CHAPTER 31

SELF-LOCATING CREDENCES

MICHAEL G. TITELBAUM

31.1 The Diachronic Problem

Just as an agent can be more or less confident that straw can be spun into gold, she can be more or less confident that today is Tuesday or that she is standing on Mount Hood. So besides representing an agent’s degrees of belief about what the world is like, we might want to represent her degrees of belief about where (or when, or who) she is in the world.

Doing so may require us to reconsider the objects over which credences are distributed. Suppose, for example, we have been taking credences to be distributed over propositions defined as sets of possible worlds. To accommodate self-locating credences we might follow Lewis (1979) in distributing credences over propositions that are sets of centered worlds—where a centered world \( \langle W, c \rangle \) is an ordered pair of a traditional possible world and a center. Propositions composed of centered worlds can be sorted into two kinds: A proposition is uncentered if for any centered world \( \langle W, c \rangle \) it includes, it also includes every other centered world indexed to the same \( W \). Otherwise a proposition is centered. “Today is Tuesday” expresses a centered proposition, because it divides centered worlds indexed to the same traditional world from each other; “lead can be spun into gold” expresses an uncentered proposition. (Henceforth I will refer to traditional possible worlds simply as “worlds.”)

Credences obeying the probability axioms can be distributed just as easily over propositions composed of centered worlds as they can over traditional propositions. So this change in the objects of credence does not create any problems for Bayesians’ usual synchronic constraints. Of course there is debate over whether propositions composed of centered

---

1 A center is typically taken to be an ordered pair of a time and an individual. If we are allowing for the possibility of time-travel, a center may need to be a triple of a time, an individual, and a location. This allows for the possibility that having time-traveled, the same individual might be in two locations at the same time.

2 The relevant technical point is that the set of propositions composed of centered worlds can form a sigma algebra just as easily as the set of propositions composed from standard worlds, which is what we require to assign a probabilistic distribution over the set.
self-locating credences

worlds are the true objects of degrees of belief, but because that is the most popular position in the self-locating credence literature I will adopt it here. Most of the arguments in the objects-of-credence debate are familiar from the objects-of-belief debate, so credences don't add much that's new here.3

What is new is the havoc self-locating credences create for Bayesians' traditional diachronic constraint: updating by conditionalization. The current consensus in the self-locating credence literature is that obtaining a general updating scheme for degrees of belief in both centered and uncentered propositions requires us to alter (or at least supplement) conditionalization in some way.

To see why, suppose we have an agent who is currently certain that it is Tuesday. Intuitively, it is apparent there are some things that agent could learn as time goes on that would make it rational for her to decrease her certainty in that proposition. (In fact, if she is carefully watching a clock she is certain is accurate, there are some things the agent could learn that would make her credence in that proposition reach zero!) We want our updating rule to capture such transitions. But if conditionalization says that upon learning proposition E, the agent should change her credence in the proposition T that it is Tuesday to

$$P_j(T) = P_i(T | E) = \frac{P_i(T \& E)}{P_i(E)} \quad (31.1)$$

(where $P_i$ and $P_j$ are her prior and posterior credences, respectively), we have a problem: Since $P_i(T) = 1$, the probability calculus tells us there is no E such that $P_i(E) > 0$ and $P_i(T \& E) \neq P_i(E)$. So for any E we feed into the conditionalization rule, it either will tell us that $P_j(T) = 1$ or will give us no guidance as to what that credence should be (because it will set that credence equal to a fraction with denominator zero).4

One might attribute this problem to the agent's (perhaps unreasonable) initial certainty in T, but similar problems arise for Jeffrey-conditionalization updating regimes that honor Regularity by forbidding certainty in empirical propositions.5 So instead of retreating to Jeffrey conditionalization, a number of authors have granted that an agent might reasonably be certain of T and have proposed new formal updating schemes to replace conditionalization. These updating schemes replicate the effects of conditionalization when

---

3 Though see Chalmers (2011) for the suggestion that credence considerations add new teeth to Frege’s Puzzle arguments against referentialism about content. For other discussions of the objects of credence see Stalnaker (2008) and Pust (2012). In Titelbaum (2013b) I suggest that Bayesians can remain neutral among competing views about the objects of credence by modeling degrees of belief using numbers assigned to natural-language sentences.

4 One might avoid this problem by making the following moves: first, read conditionalization as asserting the first equality in Equation (31.1) but not the second; next, define conditional credences in some way other than as ratios of unconditional credences (see Kenny Easwaran’s chapter (9) in this volume); and finally, tell a positive story about the values of credences conditional on propositions with unconditional-credence zero that yields substantive and plausible results for cases such as the T/E case under discussion and the cases that appear later in this essay. Chalmers (2011, section 13) floats something like this proposal, but I have never seen anyone develop it into a full-blown updating system.

5 For discussion of the problems self-location causes Jeffrey conditionalization see Kim (2009) and Titelbaum (2013b, chapter 12).
self-location is not involved, but also yield plausible results when agents update centered propositions such as $T$.

## 31.2 A Master Narrative

The main difficulty in describing these new updating schemes is that there are so many of them (more than a dozen at the time of writing!). Instead of working through the details of each, I am going to sort them into three groups. For each group, I will (at a high level of abstraction) describe the basic updating approach of the schemes in that group, then describe difficulties common to every scheme in the group.

To understand the differences between these three sorts of schemes, consider the following story:

**Rick**: So far no rain has fallen on the 4th of July, but Rick can see ominous clouds approaching on the horizon. So he assigns a 0.7 credence to the proposition that it rains today. Before the clouds arrive and settle the matter he falls asleep, and then wakes up the next day (certain that he’s slept for exactly one night and it’s now July 5th). To what proposition should Rick now assign a 0.7 credence?

The role of an updating scheme is to coordinate credences assigned at different times. Conditionalization, for instance, requires an agent to line up her unconditional credences at a later time with particular conditional credences assigned earlier on. The story about Rick above also asks a coordination question: it asks us to pick a proposition whose July 5th credence should align with Rick’s July 4th credence that it would rain that day. There are many propositions Rick might entertain on July 5th that would suit this role, among them the propositions expressed by these sentences:

- It rained yesterday.
- It rained on July 4th.
- It rained *that day* (where “that day” refers de re to July 4th).

Everyone will agree that Rick should assign a 0.7 credence to each of those three propositions on July 5th. But each of the schemes I will discuss focuses on one of the three, builds a formalism that directly calculates Rick’s July 5th credence in that proposition, then sets Rick’s credences in the other two via some process downstream. **Shifting schemes**, for example, provide a diachronic rule that links Rick’s July 5th credence in the proposition expressed by “It rained yesterday” to his July 4th credence in the proposition expressed by “It rains today.” Rick’s July 5th credences in the other two propositions are then determined using synchronic rules and his July 5th credence that it rained yesterday. **Stable base schemes**, on the other hand, determine Rick’s July 5th credence that it rained on the 4th

---

6 To simplify our discussion of these schemes I will henceforth set aside the Regularity issue, allow rational agents to assign certainty to contingent propositions, and not concern myself with how the schemes examined could be extended to Jeffrey-style updates.

7 This is easily done using a probability-calculus theorem I call “Substitution,” which says that if an agent is certain two propositions have the same truth-value, one proposition can be substituted for the
of July diachronically from his July 4th credences, then use that July 5th credence to synchronically determine the others. Demonstrative schemes work with the de re proposition before approaching the others.⁸

In the sections that follow I will present (simplified) representative examples of each sort of scheme. We will see that having its primary focus on a different one of these propositions leaves each sort of scheme with a different sort of blind spot.⁹

### 31.3 Shifting Schemes

Shifting schemes focus on the fact that in the Rick story, Rick’s July 4th 0.7 credence in the proposition expressed by “It rains today” should become a July 5th 0.7 credence in the proposition expressed by “It rained yesterday.” Intuitively, the latter proposition is generated from the former by shifting the indexicals backward one day (and changing tenses appropriately). July 4th credences assigned to propositions about the current day are aligned with July 5th credences assigned to propositions about the previous day; “today” becomes “yesterday,” “tomorrow” becomes “today,” etc.¹⁰

There are various ways of formally implementing this idea; shifting schemes have been offered by Kim (2009), Meacham (2010), Schulz (2010), Schwarz (2012), and Santorio (ms). Here we’ll take Kim’s scheme as representative.

Kim introduces an “at” operator that maps propositions to propositions. Kim tells us that for proposition φ and context τ, “φ at τ” means that φ is true at τ (Kim 2009, p. 302). For example, if RT is the centered proposition expressed by “It rains today,” [RT at Monday] is the uncentered proposition that it rains on Monday.¹¹

other salva veritate into any expression describing the agent’s credences. (For more on Substitution—and a proof—see Titelbaum (2013b, chapter 3.) This theorem entails that if an agent is certain two propositions have the same truth-value, he must assign them the same credence. Since Rick is certain on July 5th that all three bulleted propositions have the same truth-value, setting his credence in any one suffices to set his credence in the others.

⁸ A great deal of the structure of this chapter—as well as many of the examples in it—were inspired by Branquinho (2008). Branquinho’s main question is about the coordination of thoughts, not credences: in what circumstances can we say an agent has maintained the same thought (or belief) over time, or (approached slightly differently) what does it take to re-express the same thought one had earlier when that thought was earlier expressed in indexical terms? Branquinho’s work is in turn a response to proposals in Evans (1981), Kaplan (1988b), and Kaplan (1988a). (I am grateful to Peter Ludlow for introducing me to Branquinho’s work and for discussions of it.)

⁹ There are a number of discussions of diachronic self-locating credence constraints that, while excellent, do not put forward a comprehensive formal scheme for updating self-locating degrees of belief. (See, for example, Bostrom (2007), Bradley (2010), Cozic (2011), Bradley (2011), Horgan and Mahtani (2013), and Manley (ms).) For reasons of space I will not discuss those here.

¹⁰ From a completely unscientific study of exactly one subject, I can report that even relatively new language users are capable of shifting indexicals in this fashion. When her mother told my three-year-old, “Your father can give you some of my chocolate,” she immediately turned to me and reported, “You can give me some of mom’s chocolate.”

¹¹ Kim doesn’t define the “at” operator much more precisely than this, but we can give a simple definition that works at least for the case in which τ is a specific center. In this case the proposition [φ at
Now consider cases in which an agent learns nothing between times \( t_i \) and \( t_j \) except that the time has advanced from \( t_i \) to \( t_j \). In these cases, Kim argues that the agent should update according to this rule:

\[
P_j(X) = P_i(X \text{ at } t_j)
\]

The idea here is that if the agent learns nothing between \( t_i \) and \( t_j \) about how events unfold in the world, his credence at \( t_j \) that \( X \) is true then should be the credence he had at \( t_i \) that \( X \) would be true at \( t_j \).

Since Rick learns nothing during our story except that some time has passed, this rule applies to his case. So if \( \text{RY} \) is the centered proposition expressed by “It rained yesterday,” we have

\[
P_{\text{July } 8} (\text{RY}) = P_{\text{July } 7} (\text{RY at July } 8 | \text{SY at July } 8)
\]

The proposition yielded on the right-hand side by the “at” operator is the (uncentered) proposition that it rains one day before July 5th. But that, of course, is just the proposition that it rains on July 4th, which is captured on July 4th by \( \text{RT} \). So we have

\[
P_{\text{July } 8} (\text{RY}) = P_{\text{July } 4} (\text{RT})
\]

This says that Rick’s July 5th credence that it rained yesterday should equal his July 4th credence that it rains today. Kim’s “at” operator formally does the job we do informally by shifting indexicals around—transforming “today” to “yesterday,” “tomorrow” to “today,” and so forth as the days roll by.

What about cases in which agents learn more than just that time has passed? The rule we’ve just seen generalizes to Kim’s

**Shifted Conditionalization** If an agent learns proposition \( Y \) (and nothing stronger) between \( t_i \) and \( t_j \), and is certain at \( t_j \) that it is \( t_j \), then for any proposition \( X \)

\[
P_j(X) = P_i(X \text{ at } t_j | Y \text{ at } t_j)
\]

For example, suppose Rick’s story were changed so that upon awakening on July 5th he learned that it had stormed the previous evening. Then Shifted Conditionalization would give us

\[
P_{\text{July } 5} (\text{RY}) = P_{\text{July } 4} (\text{RY at July } 5 | \text{SY at July } 5) = P_{\text{July } 4} (\text{RT} | \text{ST})
\]

where \( \text{SY} \) is the proposition expressed by “It stormed yesterday evening” and \( \text{ST} \) is the proposition expressed by “It storms this evening.” Equation (31.4) tells us that if Rick wakes up on July 5th and learns that it stormed the previous evening, his credence that it rained on July 4th should be just what his credence would have been on July 4th that it would rain that

\[
\tau \] is the set of world-center pairs \( \langle W, c \rangle \) such that \( \langle W, \tau \rangle \) is a member of \( \phi \). This means that whenever \( \tau \) is a specific center, \( \phi \) at \( \tau \) will be an uncentered proposition. (It also means that applying the “at” operator with a specific center to an uncentered proposition yields the same uncentered proposition.) Since our examples all concern day-long contexts, a day will count as a specific center for our purposes. More precisely, Rick is certain on July 4th that the proposition \( \text{RT} \) has the same truth-value as the proposition \( [\text{RY at July } 5] \). So Substitution (see note 7 above) allows us to move from Equation (31.2) to Equation (31.3).
day had he learned on July 4th that it was going to storm \textit{that} evening.\textsuperscript{13} Which is exactly right.

Yet Kim’s Shifted Conditionalization has a shortcoming, revealed by stories like this:

\textbf{Rip van Winkle:} On the 4th of July Rip van Winkle has credence 0.7 in “It rains today.” He falls asleep and wakes up some time later, wildly uncertain what day it is. To what proposition should he now assign a credence of 0.7?

Intuitively, shifting schemes want to answer this question by taking Rip’s July 4th credence in the proposition expressed by “It rains today” and replacing “today” with some other indexical reflecting the relationship of July 4th to the day when Rip awakens. The trouble is that Rip isn’t certain how many days he slept, so he doesn’t know which indexical (“yesterday”? “two days ago”? “a week ago”? ) is right for the job. To shift properly, an agent needs to be certain how the context in which he’s assigning his current credences relates to the context in which he assigned the credences he’s updating.

This problem is reflected formally in Shifted Conditionalization’s requirement that the agent be certain at $t_j$ that it is $t_j$. One could (and Kim does) suggest that if the agent is uncertain at $t_j$ whether it is $t_j$ or some other time $t_k$, the agent’s $t_j$ credence in $X$ should be a weighted average of the credence she would assign were she certain it was $t_j$ and the credence she would assign were she certain it was $t_k$, with the weights determined by her unconditional $t_j$ credence that it is $t_j$ or $t_k$ (respectively). The trouble with this is that the agent’s unconditional $t_j$ credences about which day it is have to be set antecedently to applying this weighting rule, via some other device altogether. For an agent who is uncertain what time it is (or where she is, or who she is), the shifting scheme itself will provide few constraints on her credences in the relevant self-locating propositions. And so the shifting scheme will not be able to tell the agent which new credences to coordinate with her old ones. Shifting schemes are very little help telling Rip which proposition gets a 0.7 credence when he awakens.

\subsection*{31.4 \textbf{Stable Base Schemes}}

When Rip awakens he may be rationally required to distribute his credences over the days it might be in a particular way. Before he fell asleep Rip may have assigned credences to uncentered propositions expressed by “When Rip falls asleep on July 4th he awakens on July 5th,” “When Rip falls asleep on July 4th he awakens on July 6th,” etc. These credences should presumably remain intact when Rip wakes up, which will drive him to assign particular credences to the centered propositions expressed by “Today is July 5th,” “Today is July 6th,” etc.

Shifting schemes are not designed to take advantage of these credences Rip assigned on July 4th; stable base schemes, on the other hand, are built entirely around an agent’s credences in uncentered propositions. Stable base schemes rely on the fact that only centered

\textsuperscript{13} Again, the last step in Equation (31.4) follows by Substitution (note 7).
propositions make trouble for conditionalization. Their core idea is to focus on the set of uncentered propositions an agent entertains, updating the members of that stable base by standard conditionalization on the uncentered propositions the agent learns. Credences in centered propositions are then set by coordinating them with the uncentered distribution at a given time. Stable base schemes are offered by (Halpern 2005), (Meacham 2008), (Titelbaum 2008), (Briggs 2010), and (Titelbaum 2013b).

While Halpern’s and Meacham’s schemes are presented somewhat differently, they ultimately describe the same three-step process for generating a $P_j$ distribution from one’s earlier $P_i$ assignments, a process Meacham calls “Compartmentalized Conditionalization”: 14

1. Temporarily consider your $P_i$ distribution only over worlds.
2. Assign any worlds incompatible with your $t_j$ evidence a credence of 0, then renormalize the remaining (non-zero) credences over worlds.
3. Now assign any centered worlds incompatible with your $t_j$ evidence a credence of 0. Then take the credence assigned to each world and distribute it among the centered worlds indexed to that world compatible with your $t_j$ evidence.

Intuitively, Compartmentalized Conditionalization generates a new distribution over uncentered propositions by conditionalizing on the strongest uncentered proposition the agent learns (that’s step (2)), then leaving that uncentered distribution intact while sorting out credences for particular centered worlds (step (3)). In Rip’s case, Rip assigns a 0.7 July 7th credence to the uncentered proposition $R_7$ that it rains on July 7th. He falls asleep, awakens, then applies Compartmentalized Conditionalization. Since Rip hasn’t eliminated any worlds upon awakening, the conditionalization in step (2) does not alter his credences in any uncentered propositions. So he continues to assign $R_7$ a credence of 0.7, exactly as he should.

14 Titelbaum (2013b, chapter 8) argues that the true troublemakers for conditionalization are not centered propositions but instead “epistemically context-sensitive” propositions, where the epistemically context-sensitive/context-insensitive distinction cross-cuts the centered/uncentered distinction in some cases. We will ignore that complication here.

15 Titelbaum (2013b) presents a version of the formal updating framework described in Titelbaum (2008) that has been altered in response to a counterexample from Moss (2012). (For an informal presentation of the new framework, see Titelbaum (2013a).) Briggs (2010) presents two formal updating schemes, one of which matches that of Meacham (2008). Meanwhile Meacham has abandoned the stable base scheme presented in Meacham (2008) in favor of the new shifting scheme of Meacham (2010), because of counterexamples he describes in this later work.

16 It’s easiest to explain Compartmentalized Conditionalization in terms of an agent’s “credence distribution over worlds” and her “distribution over centered worlds.” By “credence in a world $W$” I mean credence in the uncentered proposition containing all and only the centered worlds indexed to $W$; by “credence in a centered world $W,c$” I mean credence in the proposition containing only that world. Once one has a credence distribution over centered worlds the probability axioms generate unique credences for all other propositions.

17 Halpern and Meacham each discuss specific proposals for apportioning one’s credence in a world among the centered worlds indexed to it—for example, one could apply an indifference principle (à la Elga (2004)) to distribute a world’s credence equally among its centered worlds compatible with the agent’s $t_j$ evidence. But the details of those proposals are irrelevant to the points I will make here.
Compartmentalized Conditionalization can also take advantage of any credences Rip assigned on July 4th about how long he'd be asleep. If Rip was 0.5 confident on July 4th in the uncentered proposition that he'd sleep until July 5th, then step (2) has him retain this credence upon awakening and step (3) makes this a 0.5 credence that it's now July 5th.\footnote{Notice, by the way, that even if Rip didn't have any July 4th credences about how long he'd be asleep, Compartmentalized Conditionalization could still find a proposition at the later time to coordinate with Rip's initial 0.7 credence in RT. That's because—unlike the shifting schemes—Compartmentalized Conditionalization doesn't rely on Rip's opinions about how long he's been asleep to locate that coordinating proposition.} We could also, if we like, add to Rip's story that he gains some uncentered information upon awakening—perhaps he learns that it stormed the evening of July 4th. If we call that uncentered proposition \( S_4 \), step (2) of Compartmentalized Conditionalization would set Rip's new credence in \( R_4 \) equal to \( P_{\text{July 4}}(R_4 | S_4) \).

Under Compartmentalized Conditionalization all the changes to an agent's uncentered distribution happen in step (2), and such changes can be made only by ruling out worlds. So Compartmentalized Conditionalization is committed to the

**Relevance-Limiting Thesis** An agent should change her credence distribution over uncentered propositions only if her new evidence eliminates worlds.\footnote{I introduced this thesis in a slightly different form in T Taelbaum (2008). See Draper (2013) for further discussion.}

The Relevance-Limiting Thesis has some intuitive appeal—it suggests that if a learning episode leaves one's space of epistemically possible worlds intact, it should leave one's credence distribution over that space unaltered as well. But the following counterexample demonstrates that the Relevance-Limiting Thesis is false:

**Mystery Bag**: You are one of ten people arranged in a circle in a room. A fair coin has been flipped to determine the contents of a bag: on heads the bag contains nine black balls and one white ball; on tails it's nine white and one black. The bag is passed around the room. Each person draws one ball, holds onto it, and passes the bag until it's empty. You can't see anyone else's ball, but the one you've drawn is black.

Everyone should agree that seeing your ball is black should increase your confidence that the coin came up heads; a standard Bayesian calculation shows your credence in heads increasing from 1/2 to 9/10 upon that discovery. But now let's add the wrinkle that in this story you have no qualitative way of discriminating between yourself and the other subjects in the room. We'll have to imagine that you don't know your own name (perhaps your memory has been erased, or perhaps you were raised by scientists and never given a name); you look exactly like everyone else in the room; the room is cylindrical so you can't describe yourself as, say, "The guy in the corner"; etc. Whatever science-fiction elements are needed to make this work, consider them added to the story.\footnote{While I discovered the Mystery Bag example independently (and presented it in T Taelbaum (2013b, chapter 10)), it is very similar to a counterexample to the Relevance-Limiting Thesis presented in Bradley (2011, section 9), which Bradley in turn attributes to Matt Kotzen.}

The point of this wrinkle is that when you see the black ball, your new evidence does not rule out any worlds you entertained earlier. You knew there would be at least one black ball in the bag no matter what the coin-flip outcome, so you knew all along there would exist one
person in the room who saw a black ball. You gain evidence upon seeing your ball, but it’s the centered proposition expressed by “I see a black ball.” And since such evidence enters into Compartmentalized Conditionalization only at step (1), it is incapable of changing your credence in any uncentered propositions. So Compartmentalized Conditionalization will leave your credence in the uncentered proposition that the coin came up heads unaltered by your new evidence, which is clearly the wrong result. This incorrect result will be duplicated by any updating scheme committed to the Relevance-Limiting Thesis.21

Most of the time, things we can pick out indexically we can also pick out qualitatively. Thus our credences in centered propositions are mirrored in uncentered propositions as well: on July 4th Rip assigns a credence of 0.7 not only to the centered proposition expressed by “It rains today” but also to the uncentered proposition expressed by “It rains on July 4th.” But in some (perhaps highly artificial) cases we lack descriptive means of uniquely picking out the targets of indexicals. In those cases particular pieces of centered evidence gained may not be mirrored in uncentered propositions learned. In Mystery Bag, your inability to qualitatively pick yourself out means that the centered evidence that you’ve drawn a black ball isn’t reflected in any uncentered proposition learned. Yet that centered evidence can (and should) influence your credences in uncentered propositions like heads. Because relevant centered evidence may go unreflected in uncentered propositions, we cannot follow Compartmentalized Conditionalization in letting only uncentered propositions learned affect one’s uncentered distribution.

Titelbaum (2013b) proposes an updating scheme (called CLF, for “Certainty-Loss Framework”) that addresses this flaw by limiting the application of Compartmentalized Conditionalization. CLF allows an agent to update by Compartmentalized Conditionalization only if for each time and each centered proposition she entertains, the agent has some uncentered proposition she is certain at that time has the same truth-value as the centered proposition. This ensures that Compartmentalized Conditionalization will be applied only when every piece of centered evidence learned is mirrored for the agent in some uncentered proposition; in such cases it seems safe to let the agent’s uncentered distribution be driven entirely by the uncentered propositions learned.

CLF prevents Compartmentalized Conditionalization from yielding incorrect verdicts for cases like Mystery Bag. However, it introduces a new set of blind spots, typified by the following story:

Recurring Ron: Ron has a serious sleep problem: He keeps falling asleep for a random, unpredictable number of days. He’s lost track of how many times this has happened, and each time he awakens he is uncertain for how many days he slept. On one particular awakening Ron remembers of an earlier day he was awake that he was 0.7 confident it would rain that day. But Ron isn’t certain that day occurred during his most recent awakening, and he’s pretty

Notice that if the agent in Mystery Bag does have a way of picking himself out descriptively, he gains uncentered evidence (something like “Bob Jones picks a black ball”) over the course of the story. In that case, Compartmentalized Conditionalization will correctly lead him to increase his confidence in heads. The fact that adding seemingly irrelevant distinguishing information can change Compartmentalized Conditionalization’s prescriptions is what leads Meacham to criticize it in Meacham (2010). But my complaint here is not about the difference between Compartmentalized Conditionalization’s recommendations in the two cases—it’s about the fact that Compartmentalized Conditionalization gets the anonymous case wrong.
confident that wasn’t the only day he’s ever been 0.7 confident of rain. To what proposition should Ron now assign a credence of 0.7?

Stable base schemes try to cash out an agent’s credences entirely in uncentered terms, which are straightforwardly manipulable by conditionalization. But when an agent has no qualitative way of identifying a day (or a place, or a person)—as Ron lacks a unique way of qualitatively describing the earlier day on which he assigned a 0.7 confidence to rain—such schemes fall short. CLF falls silent on Recurring Ron, identifying no coordinating proposition to which Ron should assign a 0.7 credence on the new day.  

### 31.5 Demonstrative Schemes

Of course there is a way for Ron to pick out the day in question; as long as he remembers it, he can refer to it *de re* as "that day." So there is a coordinating proposition available to Ron upon awakening; he can be 0.7 confident that it rained *that day*.

Demonstrative reference capabilities provide the core of updating schemes proposed by Moss (2012) and Stalnaker (2008). I’ll focus on the latter here. Stalnaker’s key move is to reject the distinction with which this chapter began, the distinction between beliefs about what world one is in and beliefs about where (or when, or who) one is in that world. As he puts it, “Belief about where one is in the world is always also belief about what world one is in.” (p. 55) When I believe that it is raining today, I also believe of the current day that it is rainy. So my belief that it’s raining today is accompanied by a belief that rules out particular worlds, worlds in which *this* day is not a rainy one. And notice that this will be true even if I lack a qualitative, non-demonstrative way of uniquely picking out the current day. As Stalnaker sees it, centered information is *always* mirrored in uncentered propositions.

To make this work, we need to understand worlds—traditional worlds, the kinds to which centers are added to make centered worlds—in a way that incorporates demonstrative information. (So that we can distinguish between worlds in which it rains *this* day and worlds in which it doesn’t.) Stalnaker has technical proposals for how to do this; assessing those proposals goes far beyond the scope of this chapter. But notice that

---

22 Titelbaum (2013b) proposes a workaround for this problem: We introduce a nearby story in which Ron can qualitatively pick out the relevant day, derive verdicts for that story, then argue that this story is enough like the original that our verdicts should apply there too. But this is still a workaround—it isn’t a direct application of CLF to the Recurring Ron case.

23 The updating scheme in Santorio (ms) incorporates *de re* elements, but is still essentially a shifting scheme.

24 Stalnaker might be uncomfortable putting the point in “centered” vs. “uncentered” terms, in part because his own formalism employs centers in a different way from Lewis’s. Moss prefers to discuss “*de se*” vs. “*de dicto*” propositions. At times in Stalnaker (2008, chapter 3), he puts the point I’m trying to make in terms of whether it’s possible for an agent to learn something without gaining any “objective” information.

25 It’s also beyond my ambit to adjudicate between two rival proposals for understanding demonstrative reference: one that interprets demonstratives as *de re* expressions vs. another that interprets them as something like abbreviated indexical expressions of the form “the day I am thinking about right now.” Since I am analyzing Stalnaker’s updating scheme I will follow him in adopting the *de re* interpretation.
if Stalnaker’s right—if centered propositions always have uncentered reflections—then Compartmentalized Conditionalization always applies. Moreover, all the action in generating updated credences happens in Compartmentalized Conditionalization’s step (2). Uncentered propositions get their updated credence assignments at that stage, which can then immediately be copied to their centered correlates. And since step (2) is just good, old-fashioned conditionalization, Stalnaker suggests that conditionalization is the only updating rule we need.

But even with demonstrative contents taken into account, conditionalization won’t suffice to coordinate every credence. Consider the following case:

Roger Foretold: On July 4th Roger knows that he’s about to begin an extended period of sleepings and awakenings (of the kind experienced by Recurring Ron). This process begins, and some number of days later Roger finds himself awake, uncertain which awakening it is or how long he’s been asleep. On this awakening Roger looks out the window, sees clouds, and becomes 0.7 confident it will rain. With which of Roger’s July 4th credences is this 0.7 credence coordinated?

This story asks a slightly different question from our previous ones: here we have a later credence value and want to know with which earlier credence it’s coordinated. But if our updating rule is just conditionalization, that’s a valid question. On a conditionalization updating scheme every future credence you assign—even after learning evidence you previously lacked—is coordinated with some credence you assigned in the past. So there ought to be some conditional credence Roger assigns on July 4th—it will rain on such-and-such day conditional on its looking cloudy out my window on such-and-such day—that, by conditionalization, becomes his 0.7 credence in rain on the later awakening.

This causes a problem for Stalnaker because conditionalization always moves from a credence distribution over a particular space to another distribution over the same space. For Roger to take his July 4th credences, update them by conditionalization upon seeing clouds, and wind up with credences about whether it will rain today, there have to be propositions he assigned credences to on July 4th that he can on this awakening recognize as being about the current day. But there are no such propositions—there is no available way for Roger to fill in the “such-and-such’es in the previous paragraph. On July 4th Roger might consider whether it will rain tomorrow, but when he awakens he is unsure whether it currently is the day he referred to as “tomorrow” on July 4th. On July 4th Roger might also consider the prospects for rain on particular days picked out by name (“July 5th,” “July 6th,” etc.) or by qualitative description (“the day I first awaken”), but again Roger doesn’t know which of those days it is when he awakens. The demonstrative scheme was supposed to solve these lack-of-description problems using de re references; Recurring Ron couldn’t describe his earlier awakening qualitatively or relate it to the current time indexically, but he could always refer to it as “that day.” But Roger can’t do that on July 4th, because you can’t refer de re to the future. Demonstrative reference requires causal contact, which notoriously works in only one temporal direction.

Demonstrative updating schemes rely on uncentered (or “de dicto,” or “objective”) propositions invoked by demonstrative reference. As we causally interact with new objects

\[26\] Since the agent is certain a centered proposition is true just in case its uncentered correlate is, Substitution copies his unconditional credence in the latter back to the former.
we gain new referential abilities, allowing us to assign credences to uncentered propositions we could not entertain before (or perhaps the same propositions under new modes of description). But conditionalization alone cannot tell us how to update credences from a previous, smaller epistemic space to a newly-expanded one. So demonstrative schemes fall short in cases like Roger Foretold.

31.6 The Sleeping Beauty Problem

I’ve now presented a number of rival updating schemes, organized into three groups. It’s important to understand that members of one group need not see members of other groups as getting things wrong. A demonstrative schemer whose updates are driven by demonstrative reference will certainly not deny the shifter’s result that on July 5th Rick should be 0.7 confident that it rained yesterday. Members of each group simply think that by approaching self-locating update in their particular fashion they can achieve further results—fill in further blind spots—unobtainable on another approach. In fact, the most heated debates about on-the-ground results are often between members of the same group. Those debates often focus on

The Sleeping Beauty Problem: A student named Beauty arrives on Sunday to volunteer for an experiment. She will be put to sleep on Sunday night, then the experimenters will flip a fair coin. If it comes up heads, they will awaken her on Monday morning, chat with her for a bit, then put her back to sleep. If the coin comes up tails, they will engage in the same Monday process, then erase all her memories of her Monday awakening, awaken her on Tuesday morning, chat with her for a bit, then put her back to sleep.

Beauty is told all of this information, then put to sleep. She awakens on Monday morning, but because of the possibility of memory erasure is uncertain whether it is Monday. At that point, how confident should she be that the coin came up heads?

Elga (2000), who adapted this problem from Piccione and Rubinstein (1997) and introduced it to the philosophical literature, argued that the answer is 1/3. Lewis (2001) responded that the answer is 1/2. Since most parties to the resulting controversy recognized that Beauty’s only new evidence on Monday morning is self-locating (something like “It is now Monday or Tuesday”), this problem spurred much of the literature on self-locating update.

27 Weatherson (2011) makes a similar criticism of Stalnaker’s scheme, prompting Stalnaker (2011) to concede that conditionalization by itself will not generate all the updates he wants.

28 Moss’s updating rule is not just traditional conditionalization, and is cleverly arranged so that demonstrative terms need to be available only from the time they are introduced and thereafter. But her scheme still has trouble with stories in which an agent forgets information and so loses track of de re ascriptions made in the past.

29 While Sleeping Beauty may seem like an odd science-fiction problem, it has been linked to a very real issue in the philosophy of quantum mechanics. It has been argued that any updating scheme yielding a 1/3 answer to Sleeping Beauty must also yield the implausible result that every quantum experiment we ever conduct will favor Everettian (“many-worlds”) interpretations of the quantum-mechanical formalism over standard (“Copenhagen”) interpretations, regardless of the particulars of the experimental result. For citations on this discussion, and further links between Sleeping Beauty and broader philosophical issues, see Titelbaum (2013c).
One might wonder why the controversy over Beauty’s heads credence—and indeed, the whole search for a self-locating updating scheme—couldn’t have been settled by the sorts of Bayesian arguments that have been used to support diachronic rules before. For example, diachronic Dutch Books and minimizing-expected-inaccuracy arguments have been invoked to support Conditionalization, Jeffrey Conditionalization, the Reflection Principle (van Fraassen 1983), and so on. Interestingly, one can generate Dutch Book arguments and minimizing-inaccuracy arguments for each solution to the Sleeping Beauty Problem. Our understanding of how such arguments apply to cases involving self-location has not advanced far enough to settle the Sleeping Beauty dispute.

Sleeping Beauty also lies within the blind spots of all three sorts of updating scheme. When Beauty awakens on Monday she is uncertain what day it is, making trouble for shifting schemes. She also does not have a unique qualitative description for the current day, undermining stable base approaches. Finally, Beauty does not on Monday have a description also available to her on Sunday night (not even a demonstrative one) that she is certain picks out the current day, so she cannot generate current-day credences by conditionalizing the possibility space that was available to her then.

Each of the updating approaches we have considered seems reasonable as far as it goes, but gives out at some point. The creators of various schemes recognize those schemes’ limitations and in many cases have proposed workarounds, some of which suggest answers to the Sleeping Beauty Problem. But few of those schemes generate a specific solution to the problem by a simple, direct application of their formalism. Since each sort of formalism represents a different intuitive approach to coordinating credences over time, their failures suggest that our standard, intuitive approaches to reasoning about self-location may be inadequate to settle Sleeping Beauty. Perhaps it is no coincidence that despite scores of proposed formal and informal solutions, the Sleeping Beauty Problem remains controversial.

30 Because the Reflection Principle is tightly tied to conditionalization, centered propositions cause just as much trouble for the former as the latter. Arntzenius (2003) describes a number of counterexamples to Reflection that involve self-location, including the Sleeping Beauty Problem. Elga (2007) responds with a modification to Reflection that essentially implements a shifting scheme. See also Schervish, Seidenfeld, and Kadane (2004) on this topic.


32 It may seem suspicious that Sleeping Beauty (and some of the stories we considered earlier, such as Recurring Ron and Roger Foretold) involve memory loss, a phenomenon that is known to cause trouble for traditional Bayesian updating rules. (Arntzenius went so far as to assert at one point (Arntzenius 2002) that “self-locating learning plays no relevant role in the Sleeping Beauty case. The real issue is how one deals with known, unavoidable, cognitive malfunction.” He backed off from that position in Arntzenius (2003).) Authors such as Meacham (2010), Moss (2012), and Titelbaum (2013b) have explicitly worked mechanisms for modeling memory loss into their updating schemes, but these maneuvers do not remove the blind spots I described for those approaches.
ACKNOWLEDGMENTS

I am grateful to participants in the 2011 Rio Workshop on de se Attitudes for discussion of this chapter, to the editors of this volume for helpful comments, and to Olav Vassend for assistance keeping up with the literature.

REFERENCES