A challenge for halfers

CIAN DORR

Let me regale you with yet another variant of the story of Sleeping Beauty. In this one, the experiment takes place in a room with a skylight, so that Beauty can see what the weather is like outside as soon as she wakes up. The weather can be in any one of *n* different states on any given day. Beauty regards each of these states as equiprobable; moreover, she takes there to be no correlation between the weather on Monday and the weather on Tuesday, or between the weather on either day and the coin-toss. The rest of the story works as usual: Beauty will be awoken on Monday, after which a coin will be tossed; if it lands Tails, she will be woken again on Tuesday having had her memories of the Monday awakening erased; otherwise, she will stay asleep until Wednesday.[1] The weather is the only source of variation in her wakings, so if it should happen that the weather is the same on Monday and Tuesday, her total evidence will be exactly the same on both wakings.

Provided that $n > 1$, Beauty learns a new proposition when she wakes up: the proposition W_m that she wakes up at least once during the course of the experiment while the weather is in state *m*. This proposition is clearly relevant to whether the coin lands

[1] As White (200*) notices, in Dorr 2002 I sneakily changed the story to one in which she wakes in any case on Tuesday, whether or not the memory-wiping drug is administered. An explanation is in order: my view is that the answer 1/3 is exactly correct only when the expected value of the total quantity of conscious, minimally rational life in the universe is independent of the coin toss. In ordinary cases the effect of this factor is negligible: but in the limiting case where Beauty knows that she is the only rational being in the universe, and that her conscious life will be twice as long if the coin lands Tails, her credence in Heads when she wakes up should be 1/2. Generating the thirder result even in cases like this would, as Bostrom (2002) points out, require an implausible skewing of prior credences in favour of more populous worlds.

Heads, since if there are two wakings, it is more likely that at least one of them will occur on a day on which the weather is in state *m.* Moreover, the degree to which the proposition supports Tails depends on the value of *n*, since if there are more possible states of the weather, it is less likely that the weather will be the same on both days.

Thus, if we reason about this case as Roger White (200*) would have us reason about his variant of the case (in which each waking has a chance $c < 1$ of succeeding), we should conclude that Beauty should become more confident in Tails when she wakes up, the higher the value of *n*. The only way I can see to obtain such a result is to claim that her new credence in Heads should be derived from her old credence in Heads by conditionalising on W_m, the strongest new proposition she has learned. Where P_- is Beauty's Sunday credence function, Beauty's new credence in Heads will be

$$P_-(\text{Heads}|W_m)$$

$$= P_-(\text{Heads})\, P_-(W_m|\text{Heads}) \,/\, [P_-(\text{Heads})\, P_-(W_m|\text{Heads}) + P_-(\text{Tails})\, P_-(W_m|\text{Tails})]$$

$$= (1 / 2)(1 / n) / [(1 / 2)(1 / n) + (1 / 2)(1 - ((n - 1) / n)^2)]$$

$$= n / (3n - 1)$$

which is equal to 1/2 for $n = 1$ but close to 1/3 when n is large.

It follows from this that Beauty's credence in Heads should be very near 1/3 in any realistic version of the original Sleeping Beauty scenario. For unless Beauty knows in advance exactly what her wakings will be like in every respect—the weather, the pattern of shadows she sees on the ceiling, the noise made by the traffic in the street outside, whether she is hungry or thirsty....—she will learn a new proposition when she wakes up, to the effect that she wakes at least once during the experiment under such-and-such circumstances. And unless she is certain that her wakings will be exactly alike in all these multifarious respects, this new uncentred proposition will support Tails. The advice to believe in Heads to degree 1/2 will only apply in the far-fetched version of the case discussed by Lewis (2001), in which Beauty is somehow made to be certain that her total subjective state will be exactly the same on both wakings.

The idea that the presence of extraneous factors like the weather could matter in this way strikes me as obviously incorrect. If we want to resist it, we will need to explain how it could be that Beauty's credence in Heads should be the same irrespective of the value of n, despite the fact that the propositions which are evidence for her count more strongly in favour of Tails the higher the value of n. The outlines of such an explanation are well-known. Evidence, like belief, is not wholly propositional, but has an irreducibly self-locating aspect. To represent this aspect, we could take evidence as, in the first instance, a relation between subjects and "centred propositions"—properties of time-slices, or relations between believers and times. Since there is more to evidence than a list of propositions, it is no more true that one should always be more confident that P the more one's total propositional evidence counts in favour of P than that one should always be more confident that P the more one's total *visual* evidence counts in favour of P. Just as an increase in one's visual evidence for P may be counterbalanced by a decrease in one's auditory evidence, an increase in propositional evidence for P may be counterbalanced by a decrease in self-locating evidence. This is what happens in my spectrum of cases as n increases. According to thirders, this is also what happens in White's spectrum of cases as c decreases.

Perhaps White will embrace the claim that there is a relevant difference between the case with variable weather and the case where Beauty is certain in advance that she will have exactly the same evidence each time she wakes up, and that credence 1/2 is appropriate only in the latter case. Such a view has recently been endorsed by Meacham (MS), who attributes it to Halpern and Tuttle (1993). To bring out the oddness of this view, consider a case where Beauty knows that when she first wakes up she will be in a state of sensory deprivation—exactly the same state on both occasions, if there are two—but will shortly thereafter be allowed to open her eyes and learn about the weather. Any argument that could support the view that her credence should be near 1/3 in the case where she knows about the weather as soon as she wakes up will also support the view that her credence should be near 1/3 in this case, after she has opened her eyes. But it seems absurdly irrational for Beauty to have credence 1/2 in Heads when she knows perfectly well that as soon as she opens her eyes, she will acquire evidence that will justify lowering that credence. In this case, the thought is almost irresistible that there is no need to bother looking, if the result will be the same no matter what she sees.[2]

References

Bostrom, N. 2002. *Anthropic bias: observation selection effects in science and philosophy*. New York: Routledge.

Dorr, C. 2002. Sleeping Beauty: in defence of Elga. *Analysis* 62: 292–95.

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

Halpern, J. and Tuttle, M. 1993. Knowledge, probability, and adversaries. *Journal of the Association for Computing Machinery* 40: 917–62.

Lewis, D. 2001. Sleeping Beauty: reply to Elga. *Analysis* 61: 171–76.

---

[2] Another important problem for this kind of view is pointed out by Bostrom (2002): it makes it impossible for evidence ever to be relevant to the choice between two hypotheses both of which entail that the world is so big that every humanly possible evidential state is realised somewhere.

Meacham, C. J. G. MS. Sleeping Beauty and the dynamics of *de se* beliefs. Online at philsci-archive.pitt.edu/archive/00002323.

White, R. 200*. The generalised Sleeping Beauty problem: a challenge for thirders. Forthcoming in *Analysis*.