Self-Locating Belief & Bayesian Update

Ryan Doody

April 6, 2015

Two Problems of de se *Content for Bayesianism*

1. **Problem 1: *de se* learning.** I am watching the clock. I learn that *it is now 1:28PM.* A couple minutes pass. I learn that *it is now 1:30PM.* I know that if it is 1:30PM, then it isn't also 1:28PM. But, if I updated by conditionalization on the earlier information, hasn't my credence in that proposition increased to 1? (And once a proposition is assigned credence 1, there is no way for it be lowered via conditionalization).
2. **Problem 2: *de se* ignorance.** I'm driving down the highway. I know that I am supposed to take the third exit. I know I've driven by one exit, but I can't remember if I've only driven past one or if I've driven by two. The next exit is fast approaching. How confident should I be that it is the exit I want to take? What is it, exactly, that I am ignorant about in this case?

Content de dicto & *Content* de se

You can think of *de dicto* content as information about the way the world is, and you can think of *de se* content as information about your *location* in the world (e.g., which person you are, what time it is, where you are). Notice that the following sentence expresses *de se* content.

Example: an old-fashioned map encodes *de dicto* information. A map with a "You are here" sticker (or, a map with GPS) also encodes *de se* information.

(1) My pants are on fire.

What that sentence expresses is not, in general, equivalent to what any of the following sentences express.

(2) Ryan Doody's pants are on fire.

(3) The disheveled person's pants, whose reflection I see in the mirror, are on fire.

(4) Mr. & Mrs. Doody's eldest child's pants are on fire.

I could know (2)-(4) without realizing (1).

Formal Framework. We can think of information as a set of possible worlds. And let $\langle c, w\rangle$ be a *centered world,* where w is a possible world, and $c = \langle \text{Person}, \text{time}, \text{location}\rangle$ is a *center.*

- *de dicto information.* A boring set of centered worlds: for any world w, and centers c, c^*, if $\langle c, w\rangle$ is in the set, then $\langle c^*, w\rangle$ is in the set, too.
- *de se information.* A non-boring set of centered worlds. (For example, the set of centered worlds that only have 2:00PM, April 6th, 2015 in the time-parameter of their centers).

Moss' Update Procedure

Moss offers an account of how to update *de se* content. First, she claims that "given a *de se* proposition, there is a *de dicto* proposition that is equivalent with that *de se* proposition, given what you believe.

(PROXY) For any *de se* proposition, there exists a *de dicto* proposition such that given what you believe with certainty, the *de se* proposition and the *de dicto* proposition are equivalent.

Moss' update procedure has two steps: (1) **Black Box Update:** generate hypothetical credences from your previous opinions and your sense of time passing; (2) **Genuine Learning:** conditionalize these hypothetical credences on what you've genuinely learned.

The Sleeping Beauty Problem

Sleeping Beauty is put in the following philosophical experiment. (She knows all the details about the set-up). On Sunday, she will be put to sleep. A fair coin will be flipped. If the coin lands HEADS, she will be woken up on Monday. Then she will put back to sleep, and woken again on Wednesday when the experiment ends. If the coin lands TAILS, she will be woken up on Monday, and then put back to sleep. Her memory of her Monday waking will be erased, and she will be woken again on Tuesday.

Question: When she awakes on Monday, what credence should Sleeping Beauty have that the coin landed HEADS?

	Monday	**Tuesday**
HEADS	$M \wedge H$	$\neg M \wedge H$
TAILS	$M \wedge T$	$\neg M \wedge T$

Let M be the information that Sleeping Beauty expresses on Monday with the utterance "The waking experience occurring *now* is happening on a Monday."

1. **Halfer Argument.** Sleeping Beauty doesn't learn anything when she wakes up on Monday. She already knew the set-up of the experiment. And she know that the coin is fair. So, if she know that the coin is fair, and she hasn't learned anything new that's relevant to how the coin toss turned out, her credence in heads should be $\frac{1}{2}$.
2. **Thirder Argument.** Sleeping Beauty thinks that were she to know it is Monday, then her credence in heads should equal her credence in tails.

$$Cr(H \mid Monday) = Cr(T \mid Monday) = \frac{1}{2} \tag{1}$$

And, if Sleeping Beauty knew that the coin had landed TAILS, she would have no more reason to think that it is Monday than it is Tuesday, and vice versa, so:

$$Cr(Monday \mid T) = Cr(Tuesday \mid T)$$

So, her credence in HEADS should be $\frac{1}{3}$.

Which of these two arguments, if either, is correct? What should Sleeping Beauty's credences be?