SLEEPING BEAUTY: EXPLORING A NEGLECTED SOLUTION

Laureano Luna

ABSTRACT

The strong law of large numbers and considerations concerning additional information strongly suggest that Beauty upon awakening has probability 1/3 to be in a heads-awakening but should still believe the probability that the coin landed heads in the Sunday toss to be 1/2. The problem is that she is in a heads-awakening if and only if the coin landed heads. So, how can she rationally assign different probabilities or credences to propositions she knows imply each other? This is the problem we address in this article. We suggest that 'p whenever q and vice versa' may be consistent with p and q having different probabilities if one of them refers to a sample space containing ordinary possible worlds and the other to a sample space containing centred possible worlds, because such spaces may fail to combine into one composite probability space and, as a consequence, 'whenever' may not be well-defined; such is the main contribution of this paper. Keywords: Sleeping Beauty problem; strong law of large numbers; co-implication; centred worlds; Groisman's relativity.

1. The Sleeping Beauty Game

A fair coin is flipped on Sunday evening. If it comes up heads, Beauty is set to sleep in the laboratory from Monday 0:00 to Wednesday 0:00 (when she wakes up on her own at home) but woken up once on Monday. If the coin comes up tails, Beauty is likewise put to sleep from Monday 0:00 to her spontaneous awakening at home on Wednesday 0:00 but awoken twice in the laboratory: on Monday and again on Tuesday. But, if the coin lands tails, when Beauty is put to sleep again after her first awakening on Monday, she is given a drug that makes her forget her first awakening, so that, in any awakening in the laboratory, Beauty will never remember a previous one. As a consequence, the different awakenings are indistinguishable to her. Beauty is perfectly aware of the conditions of the game from, say, Sunday afternoon and, although she will always realize that she is in a laboratory awakening when she is in one, she will have no clue in any of them about how the coin landed or what day it is.

The question is what degrees of belief (or credences) Beauty should assign to heads and tails in the Sunday coin toss when she is woken up to a particular awakening. Of course, Beauty's credences should be the same in any awakening since the awakenings are indistinguishable to her. We need not choose one.

Beauty's phrases like "the awakening I am currently in" are given a definite meaning by the context.

There is an ongoing debate between two initial positions. So-called 'halfers' (see, for instance, Lewis [2001]) argue that Beauty should assign 1/2 to each of the results in whatever awakening, essentially because this is the assignment she should make on Sunday before she has fallen asleep and, upon awakening, she gets no additional information about the outcome of the Sunday coin toss. 'Thirders' (see, for instance, Elga [2000]) believe that Beauty should assign 1/3 to heads and 2/3 to tails in any awakening[1].

Following Groisman ([2008]), we will distinguish the event that the coin landed heads in the Sunday coin flip from the event that Beauty, upon awakening, is in an awakening resulting from a heads outcome in the coin toss (i.e. in a 'heads-awakening' as opposed to a 'tails-awakening'). In order to turn those events into two different outcomes of two different experiments, we will slightly modify the game by introducing a second experiment conducted by Beauty, as follows: each time Beauty is awakened in the laboratory she asks the experimenters and learns what awakening she is in, either the Monday heads-awakening or the Monday tails-awakening or the Tuesday tails-awakening. For Beauty, this behaves as a random experiment, since the outcome is random for her. As Beauty's experiment takes place in all awakenings, the probability that Beauty is in a particular awakening when she performs her experiment is the same as the tendency or likelihood the Sleeping Beauty game has to produce that particular awakening as compared to other awakenings (this would not be so if Beauty, for instance, conducted the experiment only on Monday or only on Tuesday or with some other bias). According to the laws of large numbers, the relative tendency of the Sleeping Beauty game to produce a particular awakening should become apparent, upon iteration of the Sleeping Beauty game, in that awakening's relative frequency.

2. Groisman's and Peter Lewis' Approaches

Groisman ([2008]) published a paper claiming to have solved the Sleeping Beauty conundrum. In a nutshell, Groisman's proposal was this: the sentence 'the coin lands heads' expresses different propositions or denotes different events in the two different experimental setups involved in the Sleeping Beauty game: in the coin toss setup, it denotes the simple heads landing of the coin; in the wakening setup, it means 'my (i.e. Beauty's) current awakening is a heads-awakening (i.e. one resulting from a heads outcome in the coin toss)'. Groisman writes:

> However, the phrase 'the coin landed Heads' alone does not define that event completely. As I will discuss in detail in the course of this article, an experimental setup is necessary to describe an event. (Groisman [2008], p. 411)

[1]See lists of thirders and halfers in (Groisman [2008]; Ross [2010]; Pust [2012]). As both positions seem to entail counter-intuitive propositions, more nuanced stances, like Bostrom's hybrid model (see Bostrom [2007]), have been proposed.

According to Groisman the problem arises only because those two different propositions are believed to be the same. The claim that an event, as an object of probabilistic reasoning, is defined only by reference to a particular experimental setup is what we call 'Groisman's relativity'. Determining the probability Beauty should assign to heads in the Sunday coin flip and determining the probability she should assign to the event that she is in a heads-awakening seem to be one and the same problem for Beauty to solve (namely, assigning probability to the proposition 'the coin lands heads') but Groisman argues that they are two different problems.

Groisman believes that making the opportune distinction solves the apparent contradiction: Beauty should ascribe the proposition expressed by 'the coin lands heads' in the coin toss setup probability 1/2 while she should ascribe the proposition expressed by the same sentence in the wakening setup probability 1/3. The contradiction arises only from the confusion of the two meanings of the sentence. Groisman claims he has dispelled rather than solved the problem because the distinction he proposes purportedly dispels the appearance of a problem.

Groisman is surely right that there are two different experiments and two different calculations. For Beauty, computing the probability of heads in order to determine her credence in heads on Sunday and computing the probability that she is in a heads-awakening upon awakening amounts to solving two different problems. As we will see the strong law of large numbers gives two different solutions. If, however, Groisman's proposal did not put an end to the debate around the Sleeping Beauty problem, this was most probably a consequence of the fact that the two different propositions expressed, according to Groisman, by 'the coin lands heads' imply each other in an obvious way and it is hard to see how one could rationally assign different probabilities or credences to propositions that are known to imply each other. Indeed, most people feel inclined to endorse the following principle:

(PE) Any rational agent who knows that p if and only if q has equal credence in p and q.

Groisman mentions the relation between the propositions in question but does not really address it[2]. He writes:

> At first sight, these two questions might seem similar, especially because there is a one-to-one logical cause-effect correspondence between them. Nevertheless they are not. (Groisman [2008], p. 411)

This 'one-to-one logical cause-effect correspondence' amounts at least to an 'if and only if', for, given the experimental setup, the coin landed heads on Sunday if and only if Beauty, upon awakening, is in a heads-awakening. Elga ([2000]) makes this relation explicit while he resolutely endorses (PE):

[2]In fact, Groisman reasons in accordance with (PE) in later writings (see e.g. Groisman, Hallakoun, Vaidman [2013]).

> Upon first awakening, you are certain of the following: you are in predicament H_1 if and only if the outcome of the coin toss is Heads. Therefore, calculating $P(H_1)$ is sufficient to solve the Sleeping Beauty problem. (Elga [200], p. 145)

H_1 is the 'predicament' that it is Monday and the coin has landed heads. Note how Elga assumes that two propositions known to be connected through a biconditional must be accorded the same probability or credence.

Peter Lewis ([2010]) does address explicitly the problem of unequal probabilities of materially equivalent propositions. He contends, as we do, that such propositions need not always receive same credence from a rational agent and points at the impossibility to merge centred and non-centred worlds into one class of worlds, against that proposed by David Lewis ([1979]). Following Peter Lewis' footsteps, we want to address here the problem of how two propositions that imply each other could be rationally assigned different credences by an agent aware of the co-implication. Does Groisman's relativity render such a thing possible when both centred and non-centred worlds are involved, as suggested by Peter Lewis? We will argue that, in fact, it does so at least in some cases, where combining these different classes of worlds into one and the same probability space brings about inconsistency.

3. Discussing Beauty's Credences

In this section, we argue for the claim that Beauty's correct credence in the event that she is in a heads-awakening is 1/3 but that her correct credence when she is in an awakening that the coin landed heads on Sunday is 1/2.

The fact that Beauty's experiment has no bias combines with a probability theorem called the strong law of large numbers (Kolmogorov [1930]) to set the probability that Beauty is in a heads-awakening to 1/3. However, the very same theorem sets the probability that the coin landed heads to 1/2. Let O be the set of actual outcomes $\{o_1, o_2, \dots, o_n\}$ in n iterations of an experiment, where the o_i are members of the experiment's sample space Ω, let $F: \Omega \to \mathbb{R}$ be a function from the sample space to the real numbers[3], and let $P: \Omega \to \mathbb{R}$ be a probability distribution; the theorem states that F's average value over the iterations of the experiment tends to its expected value $\mu = \sum_i F(o_i)P(o_i)$ with probability 1 as the number n of iterations tends to infinity[4]:

$$P\left(\lim_{n\to\infty} {\textstyle\sum_i F(o_i)}/n = \mu\right) = 1.$$

---

[3]$F: \Omega \to \mathbb{R}$ is a random variable, i.e. a function replacing qualitative outcomes by numerical values so that we can do math with them. Instead of independent and identically distributed random variables, we use the more intuitive notions of iteration of the same experiment.

[4]It is a condition for the theorem to apply that the iterations of the experiment can be represented by independent and identically distributed random variables with finite expectation μ, all of which is in order here: making F take $Heads$ to 1 and $Tails$ to 0, we deal with a Bernoulli random variable with $\mu = 1/2$.

The following is a consequence of the strong law of large numbers:

Theorem 1. The probability of an outcome in an experiment equals the number its relative frequency tends to with probability 1 as the number of iterations of the experiment tends to infinity.

The appendix derives a suitable form of theorem 1 from the strong law of large numbers.

Theorem 1 applies to our case in that Beauty when awakened would tend to be in a heads-awakening 1/3 of the time and this together with theorem 1 sets the probability that upon awakening Beauty is in a heads-awakening to 1/3. As Beauty can perform this calculation and know theorem 1, her credence in the event that she is in a heads-awakening should be 1/3. But, of course, this is only possible if (PE) fails.

There is at least one author who, ignoring theorem 1, rejects the strong law of large numbers (hence implicitly also the usual axioms of probability; see Kolmogorov [1933]) as a tool to compute Beauty's rational credences. Bostrom ([2007], pp. 72-5) acknowledges that if the Sleeping Beauty experiment were repeated, the relative frequency of heads-awakenings would be 1/3 in the limit while he maintains a halfer position, arguing we must reason differently for merely possible awakenings than for actual awakenings. This stance is at odds either with theorem 1 or with Lewis' Principal Principle, which states that credences should agree with probability assignments. Bostrom finds a difference between reasoning from the number of heads-awakenings there would be if we repeated the Sleeping Beauty experiment a large number of times, and reasoning from the number of heads-awakenings there will be as we repeat the Sleeping Beauty experiment a large number of times. But dealing with possible counterfactuals and not only with factual cases is an earmark of probability, as opposed to statistics, not so easy to dismiss[5].

Theorem 1, however, does not imply that Beauty, upon awakening, should substitute 1/3 for her initial 1/2. The first argument[6] for the thirder position we wish to examine is the faulty application of theorem 1 that assumes that if the relative frequency of heads-awakenings tends to 1/3, then the probability of heads, as computed by Beauty upon awakening, should be 1/3. The implication would only hold if (PE) were always true: if Beauty were rationally compelled to have the same credence in heads she has in being in a heads-awakening, she would

[5]Bostrom ([2007], p. 75) seems to apply to the strong law of large numbers the criterion that makes him prefer the Self-Sampling Assumption to the Self-Indicating Assumption, namely, that one must think of oneself as a random individual from among all actual (not all merely possible ones) of the suitable class.

[6]We address Elga's elegant argument for the thirder position in section 6: we will need there some notions not yet introduced.

be rationally compelled by theorem 1 to have credence 1/3 (as well as 1/2 on another count) in heads. It is remarkable that (PE) together with theorem 1 induces inconsistent beliefs.

The faulty argument is sometimes based on the record of Beauty's correct answers or wins in a repeated game. This form of the argument ignores the difference between the misleading claim:

'Beauty, if she always answers "heads", will correctly answer the question "what side did the coin come up?" 1/3 of the time',

and the true claim:

'Beauty, if she always answers "the heads-awakening", will correctly answer the question "which awakening are you in?" 1/3 of the time'.

The former claim is misleading because if the coin falls tails, Beauty would twice give the same answer to exactly the same question (on the Monday and on the Tuesday awakenings), and two identical answers to the same question should not count as two different cases of successful or failed guess.

It is the word 'time' that is tricky in this context: when Beauty is guessing the outcome of the coin toss, the relevant number of times is the number of tosses. In contrast, the relevant number of times, if Beauty is guessing the kind of awakening she is in, is the number of awakenings. Certainly, Beauty, if she said in each awakening 'the coin landed heads', would be uttering a truth on only about 1/3 of all the occasions (i.e. awakenings) but in about 1/2 of all the coin tosses. Note that, in contrast, there is no such deceptive repetition of the same question when Beauty is asked about her current awakening: then Beauty is never asked twice the same question because the question refers to a different awakening each time.

A related counter-argument, also based on the fact that on tails Beauty confronts twice the same result versus just once if the coin lands heads, was put forward by Bradley and Leitgeb ([2006]) to argue, against (Hitchcock [2004]) and others, that Beauty can rationally have credence 1/2 in heads although she should bet as if her credence in heads were 1/3: the authors contend that betting odds and credences come apart in the Sleeping Beauty case. Peter Lewis ([2010]) also tackles this point and he too denies that Beauty's betting odds should reflect her credence.

Actually, theorem 1 implies that Beauty is rationally compelled to compute probability 1/2 for heads because, and this is crucial, in whatever awakening she is in, the coin toss leading to that awakening is for her any Sleeping Beauty game coin toss. Beauty's being in an awakening does not make the coin toss that put her in it special. If the toss of a fair coin is no special one, then the probability of heads must be 1/2. As Beauty's experiencing an awakening is a consequence of any Sleeping Beauty coin toss, contributions to the Sleeping Beauty problem that attempt a defence of the thirder position by explaining how Beauty upon awakening gains information seem doomed, for it is clear that Beauty, upon awakening, learns nothing that would turn the Sunday coin toss into a special one.

To render this more evident and explain what we mean by 'a special coin toss', let us compare Beauty's case with the following at first sight similar setup.

The two-room experiment. Two independent experiments are conducted. A fair coin is flipped in room 1. In room 2, there is a box containing three cards; one card bears only the word 'heads' on it and two bear only the word 'tails'; we choose randomly one card in room 2. Before we see the word on our card, we are informed that the outcome of the coin flip and the word on our card match. What probability should we assign at that time to heads in the coin toss in room 1? ◻

We construct the sample space (see table 1, where we label the heads-card $H1$, and the two tails-cards $T1$ and $T2$):

TABLE 1

Coin \ Card	$H1$	$T1$	$T2$
H	$\underline{H \,\&\, H1}$	~~H & $T1$~~	~~H & $T2$~~
T	~~T & $H1$~~	T & $T1$	T & $T2$

and then apply Laplace's rule.

Indeed, all possible outcomes of the double experiment are equiprobable because heads and tails are equiprobable, any card is equally likely to be chosen, and the experiments are independent of one another. So, the experiment yields 6 initially possible and equiprobable cases, of which 3 (the ones not crossed out) are consistent with the information that card and coin match; and of these, 1 (underlined) is favorable to the coin having landed heads. So we must assign probability 1/3 to heads in the coin toss. In view of the coincidence, if we assigned probability 1/2 to heads in room 1 and probability 1/3 to a heads-card in room 2, we would be assigning two different probabilities to one and the same outcome, namely, H & $H1$.

It would seem that the case is essentially the same as Beauty's. Beauty in an awakening can be thought of as doing something very similar to what we have done in room 2: randomly choosing one item (an awakening) out of a set of three (possible awakenings). From an epistemological point of view, Beauty chooses randomly by simply saying 'the awakening I am presently in'. In the two-room experiment, when we learn that our choice matches the outcome of a coin toss, we gain information about that outcome because we know the match is more probable if the coin fell tails since we are more likely to draw a tails-card; this makes tails more likely[7]. It would seem that our learning of the coincidence of coin and card is equivalent to Beauty's knowing that the awakening she has chosen occurs as a consequence of the Sunday coin toss so that the items coin-toss-outcome and type-of-awakening must match. Like in the two-room experiment, this would seem to bear information about the outcome of the Sunday toss, for her awakening is a heads-awakening if and only if the coin landed heads (just as our card is a heads-

[7]Conditional probability implies this: $P(x|y) = nP(x) \rightarrow P(y|x) = nP(y)$.

card if and only if the coin landed heads), and the probability for her to be in a heads-awakening is just 1/3, not 1/2.

There is, however, a crucial difference between the Sleeping Beauty and the two-room experiment. In the two-room experiment, the coin toss in room 1 is no longer any coin toss for us, once we know of the match, but a special one: one that has happened to match the word in our card, an event that is more likely if the coin landed tails. In contrast, nothing in Beauty's awakening makes the Sunday coin toss stop being any Sleeping Beauty coin toss: here the match between the side of the coin and the type of the awakening was guaranteed from the very beginning. As if-and-only-if has always had probability 1, it conveys no information and is therefore unable to alter the probability of heads, so as to make it possible to combine the outcome heads in the coin toss and the outcome heads-awakening in the choice of an awakening into one and the same possible outcome endowed with just one probability assignment (as H & $H1$ in the two-room experiment).

However, as the present approach requires the failure of (PE), let us address this topic.

4. $(\boldsymbol{PE})$'s Failure

We interpret the set of possible outcomes of an experiment (the sample space Ω) as a set of possible worlds:

$$\Omega = \{A_1, A_2, , \dots A_m\},$$

not in the sense of metaphysical possibility but in the sense of Aristotelian potency: these worlds ('A-worlds' henceforth) are possible futures of the actual world. We can restrict ourselves to finite sample spaces. So, we can think of a random experiment with sample space Ω as a point of ramification in the world's history: after which, the world branches out into m possible futures. For instance, when dealing with a fair coin toss, the sample space is:

$$\Omega_C = \{A_H, A_T\},$$

where the subscripts correspond to the outcomes of the toss. The sample space of Beauty's experiment is:

$$\Omega_A = \{A_{\mathbb{M}H}, A_{\mathbb{M}T}, A_{\mathbb{T}T}\},$$

where the first subscript refers to the day of the week and the second to the outcome of the coin toss. That all members of Ω_A are equiprobable follows from theorem 1 and the experimental setup: the relative frequency of each of the three members would tend to 1/3 upon iteration of the experiment. We define now property $\mathcal{A}$ for A-worlds:

Definition 1. $\mathcal{A}(A) =_{DEF} A$ becomes actual.

We define a satisfaction relation $\Vdash$ between A-worlds and propositions:

Definition 2. $A \Vdash p \ =_{DEF} \ \mathcal{A}(A) \rightarrow p$.

For any proposition p, let $P(p)$ be the probability of p and for any set x, let $|x|$ be the cardinality of x. From the assumption that our probability distribution is uniform together with definition 2, we obtain:

Laplace's Rule.

$$P(p) = \frac{|\{n\,;\ A_n \Vdash p\}|}{|\Omega|}$$

We define a trans-experiment equivalence relation $E \subseteq \Omega \times \Omega'$, where $\Omega \cap \Omega' = \emptyset$:

Definition 3. $E(A_i, A_j) =_{DEF} \mathcal{A}(A_i) \leftrightarrow \mathcal{A}(A_j)$.

We wish to characterize the equivalence relation $\leftrightarrow_{\Omega\Omega'}$ which obtains between two propositions that state actualizations of two different A-worlds[8] in two different sample spaces Ω, Ω' if and only if those A-worlds are compelled by their experimental setups either to both become actual or to both fail to become actual. Such is the case for the proposition that the coin lands heads in the Sunday coin toss (which states the actualization of A_H from Ω_C) and the proposition that the awakening Beauty is presently in is a heads-awakening, which states the actualization of $A_{\mathbb{M}H}$ from Ω_A. We can characterize relation $\leftrightarrow_{\Omega\Omega'}$ in terms of A-worlds by means of theorem 2, which follows immediately from our definitions.

Theorem 2. Let Ω, Ω' be two disjoint sample spaces; then

$$p \leftrightarrow_{\Omega\Omega'} q$$

iff

$$\forall A_i \in \Omega, \forall A_j \in \Omega' \left(E(A_i, A_j) \rightarrow (A_i \Vdash p \leftrightarrow A_j \Vdash q)\right). \ \square$$

Now let $Heads$ be the proposition that the coin landed heads in the Sleeping Beauty Sunday toss and let $H - awakening$ be the proposition expressed by 'the awakening I am presently in is a heads-awakening' as stated by Beauty on Monday.

Theorem 3. $(Heads \leftrightarrow_{\Omega_C\Omega_A} H - awakening)\ \&\ \left(P(Heads) \neq P(H - awakening)\right)$

Proof. To see the first conjunct is true, consider the unique pair in $E \subseteq \Omega_C \times \Omega_A$:

$$E(A_H, A_{\mathbb{M}H});$$

and the relevant satisfaction relations between A-worlds and propositions:

---

[8]Some propositions may state compound actualizations: 'I have just been woken to a tails-awakening' would state $\mathcal{A}(A_{MT}) \vee \mathcal{A}(A_{TT})$. This requires introducing the usual σ-algebra of events in probability spaces and we will do so occasionally.

$$A_H \Vdash Heads \,\&\, A_{\mathbb{M}H} \Vdash H - awakening.$$

As theorem 1 implies the members of Ω_A to be equiprobable in Beauty's experiment, the second conjunct follows from Laplace's rule and the fact that

$$\frac{|\{A_H\}|}{|\Omega_C|} = \frac{1}{2} \neq \frac{|\{A_{\mathbb{M}H}\}|}{|\Omega_A|} = \frac{1}{3}. \quad \square$$

Theorem 3 shows that two $\leftrightarrow_{\Omega_C \Omega_A}$-equivalent propositions may have different probabilities in their respective sample spaces. Now, suppose we are informed that Beauty is presently undergoing a Sleeping Beauty experiment and that she is currently in an awakening, and assume this is all our relevant information. Theorem 1 tells us we should assign probability 1/2 to heads in the coin flipping and probability 1/3 to Beauty's presently being in a heads-awakening. Thus, a rational agent determining his credences in accordance with his probability assignments should have different credences in those $\leftrightarrow_{\Omega_C \Omega_A}$-equivalent propositions. (PE) fails, unless one of the following two propositions is true: the axioms of probability are wrong in a relevant way or rational agents should not assign credence in this case in accordance with their probability assignments. Both caveats seem unjustifiable here.

5. Making Sense of (PE)'s Failure

The intuitive argument for (PE) is this: when you know that p if and only if q, you know that, whatever the experiment, the set of outcomes making p true is the same as the set of outcomes making q true, so that if you assigned two different probabilities to those propositions, you would be knowingly assigning two different probabilities to one and the same set of outcomes.

Tschirk ([2016]) includes among the desiderata for reasoning about plausibility the following, which is a version of (PE):

> (II) Plausible reasoning qualitatively corresponds with common sense. (…) From desideratum (II) the following rule, which we call 'implication rule' can be derived:
>
> If, given C, A implies B and B does not imply A, then, given C, B is more plausible than A.
>
> It corresponds with common sense because, given C, B is true whenever A is true, but B can even be true when A is false. (Tschirk [2016], p. 80)

Indeed, the fact that the set of outcomes rendering p true is the same as the set of outcomes rendering q true is what we mean by 'whenever p is true, so is q, and vice versa'. So, if we are dealing with random experiments, 'whenever p' should mean 'in all possible outcomes making p true', and 'whenever q' should mean 'in all possible outcomes making q true'. That two propositions implying each other (in the sense that the set of outcomes making one true is the same as the one making the other true) have the same probability is a theorem[9]: but for such to be the case the propositions must refer to one and the same sample space. For instance, the proposition 'even and greater than 3' and the proposition 'either 4 or 6' are represented by the same subset of the sample space of a die roll:

$$\{2, 4, 6\} \cap \{4, 5, 6\} = \{4\} \cup \{6\} = \{4, 6\}.$$

Here, that both propositions have the same probability is an elementary mathematical fact. But what if p and q refer to two different experiments with two different sample spaces Ω and Ω'? Then the double 'whenever' should refer to a combined sample space $\Omega'' \subseteq \Omega \times \Omega'$, such that the subset of Ω'' containing exactly the outcomes that make p true is the same as the one containing exactly the outcomes that make q true. In the two-room experiment that set is $\{(H, H1)\}$. But what if the sample spaces concerned cannot combine into one because their respective members are of a definitely different nature? In such a case 'whenever' may not be well-defined, and this would make a mathematically consistent failure of (PE) possible[10].

In the Sleeping Beauty game, the problem is that the propositions $Heads$ and $H - awakening$ appear to pick one and the same temporal continuation of the world; this is what their $\leftrightarrow_{\Omega_C \Omega_A}$-equivalence seems to mean after all: either both of them come out true or both come out false when all experiments have been conducted; thus, they lie in exactly the same possible future. So, it would seem that Beauty, in assigning different probabilities to those propositions, would be assigning two different probabilities to one and the same temporal continuation of the world. Such a thing would amount to inconsistency. But is it as it seems?

We believe not and an analysis of the case can give us a clue regarding the reason why two sample spaces could in some cases be unable to combine. Note that the outcomes of Beauty's experiment are not ordinary A-worlds, that is, A-worlds representing the branching future of the world resulting from a physical random experiment (e.g. what side the coin came up), but self-locating alternatives ensuing from an epistemic random experiment (i.e. Beauty's asking what day it is). The only physical random experiment branching the future into ordinary A-worlds in our story is the coin toss. The A-worlds in Ω_A do not represent possible futures but possible locations of Beauty and what Beauty has to compute in her experiment is not the probability that the world took one or another path into the future but the probability for each of her possible locations ($\mathrm{A}_{\mathbb{MH}}$, $\mathrm{A}_{\mathbb{MT}}$, and $\mathrm{A}_{\mathbb{TT}}$) to be the actual, as suggested in Figure 1. Her branching of the world is epistemic and self-locating.

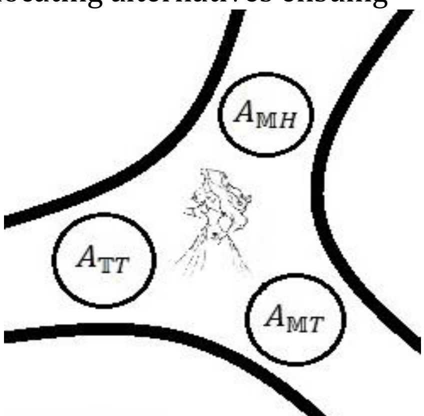

FIGURE 1

[9]From the fact that if $p \to q$, then $P(p \,\&\, q) = P(p)$, we have that if $p \leftrightarrow q$, then $P(p \,\&\, q) = P(p) = P(q)$. But this is only so if p and q dwell in one sample space.
[10]Peter Lewis ([2010]) considers a Dutch Book argument for (PE) and convincingly rejects its applicability to the Sleeping Beauty case.

Consequently, while the probability of $Heads$ concerns which possible future becomes actual, the probability of $H - awakening$ concerns self-location, namely, which awakening Beauty is currently in. When philosophers treat sample spaces as sets of possible worlds, the worlds that represent possible self-locations are called centred worlds. So, the coin toss experiment in the Sleeping Beauty game yields two equiprobable non-centred A-worlds but Beauty's experiment yields three equiprobable centred A-worlds.

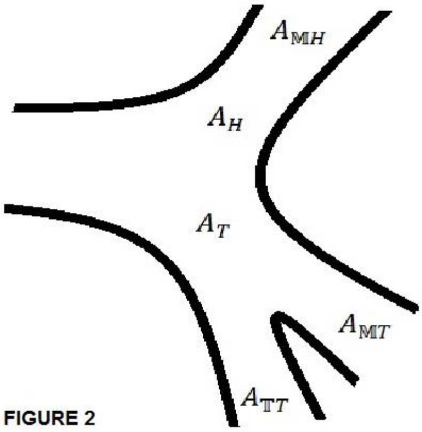

FIGURE 2

The two centred tails-awakening A-worlds are not a sub-branching of the one non-centred tails A-world, against what is suggested in Figure 2. They are part of a different structure and express a different perspective over the world. Therefore, assigning different probabilities to $Heads$ (i.e. to a possible future of the world) and to $H - awakening$ (i.e. to one of Beauty's possible locations) need not involve contradiction: the relativity of probability to sample spaces does not translate here into inconsistent probability assignments to one and the same possible branch of time, since in one case the question is about non-centred worlds but in the other it is all about centred worlds.

Beauty can set the probability that the world became a heads-world as a consequence of the Sunday toss to $1/2$ and the probability that she is in a heads-awakening to $1/3$. In fact, she relies on the former to compute the latter, for it is precisely because the main branches in Figure 2 are equiprobable that the relative frequency of heads-awakenings tends to $1/3$ when the experiment is repeated. Thus, the proposition that the world became a heads-world on Sunday is more probable for Beauty than the proposition that Beauty, upon awakening, is in a heads-awakening, because the heads-world, even if it represents half the chance of the world's future in the coin toss experiment, has only 1 in 3 equiprobable ways to accommodate Beauty in an awakening. And this is so even if the two propositions must obtain or fail together in our game[11].

[11]A case akin to ours may have been implicitly made by Carroll and Seben ([2018]) reasoning within a many-worlds interpretation (MWI) of Quantum Mechanics. In their version, Beauty's destiny is decided by an up/down spin measurement with equal amplitudes. They adopt the thirder position but argue this is not in contradiction with the Born rule because two different kinds of uncertainty are involved: between-branch uncertainty, which obeys the rule, and within-branch uncertainty, which need not. As far as we can see, their approach entails that the branches are equiprobable for Beauty to be on, though Beauty has equal probability to be in each of the awakenings, which would set the probability of being in an up-awakening equal to $2/3$. The authors may be prohibited by their implicit commitment to (PE) from deriving this consequence: indeed, Beauty is on the up-branch if and only if she is in an up-awakening. We conjecture that if partisans of MWI manage to make sense of quantum probabilities, disposing of (PE) may be their only chance to cope with quantum versions of the Sleeping Beauty problem.

If Groisman's relativity is right and an event, as an object of probabilistic reasoning, must be referred to a particular experimental setup in order to be properly defined, then mixing self-locating with non-self-locating events may lead to inconsistency because it may imply conflating different events. Figure 2 shows the only way in which Ω_C and Ω_A could combine into one probability space and it is easy to see that there is no way to consistently assign probabilities assuming theorem 1: the upper branch should receive probability 1/2 as a non-centred world but probability 1/3 as a centred one. In the following paragraphs, we show how intertwining self-locating with non-self-locating events in one and the same process of reasoning in the Sleeping Beauty problem leads to a contradiction (assuming the strong law of large numbers and Bayes' theorem).

If we define a σ-algebra on Ω_A, we can define the self-locating events or centred A-worlds $A_{\mathbb{M}}$ (meaning 'the awakening I am currently in is a Monday-awakening' as said by Beauty upon awakening), and $A_{\mathbb{T}}$ (meaning 'the awakening I am currently in is a Tuesday-awakening' as said by Beauty upon awakening). Then we can show that mixing non-centred worlds A_H and A_T with such centred worlds in the same reasoning leads to contradiction. From the Sleeping beauty game design and theorem 1, we have:

$$P(A_H) = P(A_T) = 1/2 \tag{1}$$

$$P(A_{\mathbb{M}}) = 2/3 \tag{2}$$

$$P(A_{\mathbb{M}}|A_T) = 1/2$$

$$P(A_{\mathbb{M}}|A_H) = 1 = 2P(A_{\mathbb{M}}|A_T) \tag{3}$$

From (3), by Bayes' theorem

$$P(A_H|A_{\mathbb{M}})\, {P(A_{\mathbb{M}})}\big/{P(A_H)} = 2\, P(A_T|A_{\mathbb{M}})\, {P(A_{\mathbb{M}})}\big/{P(A_T)} \tag{4}$$

From (1) and (4):

$$P(A_H|A_{\mathbb{M}}) = 2P(A_T|A_{\mathbb{M}}) \tag{5}$$

which implies

$$P(A_H|A_{\mathbb{M}}) = 2/3 \tag{6}$$

$$P(A_T|A_{\mathbb{M}}) = 1/3$$

Now, from (2) and (3):

$$P(A_{\mathbb{M}}|A_H) = 1 = {}^{3}\!/_{2}\, P(A_{\mathbb{M}}) \tag{7}$$

And from (1), (7), and conditional probability[12]

$$P(A_H|A_{\mathbb{M}}) = {}^3\!/_2\, P(A_H) = {}^3\!/_2\, {}^1\!/_2 = {}^3\!/_4 \tag{8}$$

in contradiction with (6).

Note that setting $P(A_H) = 1/3$, $P(A_T) = 2/3$, we get $P(A_H|A_{\mathbb{M}}) = 1/2$ from (4) as well as from (8). In other words, if we treat A_H as theorem 1 tells us to treat the self-locating event 'the awakening I am presently in is a heads-awakening' (i.e. as $A_{\mathbb{M}H} \in \Omega_A$), no contradiction follows.

So, in Groisman's spirit, we suggest that whenever Beauty, upon awakening, combines 'it is Monday/Tuesday' with 'the coin landed heads/tails', she automatically adopts a self-locating perspective in which 'the coin landed heads/tails' necessarily becomes 'the awakening I am currently in is a heads/tails-awakening'. Even if Beauty is able to know that heads in the Sunday coin toss has probability 1/2, everything behaves as though what makes Beauty's being in a Monday awakening more probable is not the heads outcome in the toss of the coin but her being in a heads-awakening. In other words, the sample space displayed in Table 2 appears not to exist:

TABLE 2

$\Omega_C \setminus \Omega_A$	$A_{\mathbb{M}H}$	$A_{\mathbb{M}T}$	$A_{\mathbb{T}T}$
A_H	A_H & $A_{\mathbb{M}H}$		
A_T		A_T & $A_{\mathbb{M}T}$	A_T & $A_{\mathbb{T}T}$

The reason is that there is no way, consistent with theorem 1, to assign probabilities to the squares in the table. So, the paradoxical in the Sleeping Beauty paradox would be the nonexistence of a prima facie existent mathematical object, just like in the paradoxes of set theory. The contrast with the two-room experiment is glaring. There, two experiments branching the future into ordinary non-centred A-worlds are conducted and their outcomes combine into the composite sample space shown in Table 1, according to which the propositions that imply each other, namely, $H, H1$, have both probability 1/3.

6. Elga's and Lewis' Arguments

As Bostrom ([2005]), has argued, the original arguments by Elga ([2000]) and Lewis ([2001]) encounter obvious problems in more extreme versions of the Sleeping Beauty game. If we are right that Ω_C and Ω_A cannot combine into one sample space, that they are in that sense incommensurable, then the difficulties affecting Elga's or Lewis' arguments could originate from the fact that they try to combine them. We will try to show that this is, in fact, the case.

Elga ([2000]) reasons as follows. If Beauty were to know that the coin has landed tails, she would assign the same probability to her current awakening being the

[12] See footnote 7 on p. 7.

Monday or the Tuesday one, for they are indistinguishable to her and both are sure to take place. Then, by an indifference principle, we get

$$P(A_{\mathbb{M}}|A_T) = P(A_{\mathbb{T}}|A_T)$$

whence[13]

$$P(A_{\mathbb{M}}\&A_T) = P(A_{\mathbb{T}}\&A_T) \tag{9}$$

Now Elga proposes another version of the game in which the coin is flipped on Monday night after Beauty has been awoken and put back to sleep[14]; then Beauty is awoken once again if and only if the coin falls tails. It is commonly agreed (also by Lewis) that this change does not alter the probabilities but we will have our say below on this alternate version. In this version, when Beauty learns it is Monday, she knows the coin toss is ahead, and she can only ascribe probability $1/2$ to heads in a fair coin toss that is about to take place; hence, her priors must be as shown in (10):

$$P(A_H|A_{\mathbb{M}}) = P(A_T|A_{\mathbb{M}}) \tag{10}$$

$$P(A_{\mathbb{M}}\&A_H) = P(A_{\mathbb{M}}\&A_T) \tag{11}$$

From (9) and (11):

$$P(A_{\mathbb{M}}\&A_H) = P(A_{\mathbb{M}}\&A_T) = P(A_{\mathbb{T}}\&A_T) = 1/3 \tag{12}$$

And as heads implies Monday:

$$P(A_{\mathbb{M}}\&A_H) = P(A_H) = 1/3$$

Where does Elga's reasoning go wrong? Events $A_{\mathbb{M}}$ and $A_{\mathbb{T}}$ refer unequivocally to centred worlds in the σ-algebra of Ω_A, while events A_H and A_T are interpreted by Elga as non-centred worlds. As we have seen, they cannot be meaningfully mixed in a formula such as '$P(A_H|A_{\mathbb{M}})$' in Beauty's reasoning. Consider that A_H and $A_{\mathbb{M}}$ belong in different sample spaces: '$P(A_H|A_{\mathbb{M}})$' is not akin to the probability that the die has landed on 5, given it has landed on an odd number.

In fact, (10) is problematic if A_H and A_T represent non-centred worlds. For, as we have seen in the previous section, it is clear that A_H makes $A_{\mathbb{M}}$ twice as probable as A_T does; and as A_H and A_T -as non-centred worlds- are previously equally probable (since the coin is fair), then, by conditional probability, (see footnote 7 on p. 7) $A_{\mathbb{M}}$ should make A_H twice as likely as A_T. Consequently, we should have:

$$P(A_H|A_{\mathbb{M}}) = 2P(A_T|A_{\mathbb{M}}) \tag{5}$$

[13]Because $P(A_{\mathbb{M}}|A_T) =_{DEF} P(A_{\mathbb{M}}\&A_T)/P(A_T)$; $P(A_{\mathbb{T}}|A_T) =_{DEF} P(A_{\mathbb{T}}\&A_T)/P(A_T)$.
[14]In this version, Beauty's experiment could not elicit an informed response on Monday but nothing essential changes if she receives the response sometime later.

Elga reasons for (10) in an alternative arrangement, in which Beauty is taken out of the wakening setup into the coin toss setup because heads and tails are no longer for her self-locating events but possible futures of the world. So, if we take the difference between centred and non-centred worlds seriously, it is very dubious that a result based on the alternative arrangement can be transferred to the original one.

Elga's argument makes perfect sense if 'heads' and 'tails' are interpreted all along as events in the wakening setup, that is, as centred worlds or self-locating events. If we read A_H and A_T in Elga's argument as referring to self-locating events whenever they are combined with other self-locating events, what his argument really shows is that the three possible awakenings are equally likely for Beauty to be Beauty's current awakening. But this only translates into the assignment of probability/credence 1/3 to heads in the Sunday coin toss on the assumption of (PE). Thus, 1/3 is indeed the probability/credence Beauty should assign upon being awakened to the event that she is in a heads-awakening but, if we dare drop (PE), not necessarily the probability she should assign to the Sunday coin flip having come out heads.

Lewis ([2001]), too, conflates centred with non-centred worlds. He defends the halfer position but as he equates the outcome heads in the coin toss with the outcome heads-awakening in Beauty's experiment, he is compelled to the conclusion that Beauty, informed it is Monday, should assign probability/credence 2/3 to heads, so multiplying her prior 1/2 by 4/3, because heads makes Monday 4/3 times more likely (increasing its probability from 3/4 to 1). Our proposal is that it is the self-locating event that Beauty is in a heads-awakening that should be affected by Beauty's credence in the self-locating event that she is in a Monday-awakening, increasing its credence/probability from 1/3 to 1/2, while the non-centred possible world denoted by 'heads' should retain credence 1/2. The claim that Beauty should remain a halfer even if she learns it is Monday is called the 'double-halfer' position. However, ours is not exactly the double-halfer position because such a position, as usually expounded, combines centred and non-centred worlds[15].

[15] This is why our position escapes Titelbaum's ([2012]) argument against the double-halfer, which in our view furnishes another example of the inconsistency of combining centred with non-centred worlds. In Titelbaum's scenario, the coin is first tossed on Monday night and then flipped again on Tuesday night, whatever the outcome of the Monday toss. Beauty is informed on Sunday of the additional coin toss and nothing else changes in the game. What should be Beauty's credence upon awakening in 'today's coin toss will be heads'? On one hand, the figure is obviously 1/2, because, whatever the day, the coin that is going to be tossed is fair. On the other hand, the figure should be greater than Beauty's credence that the Monday coin will land heads because it is the probability of this disjunction: Today is Monday and the Monday coin toss is heads or today is Tuesday and the Tuesday coin toss is heads; as these events are disjunct, the probability of their disjunction is the sum of their probabilities. Note that as heads in the Monday coin toss implies today's being Monday, the probability of the first disjunct is the same as the

As anticipated, Bostrom ([2005]) uses his Extreme Sleeping Beauty example to show that both Elga's and Lewis' approaches have inadmissible consequences. We will examine the case in order to show how the troubles pointed out by Bostrom are overcome by our proposal of interpreting all events in Beauty's computation as self-locating events and assigning them probabilities in accordance with theorem 1. Bostrom's example goes as follows:

> *Extreme Sleeping Beauty*
>
> This is like the original problem, except that here, if the coin falls tails, Beauty will be awakened on a million subsequent days. As before, she will be given an amnesia drug each time she is put to sleep that makes her forget any previous awakenings. When she awakes on Monday, what should be her credence in HEADS? (Bostrom [2005], p. 62)

Bostrom shows that Elga's reasoning leading to (12) would lead now to assigning heads probability $1/(10^6 + 1) \approx 0$, which is hard to admit[16]. What his argument shows, however, as we view it, is that (PE) has to fail, for assigning heads probability $1/(10^6 + 1)$ seems inadmissible but theorem 1 establishes that the probability for Beauty to be in a heads-awakening in the Extreme Sleeping Beauty game is in fact $1/(10^6 + 1)$. And there is no obvious reason for credence to deviate from probability.

To see how Lewis' approach fares in an extreme example, take k to be a large number, the number of tails-awakenings in an extreme version of the Sleeping Beauty game. Then, adopting Lewis' position, Bostrom's Extreme Sleeping Beauty would yield:

$$P(A_H) = P(A_T) = 1/2 \qquad (1)$$

---

probability that the Monday coin toss is heads. Surely, Beauty's credence that today is Tuesday and the Tuesday coin toss will be heads is greater than 0; as a consequence, Beauty's credence that today's coin toss will be heads should be greater than her credence that the Monday coin toss is heads. If the latter is 1/2, the former is greater than 1/2, which is extremely counter-intuitive (so far Titelbaum). But decreasing the latter to the end of tuning the former to 1/2 would also be counter-intuitive, for it would imply that Beauty, just upon awakening, has more information about the Monday coin toss outcome than about today's toss outcome when, in fact, all she knows, in either case, is that the coin is fair. We suggested in the preceding section that as soon as Beauty considers a compound event containing a self-locating one, she must adopt an entirely self-locating perspective; thus, we suggest that in the event 'today is Monday and the Monday coin toss will be heads' the second conjunct has to be interpreted as the self-locating 'I am in a Sleeping Beauty heads-awakening', which should receive credence 1/3. Upon plausible assumptions, this would yield probability 1/2 for today's coin toss being heads, as expected. I wish to thank an anonymous reviewer for pointing Titelbaum's paper out to me.

[16]Bostrom's figure is $1/(10^6 + 2)$, for he assumes the coin toss occurs after the Monday awakening.

As there are two equiprobable ways (heads and tails) in which it can be Monday, in one of them with probability 1 and in another with probability $1/k$:

$$P(A_{\mathbb{M}}) = \frac{1}{2} + \frac{1}{2} \cdot \frac{1}{k} = \frac{k+1}{2k} \tag{13}$$

From (13) and the game design:

$$P(A_{\mathbb{M}}|A_H) = 1 = \frac{2k}{k+1} P(A_{\mathbb{M}}) \tag{14}$$

From (1) and (14), by conditional probability[17]:

$$P(A_H|A_{\mathbb{M}}) = \frac{2k}{k+1} P(A_H) = \frac{2k}{k+1} \cdot \frac{1}{2} = \frac{k}{k+1} \approx 1,$$

which seems incredible, since Beauty, in an iterated experiment, is expected to be in a heads-awakening about half the time she is awake on Monday. Interpreting all events involved as self-locating events (as we propose to do), hence making 'A_H' denote the event (in the σ-algebra of Ω_A) that Beauty upon awakening is in a heads-awakening, and using theorem 1, the calculus would be:

$$P(A_{\mathbb{M}}) = \frac{2}{k+1} \tag{15}$$

$$P(A_H) = \frac{1}{k+1} \tag{16}$$

From (15) and the design of the game:

$$P(A_{\mathbb{M}}|A_H) = 1 = \frac{k+1}{2} P(A_{\mathbb{M}}) \tag{17}$$

From (16), (17), and again by conditional probability

$$P(A_H|A_{\mathbb{M}}) = \frac{k+1}{2} P(A_H) = \frac{k+1}{2} \cdot \frac{1}{k+1} = \frac{1}{2},$$

just as theorem 1 would have it. This suggests that both Elga's and Lewis' problems are caused by mixing centred with non-centred worlds into one and the same framework.

We wish to briefly consider the proposal by Cisewski *et al.* ([2017]), for we believe it supplies another example of how centred and non-centred worlds do not lend themselves to be combined in this context. The authors' main aim is to provide a mathematical model of the Sleeping Beauty problem capable of pinpointing the

[17]See footnote 7 on p. 7.

differences in premises or assumptions between halfers and thirders. To this end, they make Beauty update her credence, in an awakening at time t, on the event that the total evidence she currently possesses (the authors call it x) is the same as the one Beauty would have in her Monday awakening or the same as the one she would have in her Tuesday one (with inclusive 'or') at t. This event they call C_{xt}. Then the solution to the Sleeping Beauty problem, according to the authors, is the value of the formula $P(H|C_{xt})$, where H is meant to denote the event that the coin falls heads. The authors obtain a result (corollary 1) and show that the mathematics compels to this conclusion:

> According to corollary 1, the halfers' assumption entails that, with probability 1, everything that Sleeping Beauty knows on Monday at time t (including every ache, pain, and bodily function) will be known again on Tuesday if the coin lands tails. (Cisewski *et al.* [2017], p. 331)

The conclusion is rather weird: the halfer position would entail things intuitively unrelated to it; however, corollary 1 is a theorem and it does entail what the authors claim it does. The rationale for such a strange-looking conclusion is that if heads and tails must keep their Sunday probabilities unchanged for Beauty in an awakening, they must be independent of the occurrence of C_{xt}, and for this to be so, tails should not offer x double opportunity to be the same as Beauty's total evidence in some of her awakenings. As tails brings about two awakenings instead of one, the only way this can be the case is if, when the coin lands tails, Beauty's total evidence in her Monday awakening is exactly the same as her total evidence in her Tuesday awakening. But, as we have seen in section 3, if it is the total number of heads-awakenings as compared to the total number of tails-awakenings that would determine different probabilities for heads and tails, it must not really be for heads and tails but for Beauty's being in a heads-awakening and for her being in a tails-awakening. Therefore, H in $P(H|C_{xt})$ should be read as the event that Beauty is in a heads-awakening; then the mathematics makes perfect sense, on the assumption that it is on C_{xt} that Beauty must conditionalize and update.

However, Beauty upon awakening knows more than the occurrence of C_{xt}: she knows something stronger than the proposition that her evidence is the same as the one she would have in one awakening or another: she knows that she is in fact in an awakening. C_{xt} is a fact about Beauty's total knowledge (namely, that it is the same as she would have in one or another of her awakenings) but maybe not Beauty's total knowledge. So, it seems dubious that $P(H|C_{xt})$ is the correct formula. The authors show that if $P(H|C_{xt}) = 1/3$, then the probability, given tails, that Beauty's total evidence is the same in the Monday and in the Tuesday awakenings is exactly 0, which does not appear to be intuitively entailed by the thirder position, though it follows from the choice of C_{xt} for Beauty to update upon.

Therefore, the attempt of the authors at pinpointing the diverging assumptions of halfers and thirders could be unsuccessful, both because they mix self-locating and non-self-locating events in one and the same formula[18] and perhaps because their modelling violates the principle of total evidence.

---

[18]The authors seem to shun self-locating events and this appears to be the reason they introduce a set X of packs of evidence (of which x is a member) and a time t,

7. Conclusion

Considerations in the previous sections suggest that sample spaces containing centred worlds cannot always combine with samples spaces containing non-centred worlds into sets of composite worlds. In such cases, mixing worlds of different kinds in the same reasoning framework may lead one astray.

At the same time, the impossibility to meaningfully combine sample spaces with different kinds of worlds may lead to the failure of (PE): the relation between propositions established by a true biconditional may in such cases convey truth-value but not probability, hence also not credence. Ultimately, the reason is that in such cases there is no way to make the meaning of 'whenever' in 'whenever p, then q, and vice versa' precise, since the sample spaces of the different experiments to which p and q refer do not meaningfully combine into one.

Thus, we suggest that although it is true that Beauty is in a heads-awakening if and only if the coin landed heads, Beauty, upon awakening, should assign probability $1/3$ to the former and probability $1/2$ to the latter. In fact, these are the only assignments consistent with the strong law of large numbers.

Appendix

We use the strong law of large numbers to derive a suitable form of theorem 1.

Theorem. The probability $P(H)$ of the outcome H from the sample space $\Omega = \{H, T\}$ in an experiment to which the strong law of large numbers applies[19], is the number which the relative frequency $f_R(H)$ of H tends to with probability 1 as the number n of iterations of the experiment tends to infinity.

Proof. Let $F: \Omega \to \{0, 1\}$ be such that

$$F(H) = 1;$$
$$F(T) = 0.$$

Let $O = \{o_1, o_2, \dots, o_n\}$ be the set of actual outcomes in n iterations of the experiment. As the o_i are members of Ω, F returns values for them in this way:

$$F(o_i) = 1 \; if \; o_i = H;$$
$$F(o_i) = 0 \; if \; o_i = T.$$

---

and make Beauty update at t not on something like 'I am presently in an awakening' but on something like 'the member x of X happens to be at time t the same as Beauty's total evidence in some of her Sleeping-Beauty-awakenings at time t''. However, for Beauty, when she updates on x at t, t is 'now' and x is 'my current total evidence'. Thus, self-locating events are still with us. I wish to thank an anonymous reviewer for pointing (Cisewski *et al.* [2017]) out to me.

[19]Remember the condition is that its iterations can be represented by independent and identically distributed random variables with finite expectation.

As $F(o_i) = 0$ if $o_i = T$, we have:

$$number\ of\ H - outcomes\ = \sum_i F(o_i).$$

Thus, the relative frequency $f_R(H)$ of H is

$$f_R(H) =_{\text{def}} number\ of\ H - outcomes/number\ of\ outcomes = \sum_i F(o_i)/n \quad (1)$$

The expected value of any real valued function $g: X \to R$ is

$$\mu = \sum_i g(x_i) \cdot P(x_i),$$

where the x_i are the members of X. Therefore, for F, μ is

$$\mu = F(H) \cdot P(H) + F(T) \cdot P(T) = 1 \cdot P(H) + 0 \cdot P(T) = P(H) \quad (2)$$

The strong law of large numbers states that the average value of F tends to F's expected value with probability 1 as n tends to infinity:

$$P\left(\lim_{n\to\infty} {\sum_i F(o_i)}/{n} = \mu\right) = 1 \quad (3)$$

Substituting in (3) in accordance with (1) and (2), we have

$$P(\lim_{n\to\infty} f_R(H) = P(H)) = 1. \quad \square$$

References

Bostrom, N. [2007]: 'Sleeping Beauty and Self-location: A Hybrid Model', *Synthese* **157**(1), pp. 59–78.

Bradley, D. and Leitgeb, H. [2006]: 'When Betting Odds and Credences Come Apart: More Worries for Dutch Book Arguments', *Analysis* **66**(2). pp. 119–27.

Carroll, S. and Seben, C. [2018]: 'Self-locating Uncertainty and the Origin of Probability in Everettian Quantum Mechanics', *The British Journal for the Philosophy of Science* **69**(1), pp. 25–74.

Cisewski, J., Kadane, J.B., Schervish, M.J., Seidenfeld, T. and Stern, R. [2016]: 'Sleeping Beauty's Credences', *Philosophy of Science* **83**(3), pp. 324–47.

Elga, A. [2000]: 'Self-locating Belief and the Sleeping Beauty Problem', *Analysis* **60**(2), pp. 143–7.

Groisman, B. [2008]: 'The End of Sleeping Beauty's Nightmare', *The British Journal for the Philosophy of Science* **59**(2), pp. 409–16.

Groisman, B., Hallakoun, N. and Vaidman, L. [2013]: 'The Measure of Existence of a Quantum World and the Sleeping Beauty Problem', *Analysis* **73**(4), pp. 695–706.

Hitchcock, C. [2004]: 'Beauty and the Bets', *Synthese* **139**(3), pp. 405–20.

Kolmogorov, A. N. [1930]: 'Sur la Loi Forte des Grands Nombres', *Comptes Rendus Hebdomadaires des Séances de l'Académie des Sciences* **191**, pp. 910–912.

Kolmogorov, A. N. [1933]: *Grundbegriffe der Wahrscheinlichkeitsrechnung.* Berlin: Springer.

Lewis, D. [2001]: 'Sleeping Beauty: Reply to Elga', *Analysis* **61**(2): 171–176.

Lewis, P. J. [2010]: 'Credence and Self-location', *Synthese* **175**(3), pp. 369–82.

Pust, J. [2012]: 'Conditionalization and Essentially Indexical Credence', *Journal of Philosophy* **109**(3), pp. 295–315.

Ross, J. [2010]: 'Sleeping Beauty, Countable Additivity, and Rational Dilemmas', *Philosophical Review* **119**(4), pp. 411–47.

Titelbaum, M. G. [2012]: 'An Embarrassment for Double-Halfers', *Thought* **1**(2), pp. 146–51.

Tschirk, W. [2016]: 'The Principle of Indifference Does Not Lead to Contradictions', *International Journal of Statistics and Probability* **5**(4), pp. 79–83.