have to take a stand on what thoughts are, how fine- or coarse-grained they are, nor what kind of internal structure they have—beyond the minimal commitments I laid out in §3. And, given the wide diversity of views on thoughts, it would be better if my definition of the λ-surrogate of a thought did not require me to take such a stand. So, instead, I will adopt the following approach: I will assume that you have provided me with a certain set of thoughts over which your credences are defined. These thoughts could be sets of centred epistemically possible worlds, or guises, or structured propositions, or algorithms for computing truth-values, or what-have-you. So long as they satisfy the criteria I laid out in §3 above (together with some additional closure properties I'll introduce in §B.1 below), they're fair game. Then, with some minimal information about these thoughts (in particular: which are *de dicto*, which are *de se*, and which are *a priori* knowable), I will go on to say what it takes for one thought to be a λ-surrogate of another.

B.1 Scenarios

So I'll assume that there's some set of thoughts, $\mathcal{T}$, over which your credences are defined. I am going to take for granted that, while this set may be uncountably infinite, it is at least recursively generated from some underlying countable set of thoughts, which I will call the *primitive thoughts*, $\mathcal{P} \subseteq \mathcal{T}$. I'll suppose that the set $\mathcal{T}$ is closed under finite negation and conjunction (and, thereby, under finite disjunction as well). That is: if p and q are included in $\mathcal{T}$, then so too are $\sim p$ and $p \wedge q$. Thus, we can consider the set, $\mathcal{B}$, which is just the closure of $\mathcal{P}$ under finite negation and conjunction. Call the thoughts in $\mathcal{B}$ the *base* thoughts. I'll suppose that, if $\Sigma \subseteq \mathcal{B}$ is any set of base thoughts, then there is a thought, $\bigwedge \Sigma \in \mathcal{T}$, such that it's *a priori* knowable that $\bigwedge \Sigma$ is true iff every member of Σ is true. (There may be *multiple* thoughts like this, depending upon how fine- or coarse-grained thoughts are. For instance, it's *a priori* knowable that $\bigwedge \Sigma \vee \bigwedge \Sigma$ is true iff $\bigwedge \Sigma$ is. So if the latter thought is different from the former, there will be more than one thought which is true iff every thought in Σ is, *a priori*.) Given any set of base thoughts $\Sigma \subseteq \mathcal{B}$, we may then define $\bigvee \Sigma \stackrel{\text{def}}{=} \sim \bigwedge \{\sim p \mid p \in \Sigma\}$. Then, it's *a priori* knowable that $\bigvee \Sigma$ is true iff *some* thought in Σ is true. (I'll suppose that, given a set of primitive thoughts $\mathcal{P}$, the set $\mathcal{T}$ contains only the thoughts that it must contain in order to satisfy the above conditions; it contains no other thoughts than these.)

An example: suppose you will flip a coin a countably infinite number of times, and you have a countably infinite set of primitive thoughts which includes: 'the first flip lands heads', 'the first flip lands tails', 'the second flip lands heads', 'the second flip lands tails', and so on. By closing the set of primitive thoughts under finite negation and conjunction, you will not get thoughts like 'the coin lands heads every time' and 'the coin lands heads on every prime flip'. However, infinitary conjunction on the sets { 'the nth flip lands heads' | $n \in \mathbb{N}$} and { 'the nth flip lands heads' | n is prime } will give us these thoughts.

Parenthetically, you may wonder why I introduced infinitary conjunction by way of the base thoughts, $\mathcal{B}$. Why not say instead that, for any $\Sigma \subseteq \mathcal{T}$, there is a thought, $\bigwedge \Sigma \in \mathcal{T}$ such that it is *a priori* that $\bigwedge \Sigma$ is true iff every thought in Σ is? Because, if you have a fine-grained conception of thoughts, you might be inclined to think that, if two sets Σ and Σ^* contain different thoughts, then their conjunctions, $\bigwedge \Sigma$ and $\bigwedge \Sigma^*$, are different thoughts. But then, we would have said that there is an injective function from the powerset of $\mathcal{T}$ to $\mathcal{T}$. And Cantor's theorem shows that this is impossible. Closing under infinitary conjunction only for sets of base thoughts avoids this potential problem for fine-grainers.

These closure assumptions may be implausible in some applications. In those applications, the procedure I'll introduce below may fail to define locations and locational surrogates for your thoughts. However, the definition will still be general enough to apply in many of the most interesting kinds of cases from the philosophical literature. For instance, whenever the set of thoughts forms a complete, atomic Boolean algebra, these closure conditions will be satisfied. Since most of the philosophical discussion has taken place under the assumption that your credences are defined over complete, atomic Boolean algebras, these conditions should not be seen as unduly restrictive—though I acknowledge there may be some interesting cases where we shouldn't expect them to be satisfied.

I will take for granted a distinction between thoughts which are *a priori* knowable and those which are not. In broad outline, a thought is *a priori* knowable iff you can know it independent of experience. Take any set of base thoughts, $\Sigma \subseteq \mathcal{B}$. If it's not *a priori* knowable that some thought in Σ is false—that is, if $\sim \bigwedge \Sigma$ is not *a priori* knowable—then I will say that the set of thoughts Σ is *epistemically satisfiable*. Epistemic satisfiability is not the same as metaphysical satisfiability. For instance, the sets { 'Twain is gifted', 'Clemens is not gifted' } and { 'water is not H_2O' } are epistemically satisfiable but not metaphysically satisfiable. And { 'I am Fred', 'Fred is not here' } and { 'Exactly one person invented the zipper', 'Julius didn't invest the zipper' } are metaphysically satisfiable, but not epistemically satisfiable.

If a set of base thoughts $\Sigma \subset \mathcal{B}$ is epistemically satisfiable, but there is no set, Σ^*, such that $\Sigma \subset \Sigma^* \subset \mathcal{B}$ and Σ^* is epistemically satisfiable, then I'll say that Σ is a *maximal* epistemically satisfiable set of base thoughts. Let Σ be a maximal epistemically satisfiable set of base thoughts. Then, by assumption, there is a thought, $\sigma \stackrel{\text{def}}{=} \bigwedge \Sigma$, such that it is *a priori* knowable that σ is true iff every thought in Σ is true. I will call the thought σ a *scenario*.[19] A scenario is a thought which epistemically settles the truth-value of every base thought. As a notational matter, I will reserve the lowercase Greek letter 'σ' for scenarios. (Notice that there may be multiple scenarios for each maximal epistemically satisfiable set of base thoughts. After all, it's *a priori* that σ is true iff $\sigma \vee \sigma$ is, too. So if $\sigma \vee \sigma$ is a different thought from σ, then both will be scenarios corresponding to the same maximal epistemically satisfiable set of base thoughts.)

In general, given any two thoughts $p, q \in \mathcal{T}$, if the material conditional $p \to$

19. I borrow this term from Chalmers (2011*a*).

q is *a priori* knowable, then say that p *verifies* q.[20] Verification is an epistemic analogue of metaphysical necessitation. Whereas p metaphysically necessitates q iff the material conditional $p \rightarrow q$ is *necessary*, p *verifies* q iff the material conditional $p \rightarrow q$ is *a priori*. If p verifies q and q verifies p, then I'll say that p and q are *a priori equivalent*. If your credences are probabilistic,[21] then they will not distinguish between any two *a priori* equivalent thoughts. That is: if p and q are *a priori* equivalent and C is a probability, then $C(p)$ must equal $C(q)$.

For every thought, p, a scenario will either verify p or it will verify $\sim p$. To see this, suppose that Σ is a maximal epistemically satisfiable set of base thoughts, and let $\sigma \stackrel{\text{def}}{=} \bigwedge \Sigma$. By the maximality of Σ, for every base thought q, either $q \in \Sigma$ or $\sim q \in \Sigma$. If $q \in \Sigma$, then σ must verify q; and if $\sim q \in \Sigma$, then σ must verify $\sim q$. So a scenario will epistemically settle the truth-value of every base thought by either verifying it or its negation. It will thereby settle the truth-value of every thought in $\mathcal{T}$, since the truth-value of every thought in $\mathcal{T}$ is *a priori* knowable, given the truth-values of all of the base thoughts.

Scenarios may be seen as the epistemic *atoms* of thoughts; for every thought is *a priori* equivalent to some disjunction of scenarios. For instance, consider the set of scenarios $\{\sigma \mid \sigma \text{ verifies } p\}$. It is *a priori* knowable that p is true iff $\bigvee\{\sigma \mid \sigma \text{ verifies } p\}$ is. So every thought may be decomposed into a disjunction of the scenarios which verify it.

B.2 Worlds and Locations

Some of the thoughts in $\mathcal{T}$ will only concern what the world is like. They won't say anything about who you are, or when and where you are located in the world. Call these thoughts *de dicto*. Others will say something about who you are, where you are, or what time it is. Call these thoughts *de se*. Beyond these remarks, I won't be saying what it takes for a thought to be *de dicto* or *de se*. In particular, I won't be offering any formal characterisation of which thoughts are *de dicto* and which are *de se*. Instead, I hope to rely on your intuitive understanding of what it is for a thought to say something about who you are, where you are, or what time it is, and what it is for a thought to say nothing about any of those questions. If you grant me this distinction, then I'll use it to tell you what a *location* is, and to define a locational surrogate, p_λ, for each thought p and each location λ.

20. The material conditional $p \rightarrow q$ is just $\sim(p \wedge \sim q)$. I borrow the term 'verifies' from Chalmers (2011*a*).

21. When I say that your credences are probabilistic, I mean that, for any thoughts $p, q_1, q_2, \ldots$: 1) $C(p) \geqslant 0$, 2) if p is *a priori* knowable, then $C(p) = 1$; and 3) if it is *a priori* knowable that at most one of $q_1, q_2, \ldots$ are true, then $C(q_1 \vee q_2 \vee \ldots) = C(q_1) + C(q_2) + \ldots$.

Notice that a *de se* thought may tell you *more* than just when, where, or who you are. It may also tell you something about what the world is like. For instance, the thought 'I am Beyoncé and grass is green' provides both the *de se* information that you are Beyoncé and the *de dicto* information that grass is green. In this subsection, I will want to pull apart these two kinds of information. I'll do that by first defining what I'll call a *world*, and then going on to define what I'll call a *location*. Intuitively, a world will tell you everything about what things are like in as rich a detail as your *de dicto* thoughts will permit. And a location will tell you who, when, and where you are, in as rich a detail as your *purely de se* thoughts will permit.

If the set of thoughts $\mathcal{T}$ is rich enough, we will be able to factorise every scenario into a pair of a world and a location. This will allow us to reconstruct, out of the thoughts in $\mathcal{T}$, a structure similar to the framework of centred possible worlds from §4.1.

Let's start with worlds. The guiding idea is that a world is a kind of thought which tells you everything there is to tell about what the world is like, and doesn't tell you anything about who you are, what time it is, or where you are located in the world. To construct a world, take any scenario, σ, and consider the set of *de dicto* base thoughts which σ verifies. Call this set 'Ω_σ'. That is, $\Omega_\sigma = \{p \in \mathcal{B} \mid p \text{ is } \textit{de dicto} \wedge \sigma \text{ verifies } p\}$. By assumption, there is a thought $\omega_\sigma \overset{\text{def}}{=} \bigwedge \Omega_\sigma$ such that it's *a priori* knowable that ω_σ is true iff every thought in Ω_σ is true. Let us call ω_σ the *world* of the scenario σ. The world of σ verifies all of the *de dicto* thoughts which σ verifies, and it verifies nothing else. It perfectly encapsulates the *de dicto* information included in the scenario σ. As a notational matter, I'll reserve the lowercase Greek 'ω' for worlds. And, if $\omega_\sigma = \omega_{\sigma^*}$, then I will say that the scenarios σ and σ^* are *world-mates*.

On to locations. The guiding idea is that a location is a kind of thought which tells you everything there is to tell about who, when, and where you are, but doesn't tell you anything else about the world. I will build up to my definition of a 'location' by starting with the notion of a *purely de se* thought. A purely *de se* thought is a thought which does not verify any (non-trivial) *de dicto* thoughts. From this it follows that, if p is purely *de se*, then, for any scenario, σ, there will be some scenario σ^* which is a world-mate of σ's, and which verifies p. In other words, a purely *de se* thought does not rule out any worlds.

Notice that a *de se* thought like 'I am Beyoncé' needn't be purely *de se*—depending upon how rich the set of thoughts $\mathcal{T}$ is. If $\mathcal{T}$ includes thoughts about whether Beyoncé exists, then 'I am Beyoncé' will verify the *de dicto* thought that Beyoncé exists. There will nonetheless be a nearby thought which *is* purely *de se*—namely, the thought 'I am Beyoncé, *if she exists*'. (There are similar issues with distinctness claims like 'I am not Beyoncé'. 'I am not Beyoncé' verifies the *de dicto* thought 'Someone is distinct from Beyoncé'. Here, too, there is a nearby purely *de se* thought, namely 'I am distinct from Beyoncé if anyone is'.)

To construct a location, take any scenario, σ, and consider the set of purely *de se* base thoughts which σ verifies. Call this set Λ_σ. That is, $\Lambda_\sigma = \{p \in \mathcal{B} \mid p \text{ is purely } de\ se \wedge \sigma \text{ verifies } p\}$. By assumption, there is a thought $\lambda_\sigma \stackrel{\text{def}}{=} \bigwedge \Lambda_\sigma$ such that it is *a priori* knowable that λ_σ is true iff every thought in Λ_σ is true. Let's call λ_σ the *location* of the scenario σ. The location of σ verifies all of the purely *de se* thoughts which σ verifies, and it verifies nothing else. It exactly expresses all of the purely *de se* information included in the scenario σ. As a notational matter, I will use lowercase Greek letters for locations, with the exception of σ and ω, which I reserve for scenarios and worlds, respectively.

B.2.1 An Example. For illustration, consider the following three thoughts: 'I am Beyoncé' (β), 'Beyoncé is sick' (b), and 'I am sick' (i). If we start with these three primitive thoughts, $\mathcal{P} = \{\beta, b, i\}$, then there will be six scenarios, which we can call '$\sigma_{\beta bi}$', '$\sigma_{\beta\overline{b}\overline{i}}$', '$\sigma_{\overline{\beta}bi}$', '$\sigma_{\overline{\beta}b\overline{i}}$', '$\sigma_{\overline{\beta}\overline{b}i}$', and '$\sigma_{\overline{\beta}\overline{b}\overline{i}}$'. In the scenario $\sigma_{\beta bi}$, you are Beyoncé, Beyoncé is sick, and (therefore) you are sick. In the scenario $\sigma_{\overline{\beta}b\overline{i}}$, you are not Beyoncé, Beyoncé is sick, but you are not. And so on for the other scenarios, in the natural way. Each of the primitive thoughts is *a priori* equivalent to some disjunction of scenarios. For instance, β is *a priori* equivalent to $\sigma_{\beta bi} \vee \sigma_{\beta\overline{b}\overline{i}}$, and i is *a priori* equivalent to $\sigma_{\beta bi} \vee \sigma_{\overline{\beta}bi} \vee \sigma_{\overline{\beta}\overline{b}i}$.

If we begin with just this set of thoughts, then b and $\sim b$ will be the only non-trivial *de dicto* thoughts, up to *a priori* equivalence. (That is to say: any non-trivial *de dicto* thought will be *a priori* equivalent to either b or $\sim b$.) Then, there will be only two worlds. For every scenario will either verify b or it will verify $\sim b$. If a scenario, σ, verifies b, then any other (non-trivial) *de dicto* thought it verifies will be equivalent to b, so the set Ω_σ of the *de dicto* thoughts verified by σ will contain only trivial thoughts like $b \vee \sim b$ and thoughts equivalent to b. Then, the conjunction $\bigwedge \Omega_\sigma$ will be *a priori* equivalent to b. So b is a world. And similar reasoning shows that, for any scenario, σ, which verifies $\sim b$, $\bigwedge \Omega_\sigma$ will be *a priori* equivalent to $\sim b$. So $\sim b$ is a world. And these are the only two worlds (up to *a priori* equivalence).

Likewise, β and $\sim \beta$ will be the only purely *de se* thoughts (up to *a priori* equivalence). If you had opinions about thoughts like 'Beyoncé exists', then β would not count as purely *de se*, since it would verify the *de dicto* 'Beyoncé exists'. But, given that we're beginning with the expressively impoverished set of primitive thoughts $\{\beta, b, i\}$, β will not verify any *de dicto* thoughts about which you have opinions.

Then, there will be only two locations (up to *a priori* equivalence): β and

$\sim\beta$. Every scenario will either verify β or it will verify $\sim\beta$. If a scenario, σ, verifies β, then any other (non-trivial) purely *de se* thought it verifies will be equivalent to β. And so the set Λ_σ of purely *de se* thoughts it verifies will contain only trivial thoughts like $\beta \vee \sim\beta$ and thoughts which are *a priori* equivalent to β. Then, the conjunction $\bigwedge \Lambda_\sigma$ will be *a priori* equivalent to β. So β will be a location. Similar reasoning shows that, for any scenario, σ, which verifies $\sim\beta$, $\bigwedge \Lambda_\sigma$ will be *a priori* equivalent to $\sim\beta$. So $\sim\beta$ will be a location. And these are the only two locations (up to *a priori* equivalence).

B.2.2 Centred Worlds. Even though locations are defined in terms of purely *de se* thoughts, they need not be purely *de se* themselves. To illustrate, notice that both 'I am Twain, if he exists' and 'I am Clemens, if he exists' are purely *de se*. While neither of these *individually* verifies any *de dicto* thought, *together* they verify the *de dicto* thought that, if both Twain and Clemens exist, then Twain is Clemens. Therefore, the location of any scenario which verifies both 'I am Twain, if he exists' and 'I am Clemens, if he exists' will verify the *de dicto* thought that Twain is Clemens, unless one of them doesn't exist. And this thought is not be purely *de se*.

For this reason, it could be that a location is not compatible with every world. That is, there could be a world ω and a location λ such that ω verifies $\sim\lambda$—or, equivalently, $\{\omega, \lambda\}$ is not epistemically satisfiable. For instance, let ω be a world at which Twain does not exist, and let λ be a location which verifies 'I am Twain'. Then, ω will verify $\sim\lambda$. If ω does not verify $\sim\lambda$, then I will call the ordered pair (ω, λ) a *centred world*. (To ward off a potential confusion: I've borrowed the phrases 'world' and 'centred world' to bring out structural similarities with the more familiar framework of centred metaphysically possible worlds, but the thing I'm calling a 'centred world' is just a pair of thoughts—it is not a pair of a metaphysically possible world and a person, place, and time.)

For every scenario, σ, there will be exactly one centred world $(\omega_\sigma, \lambda_\sigma)$ such that σ verifies $\omega_\sigma \wedge \lambda_\sigma$. So to each scenario there corresponds exactly one centred world. However, it could turn out that two different scenarios correspond to the same centred world. For instance, in the example from §B.2.1, there are two scenarios ($\sigma_{\overline{\beta}bi}$ and $\sigma_{\overline{\beta}b\overline{i}}$) which verify both the world b and the location $\sim\beta$. So there are two scenarios corresponding to the one centred world $(b, \sim\beta)$. The reason for this is that the conjunction $b \wedge \sim\beta$ does not verify either i or $\sim i$—that is 'Beyoncé is sick and I am not Beyoncé' does not verify either 'I am sick' or 'I am not sick'. For this reason, the centred world $(b, \sim\beta)$ does not epistemically settle every thought; there are multiple ways of 'filling out' the the scenario, consistent with the information provided by $b \wedge \sim\beta$.

Nonetheless, in many cases, the correspondence between scenarios and *locations*

worlds	β	$\sim\beta$	
b	$\sigma_{\beta bi}$	$\sigma_{\bar{\beta}bi}$	$\sigma_{\bar{\beta}b\bar{i}}$
$\sim b$	$\sigma_{\beta\overline{bi}}$	$\sigma_{\overline{\beta b}i}$	$\sigma_{\overline{\beta bi}}$

FIGURE 1: With the three primitive thoughts 'I am Beyoncé' (β), 'Beyoncé is sick' (b) and 'I am sick' (i), there are six scenarios, $\sigma_{\beta bi}, \sigma_{\beta\overline{bi}}, \sigma_{\bar{\beta}bi}, \sigma_{\bar{\beta}b\bar{i}}, \sigma_{\overline{\beta b}i}$, and $\sigma_{\overline{\beta bi}}$. There are two worlds, b and $\sim b$. b is verified by the scenarios in the first row, and $\sim b$ is verified by the scenarios in the second row. And there are two locations, β and $\sim\beta$. β is verified by the scenarios in the first column, and $\sim\beta$ is verified by the scenarios in the second column. The location β is atomic, and the location $\sim\beta$ is not atomic.

centred worlds will be one-to-one. I'll say that a location λ is *atomic* iff, for every world ω, there is no more than one scenario which verifies $\lambda \wedge \omega$. (In the example from §B.2.1, the location β is atomic. For every world, β is verified by a single scenario in that world: $\beta \wedge b$ is only verified by $\sigma_{\beta bi}$, and $\beta \wedge \sim b$ is only verified by $\sigma_{\beta\overline{bi}}$. However, the location $\sim\beta$ is not atomic. For some world, $\sim\beta$ is verified by multiple scenarios in that world: for instance, $\sim\beta \wedge b$ is verified by both $\sigma_{\bar{\beta}bi}$ and by $\sigma_{\bar{\beta}b\bar{i}}$.) If *every* location is atomic, then the the correspondence between scenarios and centred worlds will be one-to-one. That is: for every centred world (ω, λ), there will be a scenario σ such that $\omega \wedge \lambda$ is *a priori* equivalent to σ.

B.3 Locational Surrogates

Take any thought, p, and any location, λ. With these, I wish to find a surrogate for p which, roughly and metaphorically, is true *anywhere* in a world so long as p is true at the location λ in that world. I'll call this the λ-surrogate of p, and I'll write it 'p_λ'.

In general, I will say that the λ-surrogate of a thought, p, is a thought, p_λ, which is verified by any scenario which is a world-mate of a scenario which verifies both p and λ. For, if you are in a scenario which is a world-mate of a scenario which verifies both p and λ, then you are in a world such that $\lambda \wedge p$ expresses a truth *somewhere* in that world. So let us define p_λ as follows:

$$(p_\lambda) \qquad p_\lambda \stackrel{\text{def}}{=} \bigvee \{\sigma \mid \exists \sigma^* : \omega_\sigma = \omega_{\sigma^*} \wedge \sigma^* \text{ verifies } \lambda \wedge p\}$$

(We needn't concern ourselves with distinctions between *a priori* equivalent thoughts, so I'll call any thought which is *a priori* equivalent to this disjunction 'p_λ'.) In the example from §B.2.1, the β-surrogate of the thought i, i_β, will be the disjunction $\sigma_{\beta bi} \vee \sigma_{\bar{\beta}bi} \vee \sigma_{\bar{\beta}b\bar{i}}$, which is *a priori* equivalent to b, 'Beyoncé is sick'.

(The scenario $\sigma_{\beta b i}$ is included in the disjunction because it is a world-mate of itself, and it verifies both β and i. And both $\sigma_{\overline{\beta} b i}$ and $\sigma_{\overline{\beta} b \overline{i}}$ are world-mates of $\sigma_{\beta b i}$, which verifies both β and i.)

Notice that, by construction, p_λ will be a *de dicto* thought, whether or not p is *de dicto*. For p_λ will be *a priori* equivalent to a disjunction of worlds. To appreciate this, notice that, if one scenario which verifies the world ω is included in $\{\sigma \mid \exists \sigma^* : \omega_\sigma = \omega_{\sigma^*} \wedge \sigma^* \text{ verifies } \lambda \wedge p\}$, then so too are all of its world-mates. So p_λ does not distinguish between scenarios which are in the same world. Notice also that the λ-surrogate of λ itself, λ_λ, will be a *de dicto* thought which is true so long as the thought λ expresses a truth for somebody, somewhere and somewhen. So λ_λ says (roughly, and metaphorically) that the location λ is *occupied*.

This locational surrogate behaves well so long as the location λ is atomic. However, if λ is non-atomic, it will give some odd results. For instance, consider the non-atomic location $\sim\beta$. The $\sim\beta$-surrogate of i, $i_{\sim\beta}$, will be the *a priori* knowable disjunction $\sigma_{\beta b i} \vee \sigma_{\overline{\beta} b i} \vee \sigma_{\overline{\beta} b \overline{i}} \vee \sigma_{\beta \overline{b i}} \vee \sigma_{\overline{\beta b} i} \vee \sigma_{\overline{\beta b i}}$. (The first three scenarios are included in the disjunction because they are all world-mates of the scenario $\sigma_{\overline{\beta} b i}$, which verifies $\sim\beta \wedge i$. The last three are included because they are world-mates of $\sigma_{\overline{\beta b} i}$, which verifies $\sim\beta \wedge i$.)

In §5 below, I will (in rough outline) advise you to defer to your doctor by setting your credence that you are sick, i, given that you are at some location, λ, equal to your doctor's credence in i_λ. It will be important, then, that the location λ is *atomic*. Think about what happens if we apply this advice to the non-atomic location $\sim\beta$ in our example. Then, it will tell you: give that you are not Beyoncé, you should be certain that you are sick. (Since $i_{\sim\beta}$ is *a priori* knowable, your doctor will be certain of it.) This is bad enough, but it would also tell you that, given that you are not Beyoncé, you should be certain that you are *not* sick—since $(\sim i)_{\sim\beta}$ is *also* the *a priori* knowable disjunction of every scenario.

If we suppose that every location is atomic, we may provide an alternative characterisation of the thought p_λ. For, if every location is atomic, then each scenario corresponds one-to-one with a centred world, (ω, α). Then, we may say that a centred world verifies a thought iff the thought is verified by the centred world's corresponding scenario. And we may then characterise the thought p_λ in terms of which centred worlds verify it. In particular, we may say that (ω, α) verifies p_λ iff (ω, λ) is a centred world and (ω, λ) verifies p. So long as every location is atomic, this implicit definition of p_λ will be equivalent to the one given in (p_λ) above.

Braun, David. 2016. “The Objects of Belief and Credence.” In *Mind*, **125** (498): 469–497. [i]

Chalmers, David. 2011*a*. “The Nature of Epistemic Space.” In *Epistemic Modality*, edited by Andy Egan & Brian Weatherson, Oxford: Oxford University Press. [v], [vi]

Chalmers, David J. 2011*b*. “Frege’s Puzzle and the Objects of Credence.” In *Mind*, **120** (479): 587–635. [11], [i]

Dorst, Kevin. 2020. “Evidence: A Guide for the Uncertain.” In *Philosophy and Phenomenological Research*, **100** (3): 586–632. [6]

Dorst, Kevin, Levinstein, Benjamin A., Salow, Bernhard, Husic, Brooke E., & Fitelson, Branden. ms. “Deference Done Better.” [1], [6]

Elga, Adam. 2007. “Reflection and Disagreement.” In *Noûs*, **41** (3): 478–502. [5]

Elga, Adam. 2013. “The puzzle of the unmarked clock and the new rational reflection principle.” In *Philosophical Studies*, **164**: 127–139. [6]

Fitts, Jesse. 2014. “Chalmers on the Objects of Credence.” In *Philosophical Studies*, **170** (2): 343–358. [i]

Gaifman, Haim. 1988. “A Theory of Higher Order Probabilities.” In *Causation, Chance, and Credence: Proceedings of the Irvine Conference on Probability and Causation*, edited by Brian Skyrms & William L. Harper, Dordrecht: Kluwer Academic Publishers, volume 1, 191–220. [5]

Gibbard, Allan. 2012. *Meaning and Normativity*. Oxford: Oxford University Press. [ii]

Hall, Ned. 1994. “Correcting the Guide to Objective Chance.” In *Mind*, **103** (412): 505–517. [6]

Ismael, Jeann. 2008. “Raid! Dissolving the Big, Bad Bug.” In *Noûs*, **42** (2): 292–307. [5]

Ismael, Jeann. 2015. “In Defense of IP: A Response to Pettigrew.” In *Noûs*, **49** (1): 197–200. [5]

Kripke, Saul. 1979. “A Puzzle About Belief.” In *Meaning and Use*, edited by A. Margalit, Dordrecht: Springer, volume 3, 239–283. [3]

Lewis, David K. 1980. "A Subjectivist's Guide to Objective Chance." In *Studies in Inductive Logic and Probability*, edited by Richard C. Jeffrey, Berkeley: University of California Press, volume II, 263–293. [2]

Lewis, David K. 1994. "Humean Supervenience Debugged." In *Mind*, **103** (412): 473–490. [6]

Pettigrew, Richard & Titelbaum, Michael G. 2014. "Deference Done Right." In *Philosophers' Imprint*, **14** (35). [5]

Skyrms, Brian. 1980. "Higher Order Degrees of Belief." In *Prospects for Pragmatism*, edited by D. H. Mellor, Cambridge University Press, chapter 6, 109–137. [5]

van Fraassen, Bas C. 1984. "Belief and the Will." In *The Journal of Philosophy*, **81** (5): 235–256. [2], [15]

van Fraassen, Bas C. 1995. "Belief and the Problem of Ulysses and the Sirens." In *Philosophical Studies*, 77: 7–37. [15]

Yli-Vakkuri, Juhani & Hawthorne, John. forthcoming. "Intensionalism and Propositional Attitudes." In *Oxford Studies in the Philosophy of Mind*. [ii]