Chang, M. et al. (Eds.) (2019). Proceedings of the 27th International Conference on Computers in Education. Taiwan: Asia-Pacific Society for Computers in Education

Another Perspective of the Sleeping Beauty Problem: What Lessons Can We Learn from the Sleeping Beauty Problem?

Shih-Hsun CHEN
Faculty of Philosophy and Arts at the University of West Bohemia, Czech Republic
* bb9124037@hotmail.com

Abstract: The Sleeping Beauty problem is still an important issue in decision theory. Since Elga raised this problem in 2000, there are two opinions as to what the solution is, with some thinking the answer is 1/2 and others thinking the answer is 1/3. Even though both sides seem to have reasonable reasons to support their answer, there is still no convincing conclusion. In addition to attempting to solve this debate, this paper provides another perspective on the Sleeping Beauty problem: no matter whether the final answer is 1/2 or 1/3, it will not challenge our intuition to a fair coin.

Keywords: decision theory, the Sleeping Beauty problem, paradox

1. Introduction

The Sleeping Beauty problem is as follows: researchers tell Sleeping Beauty about the following experimental procedure: Sleeping Beauty will be put to sleep on Sunday. After she falls asleep, the researchers will toss a fair coin. If the coin toss lands heads up, Sleeping Beauty will be woken up on Monday, and then will answer the question "What is your credence now for the proposition that the coin landed heads up?" After she answers, Sleeping Beauty is put back to sleep with an amnesia-inducing drug which means she has no memory of being woken. If the coin toss is tails, Sleeping Beauty is woken up on Monday and Tuesday and is asked, "What is your credence now for the proposition that the coin landed heads?

EIga (2000) first proposed the Sleeping Beauty problem which has generated great debate amongst the decision theory community. The candidate answers can be divided into two options: 1/2 and 1/3. Here we call supporters of 1/2 Halfers ((Lewis (2001), White (2006), Hawley (2013)), and supporters of 1/3 Thirders ((EIga (2000), Arntzenius (2003), Horgan (2004, 2007), Weintraub (2004), Bovens (2010)). Even though the proponents of the same answer may base their thinking on different reasons and even cause some inconsistent situations, this does not pose any problem in this paper.

This paper is divided into four sections. The first section is the introduction. The second section roughly explains the difference between Halfers and Thirders and uses Bertrand's box paradox as a comparison. The third section argues that when a Thirder claims that the answer to the Sleeping Beauty problem is 1/3, this may be due to their misunderstanding of the results of the experiment. The fourth section attempts to analyse the Sleeping Beauty problem from a more comprehensive perspective, indicating that no matter what the final answer is to the Sleeping Beauty problem, we must be careful whether such an answer challenges our intuition about a fair coin. The fifth section is the conclusion.

2. Halfers and Thirders

In this section, I briefly explain the differences between Halfers and Thirders. The reason why Halfers believe the answer is 1/2 is that Sleeping Beauty knows the details of the test on Sunday and when she wakes up, she does not gain any extra information. Therefore, when Sleeping Beauty is woken up during the experiment, this does not change the probability of the coin toss.

The reason why Thirders believe the answer is 1/3 is because there are three situations in which Sleeping Beauty could have been woken up: C1(Heads, Wake on Monday), C2(Tails, Wake on Monday) and C3(Tails, Wake on Tuesday). The probability of these three events is the same, and the three events constitute the whole situation. Hence, the probability of these three events occurring is 1/3. In this case, the head event only occupies one of three situations. Therefore, under the premise that Sleeping Beauty has been woken up, the probability of the coin toss landing heads up is 1/3.

The Halfers argue that Sleeping Beauty has received no new information upon waking, so there is no difference in the information Sleeping Beauty has before the experiment and the information she has during the experiment. As the fourth section in this paper raises similar arguments, I discuss this later.

I compare the Thirders' argument to Bertrand's box paradox. Joseph Bertrand first posed Bertrand's box paradox in his 1889 book "Calcul des probabilités". Bertrand's box paradox is as follows: there are three boxes. The first box contains a gold coin and a silver coin. The second box contains two silver coins. The third box contains two gold coins. After randomly selecting a box and taking out a coin, what is the probability that the coin will belong to the first box given that this coin is gold?

Bertrand's box paradox can be regarded as a classic exercise in teaching probability. I use a structural analogy between Bertrand's box paradox and the Sleeping Beauty problem to highlight the difference. Given a gold coin has being drawn (Sleeping Beauty has been woken up), the chance that the gold coin belongs to the first box is 1/3 (the probability that the coin toss landed heads up is 1/3). Because there are only three possibilities for getting a gold coin (there are three possibilities for Sleeping Beauty to be woken up), namely G1 in the first box (C1), G2 in the third box (C2), and G3 in the third box (C3), G1, G2, and G3 are three possible events given one gold coin has been taken out (C1, C2, and C3 are three possible events given Sleeping Beauty has been woken up). Hence, the probability of the gold coin belonging to the first box is 1/3 (so the probability of the coin toss landing heads up is 1/3). Figure 1 shows the probability branch diagram of Bertrand's box paradox.

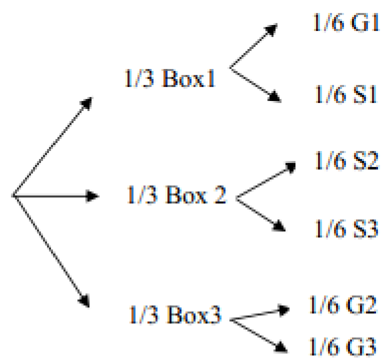

Figure 1.

However, we should be careful, even though the above analogy looks reasonable. The difference lies in the fact that G1, G2, and G3 are mutually exclusive, that is, $P(G1 \cap G2) = 0 = P(G1 \cap G3) = P(G2 \cap G3)$. After all, we know that only one coin will be taken out during the entire experiment, so the probability of drawing out two (or more) coins is zero.

However, C1 and C2 are not mutually exclusive. When one of C1 or C2 occurs, the other must happen. In other words, regardless of the individual probability of these three events, $Pr\ (C2 \cap C3)$ is not equal to zero. This is a reasonable. When the coin toss lands tails up, Sleeping Beauty will be woken up on Monday and Tuesday. Based on this, even if $Pr(C1) = Pr(C2) = Pr(C3)$, we can question whether $Pr(C1) + Pr(C2) + Pr(C3) = 1$.

From White's paper (2006), we can see some of the differences between Halfers and Thirders. White proposed a generalized Sleeping Beauty problem. In the generalized Sleeping Beauty problem, White introduced c, $c < 0 \leq 1$. In this version, c refers to the probability that Sleeping Beauty will be actually woken up during the experiment. When $c = 1$, it is the original Sleeping Beauty problem. In the generalized Sleeping Beauty problem, in addition to C1 (Heads, Wake on Monday), C2 (Tails, Wake on Monday) and C3 (Tails, Wake on Tuesday), we also have C1*(Heads, Not Wake on Monday) , C2*(Tails, Not Wake on Monday) and C3*(Tails, Not Wake on Tuesday). Through the calculation of conditional probability (details can be seen in White 2006), we can know that the probability of the coin landing heads up is 1/ (3-c). Here, we can see the difference between the probability branch diagrams of both sides from Figure 2 and Figure 3. Figure 2 represents the perspective of Thirders and Figure 3 represents the perspective of Halfers.

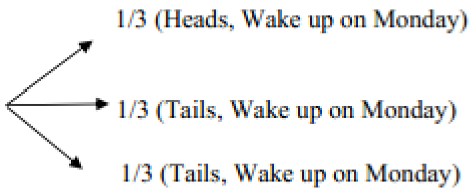

Figure 2

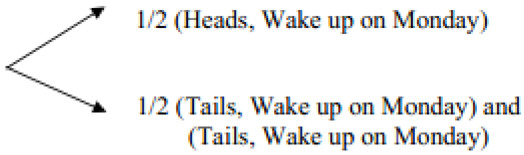

Figure 3

In this regard, if the Thirders attempt to support their answer using a similar proof, they may need to give other reasons to justify it.

3. Experimental point of view

For the Sleeping Beauty problem, Thirders (who believe the answer is 1/3) sometimes attempt to justify their point of view by changing some of the conditions of the experiment. For example, if the coin toss lands tails up, Sleeping Beauty will be woken up in all the following thousand days. This is an attempt to capture the intuition that "1/2 is unreasonable" by amplifying the number of days that Sleeping Beauty is awakened. However, these authors do not give a thorough explanation for this kind of experiment. I argue that it is because of the misunderstanding of the experimental details such that Thirders misused the experimental results. This is why the Monty Hall problem is a classic example of probability teaching, but the Sleeping Beauty problem is still unsolved in the field of decision theory.

Thirders explain that there are only three situations in which Sleeping Beauty will be woken up: (Heads, Monday) (Tails, Monday) (Tails, Tuesday). Of the three possibilities, because each probability of occurrence is equal, the likelihood of the coin toss being heads up is 1/3. Regardless of whether the Thirders hold some philosophical position to defend their belief that the answer is 1/3, I argue from an experimental aspect that 1/3 is not correct.

Suppose there are a hundred Sleeping Beauties to participate in this experiment, and when any Sleeping Beauty is woken up, we ask her "what do you think is the probability that the coin toss landed heads?" In typical statistical calculations, about 50 people say the coin toss is heads, and the remaining 50 people say tails. In this way, we can get 150 pieces of data roughly: 50 pieces (Heads, Monday), 50 pieces (Tails, Monday), and 50 pieces (Tails, Tuesday). Thirders claim that the heads situation is one-third of the total. Therefore, under the premise that Sleeping Beauty has been woken up, the probability of the coin toss being heads up is 1/3.

All this sounds reasonable, but it is an incorrect claim. It can be claimed that if someone randomly picks one of the 150 samples, the chance of it being heads is indeed 1/3. However, if we ask the awakened subjects what is the probability of the coin toss being heads, the answer is 1/2. We have to be careful of the difference between the two claims. There are 150 samples in this experiment, but this experiment only has 100 subjects. In other words, not every subject will provide the same number of samples: each Sleeping Beauty who thinks the coin toss is heads only offers one piece of data, but each Sleeping Beauty who thinks the coin toss is tails provides two pieces of data. Thirders misread the experiment results.

I contend that the Thirders’ argument can be refuted from an experimental point of view, but as there are many other arguments in favour of the Thirders’ viewpoint, this thesis cannot provide a more detailed explanation. Even so, we can draw an analogy between the Monty Hall problem and the Sleeping Beauty problem. If Thirders' arguments about the experiment are correct, then similar to the Monty Hall problem, when the statistics tell us how to choose in this game, it is rational to follow this advice. However, as previously mentioned, it is because of the details of the experiment that Thirders omit, so that the actual meaning of the experiment, that is, the statistical results, cannot play a real effect.

Of course, there may be some other experimental designs. But if Thirders want to convince everyone that 1/3 is the correct answer in the practical sense, they need to show that the experiment I discussed in this section has reasoning errors, or else suggest another experiment that really captures the meaning of the Sleeping Beauty problem.

4. Another Perspective for The Sleeping Beauty Problem

This section looks at the Sleeping Beauty problem from another perspective. First, we discuss the classic problem of probability: the Monty Hall problem.

The Monty Hall problem is as follows: a participant is given the choice of selecting one of three doors. Behind one door is a car; behind the others, goats. When the participant chooses a door, the host who knows what is behind every door opens one door with a goat behind it. Then, the host asks the participant if they want to change to the other closed door. Or in other words, will changing to another closed door increase their chances of winning the car?

The proof of the Monty Hall problem is very similar to that of Bertrand's box paradox. First, there are three possible things that can be selected, C (car), G1 (Goat 1), G2 (Goat). At the beginning of the game, Pr(C) = Pr(G1) = Pr(G2) = 1/3. After the host opens one of the doors with the goat behind it, if the participant chooses to change their choice of door, there are three possibilities. The first possibility is that the participant had firstly selected C at the beginning, but after switching to a different door, the participant gets G1 or G2. The second possibility is that the participant had firstly selected G1 but got C after the switch. The third possibility is that the participant had first selected G2 and got C after the switch. Based on the assumption that these three possibilities are equal, the chance of choosing the car is 2/3. Figure 4 shows a probability tree diagram of the Monty Hall problem.

What can we learn from the Monty Hall problem? At the end of Elga's paper (Elga, 2000), Elga states: "at least one new question arises about how a rational agent ought to update her beliefs over time." It is essential that when information changes, the original probability configuration may also change accordingly. We do not doubt this, so at the beginning of the game, the chance of selecting a car is 1/3. We also do not doubt that knowing that there are two closed doors and only one car is behind them, the probability of winning a car is 1/2. However, we should think about it: after the first choice we make, and then the host opens another door in a non-random manner, does the behavior of the host affect the original probability distribution? If the host's behavior and the participant's behavior are independent, the behavior of the host will not affect the chances of this participant getting a car. However, this is not an independent case. Therefore, from the example of the Monty Hall problem, we know that when new information is generated, we should check whether the new information affects the original probability distribution of beliefs. Moreover, we can test whether the new information is independent of the probability distribution of old beliefs.

Taking the example of the Monty Hall problem to compare with the Sleeping Beauty problem, we can think in the following way. The coin toss not only determines whether the Sleeping Beauty will be woken up on Tuesday, it also determines the number of days that Sleeping Beauty will not be woken up. When we asked: "When Sleeping Beauty wakes up for the first time, what is the probability of the coin toss being tails?" or we can also ask: "When Sleeping Beauty falls asleep on the third day, what is the probability of the coin toss being tails?" Of course, when Sleeping Beauty falls asleep, she cannot answer this question, but we can solve this problem in another way, such as asking Sleeping Beauty this hypothetical question before the experiment. If we agree that "when Sleeping Beauty falls asleep on the third day, what is the probability of the coin toss being tails?" is a legitimate question, then we can easily conceive similar questions, such as changing the third day to the n-th day.

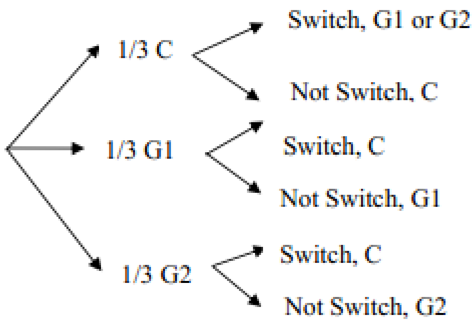

Figure 4.

In addition, we can see that the ratio of "being sleeping" and "been woken up" in the probability space is very huge. Therefore, the Sleeping Beauty problem is a problem of conditional probability and the proportion of “been woken up” in the entire probability space is extremely small. Therefore, no matter whether the final answer to the Sleeping Beauty problem is 1/2 or 1/3, this should not challenge our intuition about a fair coin. When we are thinking about the problem of Sleeping Beauty, it is easy to overestimate the significance of Halfers’ and Thirders’ arguments which will challenge our daily intuition about a fair coin.

From the above analysis, in the situation where Sleeping Beauty is awakened relative to the situation of falling asleep, the proportion in the probability space is relatively very small, so the Sleeping Beauty problem, regardless of the answer, is essentially irrelevant to the probability of the coin toss being heads.

5. Conclusion

This paper does not solve the problem of Sleeping Beauty and does not offer proof or counterevidence in relation to Halfers' and Thirders' arguments. In section 3, I challenge Thirders using an experimental point of view and argue that Thirders misunderstand the results of the experiment. The challenge is mainly to indicate that if the results in the actual sense of the experiment are clear, then the problem may not cause such a big controversy. Taking the Monty Hall problem as an example, even though some people may not understand the mathematical demonstration immediately, in a practical sense, we should give the participants the chance to "switch the door".

Section 4 explores the lessons we can learn from the Sleeping Beauty problem from a broader perspective. I argue that no matter whether the final answer is 1/2 or 1/3, it cannot challenge our intuition about a fair coin. The Sleeping Beauty problem is a problem of conditional probability. As for whether the Sleeping Beauty problem can challenge our other intuitions, and whether the answer is 1/2 or 1/3, or even other answers, this requires more research.

References

Arntzenius, F. (2003). Some problems for conditionalization and reflection. *Journal of Philosophy* 100: 356–70.
Bovens, l. (2010) Judy Benjamin is a Sleeping Beauty. *Analysis*, 70(l):23-26.
Elga, A. (2000). Self-locating belief and the Sleeping Beauty problem. *Analysis* 60: 143–147.
Hawley, P. (2013). Inertia, Optimism and Beauty. *Nous* 47(1), 85-103.
Horgan, T. (2004) Sleeping Beauty awakened: new odds at the dawn of the new day. *Analysis* 64: 10–21.
Horgan, T. (2007) Synchronic Bayesian updating and the generalized sleeping beauty problem. *Analysis*, 67(293):50- 59.
Lewis, D. (2001). Sleeping Beauty: reply to Elga. *Analysis* 61: 171–76.
Weintraub, R. (2004). Sleeping Beauty: a simple solution. *Analysis* 64: 8–10.
White, R. (2006) The generalized sleeping beauty problem: a challenge for thirders. *Analysis*, 66(290): 114- 119.