When Sleeping Beauty First Awakes

Atkinson, David; Peijnenburg, Jeanne

Document Version
Final author's version (accepted by publisher, after peer review)

When Sleeping Beauty First Awakes

David Atkinson and Jeanne Peijnenburg

Faculty of Philosophy, University of Groningen, The Netherlands

Abstract

Ever since its introduction, the Sleeping Beauty Problem has been fought over by the halfers against the thirders. We distinguish three interpretations of the original problem as described in Adam Elga's seminal paper on the subject. Elga's intended interpretation leads to the position of the thirders; but the other readings result in that of the halfers. We show that all three of these results can be obtained by making use of objective probabilities in a four-dimensional rather than a three-dimensional space. Our reasoning avoids various problems, not only of Dutch Book and other subjectivist approaches, but also of earlier treatments in terms of objective probabilities.

Keywords
Sleeping Beauty, probability, credence, hypothetical relative frequency.

1   Introduction

The story is well known: on Sunday evening, Sleeping Beauty, who is mathematically well schooled, is informed that a fair coin has just been tossed. She is further told that she will shortly be put to sleep for two days. If the coin landed heads, she will be briefly awakened on Monday, and put to sleep again; if it landed tails, she will be briefly woken up on Monday and on Tuesday. Moreover, all her memories of the Monday awakening will be erased from her brain. In the event that she is awakened on Tuesday, she will be unaware that this is her second awakening: for all she knows, it might be Monday. She will however remember the rules of the game. After Beauty has received this information on Sunday evening, and before she is put to sleep, she is asked

> "When you are first awakened, to what degree ought you to believe that the outcome of the coin toss is heads?" (Elga 2000, 143).

The question has spurred an ongoing debate between thirders and halfers, with no agreement in sight.[1]

In this paper we argue that the thirder position constitutes the answer to the question as Elga intended it. There are however at least two different interpretations of Elga's question, and under both the answer of the halfers is correct. Moreover, all three of these answers can be obtained by reasoning based on objective probabilities or chances.

Elga describes the Beauty scenario in terms of centred propositions, defined as sets of centred worlds. He arrives at his thirder position by using the subjectivist approach, based on credences rather than on objective probabilities or chances. But while it may be true that many centred propositions cannot be assigned an objective probability, the Sleeping Beauty scenario is rather different. For it can be repeated, either in fact or in thought, thus opening the door to an objective treatment in terms of hypothetical relative frequencies. The key to this treatment is simpy an application of David Lewis's Principal Principle (Lewis 1980), according to which Sleeping Beauty's rational credences match the calculated chances. The objective approach is also applicable to two different interpretations of Elga's question, both of which yield one half as an answer.

We will start our discussion of the Sleeping Beauty problem in subjectivist terms, notably in terms of Dutch Book considerations. This culminates in a particularly interesting contribution to the debate, namely that by Peter Lewis (Lewis 2010). As Lewis sees it, the problem springs from the unquestioned assumption that Beauty's degree of belief that the coin came up heads is the same as her degree of belief that the coin came up heads and that it is Monday. Both Adam Elga and David Lewis, founding fathers of thirders and halfers respectively, endorsed this assumption (Elga 2000; Lewis 2001), and practically all participants in the debate have followed suit. Peter Lewis

---

[1]Thirders are for example Arntzenius 2003, Bovens 2010, Bovens and Ferreira 2010, Bradley 2003, Dieks 2007, Dorr 2002, Draper and Pust 2008, Elga 2000, Groisman et al. 2013, Hitchcock 2004, The Oscar Seminar 2008, Horgan and Mahtani 2013, Monton 2002, Titelbaum 2008, Titelbaum 2012, Wilson 2013. Among the halfers are Bradley and Leitgeb 2006, Cozic 2011, Hawley 2013, D. Lewis 2001, P. Lewis 2007, P. Lewis 2010, Pust 2012, Pust 2013, Ross 2010, White 2006.

however claims that the assumption is false: Beauty's credence in the uncentred 'heads came up' differs from her credence in the centred 'heads came up and it is Monday'.

We discuss the Dutch book arguments, as well as Peter Lewis's claim that Beauty's credence in 'heads' differs from her credence in 'heads and it is Monday', in Sect. 2. In Sect. 3 we describe a problem that Lewis has identified in his own approach: the conclusion he wants to defend seems to spark a counterintuitive reaction. Lewis himself tries to solve this problem by taking further recourse to Dutch books. However, as we point out, Dutch book reasonings sometimes fail to give the final word, due to a lack of clarity about which fair betting odds determine which credences. Thus we tend to agree with Peter Lewis when he concludes by saying that perhaps "no Dutch book can be constructed" for Beauty's credence in 'heads' which indubitably shows that it differs from her credence in 'heads and it is Monday' (Lewis 2010, 377).

We then propose a more satisfactory way to handle the problem than by using Dutch books, namely by using objective probabilities, based on the method of hypothetical relative frequencies. This objectivist argument is first described informally in Sect. 4 and then spelled out formally in Sect. 5. At the heart of the argument is the idea that the Beauty scenario involves a probability space of four rather than three dimensions. Consequently, as we will see, there are at least three ways in which Elga's question can be interpreted: we informally explain the three readings in Sect. 4, and give a more formal treatment in Sect. 5.

Our reasoning concerning the first of these interpretations resembles that of Terry Horgan (2004, 2007, 2008), but in Sect. 6 we explain how it differs: where Horgan relies on what he calls 'preliminary probabilities', we use relative frequencies. As a result, Joel Pust's objections to Horgan's argument are not relevant to our approach (Pust 2008, 2011, 2012). Sect. 7 is devoted to OSCAR's objectivist argument in favour of thirders. We identify a questionable assumption that is intrinsic to the OSCAR approach, and explain how we can circumvent it.

In the debate about Sleeping Beauty it is customary to make use of the following abbreviations:

H : The coin landed heads.

T : The coin landed tails.

MON: It is Monday.

TUES: It is Tuesday.

H_1 : The coin landed heads and it is Monday.

T_1 : The coin landed tails and it is Monday.

T_2 : The coin landed tails and it is Tuesday.

Here H and T are uncentred propositions, and MON, TUES, H_1, T_1, and T_2 are all centred ones. In the sequel we will adopt these abbreviations. In Sect. 5, where we explain our argument in formal detail and argue that we need a probability space of four rather than three dimensions, we will introduce two more propositions and we will make further distinctions in the class of centred propositions.

2 Dutch Books

Soon after Sleeping Beauty's plight became known to the philosophical world, attempts were made to palliate it by making use of Dutch book arguments. The basic idea here is that, since Beauty is a perfectly rational agent, she will not be liable to a (synchronic or diachronic) Dutch book. Her credences, reflected in the odds at which she considers a bet on H to be fair, are such that no bookie will ever be able to turn her into a money pump. Hence any answer to the original question that *does* transform Beauty into a money pump does not reveal her credences. Conversely, any answer that demonstrably protects her from such a monetary mishap is *prima facie* a reply to Elga's question "When you are first awakened, to what degree ought you to believe that the outcome of the coin toss is heads?". The problem however is that people have different ideas about what constitutes a fair bet for Beauty, and thus disagree about what her credences are. We shall illustrate this difficulty by sketching very briefly a number of Dutch book arguments.

Everyone in the Sleeping Beauty debate agrees that on Sunday, before she is put to sleep, Beauty's credence in H is one half, $P_-(H) = \frac{1}{2}$.[2] The

---

[2]Following David Lewis, we use the symbol P_- for Beauty's credence function on Sunday, P for her credence on Monday after her first awakening, and P_+ for her credence, as it is modified after she is told that it is Monday.

question is whether her credence on Monday after her first awakening, $P(H)$, is still one half. The following Dutch book argument suggests that it is.

Imagine that Beauty is offered a bet on H on Sunday, before she is put to sleep. Given that $P_-(H) = \frac{1}{2}$, Beauty will consider odds of 1:1 to be fair. Now imagine that she is offered a second bet on H on Monday, after her awakening. Then the only way for her to avoid a Dutch book is by again adopting odds of 1:1. That is, $P(H)$ must still be one half, for otherwise she could be turned into a money pump (Lewis 2010, 374).

Christopher Hitchcock (2004) disagrees. He notes that the above argument is not a genuine Dutch book. The bookie can only make the second book if he knows that it is Monday, and not Tuesday. Sleeping Beauty however cannot distinguish between those days, so there is an epistemic asymmetry which in Dutch book arguments is not allowed. In order to make it a legitimate Dutch book argument, the bookie should be subjected to the same routine of sleeping, waking, and amnesia as Sleeping Beauty. But if that is done, Beauty will consider the second bet to be fair only at odds of 2:1 on heads, and not 1:1. For the bookie will now offer Beauty a bet each time that he is woken up. If the coin lands tails, he is woken up twice (on Monday and on Tuesday) and so he twice offers the bet; if the coin lands heads, he offers the bet only once (on Monday). Since he does not know whether it is Monday or Tuesday, the bets he offers on those days will be the same. If Beauty on awakening were to accept these bets at odds of 1:1, she would be Dutch booked. Hitchcock concludes that the only way for Beauty to avoid a Dutch book is by adopting odds of 2:1, and thus to have on awakening a credence in H of one third: $P(H) = \frac{1}{3}$. In this conclusion he is followed by the majority of the philosophers who have written about Sleeping Beauty's predicament, including those who do not use Dutch book arguments (see footnote 1 for a few leading examples).

Darren Bradley and Hannes Leitgeb have critized Hitchcock's argument on the grounds that Beauty need not be a thirder in order to avoid a Dutch book (Bradley and Leitgeb 2006). She can be a halfer, for being a halfer does not mean that, when woken up, she will accept a bet on heads at odds of 1:1. In fact, she will not. She realizes that in the second bet there are twice as many tails awakenings as head awakenings, so if a tail has come up, she actually bets twice, whereas she bets only once if the coin landed heads. If the entire game were repeated 1000 times, and she were to lay down (each time) a stake of $1 on the thesis that she is in a heads awakening, she would, on the average, win 500 times, being paid $3 at odds of 2:1, so she would win a net

$2 each time, a total winnings of $1000. On the average she would however lose the bet 500 times on Monday (because a tail came up on Sunday), at a total cost of $500, but also 500 times on Tuesday, at a further cost of $500. On the average, then, she would win $1000 and lose $1000. This implies that she considers betting odds 2:1 as fair, and 1:1 as unfair. However, and this is Bradley and Leitgeb's point, it does not imply that her credence in heads is one third. In the happy locution of Bradley and Leitgeb, odds and credences 'come apart' in scenarios like that of Sleeping Beauty: the number of bets on heads here differs from the number of bets on tails. Precisely because Hitchcock's bookie does not offer the same number of books in the event that the coin falls heads or tails, the fair odds for the second book do not reflect Sleeping Beauty's credence in H.

Kai Draper and Joel Pust find Bradley and Leitgeb's criticism of Hitchcock unconvincing (Draper and Pust 2008). Their argument hinges on the value of $P(H|MON)$, i.e. the credence that the coin landed heads conditioned on Beauty's being told that today is Monday. The standard way of calculating $P(H|MON)$ is on the basis of the assumption $P(H_1) = P(T_1) = \frac{1}{3}$, which results in a value of one half.[3] However, if $P(H) = \frac{1}{2}$, as Bradley and Leitgeb want it to be, then $P(H|MON) = \frac{2}{3}$, in accordance with the well-known reasoning as given by David Lewis (2001). David Lewis himself was not too happy with the conclusion, and in general it is considered to be the Achilles heel of halfers.[4]

Note, however, that Draper and Pust's criticism of Bradley and Leitgeb rests on the assumption that $P(H) = P(H_1)$. In endorsing that assumption they are by no means in the minority: almost everybody in the Sleeping Beauty debate underwrites it. For thirders, who maintain that $P(H_1)$ is a third, the assumption leads to the claim that $P(H)$ is one third. For halfers, who hold that $P(H)$ is a half, it is a compelling reason to maintain that $P(H_1)$ is one half. But if one allows that $P(H) \neq P(H_1)$, then it becomes possible to have one's cake and eat it too: $P(H) = \frac{1}{2}$ and $P(H|MON) = \frac{1}{2}$.

Peter Lewis is one of the few philosophers who contend that $P(H)$ and

---

[3] $P(H|\text{MON}) = \dfrac{P(H \wedge \text{MON})}{P(\text{MON})} = \dfrac{P(H_1)}{P(H_1 \vee T_1)} = \dfrac{P(H_1)}{P(H_1) + P(T_1)} = \dfrac{\frac{1}{3}}{\frac{1}{3} + \frac{1}{3}} = \frac{1}{2}$.

[4] Though not of double halfers such as Peter Lewis, who maintain that $P(H) = P_+(H) = \frac{1}{2}$. Moreover Peter Lewis, unlike other double halfers, also holds that $P_+(H)$ is the same as $P(H|MON)$.

$P(H_1)$ are not equal (Lewis 2010).[5] He observes that in the debate on Sleeping Beauty the equality $P(H) = P(H_1)$ has been assumed without question: it has never been proved. If one does question the equality, and allows that $P(H)$ and $P(H_1)$ have different values, then plausible intuitions of both thirders and halfers can be preserved. Peter Lewis then tries to prove that $P(H) \neq P(H_1)$ by constructing Dutch book arguments, first for $P(H_1) = \frac{1}{3}$, and then for $P(H) = \frac{1}{2}$.

Lewis's first Dutch book argument is a clever refurbishment of Hitchcock's second bet. He turns it into a bet on $\neg H$, i.e. on tails, at odds of 1:1 on Sunday, and on H_1, rather than H, at odds of 2:1 on subsequent awakenings. The resulting bet is not only fair, but is also what Lewis calls *balanced*, by which he means that the bet on H_1 occurs with the same frequency, irrespective of whether H_1, T_1 or T_2 is true at Beauty's world (Lewis 2010, 375–376). Moreover the bet is, as we shall call it, *transparent*, meaning that Beauty is in an epistemic situation in which she can *see* that the bet is balanced. On this fair, balanced, and transparent bet, Lewis demonstrates that Beauty's credence in H_1 is one third.

The second Dutch book argument that Lewis constructs is less convincing. Although he does succeed in presenting a fair and balanced book for the conclusion that, on awakening, Sleeping Beauty's credence in H should be one half (Lewis 2010, 377), the book is not transparent: Beauty cannot see that it is balanced when she wakes up. Lewis realizes this and makes no attempt to hide the deficiency:

> "...it can be argued that the bets that make up a Dutch book must be such that they can be seen to be fair and balanced at the time that they are offered. That condition cannot be met here, and I suspect that if it is applicable, then no Dutch book can be constructed to constrain $P(H)$. If that is the case, $P(H) = \frac{1}{2}$ is not entailed by $P_-(H) = \frac{1}{2}$..." (Lewis 2010, 377).[6]

We conclude that Peter Lewis's Dutch book argument for the inequality of $P(H)$ and $P(H_1)$ falters. Although he does succeed in constructing a

---

[5]Another one is Berry Groisman (Groisman 2008). In his view, $P(H)$ corresponds to what he calls the 'set-up of coin tossing' and $P(H_1)$ to the 'set-up of wakening'. This resembles our position in that Groisman conceives the thirder and halfer answers as relating to different questions, but his detailed reasoning is different from ours. In particular he does not distinguish between the three- and the four-dimensional probability spaces.

[6]Lewis has HEADS instead of H.

persuasive Dutch book argument for $P(H_1) = \frac{1}{3}$, he is not so convincing when he tries to show that $P(H) = \frac{1}{2}$. Thus he fails to demonstrate that the two credences differ.

Dutch book reasoning is uncommonly prone to indiscriminate sniping. Why exactly should we require the bookie to be in the same epistemic situation as Sleeping Beauty? Should one demand that a book be fair, balanced, and transparent, or only fair and balanced, or simply fair? These are difficult questions, illustrating the unreliability of arguments based on Dutch book considerations.

Before proceeding to our main reasoning in Sects. 4 and 5, we will first address a difficulty that Peter Lewis has detected if one insists that $P(H)$ and $P(H_1)$ are not equal. Peter Lewis again calls on Dutch books to solve this problem, but we show that here, too, Dutch books are not needed. Our exposition in Sect. 3 serves as a stepping stone to the argument in Sects. 4 and 5.

3   Is H iff H_1 True?

While Peter Lewis concludes that $P(H)$ is unequal to $P(H_1)$, he acknowledges that there is a difficulty (Lewis 2010, 376). According to the probability calculus, A iff B entails $P(A) = P(B)$.[7] So if his conclusion that $P(H) \neq P(H_1)$ is true, then H iff H_1 is false. But H iff H_1 seems to be intuitively obvious in the Sleeping Beauty scenario. After all, 'the coin landed heads and it is Monday' clearly implies that the coin landed heads, and if the coin landed heads, then 'it is Monday' will be true of Beauty when she is first awakened (although she herself does not know what day it is).

Lewis, in fact, would like to retain the truth of H iff H_1 without giving up his conclusion. His way out of the predicament is to take further recourse to Dutch book arguments. He begins by casting doubts on the universal applicability of the rule that A iff B entails $P(A) = P(B)$, even claiming that Bradley and Leitgeb had similar reservations: "...the argument strategy that Bradley and Leitgeb used against Hitchcock's Dutch book also calls this rule into question" (ibid., 376). He then notes: "The usual rationale for the rule that A iff B entails $P(A) = P(B)$ is a (synchronous) Dutch book argument" (ibid.). He goes on to construct a Dutch book argument for this rule in line

---

[7]Indeed, if $A \rightarrow B$, then $P(B|A) = 1$, so $P(A \wedge B) = P(A)$. Similarly, if $A \leftarrow B$, then $P(A \wedge B) = P(B)$. So if $A \longleftrightarrow B$ (i.e. A iff B), then $P(A \wedge B) = P(A) = P(B)$.

with the Sleeping Beauty protocol. This argument succeeds: it shows that an agent who accepts H iff H_1, but clings to Lewis's conclusion that $P(H) = \frac{1}{2}$ and $P(H_1) = \frac{1}{3}$, can be Dutch booked. That seems to be bad news for Lewis, but he correctly notes that the Dutch book in question is not *balanced*.

The balanced Dutch book that he subsequently offers produces a different result: it *leaves open* the question as to whether or not $P(H)$ and $P(H_1)$ are equal. Lewis claims that no balanced Dutch book can be made against an agent who accepts H iff H_1 and $P(H) = P(H_1)$, but neither can a Dutch book been made against an agent who accepts H iff H_1 and $P(H) \neq P(H_1)$. An obvious conclusion would be that reasoning based on balanced Dutch book considerations is weaker than reasoning based on the probability calculus: when all is said and done, here is an assertion that the latter can prove and about which the former has nothing to say. Peter Lewis, however, heroically shuns this conclusion. Rather than yielding to the calculus, he sets up two Dutch book arguments, first for $P(H_1) = \frac{1}{3}$ and then for $P(H) = \frac{1}{2}$. As we have seen in the previous section, the former argument is convincing, but the latter less so: although the book is fair and balanced, it fails to be transparent.

In the next section we will explain how it could happen that $P(H) \neq P(H_1)$. But rather than doubting the applicability of the probability calculus on the basis of Dutch book reasoning, as Peter Lewis's argument does, this argument is completely in line with the calculus. It is based on the observation that, within the Sleeping Beauty framework, there are exceptions to H iff H_1. Given the protocol, H_1 entails H, but H does not entail H_1: it could be the case that heads have landed and that it is Tuesday.

The reason why most authors assume that H does entail H_1, and thus H iff H_1, is not difficult to discern. For *under the restricted condition that Beauty is awake*, it *does* follow that H iff H_1. After all, Beauty is awake if and only if $H_1 \vee T_1 \vee T_2$ is true. Moreover, $H_1 \vee T_1 \vee T_2$ and $H \longleftrightarrow H_1$ are both equivalent to the negation of the proposition 'the coin landed heads and it is Tuesday', i.e. $(H_1 \vee T_1 \vee T_2) \longleftrightarrow (H \longleftrightarrow H_1)$. Thus the calculus tells us only that the conditional probability of H is the same as the conditional probability of H_1:

$$P(H|H_1 \vee T_1 \vee T_2) = P(H_1|H_1 \vee T_1 \vee T_2). \tag{1}$$

Since H_1 entails H, (1) is a mathematical identity, for

$$H \wedge (H_1 \vee T_1 \vee T_2) = H_1 = H_1 \wedge (H_1 \vee T_1 \vee T_2).$$

The imagined equality of the unconditional probabilities $P(H)$ and $P(H_1)$ does not follow from (1); and although (1) is true it could be that $P(H)$ and $P(H_1)$ are not the same.

Why do most authors reason about Beauty's credences under the restriction that Beauty is awake? Again it seems the answer is not difficult to discover. It is presumably because people can only have credences when they are awake.[8] However, it is important to note that this fact poses no constraints on the content of the credences. From the fact that people can only have credences when they are awake, it does not follow that the credences they have when they are awake cannot be about the state of the world when they are asleep. As we will explain in the next section, Beauty can indeed tailor her credences to what she knows the world will be like when she herself is asleep. And this is precisely what one would expect of a rational agent.

An objection to the above reasoning might be that P stands for Beauty's credence function on Monday, and so it would be more to the point to conditionalize on MON$= H_1 \vee T_1$, rather than on her being awake, $H_1 \vee T_1 \vee T_2$. But this does not help, for

$$P(H|H_1 \vee T_1) = P(H_1|H_1 \vee T_1) \tag{2}$$

is also an identity, and for much the same reason. Thus (2) can be true while at the same time $P(H) \neq P(H_1)$. Peter Lewis's difficulty has been removed: since H iff H_1 is not true, $P(H) \neq P(H_1)$ does not involve a contradiction.

4 Sleeping Beauty's Reasoning: Three Options

The aim of the Sleeping Beauty scenario is the determination of rational credences in a situation where uncentred and centred propositions are entangled. One of the ways of discovering Beauty's rational credences is via her betting behaviour. As we have seen, it is difficult to decide when Beauty regards an aggregate of bets as fair, and balanced, in such a way that these

---

[8]And when they are awake they know that they are awake. We neglect the possibility that Sleeping Beauty could be merely dreaming that she is awake. Such outré alternatives, including the thought that she could have credences while dreaming, are beyond our remit. We shall suppose that her sleeping draught ensures dreamless sleep.

odds really do determine her credences. We therefore now propose a different way of working out Beauty's rational credences, one that makes no use of betting arguments, but rather relies on relative frequencies.

On Sunday, Sleeping Beauty is free to ruminate upon the different situations in which she may shortly find herself. She contemplates the following thought experiment. "Imagine", she says, "that the experiment were repeated every week for about four years. So every Sunday I would be put to sleep, a fair coin would be tossed, and then I would be woken up on the next day, put to sleep again, have my short-term memories of this Monday awakening erased from my brain, and be woken up again on Tuesday if the coin had landed tails, but allowed to sleep through till Wednesday if the coin had landed heads."

Suppose that the experiment were performed 2000 times in this way. Beauty continues: "I expect the coin to fall heads 1000 times during the 2000 tosses." Of course, being no mathematical neophyte, she knows all about fluctuations; her use of the word 'expect' is in the sense of the statistician's 'expectation', i.e. she means that, *on the average*, there would be 1000 heads and 1000 tails. Beauty, who is a perfectly rational agent, sets her credence in heads on Sunday, before she is put to sleep, equal to the expected relative frequency (or chance) of heads during the 2000 repeats of the experiment, that is, $p(H) = 1000/2000 = \frac{1}{2}$, where p stands for objective probability, as calculated from the expected relative frequencies.

However, the question that Adam Elga asks is not for her credence in heads on Sunday, but rather for her credence in heads on *first* awakening, that is, on Monday:

> "When you are first awakened, to what degree ought you to believe that the outcome of the coin toss is heads?" (Elga 2000, 143).

The complication is of course that after waking up on Monday she does not know what day it is: it could be Tuesday, for all she knows. It is therefore not clear that her credence in heads should still be one half when she is awakened.

There are in fact three ways in which Elga's question could be interpreted, depending on what one takes to be the meaning of "first awakened". A first possible meaning puts the stress on 'awakened'. Here Elga wants to know about Beauty's credence when she awakes. Beauty reasons: "I can calculate

the chance of heads, conditioned on my being awake; and so I will set my credence in heads equal to it, since I know that I am awake, but not that it is Monday." In this case Beauty's answer is the conditional probability of heads, given that she is awake.[9]

In a different reading, the stress is on 'first'. Here Elga is asking what Beauty's credence in heads should be, conditional on its being her first awakening. In this case Beauty's answer is the conditional probability of heads, given that it is Monday. Although Beauty does not know that it is Monday when it is Monday, she can of course still say what her credence in heads would be if she *did* have this knowledge.

A third way of understanding Elga's question is simply "What is the probability of heads on Monday?". Here Beauty's answer is supposed to reveal her rational credence in heads, as determined by the *unconditional* chance of heads, and independently of whether she is awake or asleep. By taking account only of the state of the world, leaving aside her own condition, she can calculate the chance in question. Under the third interpretation of Elga's question, Peter Lewis is right that $P(H) \neq P(H_1)$.

The salient point is that in all readings the query can be construed in terms of hypothetical relative frequencies, under the application of David Lewis's Principal Principle. The fact that Beauty is dealing with centred propositions does not prevent her from setting up an argument about relative frequencies.[10]

5 Sleeping Beauty's Calculations

Beauty's reasoning in the previous section is supported by detailed calculations. In this section we will spell them out, and give a formal analogue for each of the three interpretations of Elga's question. We will show that part of the confusion between halfers and thirders is caused by the fact that most

---

[9]Interestingly, Elga himself alludes to this objectivist reasoning (Elga 2000, 143–144). He does not develop it, presumably because he assumes that centred propositions require reasoning in subjectivist terms.

[10]Recently, Cisewski et al. have likewise argued that the debate between thirders and halfers results from conflicting assumptions that each group makes, specifying necessary and sufficient conditions for these assumptions (Cisewski et al. 2016). Their reasoning differs however from ours. The same applies to the argument of Namjoong Kim, who also argues that Elga's question is ambiguous (Kim 2015); for a reply to Kim, see Titelbaum 2015.

people assume that Beauty must use a three-dimensional probability space rather than a four-dimensional one.

Consider the 4000 different possible 'events', corresponding to Beauty's being awake 1000 times on Monday when the coin lands heads, her being awake 1000 times on Monday when the coin lands tails, her being asleep 1000 times on Tuesday when the coin lands heads, and finally her being awake 1000 times on Tuesday when the coin lands tails. The chances we shall calculate correspond to treating these possibilities as random events.

We shall supplement the propositions H, T, MON and TUES by

W: Beauty is awake

S: Beauty is asleep.

The six propositions H, T, MON, TUES, W and S constitute the events in the model.

The Cartesian product $\{H, T\} \times \{W, S\} \times \{MON, TUES\}$ yields 8 centred propositions, of which only 4 are not null, according to the rules of the Sleeping Beauty experiment. These are the possible outcomes of the model:

H_1^W : 'The coin lands heads, Beauty is awake and it is Monday'.

T_1^W : 'The coin lands tails, Beauty is awake and it is Monday'.

H_2^S : 'The coin lands heads, Beauty is asleep and it is Tuesday'.

T_2^W : 'The coin lands tails, Beauty is awake and it is Tuesday'.

By definition $H_1^W = H \wedge W \wedge MON$, but also $H_1^W = H \wedge MON$.[11] Note that the last expression looks like the definition of H_1, as given at the end of the introduction: the superscript indicates however that H_1^W is to be considered as a vector in the space spanned by the four vectors H_1^W, T_1^W, H_1^S, T_2^W. In the previous sections, as in most of the literature, $H_1 = H \wedge MON$ is tacitly supposed to be a vector in the smaller space spanned by the three vectors H_1, T_1 and T_2. The background assumption was that, when Beauty has any credences at all, she is awake, and when she is awake, she knows this (see footnote 8).

---

[11]$H \wedge MON = (H \wedge W \wedge MON) \vee (H \wedge S \wedge MON) = H \wedge W \wedge MON$, since $H \wedge S \wedge MON$, the option that the coin has landed heads, Beauty is asleep and it is Monday, is contradicted by the rules of the experiment.

By similar reasoning, we find that $T_1^{\mathrm{W}} = T \wedge \mathrm{MON}$ and $T_2^{\mathrm{W}} = T \wedge \mathrm{TUES}$. It is also true that $H_2^{\mathrm{S}} = H \wedge \mathrm{TUES}$, but in this case it is $H \wedge \mathrm{W} \wedge \mathrm{TUES}$ that is a contradiction. Further, $H_1^{\mathrm{W}} = H \wedge \mathrm{W}$,[12] and, by analogous reasoning, $H_2^{\mathrm{S}} = H \wedge \mathrm{S}$. Some of these equivalences will be used later.

Beauty calculates the chance of H_1^{W} to be equal to the number of times that she is awake on Monday when the coin lands heads (1000), divided by the total number of possible events (4000), so $p(H_1^{\mathrm{W}}) = 1000/4000 = \frac{1}{4}$. Similarly the chances of T_1^{W}, H_2^{S} and T_2^{W} are all also equal to one quarter:

$$p(H_1^{\mathrm{W}}) = p(T_1^{\mathrm{W}}) = p(H_2^{\mathrm{S}}) = p(T_2^{\mathrm{W}}) = \tfrac{1}{4} \,. \tag{3}$$

This specification of probabilities completes the definition of our model. It should be stressed that no arbitrary assumption of symmetry has been made here: the conclusion (3) follows simply from the rules of the Sleeping Beauty experiment, and the fact that the coin is fair.

Sleeping Beauty sets her credences equal to the relevant hypothetical relative frequencies, in accordance with David Lewis's Principal Principle. The relative frequency of heads, conditioned on Beauty's being awake, is

$$p(H|\mathrm{W}) = \tfrac{1}{3} \,, \tag{4}$$

(see item (a) in the appendix for this calculation). This is the answer to the *first possible meaning* of Elga's question, which is about Beauty's credence at the time of her first awakening, conditional on her being awake (and indeed she is certainly woken up for the first time on Monday).

Since $H_1^{\mathrm{W}} \wedge \mathrm{W} = H \wedge \mathrm{W}$, it is the case that

$$p(H_1^{\mathrm{W}}|\mathrm{W}) = p(H|\mathrm{W}) \,; \tag{5}$$

and from (4) we see that therefore

$$p(H_1^{\mathrm{W}}|\mathrm{W}) = \tfrac{1}{3} \,. \tag{6}$$

Notice that (4), (5) and (6) look very much like the thirders' stance, but with this crucial condition: the chances of H_1^{W} and H are equal to one another, both being one third, but only if they are conditioned on Beauty's being awake.

---

[12] $H \wedge \mathrm{W} = (H \wedge \mathrm{W} \wedge \mathrm{MON}) \vee (H \wedge \mathrm{W} \wedge \mathrm{TUES}) = H \wedge \mathrm{W} \wedge \mathrm{MON}$, since $H \wedge \mathrm{W} \wedge \mathrm{TUES}$, the option that the coin has landed heads, Beauty is awake and it is Tuesday, is contradicted by the rules of the experiment.

A puzzling feature of the first interpretation is that the word 'first' in 'first awakened' is redundant. Elga could just have easily asked "When you are awakened, to what degree ought you to believe that the outcome of the coin toss is heads?' The fact that he did not opens the possibility of a *second interpretation* of his question, where he asks for Beauty's credence at the time of her *first* awakening, that is on Monday. The relative frequency of heads, conditioned on its being Monday, is

$$p(H|\text{MON}) = \tfrac{1}{2} \, , \tag{7}$$

(see item (b) in the appendix for this calculation). This interpretation was obviously not intended by Elga, since he is a thirder and the second interpretation yields the halfer position.

What about the *third possible meaning* of Elga's question? Here Beauty is supposed to give her unconditional credence at the time of her first awakening. On Sunday she realizes that on Monday she will not know what day it is, and she also realizes that the fairness of the coin is independent of her being awake or asleep. Thus she also takes into account all the situations in which she is awake or asleep, including the situation in which she is asleep on the Tuesdays before which the coin has landed tails. Thus she calculates the unconditional relative frequency of heads to be

$$p(H) = p(H_1^{\text{W}}) + p(H_2^{\text{S}}) = \tfrac{1}{4} + \tfrac{1}{4} = \tfrac{1}{2} \, , \tag{8}$$

as should be the case, since the coin is fair.

At first sight it might appear that, since H_1^{W}, T_1^{W}, H_1^{S} and T_2^{W} are centred propositions, only credences could be attached to them, not chances. A centred proposition is after all about a specific event, in this case its being Monday or Tuesday at Beauty's temporal location; and it might seem questionable to consider such an event as an element in a series. On closer inspection, as we explained in the previous section, the propositions in the Sleeping Beauty scenario can indeed be assigned objective probabilities. For while there are well-documented difficulties in attaching a chance to such single-case events as my succumbing to cancer within the next five years, there are also single-case events for which these difficulties are less pressing or even absent. The events described by the propositions in the Sleeping Beauty experiment are a case in point. As we have seen, Sleeping Beauty's ruminations about what would happen if the experiment were repeated two thousand times engender propositions describing events (e.g. its being Monday) that are repeatable, despite their single-case appearance. Of course,

the Mondays and Tuesdays in the series of events are all on different dates, but that is irrelevant to the truth or falsity of the proposition 'it is Monday (Tuesday)'. The centricity of H_1^{W} does not prevent Beauty (in thought) from repeating the whole experiment.

6   Horgan and Pust

Terry Horgan has published a number of papers defending the thirder position (Horgan 2004, 2007, 2008; Horgan and Mahtani 2013). These papers have been criticized by Pust, who takes issue with certain symmetry assumptions that Horgan makes (Pust 2008, 2011, 2012, 2014). Although some of Horgan's findings resemble the reasoning under what we have called the first interpretation of Elga's question, his approach is quite different. Horgan is interested in *epistemic* probabilities, i.e. probabilities that are "tied to available evidence" (Horgan 2004, 13); an evidential probability essentially is "a degree of evidential support" and therefore "a rational degree of belief" (ibid.). In contrast, the reasoning under the first interpretation is based not on epistemic probabilities, but on hypothetical relative frequencies. Consequently, as we shall explain below, we can do without Horgan's symmetry principles, so that our approach is not touched by Pust's criticisms.

Horgan introduces what he calls *preliminary probabilities* of the propositions H_1, H_2, T_1 and T_2, each of which he says is equal to $\frac{1}{4}$ (Horgan 2004). These probabilities are numerically equal to our chances $p(H_1^{\mathrm{W}})$, $p(H_2^{\mathrm{S}})$, $p(T_1^{\mathrm{W}})$, and $p(T_2^{\mathrm{W}})$, see Eq.(3), but they are conceptually different. On the evidence that she has after her first awakening (not knowing that it is her first awakening), Beauty updates the preliminary probability of H_2, namely that the coin landed heads and it is Tuesday, to 0. As a result, Horgan argues, her credences in the remaining propositions are all renormalized to $\frac{1}{3}$. Horgan therefore concludes $P(\mathrm{HEADS}) = P(H_1) = \frac{1}{3}$.

Pust rejects Horgan's analysis on the grounds that the preliminary probability of H_2 before the update should be 0, not $\frac{1}{4}$ (Pust 2008). His reasoning is that it does not make sense to imagine, as Horgan does, what Beauty's preliminary probability in H_2 would be if she were to lack the information that H_2 can never be true. Horgan, in turn, replies by explaining that what he means by Beauty's preliminary probability of H_2 is determined from a third-person viewpoint:

"Beauty should assign preliminary probabilities not by contem-

plating how she would assign non-preliminary probabilities in a certain epistemic situation in which she is not conscious ... but rather by contemplating herself from a detached, third-person, perspective." (Horgan 2008, 158).

This 'third-person perspective' is quite close in spirit to our reading of Sleeping Beauty's thought experiment about what would happen if the experiment were to be repeated many times, especially to what we called the first interpretation. The important difference however is that, where Horgan has to rely on some notion of symmetry in order to assign equal preliminary probabilities to H_1, H_2, T_1 and T_2, we appeal directly to the hypothetical relative frequencies. No *a priori* logical assumption is needed in our approach: the fairness of the coin and the rules of the experiment are enough to prescribe the equality of the expected relative frequencies of H_1^W, H_2^S, T_1^W, and T_2^W.

A more recent paper of Horgan and Mahtani (2013) elaborates on Horgan's idea of generalized conditionalization: the idea is to consider, besides the standard Sleeping Beauty scenario, the three variants that are obtained (i) by switching the rôles of heads and tails, so Beauty is woken up twice if the coin lands heads, and only once if it lands tails, (ii) by switching the rôles of Monday and Tuesday, so Beauty is awakened on Monday and Tuesday if the coin lands tails, but only on *Tuesday* if it lands heads, and (iii) by switching the rôles of heads and tails, and of Monday and Tuesday, so Beauty is awakened on Monday and Tuesday if the coin lands heads, but only on Tuesday if it lands tails. Horgan's (logical) preliminary probability distribution relates to Beauty's third-person epistemic state, on the basis of the *disjunction* of the four scenarios (the original one plus the three variants (i) to (iii)). This probability distribution is strongly symmetric under exchanges of heads and tails, Monday and Tuesday, and some of the four scenarios. After generalized (synchronic) updating by conditionalization on the standard Sleeping Beauty scenario (rather than on one of the three variants), Horgan and Mahtani arrive at the thirder position.

Pust (2014) first objects by questioning Horgan and Mahtani's assumed preliminary probability distribution, which, as he shows, could quite reasonably be changed in such a way as to produce, after conditionalization, the halfer position. He then attacks the whole idea of assigning preliminary probabilities on the basis of purely epistemic notions. For if Horgan and Mahtani's method were valid, Pust argues, it should be generalizable to the situation in which the coin is biased. But then the symmetry at the heads/tails level

is broken; and, as Pust writes, Horgan and Mahtani would have to "appeal to a suitable chance-credence principle to justify the assignment of credence" (Pust 2014, 693). As Pust rightly indicates, such an appeal is foreign to the Horgan approach, which after all is essentially epistemic.

The use of a chance-credence principle is however essential to our method. Pust's criticisms have no relevance to our hypothetical relative-frequency probabilities, since we assume no probability distribution on the basis of questionable appeals to symmetry: the only input is the datum that the coin is fair, so the expectation of the number of heads is the same as that of tails. As we will see in the next section, our method also enables us easily to handle the case in which the coin is not fair.

7 The OSCAR Seminar

The OSCAR Seminar paper (2008) carries the names of sixteen authors, among them that of Horgan. It purports to show, on the basis of objective probability considerations, that Sleeping Beauty's credence on awakening that the coin landed heads is $\frac{1}{3}$. This approach is the closest to ours, but, as we shall see, it is different in that it contains another questionable assumption of an *a priori* nature, an assumption that we avoid. We will first give a simplified version of the OSCAR argument. Then we compare it with our method, and finally we review a criticism of the OSCAR paper by Pust.

From an application of the Bayes' formula, the OSCAR authors deduce

$$\frac{p(T|W)}{p(H|W)} = \frac{p(W|T)}{p(W|H)} \,, \tag{9}$$

that is, the ratio of the probabilities of tails and heads, conditional on Beauty's being awake, is equal to the ratio of the corresponding likelihoods (see item (c) in the appendix for this calculation). The authors point out that Beauty would be awake twice as long in the event that the coin were to fall tails than she would if it were to fall heads, for in the former case she would be awake for some time on two days, but in the latter case on only one day, namely Monday. "Assuming a uniform distribution over times" (OSCAR 2008, 152), the authors write disarmingly, $p(W|T)$ should be twice as large as $p(W|H)$, so from (9) we see that $p(T|W) = 2\,p(H|W)$, from which it is but a short step to the conclusion

$$p(H|W) = \tfrac{1}{3} \qquad \text{and} \qquad p(T|W) = \tfrac{2}{3} \,.$$

While we agree with this result, we find the crucial assumption of a uniform distribution over times to be gratuitous: why should the probability that the coin has landed tails be proportional to the time that Beauty is awake if tails have come up? What if the experimenters have decided (without telling Beauty) to wake her up for two hours on Monday if heads come up, but for one hour on Monday and one hour on Tuesday if tails come up? The OSCAR authors would seem in this case to be committed to $p(W|T) = p(W|H)$ and thus to the halfer position! By the method of our Sect. 5 we have established the thirder position without any assumption of uniformity over times.

The criticism that Pust (2011) has levelled at OSCAR is different from the objection we have just made. To explain the point, and to show how our method is not sullied by it, we must first briefly put aside the simplifications that we have introduced in the above explication of the OSCAR method. While OSCAR deals with objective probabilities, the authors do not consider them relative-frequency chances, but rather 'indefinite probabilities', for which they use the symbol 'prob', rather than 'p'. Thus they claim

$$\text{prob}\left(Hx|\text{B}(t,s)\,\&\,\text{Toss}(x,s)\right) = \tfrac{1}{2}\,, \tag{10}$$

in words: the indefinite probability that the coin lands heads, Hx, on condition $\text{B}(t,s)$, namely that Beauty is in a Sleeping Beauty scenario s, at time t during s, and on condition $\text{Toss}(x,s)$, namely that the coin toss, x, that is involved in s, has been effected, is equal to one half. Further, as a result of the argument that we sketched above, they also claim

$$\text{prob}\left(Hx|\text{W}(t,s)\&\text{B}(t,s)\,\&\,\text{Toss}(x,s)\right) = \tfrac{1}{3}\,, \tag{11}$$

where the extra condition, '$\text{W}(t,s)$', means that Beauty awoke in the scenario s at time t, and did not remember any previous awakenings during s.

The symbols t, s and x are to be regarded as variables, and it is for this reason that the probabilities in (10) and (11) are said to be indefinite. They can be related to 'definite probabilities', written PROB, by means of direct inference. Pust explains the OSCAR reasoning as follows:

> Let σ be a particular Sleeping Beauty scenario and let τ be the coin toss in σ ... on Sunday Beauty knows $\text{B(now},\sigma)\&\text{Toss}(\tau,\sigma)$ and so can conclude by direct inference from [(10)] that $\text{PROB}(H\tau) = \tfrac{1}{2}$. However, upon awakening during the experiment, Beauty comes to know $\text{W(now}, \sigma)\&\ \text{B(now}, \sigma)\&\text{Toss}(\tau,\sigma)$,

and as [(11)] involves 'a more specific reference property' than [(10)], Beauty should ...base her direct inference on [(11)] and conclude that $\text{PROB}(H\tau) = \frac{1}{3}$. " (Pust 2011, 291–292)

Pust objects that one could equally well base direct inference on the following indefinite probability: $\text{prob}\left(Hx|\text{Toss}(x,s)\right) = \frac{1}{2}$. According to Pust, this is not trumped by (11), since it contains only two variables, while (11) contains three. He concludes that direct inference is powerless to justify adoption of a definite probability and so direct inference alone cannot solve the Sleeping Beauty problem.

Be that as it may, our unadorned method of hypothetical relative frequencies is perfectly adequate to justify Sleeping Beauty's rational credence in heads upon awakening. As we have explained, she can calculate the expected relative frequency of the events corresponding to an awakening when the coin lands heads, and she tunes her credence in heads to this relative frequency, since she knows that she is awake and that the experiment is in progress.

Incidentally, at the end of their paper the OSCAR authors generalize the terms of the Sleeping Beauty scenario, supposing that the tossed coin is not necessarily fair, but rather has a chance α of landing heads.[13] Their method of direct inference shows that, on awakening, Beauty's credence in heads should be $\alpha/(2-\alpha)$, and they ask:

"Can any of the other arguments in the literature handle this variant of the Sleeping Beauty problem with equal aplomb?" (OSCAR 2008, 154).

The answer is in the affirmative. In fact, our method of hypothetical relative frequencies does the job with even greater ease, for in 2000 runs of the experiment, the expected number of Monday heads awakenings is 2000α, and the number of Monday tails awakenings is $2000(1-\alpha)$, the same as the number of Tuesday tails awakenings. The number of Monday heads awakenings, divided by the total number of awakenings is therefore

$$\frac{2000\alpha}{2000\alpha + 2000(1-\alpha) + 2000(1-\alpha)} = \frac{\alpha}{2-\alpha} \ .$$

---

[13]It is ironic that this generalization is the Achilles' heel of the Horgan and Mahtani paper, as we noted in the previous section.

This relative frequency determines the rational credence of Beauty, when she awakes, that the coin landed heads, in agreement with OSCAR. As a matter of fact Elga's original reasoning can also be generalized to give the same result, although admittedly it is rather long-winded (see item (d) in the appendix for this calculation).

Appendix

(a) **Proof that $p(H|\text{W}) = \frac{1}{3}$:**

$$p(H|\text{W}) = \frac{p(H \wedge \text{W})}{p(\text{W})} = \frac{p(H \wedge \text{W} \wedge \text{MON})}{p(\text{W})} = \frac{p(H_1^{\text{W}})}{p(\text{W})}.$$

The numerator here is just $\frac{1}{4}$, while the denominator is

$$\begin{aligned} p(\text{W}) &= p(H \wedge \text{W} \wedge \text{MON}) + p(T \wedge \text{W} \wedge \text{MON}) + p(T \wedge \text{W} \wedge \text{TUES}) \\ &= \tfrac{1}{4} + \tfrac{1}{4} + \tfrac{1}{4} = \tfrac{3}{4}, \end{aligned}$$

since $p(H \wedge \text{W} \wedge \text{TUES})$ is zero. Therefore

$$p(H|\text{W}) = \tfrac{1}{4}/\tfrac{3}{4} = \tfrac{1}{3}.$$

(b) **Proof that $p(H|\text{MON}) = \frac{1}{2}$:**

$$p(H|\text{MON}) = \frac{p(H \wedge \text{MON})}{p(\text{MON})} = \frac{p(H \wedge \text{W} \wedge \text{MON})}{p(\text{MON})} = \frac{p(H_1^{\text{W}})}{p(\text{MON})}.$$

The numerator is again $\frac{1}{4}$, while the denominator is now

$$\begin{aligned} p(\text{MON}) &= p(H \wedge \text{W} \wedge \text{MON}) + p(T \wedge \text{W} \wedge \text{MON}) \\ &= \tfrac{1}{4} + \tfrac{1}{4} = \tfrac{1}{2}. \end{aligned}$$

Therefore
$$p(H|\text{MON}) = \tfrac{1}{4}/\tfrac{1}{2} = \tfrac{1}{2}.$$

(c) **Proof that $\frac{p(T|W)}{p(H|W)} = \frac{p(W|T)}{p(W|H)}$:**
By Bayes' theorem,

$$p(T|W) = \frac{p(W|T)p(T)}{p(W)} \qquad \text{and} \qquad p(H|W) = \frac{p(W|H)p(H)}{p(W)}$$

21

and therefore
$$\frac{p(T|W)}{p(H|W)} = \frac{p(W|T)\, p(T)}{p(W|H)\, p(H)}\,.$$

Since the chance of a tail is the same as that of a head, $p(T)$ and $p(H)$ cancel in this equation, and we are left with
$$\frac{p(T|W)}{p(H|W)} = \frac{p(W|T)}{p(W|H)}\,.$$

(d) **Proof that $P(H) = \frac{\alpha}{2-\alpha}$ by generalized Elga method:**

Assume $P_-(H) = \alpha$. As Elga explained, it does not matter whether we (i) first toss a coin on Sunday and then wake up Beauty either once or twice depending on the outcome, or (ii) first wake up Beauty on Monday and then toss a coin to determine whether to wake her up a second time. Imagine that method (ii) is used, and that Beauty, upon awakening, were told that it is Monday. Then she would know that she is in either an H_1 or a T_1 world. Given this knowledge, her credence that she is in H_1 is her credence that an α-biased coin, soon to be tossed, will land heads:
$$P(H_1|H_1 \vee T_1) = \alpha\,.$$

This is equivalent to
$$\frac{P(H_1 \wedge (H_1 \vee T_1))}{P(H_1 \vee T_1)} = \alpha\,.$$

Since H_1 and T_1 are mutually exclusive, we find
$$\frac{P(H_1)}{P(H_1) + P(T_1)} = \alpha\,.$$

This means we have only one equation to determine two objects, viz. $P(H_1)$ and $P(T_1)$. If the coin is fair, as it is in Elga's original formulation, then this is not a problem: from $\alpha = \frac{1}{2}$ it follows immediately that $P(H_1)$ must be equal to $P(T_1)$. However, when α is any number between 0 and 1, the relation between $P(H_1)$ and $P(T_1)$ is not so clear. The first step towards discovering this relation is the inversion of both sides of the previous equation:
$$\frac{P(H_1) + P(T_1)}{P(H_1)} = \frac{1}{\alpha}\,.$$

The left-hand side can be written $1 + P(T_1)/P(H_1)$, and so

$$\frac{P(T_1)}{P(H_1)} = \frac{1}{\alpha} - 1 = \frac{1-\alpha}{\alpha}\,.$$

This means that

$$P(T_1) = \frac{1-\alpha}{\alpha} \times P(H_1)\,.$$

We follow Elga in assuming $P(T_1) = P(T_2)$, so

$$P(T_2) = \frac{1-\alpha}{\alpha} \times P(H_1)\,.$$

too. Since H_1, T_1 and T_2 are disjunct and one of them must be true, $P(H_1) + P(T_1) + P(T_2) = 1$. Thus, taking out the common factor $P(H_1)$, we obtain

$$\left[1 + \frac{1-\alpha}{\alpha} + \frac{1-\alpha}{\alpha}\right] \times P(H_1) = 1\,.$$

Gathering like terms together, we can simplify this to

$$\frac{2-\alpha}{\alpha} \times P(H_1) = 1\,.$$

and therefore

$$P(H_1) = \frac{\alpha}{2-\alpha}\,.$$

With Elga's assumption that $P(H_1)$ and $P(H)$ have the same value, we arrive at OSCAR's result. Being able to handle this generalized Sleeping Beauty is not exclusive to the objectivist approach.

Acknowledgements

References

[1] Arntzenius, F. (2003). Some problems for conditionalization and reflection. *Journal of Philosophy* 100: 356–70.

[2] Bovens, L. (2010). Judy Benjamin is a Sleeping Beauty. *Analysis* 70: 23–6.

[3] Bovens, L. and Ferreira, J.L. (2010). Monty Hall drives a wedge between Judy Benjamin and the Sleeping Beauty. *Analysis* 70: 473–81.

[4] Bradley, D. and Leitgeb, H. (2006). When betting odds and credences come apart: more worries for Dutch Book arguments. *Analysis* 66: 119–26.

[5] Bradley, D. (2003). A note on Dorr's argument for $\frac{1}{3}$. *Analysis* 63: 266–68.

[6] Bradley, D. (2011). Confirmation in a branching world: the Everett interpretation and Sleeping Beauty. *British Journal for the Philosophy of Science* 62: 323–43.

[7] Cisewski, J., Kadane, J.B., Schervish, M.J., Seidenfeld, T. and Stern, R. (2016). Sleeping Beauty's credences. *Forthcoming in Philosophy of Science. http://www.journals.uchicago.edu/doi/abs/10.1086/685741*

[8] Cozic, M. (2011). Imaging and Sleeping Beauty: a case for double-halfers. *International Journal of Approximate Reasoning* 52: 137–43.

[9] Dieks, D. (2007). Reasoning about the future: doom and Beauty. *Synthese* 156: 427–39.

[10] Dorr, C. (2002). Sleeping Beauty: in defense of Elga. *Analysis* 62: 292–96.

[11] Draper, K. and Pust, J. (2008). Diachronic Dutch books and Sleeping Beauty. *Synthese* 164: 281–87.

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

[13] Groisman, B. (2008). The end of Sleeping Beauty's nightmare. *British Journal for the Philosophy of Science*: 409–16.

[14] Groisman, B., Hallakoun, N. and Vaidman, L. (2013). The measure of existence of a quantum world and the Sleeping Beauty Problem. *Analysis* 73: 695–706: .

[15] Hawley, P. (2013). Inertia, optimism and Beauty. *Noûs* 47: 85–103.

[16] Hitchcock, C. (2004). Beauty and the bets. *Synthese* 139: 405–20.

[17] Horgan, T. (2004). Sleeping Beauty awakened: new odds at the dawn of the new day. *Analysis* 64: 10–21.

[18] Horgan, T. (2007). Synchronic Bayesian updating and the generalized Sleeping Beauty problem. *Analysis* 67: 50–59.

[19] Horgan, T. (2008). Synchronic Bayesian updating and the generalized Sleeping Beauty problem: reply to Pust. *Synthese* 160: 155–159.

[20] Horgan, T. and Mahtani, A. (2013). Generalized conditionalization and the Sleeping Beauty problem. *Erkenntnis* 78: 333–351.

[21] Kim, N. (2015). Titelbaum's theory of *de se* updating and two versions of Sleeping Beauty. *Erkenntnis* 80: 1217–1236.

[22] Lewis, D. (1980). A Subjectivist's Guide to Objective Chance, in Richard C. Jeffrey (ed.), *Studies in Inductive Logic and Probability, Volume II.* Berkeley: University of California Press, 263–293.

[23] Lewis, D. (2001). Sleeping Beauty: Reply to Elga. *Analysis* 61: 171–76.

[24] Lewis, P. (2007). Quantum Sleeping Beauty. *Analysis* 67: 59–65.

[25] Lewis, P. (2010). Credence and self-location. *Synthese* 175: 369–82.

[26] Monton, B. (2002). Sleeping Beauty and the forgetful Bayesian. *Analysis* 62: 47–53.

[27] OSCAR Seminar, The (2008). An objectivist argument for thirdism. *Analysis* 68: 149–55.

[28] Pust, J. (2008). Horgan on Sleeping Beauty. *Synthese.* 160: 97–101.

[29] Pust, J. (2011). Sleeping Beauty and direct inference. *Analysis* 71: 290–293.

[30] Pust, J. (2012). Conditionalization and essentially indexical credence. *Analysis* 109: 295–315.

[31] Pust, J. (2014). Beauty and Generalized Conditionalization: Reply to Horgan and Mahtani. *Erkenntnis* 79: 687–700.

[32] Ross, J. (2010). Sleeping Beauty, countable additivity, and rational dilemma. *Philosophical Review* 119: 411–47.

[33] Titelbaum, M.G. (2015). Reply to Kim's two versions of Sleeping Beauty. *Erkenntnis* 80:1237–1243.

[34] White, R. (2006). The generalized Sleeping Beauty problem: a challenge for thirders. *Analysis* 2006: 114–19.

[35] Wilson, A. (2013). Everettian confirmation and Sleeping Beauty. *British Journal for the Philosophy of Science* 64: 709–37.