A double-halfer embarrassment: Response to Pust

Lennart Ackermans

12th December 2024

1. Introduction

Sleeping Beauty, the renowned Bayesian reasoner, has enrolled in an experiment at the Experimental Philosophy Lab. On Sunday evening, she is put to sleep. On Monday, the experimenters awaken her. After a short chat, the experimenters tell her that it is Monday. She is then put to sleep again, and her memories of everything that happened on Monday are erased. The experimenters then toss a coin. If and only if the coin lands tails, the experimenters awaken her again on Tuesday. Beauty is told all this on Sunday. When she awakens on Monday – unsure of what day it is – what should her credence be that the coin toss on Monday lands heads?

The *thirders* argue that it should be 1/3 (Dorr, 2002; Elga, 2000; Horgan, 2004, 2007; Kim, 2022a, 2022b; Milano, 2022; Titelbaum, 2008), but some mischievous philosophers claim that it is 1/2. These *halfers* come in two kinds. *Lewisian halfers* further argue that Beauty's credence that the coin lands heads increases to 2/3 after she is told it is Monday (Lewis, 2001). *Double halfers* argue that her credence is 1/2 both before and after she is told it is Monday (Bostrom, 2007; Briggs, 2010; Meacham, 2008; Pust, 2012). This requires a denial that Beauty uses Bayesian conditionalization when she receives the evidence that it is Monday.

Titelbaum (2012) adds another detail to the experiment. The experimenters will also toss a coin on Tuesday, regardless of whether Beauty is awakened that day. When Beauty awakens on Monday, what should her credence be that *today's coin toss* lands heads?

Titelbaum's variant case is crafted to embarrass double halfers. The central intuition of the double-halfer position is that Beauty has no evidence that would justify a credence other than 1/2, both before and after she is told it is Monday. However, as Titelbaum shows, if Beauty's credence is 1/2 that the Monday coin toss lands heads (before being told it is Monday), then her credence that today's coin toss lands heads must be greater than 1/2. Besides being obviously wrong, such a credence conflicts with the central intuition of double halfism.

Pust (2023) agrees with Titelbaum that Beauty's credence in today's coin landing heads must be greater than 1/2 for halfers. However, he disagrees that this is a *distinctive* embarrassment for double halfers. Pust argues that two conflicting ways of using direct inference, both applying a direct inference principle akin to the *Principal Principle*, lead Beauty respectively to a credence of 1/2 and a credence greater than 1/2 that today's coin lands heads. Hence, according to Pust, both thirders and halfers have a problem. Neither the thirder nor the halfer has a straightforward undefeated direct inference leading to the credence that their position requires.

I show that Pust's second derivation, which concludes that Beauty's credence that today's coin lands heads is greater than 1/2, relies on a premise that is equivalent to the position of halfism itself: that Beauty assigns a credence of 1/2 to *Monday's* coin landing heads. There is a plausible and intuitive argument that Beauty has *inadmissible evidence* with respect to this proposition, which is evidence that precludes the use of direct inference. Hence, there is no embarrassment for thirders. The double halfer, on the other hand, is forced to accept Pust's second derivation. By showing that the halfer must accept conflicting probability statements – unless she randomly screams "inadmissible evidence!" in some cases, but not others – Pust's argument only deepens the state of embarrassment that double halfers ought to be in.

2. Pust's second derivation

I summarize Pust's (2023) argument using his own formalism (slightly adapted). Let *ch* be the objective probability function[1] and let P be Beauty's credence function after she awakens on Monday in Titelbaum's variant of the problem. Pust uses the following direct inference principle.

> **The Generic Direct Inference Principle (G-DIP).** Let E be evidence that is consistent with and *admissible* with respect to the proposition Ta and the objective probability evidence $ch(Tx \mid Rx) = r \,\&\, Ra$. If one's total evidence is given by $ch(Tx \mid Rx) = r \,\&\, Ra \,\&\, E$, then it is rationally required to have $P(Ta) = r$.[2]

[1] This can be a chance function or objective probability in another sense.

[2] This is based on the "Generic Direct Inference Principle" from Wallmann and Hawthorne (2020, p. 958).

It is controversial what it means for evidence to be *admissible*, but there are some straightforward cases of inadmissibility, such as non-certain disjunctions with the outcome, to which I will turn below (Wallmann & Hawthorne, 2020).

Let $\mathrm{TSB}(s)$ be the proposition that s is a Two-Toss Sleeping Beauty experiment (that is, Titelbaum's variant). Let $\mathrm{Toss}(x,s)$ be the proposition that x is a toss occurring on a day that Beauty is awake (an awakening day) in s. Let $\mathrm{H}(x)$ be the proposition that x lands heads. After Beauty has just awakened, she uses τ to refer to today's coin toss and σ to refer to the current experiment. Hence, she knows that $\mathrm{TSB}(\sigma)$ and $\mathrm{Toss}(\tau,\sigma)$.

The first derivation (endorsed by myself and implicitly by Titelbaum) uses only one step of direct inference based on G-DIP. Beauty knows that

$$ch(\mathrm{H}(x) \mid \mathrm{Toss}(x,s) \;\&\; \mathrm{TSB}(s)) = 1/2. \tag{1}$$

Supposing she has no inadmissible evidence, she should set

$$P(\mathrm{H}(\tau)) = 1/2. \tag{2}$$

Pust's second derivation works by considering three possibilities that are mutually exclusive and exhaustive. The probability of the three scenarios is obtained using direct inference. He then gives some plausible probabilities for today's coin landing heads conditional on each possibility. These probabilities are combined using the law of total probability to obtain Beauty's credence that today's coin lands heads. The three scenarios are as follows.

1. The experiment has only a single toss occurring on an awakening day, which lands heads. This possibility has an objective probability of $1/2$, since it happens just in case the Monday coin lands heads. Formally:

$$\mathrm{S}_1(s) := \exists! x(\mathrm{Toss}(x,s) \;\&\; \mathrm{H}(x) \;\&\; \neg\exists y(y \neq x \;\&\; \mathrm{Toss}(y,s)))$$
$$ch(\mathrm{S}_1(s) \mid \mathrm{TSB}(s)) = 1/2. \tag{3}$$

2. The experiment has one awakening day toss landing heads and another awakening day toss landing tails. This possibility has an objective probability of $1/4$, since the first toss must land tails and the second heads. Formally:

$$\mathrm{S}_2(s) := \exists! x(\mathrm{Toss}(x,s) \;\&\; \mathrm{H}(x)) \;\&\; \exists! y(\mathrm{Toss}(y,s) \;\&\; \neg\,\mathrm{H}(y))$$
$$ch(\mathrm{S}_2(s) \mid \mathrm{TSB}(s)) = 1/4. \tag{4}$$

3. There is no awakening day toss that lands heads. This possibility has an objective probability of 1/4 since it requires that the toss lands tails on both days. Formally:

$$\mathrm{S}_3(s) := \neg\exists x(\mathrm{Toss}(x,s) \,\&\, \mathrm{H}(x))$$
$$ch(\mathrm{S}_3(s) \mid \mathrm{TSB}(s)) = 1/4 \tag{5}$$

Conditional on S_1, since there is only one toss, today's toss must be it. Given S_1, this only toss lands heads, so it has a credence of 1. Conditional on S_2, today's coin lands heads if and only if it is Tuesday. This is clearly possible, so this probability must be greater than 0. Conditional on S_3, the probability that today's toss lands heads is 0. Formally, we have these conditional credences:

$$P(\mathrm{H}(\tau) \mid \mathrm{S}_1(\sigma)) = 1 \tag{6}$$
$$P(\mathrm{H}(\tau) \mid \mathrm{S}_2(\sigma)) > 0 \tag{7}$$
$$P(\mathrm{H}(\tau) \mid \mathrm{S}_3(\sigma)) = 0. \tag{8}$$

Applying G-DIP to each scenario, we get:

$$P(\mathrm{S}_1(\sigma)) = 1/2 \tag{9}$$
$$P(\mathrm{S}_2(\sigma)) = 1/4 \tag{10}$$
$$P(\mathrm{S}_3(\sigma)) = 1/4. \tag{11}$$

Finally, applying the law of total probability to (6)-(11), we get:

$$P(\mathrm{H}(\tau)) = 1/2 + 1/4 \cdot P(\mathrm{H}(\tau) \mid \mathrm{S}_2(\sigma)) > 1/2. \tag{12}$$

Pust takes the conflict between the two derivations to show that there is a failure of admissibility in *both* derivations. This conclusion is clearly unwarranted. Admissibility of evidence is defined with respect to a proposition and its objective probability (Wallmann & Hawthorne, 2020), and the above direct inferences involve different propositions and chances. It is thus possible that there is only a single failure of admissibility. Since the applications of G-DIP yielding (10) and (11) are not needed for the conclusion (see below), either (2) or (9) must involve inadmissible evidence.

Nevertheless, Pust's final conclusion is correct: lacking an argument that there is a failure of admissibility in one case but not the other, the above poses a problem for thirders and halfers alike. However, there *is* such an argument.

3. The inadmissible evidence

The first scenario, $S_1(\sigma)$, occurs if and only if the Monday coin in σ lands heads. Hence, $S_1(\sigma)$ is just the proposition of interest in the original Sleeping Beauty problem, to which thirders assign a probability of 1/3. I argue that Beauty has inadmissible evidence with respect to this proposition and the objective probability evidence (3).

Upon waking up, Beauty knows that the experimenters have awakened her today. As thirders have often argued, this is evidence that affects her credences. The fact that she is awakened today entails that the coin landed tails or it is Monday (Mon). Hence, Beauty knows the logical disjunction $\neg S_1(\sigma) \vee \text{Mon}$. This evidence is relevant for the proposition that the Monday coin landed tails, and therefore it is relevant for heads. Since it is relevant, it is inadmissible. This is clear intuitively, but there is also an extremely plausible argument to that effect.

The admissibility of logical disjunctions involving the outcome has been proven by Wallmann and Hawthorne (2020, p. 963). The following is a simplified version of their theorem 3.

> **Inadmissible disjunctions.** Let P_0 be an initial credence function. Let F be any proposition, and let E be evidence that is consistent with and admissible with respect to the proposition Ta and the objective probability evidence $ch(Tx \mid Rx) = r \,\&\, Ra$, with $0 < r < 1$. Hence, by G-DIP, we have $P_0[Ta \mid ch(Tx \mid Rx) = r \,\&\, Ra \,\&\, E] = r$. Suppose also that $P_0[Ta \vee F \mid ch(Tx \mid Rx) = r \,\&\, Ra \,\&\, E] < 1$. Then we have
>
> $$P_0[Ta \mid ch(Tx \mid Rx) = r \,\&\, Ra \,\&\, E \,\&\, (Ta \vee F)] > r.$$

To apply this theorem to the case of Beauty, we could imagine for a moment that she doesn't know $\neg S_1(\sigma) \vee \text{Mon}$. One way to do this is to suppose she has a lucid dream on both Monday and Tuesday before the point at which the experimenters may awaken her. As I argue in Ackermans (2024), Beauty's evidence while dreaming is admissible with respect to the proposition that the Monday coin lands heads. Moreover, while dreaming, it is clear that the probability of $\neg S_1(\sigma) \vee \text{Mon}$ is lower than 1.[3] Hence, the conditions of *Inadmissible disjunctions* are satisfied. Adding the evidence $\neg S_1(\sigma) \vee \text{Mon}$ would increase the probability in $\neg S_1(\sigma)$, and so it would decrease the probability in $S_1(\sigma)$. This means that $\neg S_1(\sigma) \vee \text{Mon}$ is a *defeater*, and on any plausible

[3] In the dream, Beauty does not know whether it is Monday or Tuesday. She also doesn't know the outcome of the Monday coin. Hence, it is possible that it is Tuesday and that the Monday coin landed heads. It follows that $\neg S_1(\sigma) \vee \text{Mon}$ has a probability less than 1.

account of admissibility, evidence containing a defeater is inadmissible (Wallmann & Hawthorne, 2020, p. 959).

Hence, there is a very strong argument available for the inadmissibility of Beauty's evidence with respect to the direct inference in (9). At the same time, the thirder has good reason to accept the admissibility of Beauty's evidence with respect to the proposition that today's coin lands heads, which she can therefore set to the chance (1). This is simply overwhelmingly plausibly. As Titelbaum puts it:

> Imagine the experimenters put the coin in Beauty's hand and say, "This is the coin we're going to flip in ten minutes. It's fair – it has a 1/2 objective probability of coming up heads. And however the flip comes out, its outcome has no influence on your present condition. Heck, if you like you can be the one to flip it." Standing there with the coin in her hands, Beauty is supposed to be more than fifty percent confident that it'll come up heads? (Titelbaum, 2012, p. 149)

In summary, Pust's derivations do not show that thirders have any reason to doubt the admissibility of Beauty's evidence in the proposition that today's coin lands heads. Titelbaum's embarrassment, therefore, is unique to double halfers.

4. What the embarrassment consists of

The central intuition for double halfers is that Sleeping Beauty *can* use direct inference to set her probability in the Monday coin landing heads to 1/2. Pust's second derivation, however, can be simplified to use only this direct inference, that is, the inference yielding (9). This is just another way of making Titelbaum's argument to which Pust responded.[4]

Consider that $S_2(\sigma)$, the scenario that the current experiment has one awakening day toss landing heads and another awakening day toss landing tails, is possible. Hence, we can conclude $P(S_2(\sigma)) > 0$ without using direct inference.

Consider also that Beauty's credence in $S_3(\sigma)$ is irrelevant for the final step (12), since we have $P(H(\tau) \mid S_3(\sigma)) = 0$.

Hence, just using (6)-(9), $P(S_2(\sigma)) > 0$, and the law of total probability, we get:

$$P(H(\tau)) = 1/2 + P(H(\tau) \mid S_2(\sigma)) \cdot P(S_2(\sigma)) > 1/2. \qquad (13)$$

[4] Titelbaum's argument also starts out from the assumption (9), leading to the conclusion (13), using just the probability axioms.

The double halfer has to accept the above derivation, since the only direct inference used is the direct inference that is central to her own position. At the same time, the double halfer must reject the intuitively plausible direct inference based on (1). The halfer must hold that Beauty has inadmissible evidence with respect to at least one of these two direct inferences. Since the double halfer must accept admissibility in the former case, he must deny it in the latter, despite the overwhelming plausibility. As long as double halfers do not provide any reasons for this denial, this is embarrassing.

References

Ackermans, L. B. (2024). *Sleeping beauty: Why everyone should be a thirder*. PhilSci-Archive. https://philsci-archive.pitt.edu/id/eprint/23227

Bostrom, N. (2007). Sleeping beauty and self-location: A hybrid model. *Synthese*, *157*(1), 59–78. https://doi.org/10.1007/s11229-006-9010-7

Briggs, R. (2010). Putting a value on beauty. In T. S. Gendler & J. Hawthorne (Eds.), *Oxford studies in epistemology* (pp. 3–34, Vol. 3). Oxford University Press.

Dorr, C. (2002). Sleeping Beauty: in defence of Elga. *Analysis*, *62*(4), 292–296. https://doi.org/10.1093/analys/62.4.292

Elga, A. (2000). Self-locating belief and the Sleeping Beauty problem. *Analysis*, *60*(2), 143–147. https://doi.org/10.1093/analys/60.2.143

Horgan, T. (2004). Sleeping Beauty awakened: new odds at the dawn of the new day. *Analysis*, *64*(1), 10–21. https://doi.org/10.1093/analys/64.1.10

Horgan, T. (2007). Synchronic Bayesian updating and the generalized Sleeping Beauty problem. *Analysis*, *67*(1), 50–59. https://doi.org/10.1093/analys/67.1.50

Kim, N. (2022a). Sleeping beauty and the evidential centered principle. *Erkenntnis*, 1–23. https://doi.org/10.1007/s10670-022-00619-6

Kim, N. (2022b). Sleeping beauty and the current chance evidential immodest dominance axiom. *Synthese*, *200*(6), 458. https://doi.org/10.1007/s11229-022-03926-1

Lewis, D. (2001). Sleeping Beauty: reply to Elga. *Analysis*, *61*(3), 171–176. https://doi.org/10.1093/analys/61.3.171

Meacham, C. J. (2008). Sleeping beauty and the dynamics of de se beliefs. *Philosophical Studies*, *138*(2), 245–269. https://doi.org/10.1007/s11098-006-9036-1

Milano, S. (2022). Bayesian beauty. *Erkenntnis*, *87*(2), 657–676. https://doi.org/10.1007/s10670-019-00212-4

Pust, J. (2012). Conditionalization and essentially indexical credence. *The Journal of philosophy*, *109*(4), 295–315. https://doi.org/10.5840/jphil2012109411

Pust, J. (2023). No double-halfer embarrassment: A reply to titelbaum. *Analytic Philosophy*, *64*(3), 346–354. https://doi.org/10.1111/phib.12257

Titelbaum, M. G. (2008). The Relevance of Self-Locating Beliefs. *The Philosophical Review*, *117*(4), 555–606. https://doi.org/10.1215/00318108-2008-016

Titelbaum, M. G. (2012). An embarrassment for double-halfers. *Thought: A Journal of Philosophy*, *1*(2), 146–151. https://doi.org/10.1002/tht3.21

Wallmann, C., & Hawthorne, J. (2020). Admissibility troubles for bayesian direct inference principles. *Erkenntnis*, *85*(4), 957–993. https://doi.org/10.1007/s10670-018-0070-0