The Sleeping Beauty Problem

Abstract

This text examines the Sleeping Beauty problem, a decision theory enigma. It shows the positions of "halfers" and "thirders" in assigning probabilities. The halfers assert symmetrical initial probabilities based on the fact that the tossed coin is fair. The main arguments here are from David Lewis. In contrast, the thirder's viewpoint, championed by Adam Elga, argues for a probability of $\frac{1}{3}$ for heads due to experiment symmetry. The solution attempt tries to clarify the paradox by deflating the problem and splitting it into two central arguments to unify the "halfers" and "thirders" positions into two correct answers for different aspects of the question and the surrounding environment. The text underscores the ongoing debate's complexity.

1 History

The Problem was originally stated from Arnold Zuboff in the 1980s and later published in the work "One Self: The Logic of Experiment" (Zuboff, 1990). As a part of a collection of different thought experiments. In the year 2000, Adam Elga responded by publishing an analysis titled "Self-locating Belief and the Sleeping Beauty Problem," (Elga, 2000) in which he offered insights and rephrased the original Sleeping Beauty Problem. In 2001, David Lewis contributed to the discourse with the analysis "Sleeping Beauty: reply to Elga," (Lewis, 2001). He stated an alternative perspective on the solution to the posed question within the experimental context.

2 The Problem

A Participant (Sleeping Beauty) volunteers to undergo the following experiment and is told all of the following details: Some researchers are going to put you to sleep. During the two days that your sleep will last, they will briefly wake you up either once or twice, depending on the toss of a fair coin (Heads: once; Tails: twice). After each waking, they will put you to back to sleep with a drug that makes you forget that waking. When you are first awakened, to what degree ought you believe that the outcome of the coin toss is Heads?

	Monday	Tuesday
Heads	woken up	Asleep
Tails	woken up	woken up

Table 1: Overview over all possible events that can occur.

3 The thirder's perspective

3.1 Adam Elga's Interpretation

At a moment of Sleeping Beauty's awakening, there are three possible relevant states:

$$T1 \ : \text{Heads (H) and Monday (M)}$$
$$H1 \ : \text{Tails (T) and Tuesday (U)}$$
$$H2 \ : \text{Tails (T) and Monday (M)}$$

To solve the problem, the focus is on determining the credence function P for each state.

By considering the cases where Beauty learns certain information upon awakening. It's shown that $P(T1) = P(T2)$ because they are indistinguishable predicaments. The choice of awakening method doesn't affect Beauty's initial credence in Heads.

When Beauty learns it's Monday, her credence in Heads is $\frac{1}{2}$, unaffected by the possibility of awakening on Tuesday. This leads to the conclusion that $P(H1) = P(T1)$, and combining all results, it's found that $P(H1) = P(T1) = P(T2) = \frac{1}{3}$.

This conclusion is reached by analyzing Beauty's changing knowledge as she wakes up and learns the outcomes, showing that each predicament's probability is $\frac{1}{3}$. (Elga, 2000)

4 The Halfers' Perspective

According to this viewpoint, Sleeping Beauty's answer on the question should be $\frac{1}{2}$.

4.1 Central Argument

The central argument of the halfers' perspective is rooted in the principle of indifference. They argue that at the beginning of the experiment, Sleeping Beauty has no reason to favor one outcome over the other. Thus, her initial credence on a fair coin landing heads should be $\frac{1}{2}$.

4.2 Symmetry Revisited

Elga's symmetry argument suggests that the probabilities of certain events should be equal due to the symmetry of the experiment. However, the halfers' perspective contends that this symmetry does not necessarily require equal probabilities. Instead, the principle of indifference prevails, resulting in balanced initial credences.

4.3 Conclusion

The halfers' perspective questions the complexity of the Sleeping Beauty experiment by asserting that symmetry and evidence should not alter Sleeping Beauty's initial assumptions about the coin toss probabilities. This interpretation underlines the importance of maintaining indifference because there is no new information given after the first introduction of the experiment.

5 Solution Attempts

The end of Sleeping Beauty's Nightmare
In the paper by Berry Groisman: "The End of Sleeping Beauty's Nightmare," two central arguments are given to solve the problem. These are based on breaking down the problem into its parts.

The first part is called "from contradiction to consistency":
The argument is based on the fact that half and third are in an apparent contradiction that can be overcome. Berry Croismen cites that there is no actual contradiction between Sleeping Beauty's belief in the coin's fairness and its $\frac{1}{3}$ wake-up credibility. He says that both answers, $\frac{1}{3}$ and $\frac{1}{2}$, are correct, but answers to two different interpretations of the question. The phrase "the coin landed heads" alone does not define the event completely. Thus, the question posed to Sleeping Beauty is: 'What is your credence the coin landed Heads under the setup of wakening?' It can be rephrased as 'What is your credence that this awakening is a Head-awakening under the setup of wakening?' And the correct answer should be $\frac{1}{3}$. However, "the coin landed Heads under the setup of coin tossing" is another event. If we asked Sleeping Beauty, 'What is your credence that the coin landed Heads under the setup of coin tossing?', the correct answer would be $\frac{1}{2}$, as Sleeping Beauty still believes the coin is fair.

The second one is called "The inanimate version":
The text presents an analytical approach to clarify the paradox and eliminate subjective factors. It introduces an automated scenario using a coin-tossing device with colored balls to determine the probability of drawing a green ball from a mixed box. Based on relative frequency, theoretical calculations lead to a probability approximation of $\frac{1}{3}$ for this event. The discussion delves into the misconception of equivalence, highlighting the distinction between placing and removing a green ball. The application of Bayes' rule is explored, revealing the fallacy in associating these events. The emphasis is on the divergence between the circumstances and the explanation for non-equivalent probabilities. Despite initial perplexity, the text concludes that there is no contradiction, thus shedding light on the paradox.

6 Conclusion

In conclusion, this analysis explores the Sleeping Beauty problem by showing the contrasting positions held by "halfers" and "thirders" regarding probability assignments. David Lewis and Adam Elga's perspectives diverge significantly on whether Sleeping Beauty should alter her belief on the probability of the fair coin toss due to her altered physical state during the questioning, despite the absence of any new informational input.

The attempt to untangle this enigma introduces two significant solution strategies that offer distinct vantage points to address the paradox. In "The End of Sleeping Beauty's Nightmare," Berry Groisman presents two key arguments: "from contradiction to consistency" contends that both 1/3 and 1/2 answers are valid, reflecting distinct interpretations of the question's context; "The inanimate version" introduces an automated scenario and theoretical calculations to resolve the paradox, revealing non-equivalent probabilities by distinguishing events and shedding light on the underlying complexities. This approach emphasizes the nuanced nature of the question's phrasing and its implications on the answer. It's important to note that the discussion surrounding the Sleeping Beauty problem extends beyond the confines of this analysis. Many scholars, philosophers, and researchers have contributed their viewpoints, enriching the discourse with diverse insights. As new perspectives emerge and novel experiments are crafted, new ideas and methods draw a more narrow circle around a possible correct answer or explanation of this problem.

References

Elga, A. (2000). "self-locating belief and the sleeping beauty problem". *Analysis. 60 (2): 143–147. CiteSeerX 10.1.1.32.3107. doi:10.1093/analys/60.2.143. JSTOR 3329167.*.

Lewis, D. (2001). "sleeping beauty: reply to elga". *Analysis 51: https://doi.org/10.1093/analys/61.3.171*.

Zuboff, A. (1990). "one self: The logic of experience". *Inquiry: An Interdisciplinary Journal of Philosophy. 33 (1): 39–68. doi:10.1080/00201749008602210.*.