

Working Draft – Comments Welcome! Last Updated 2006.01.07
By Justin C. Fisher – University of Arizona, Department of Philosophy
Homepage: <http://www.u.arizona.edu/~jcfisher> E-mail: jcfisher@u.arizona.edu

The Java Problem – Another Puzzle about Self-Locating Beliefs.

Abstract.

I present what I call the Java problem – an everyday case that raises many of the same interesting issues that are raised by the recently much-discussed Sleeping Beauty problem. Java has recently drunk a cup of coffee that may or may not – depending on a fair coin-flip – have contained thought-accelerating caffeine, and she is wondering what probability she should assign to the claim that her coffee did indeed contain caffeine. The intuitively correct answer to give in the Java problem is clearly $\frac{1}{2}$, even though most arguments that have been mustered for the $\frac{1}{3}$ answer in the Sleeping Beauty problem have analogs that point to the clearly wrong $\frac{1}{3}$ answer in the Java problem. This, I argue, is good reason to call these ‘thirder’ arguments into question, and to embrace the $\frac{1}{2}$ answer in both cases.

Bleary-eyed, Java stumbled into her kitchen and poured a cup of coffee. She found her way to her armchair and sank into it. As she began to sip her coffee, she wondered whether it contained caffeine. Her husband has the peculiar habit of flipping a fair coin each morning, and brewing caffeinated coffee if the flip is heads, and decaf otherwise. Java is not fortunate enough to be able to tell the difference between these by taste or smell. So, quite reasonably, she assigns probability $\frac{1}{2}$ to the claim HEADS that the flip was heads, and hence that her coffee contains caffeine.

But, Java recalls, caffeine makes her thought processes work twice as fast – trains of thought that would take her 2 minutes to do absent caffeine she can do in 1 minute under the influence of this potent stimulant. Java knows she can’t tell ‘from the inside’ whether her thoughts are proceeding at a caffeinated clip or at a decaffeinated crawl – things would seem the same to her either way. (She’s not currently wearing her watch, nor does she have any other external processes available for useful comparison.) By now she’s quite sure that, if there was caffeine in her coffee, it has had time to take effect. What probability should she *now* assign to the claim HEADS that her coffee contained caffeine?

I, for one, think it's clear that Java should still assign probability $\frac{1}{2}$ to HEADS. After all, Java's case is not all that different from everyday cases that many of us face, and it would be quite shocking to discover that we rationally ought to adjust our probabilities for ordinary events as we come to think that stimulants or depressants may have entered our bloodstreams. It is *prima facie* absurd to say, "A moment ago I assigned probability $\frac{1}{2}$ to the claim that my coffee had caffeine in it; I haven't learned anything new about my coffee, but now I think that probability is only $\frac{1}{3}$." However, a number of recent papers regarding the Sleeping Beauty problem have proposed lines of reasoning that seem to be committed to precisely this counter-intuitive conclusion. This, I will argue, is good reason to call these lines of reasoning into question.

The Sleeping Beauty problem goes as follows. Beauty knows that experimenters will put her to sleep for two days. On Sunday they will flip a fair coin. If the coin comes up heads, they will waken Beauty just for a few hours on Monday. If the coin comes up tails, they will first waken her for a few hours on Monday, then use an amnesia drug to erase her memory of this waking, and finally waken her again on Tuesday in a circumstance subjectively indistinguishable from the circumstance in which she awoke on Monday. The commonly-asked question is, when Sleeping Beauty awakens what probability should she assign to the claim, HEADS, that the flip was heads?

In my view the correct answer is $\frac{1}{2}$. One intuitive reason for thinking this is that it is clear that Beauty should assign probability $\frac{1}{2}$ to HEADS on Sunday night, and it seems that nothing happens between then and Monday morning that should change this assignment (see David Lewis, 2001). My full reasons for being a 'halfer' are that I think that the most clearly probability-like role in a good comprehensive theory of rational belief-formation and decision-making will be played by a function that assigns the value $\frac{1}{2}$ to HEADS in this case. Sarah Wright and I (in preparation) spell out this case elsewhere and I won't attempt to present it here. Instead, I will focus on the relation between the Sleeping Beauty problem and the Java problem described above.

A number of authors have argued that the correct answer to the Sleeping Beauty problem is $\frac{1}{3}$. Before directly considering such 'thirder' arguments, I will say a little about the source of this $\frac{1}{3}$ answer, and why similar reasoning would apply to Java.

Each problem involves self-locating belief. Beauty and Java each know themselves to occupy one of two sorts of possible worlds – either a world in which the relevant flip came up heads or a world in which the flip came up tails – but each lacks knowledge about *when* exactly she is located in such a world. Beauty doesn't know whether *now* is Monday or Tuesday. Java doesn't know whether *now* is relatively soon after the caffeine took effect yielding a high speed train of thought, or instead relatively long after the decaf failed to kick in, leaving her in a plodding train of thought. For both Beauty and Java, the total amount of time in a TAILS-world that is consistent with her current epistemic state is *twice as much* as the total amount of such time in the corresponding HEADS-world. For Beauty, it could be either Monday or Tuesday if TAILS, and either of these two predicaments would be subjectively indistinguishable from the corresponding HEADS & Monday predicament. For all Java knows, she could *either* be engaged in a plodding decaffeinated train of thought, which narrows her temporal location to a certain stretch of time in a TAILS-world, *or* she could be in a twice-as-fast caffeine-accelerated train of thought, which narrows her current temporal location to *half as long* a stretch of time in a HEADS-world. It is this 2:1 ratio that, in one way or another, leads thirders to say that Beauty should assign twice as much probability to TAILS as HEADS ($\frac{2}{3}$ and $\frac{1}{3}$ respectively). Since a very similar 2:1 ratio exists in Java's case, close analogs of these arguments lead to the conclusion that Java, too, should assign twice as much probability to TAILS as HEADS.

Let us now consider four thirder arguments regarding Sleeping Beauty to confirm that these do indeed have parallels for Java. First there is the long-run frequencies argument proposed by Elga (2000, 144):

Imagine the experiment repeated many times. Then in the long run, about $\frac{1}{3}$ of the wakings would be *Heads-wakings* — wakings that happen on trials in which the coin lands Heads. So on any particular waking, you should have credence $\frac{1}{3}$ that that waking is a Heads-waking, and hence have credence $\frac{1}{3}$ in the coin's landing Heads on that trial. This consideration remains in force in the present circumstance, in which the experiment is performed just once.

We may make a parallel argument regarding Java. For simplicity, let's assume that Java believes that time is quantized, so that there are 'clicks' of time, and that there are no

changes in the world between clicks.¹ Imagine that Java's scenario is repeated week after week. On half the weeks, the flips would be heads, so she would drink caffeine and spend some number C of clicks breezing her way through an epistemic predicament like the one she currently is in. The other weeks the flips would be tails, she would drink decaf, and she would spend $2C$ clicks plodding through such a predicament. After many repetitions, we would expect only $\frac{1}{3}$ of all the clicks spent in such an epistemic predicament to be caffeinated HEADS-clicks. So, if agents are supposed to conform their probability-assignments to expected long-run frequencies, then it follows that Java should assign probability $\frac{1}{3}$ to the claim, HEADS, that her current click is a caffeinated HEADS-click.

I will consider a potential objection in a moment, but first let me present a second sort of thirder argument which raises very similar issues. Kierland and Monton (2005) propose that what really motivates thirders is the intuition that agents should assign probabilities in a way that minimizes expected *total error* accumulated over time spent awake. Beauty knows she will be awake and accumulating error for twice as much time if TAILS as she would if HEADS, so she should skew her probability assignment towards TAILS to reflect this. The same sorts of considerations would apply to Java – if she is plodding along on a slow decaf train of thought, she will spend twice as long at it as she would if caffeine-accelerated, so to reduce her expected total error over time, she should skew her probability assignments towards TAILS and decaf. Hence, this sort of argument also yields the counter-intuitive $\frac{1}{3}$ answer.

Against both these arguments, a thirder might object that her arguments should be construed, not in terms of *periods of time spent in various ways*, but instead in terms of *distinct epistemic states* that an agent enters into. Over a run of many weeks, out of all the occasions where recently-wakened Java first enters into an epistemic state of wondering about the contents of her coffee cup, only *half* will be states that occur at a decaffeinated crawl, even if these do take up twice as much time as their caffeinated counterparts. And when the flip is TAILS, Java enters into just as many potentially-

¹ My argument could be reformulated without this assumption – e.g., by defining a 'click' as some period of time that is so short that no significant amount of (caffeinated) cognitive processing could happen within a click. Even if one wanted to reject such maneuvers and insist that my assumption of quantization is significant, one still would need to face the troubles that the *quantized* Java problem poses for thirders.

erroneous epistemic states as she would enter into if the flip is HEADS, even if she spends twice as long doing it. I agree that, *construed in terms of distinct epistemic states* rather than in terms of *lengths of time*, the above thirder-arguments would yield the $\frac{1}{3}$ answer for Beauty and the $\frac{1}{2}$ answer for Java. But the challenge that faces thirders is explaining why their arguments should be construed *this way*, rather than the way I construed them above. Without a principled reason to think that the new construal is immune to the difficulties that clearly plagued the old one, we should worry that the new construal may be an unreliable guide, just as the old one was. I will leave this challenge for thirders, and move on to a third sort of thirder argument.

This third argument asks a question about what bets an agent rationally should make, and attempts to draw conclusions regarding what probabilities that agent should assign. Suppose Beauty knows that whenever she is awakened she will be offered the opportunity to ante \$2 for the chance to win \$5 if HEADS.² It is sub-optimal over the long run to be disposed to accept such bets. For each expected occasion of netting \$3 on a winning bet, you must expect to lose \$2 *twice* – once Monday morning and again Tuesday morning. Hence, there is a strong intuition that recently-wakened Beauty should reject this bet, or any bet on HEADS whose ante costs more than $\frac{1}{3}$ its gross payoff. And hence, one might conclude, Beauty should assign probability $\frac{1}{3}$ to HEADS to reflect this rational betting threshold.³

To build a parallel argument regarding Java, let's add one further stipulation whose legitimacy we will question in a moment. Let's stipulate that the way Beauty signifies her acceptance of the bet is by holding down a button, where she knows that for every second of real time that she holds down the button she will commit herself to ante an additional \$2 to win an additional \$5 if heads.⁴ This stipulation does not affect the reasoning above – Beauty still should reject the bet so long as the ante (per second) costs more than $\frac{1}{3}$ the gross payoff (per second).

Notice that the very same reasoning applies if *Java* is offered such a bet. Suppose Java holds down the button for what seems to her to be two seconds. Then, if she is buzzing along in a caffeinated HEADS-world, she will have held down the button for

² Throughout, I will assume that the agents value only dollars, and that they value each dollar equally.

³ Arguments in this ballpark are proposed by Arntzenious (2002) and Hitchcock (2004).

⁴ Suppose this rate also applies to fractional seconds that the button is held down.

only one second of *real* time, so she will net only \$3; whereas, if she is plodding along at her normal pace in a decaf TAILS-world, she will have held down the button for two seconds, and thus will lose \$4. Given the fairness of the coin, the prospect of losing \$4 should outweigh the prospect of netting \$3, so Java should reject the bet.⁵ More generally, like Beauty, Java should reject such bets if the ante (per second) costs more than $\frac{1}{3}$ the gross payoff (per second).⁶

To show that the thirder must reach this conclusion in the Java problem, I stipulated that the betting mechanism allows for more accumulation of bets the more time you hold the button down. One might object: *Surely this isn't the right kind of bet – instead we should talk about bets such that, regardless of how long you spend accepting them, you end up being committed to them only once!*⁷ Suppose Java is offered the opportunity to ante \$2 to win \$5 if HEADS in the way the objector describes. Now it doesn't matter how long it takes her to accept the bet; she'll lose only \$2 if TAILS, and net \$3 if HEADS, so (given the fairness of the coin) she should accept. Indeed, she should accept any such bet whose ante costs less than $\frac{1}{2}$ the gross payoff. This reflects the intuitively correct claim that Java should assign probability $\frac{1}{2}$ to HEADS.

I'm very sympathetic to this objection, but I don't think it will help the thirder. Suppose recently-awakened Beauty is offered the opportunity to ante \$2 to win back \$5 if HEADS, with the guarantee the objector demanded – that no matter how much time she spends accepting this bet during the scenario, she will be committed to it only once. Given this guarantee, it doesn't matter that, if TAILS, Beauty will accept the bet both on Monday and Tuesday, she'll still be committed to the bet only once and lose only \$2. Given the fairness of the coin, this potential loss is outweighed by the prospect of

⁵ This is assuming that Java doesn't have some external way of judging how long she will be holding down the button. If, for example, she could set a timer that would ensure that the button would be held for exactly one second of real time, then she should do so, as the prospect of winning \$3 would outweigh the prospect of losing \$2 (given the fairness of the coin). But let us suppose that she has no such timer available.

⁶ It may be worth noting that this reasoning exactly parallels the argument from accumulated error above – here we're accumulating additional *wagers* each second rather than additional *error*, but the principle is the same.

⁷ Technically, I think the more relevant sort of bet would be one that randomly chose some time that you were accepting that sort of bet, and makes you committed to just the bet at that time, and not at any of the other times that you were accepting it. This reformulation is necessary to handle, for example, Sleeping Beauty's betting on whether 'now' is Monday or Tuesday. However, this reformulation makes no difference to the cases discussed in the main text, so I will continue with the simpler formulation that I put in the mouth of the objector.

winning \$3, so Beauty should *accept* this bet. More generally, when offered this sort of bet, Beauty should accept if the ante costs less than $\frac{1}{2}$ the gross payoff. This rational betting threshold reflects what *I think* is Beauty's rational probability assignment: $\frac{1}{2}$ to HEADS. Since the thirder wants to reject this conclusion, he must resist the suggestion that Beauty's probabilities should reflect rational thresholds for *this sort* of bet.

The thirder is left with a trilemma. If she endorses arguments involving bets like the one involving the button that commit you to anteing more the more time you spend accepting them, then she is stuck endorsing the counter-intuitive $\frac{1}{3}$ answer in the Java problem. If she instead moves to arguments involving some other sort of bets, then she will need to both (a) come up with a sort of bet⁸ that yields her desired threshold in each case – $\frac{1}{3}$ for Beauty, $\frac{1}{2}$ for Java – and (b) give a principled argument for thinking that agents' probabilities should reflect rational thresholds for *this new sort of bet*, rather than for either of the sorts that I mentioned above. And if the thirder instead abandons betting arguments entirely, then she needs some *non-betting* argument for why we should be thirders.

⁸ Perhaps most plausibly, the thirder might hold that a rational probability assignment should reflect the ratio between “the real ante” and “the real payoff” on a rationally neutral bet. In the case where holding the button commits Java to ante \$A per second and offers a payoff of \$P per second if HEADS, Java's pressing the button for an apparent 2 seconds may potentially cause her to ante as much as 2A (if TAILS), so we might construe this as “the real ante” that she risks losing if she decides to accept the bet. Similarly, her deciding to accept may potentially net her only as much as P - A (if HEADS), which would be a P + A gross payoff beyond the “real” ante, so she might construe this as “the real payoff”. As noted above, this bet will be rationally neutral when $\frac{A}{P} = \frac{1}{3}$. When this happens, the ratio between “real ante” and “real payoff” will be $\frac{2A}{3A+A} = \frac{1}{2}$, which is the intuitively correct probability for Java to assign to HEADS. What about the case where Java will be equally committed no matter how long she holds the button down? In this case, the “real” ante and payoff, just are the stated A and P. As noted above, this bet is rationally neutral when $\frac{A}{P} = \frac{1}{2}$, so again this proposal yields the intuitively correct probability $\frac{1}{2}$.

I think this proposal is fine for Java. However, it's not at all clear how to apply this sort of proposal to the case where Beauty will be only singly committed to the bet even if she accepts it both on Monday and on Tuesday. If we consider the “real” ante and payoff in this case to be the stated values A and P, then this reasoning will yield the $\frac{1}{2}$ answer for Beauty as well. I happily endorse this answer, and, indeed, (something close enough to) this sort of reasoning is endorsed by my Disposition-Based Decision Theory (Fisher in prep; c.f., Gauthier 1985, 1986). However, many people are instead inclined to accept Causal Decision Theory, which (effectively) holds that we should construe as “real” antes and payoffs, only those changes that might be *caused* by the choice in question, with factors outside the agent's current causal control considered as fixed constraints. If the flip is TAILS, then whether or not Beauty's choice today will cause any difference in her payoff will depend upon whether or not she will end up having committed herself to the bet on the other day of the experiment, something that is outside her current causal control. This yields very puzzling and uncomfortable results for Causal Decision Theory, as laid out in (Fisher & Wright #####). For present purposes, I will leave it as a challenge to the thirder to try to find a principled way of pulling the $\frac{1}{3}$ answer out of this mess.

Let us now consider a final sort of argument that thirderers make: arguments from Bayesian updating. Several such arguments have been proposed (including Elga 2000), but I will concentrate on a particular strategy employed (in slightly different ways) by Cian Dorr (2002) and Terry Horgan (2004). They note that Beauty's situation is not importantly different from one in which she thinks that, even if the flip was HEADS, she would be given amnesia and awakened on Tuesday, but soon after she awakens each day she is informed whether or not it is Heads & Tuesday. It is fairly clear that, when she awakens, Beauty should assign equal probability ($\frac{1}{4}$) to four predicaments:

- H_1 – Heads & Monday.
- H_2 – Heads & Tuesday.
- T_1 – Tails & Monday.
- T_2 – Tails & Tuesday.

When she then learns H_2 is not the case, Beauty must somehow redistribute her probability to the remaining three possibilities (H_1 , T_1 , and T_2). Given her earlier flat probability distribution and commonly accepted principles of Bayesian updating,⁹ she must redistribute this probability equally; hence assigning probability $\frac{1}{3}$ to HEADS.

A roughly parallel argument applies to a slight variant of the Java problem, and illustrates how Bayesian updating can run awry in these cases. Suppose that, after drinking her coffee, Java glimpses the hour-indicator on a digital clock, and hence knows that it is between 8:00 and 9:00.¹⁰ The following 120 possibilities are open to Java:

- H_{00} – HEADS & now is 8:00 and some seconds.
- T_{00} – TAILS & now is 8:00 and some seconds.
- H_{01} – HEADS & now is 8:01 and some seconds.
- .
- .
- .
- T_{59} – TAILS & now is 8:59 and some seconds.

⁹ Bayesian updating calls upon an agent, upon learning E, to assign to claim C the same conditional probability that she previously assigned to C given E. Given that previously $\Pr(H_1) = \Pr(H_2) = \Pr(T_1) = \Pr(T_2)$, and given the axioms of probability, it follows that the previous conditional probability $\Pr(H_1 | H_1 \text{ or } T_1 \text{ or } T_2)$ must have been $\frac{1}{3}$. Hence, Bayesian updating requires that someone with these previous probabilities assign probability $\frac{1}{3}$ to H_1 upon learning that either H_1 , T_1 or T_2 is the case.

¹⁰ One allow for a slight delay for her recognition of the significance of what she glimpsed. E.g., if this recognition would take 2 minutes in decaf-mode, then Java knows that either TAILS and the time is between 8:02 and 9:02, or HEADS and the time is between 8:01 and 9:01. This offset of 1 or 2 minutes would be irrelevant to the argument below – what matters is that, either way, she has exactly one hour's worth of possibilities. Since this offset is irrelevant, I will ignore it in what follows.

Since each of the 60 HEADS-possibilities ($H_{00}, H_{01}, \dots, H_{59}$) involves one minute, and since there is no reason to favor any one over the others, it seems clear that Java should assign equal probability to each. (Dorr and Horgan embrace this sort of move as an application of what Elga (2000) calls the “limited principle of indifference”.) Similarly, Java should assign equal probability to each of the 60 TAILS-possibilities. There is also a fairly strong intuition – though my argument need not rely upon it¹¹ – that her probabilities for each of these 120 possibilities should be the same (namely $1/120$), and hence that she should assign total probability $1/2$ to HEADS.

This flat probability distribution is analogous to the flat probability distribution Dorr and Horgan think Beauty should assign before she gains evidence eliminating all but three possibilities. Now we need an analog of this evidence for Java. Suppose Java hears the half-hour chime from a clock, and hence knows that, at the time of the chime, it was exactly 8:30. Suppose she also knows that it would have taken more or less time to recognize the chime, depending upon whether her cognition was caffeine-accelerated. For the sake of simple numbers, let us say she would recognize it at 8:32 if TAILS and at 8:31 if HEADS. Now, if TAILS, she would stay in her current epistemic predicament for twice as long as she would if HEADS, before recognizing that she had moved on to other thoughts. Again for the sake of simple numbers, let us say she would stay in this predicament for two minutes if TAILS and for one minute if HEADS. (Doing all this in terms of whole minutes is quite arbitrary. We could substitute shorter lengths of time, to ensure our partition is at least as fine-grained as Java’s capacity to detect changes in her epistemic predicament. So long as we respect the 2:1 ratio imposed by the caffeine, the end result will be the same. Without loss of generality, I will stick with whole-minute increments.) This leaves three possibilities for Java:

- H₃₁ – HEADS & now is 8:31 and some seconds.
- T₃₂ – TAILS & now is 8:32 and some seconds.
- T₃₃ – TAILS & now is 8:33 and some seconds.

¹¹ One might resist this intuition and say, for example, that Java should assign probability $1/180$ to each of the 60 HEADS-minutes and probability $1/90$ to each of the TAILS-minutes, yielding a total probability of $1/3$ for HEADS, and $2/3$ for TAILS. Even so, you will get the very strange result that after Bayesian-updating in response to the information described below, Java should *change* these probabilities, and assign probability $1/5$ to HEADS and $4/5$ to TAILS. As I note below, this *change* is problematic enough.

Having recognized the chime, Java must redistribute her probability among these remaining three possibilities. Given her earlier flat probability distribution, and the same principles of Bayesian updating that the thirder employed for Beauty, Java must assign probability $\frac{1}{3}$ to each of these remaining possibilities, and in particular, she must assign probability $\frac{1}{3}$ to HEADS.

This conclusion is problematic in several ways, but I will focus on just one. If this Bayesian line of reasoning is correct, then, just by learning that it was recently 8:30 – something that *prima facie* has nothing at all to do with the coin or the coffee – Java must *change* the probability that she assigns to the claim that the coin flipped heads and her coffee contained caffeine (from $\frac{1}{2}$ to $\frac{1}{3}$).

Worse, nothing in this argument even required that Java learn the time. What is doing the work is that she notices herself to be in a particular epistemic predicament that, due to the effects of caffeine, she would spend twice as long in if TAILS than she would if HEADS. This 2:1 ratio, together with the flatness of her earlier distribution, is enough to force her to change her probabilities in favor of TAILS. We might just as easily have framed the argument in terms of her noticing that she is imagining a pink elephant. When she notices that she is in this epistemic predicament, and knows that she would spend twice as much time in it if TAILS than if HEADS, Bayesian updating will lead her to change her probabilities in favor of TAILS.

The thought that *such seemingly irrelevant evidence should force changes in these probabilities* is so counter-intuitive as to make one suspect that Bayesian updating must be running awry in this case. I attempt elsewhere¹² to diagnose exactly what the problem is with using Bayesian updating in such cases, but I will rest here, having shown that it does lead to serious problems.

It is worth stressing that the features that make Bayesian updating yield counter-intuitive results for Java exactly parallel the features that that do the work in yielding the $\frac{1}{3}$ answer in Dorr and Horgan's modified Sleeping Beauty scenario. In both cases, there is a point where it is intuitively clear that an agent should assign a flat probability distribution to numerous options. Then, the agent receives information that eliminates

¹² See, Fisher & Wright (in preparation). For a related halfer proposal that offers a good alternative to Bayesian updating in this case, see Meacham (manuscript).

some of these possibilities, leaving open twice as many Tails-possibilities as Heads-possibilities. Given this 2:1 ratio and the flatness of the prior distribution, Bayesian updating prescribes the $\frac{1}{3}$ answer in each case. Since Bayesian updating is working in essentially the same way in both cases, insofar as we're quite sure it has given us the wrong answer in one case (Java's) we should doubt its trustworthiness as a guide in the other (Beauty's). This is enough to call into question thirder's Bayesian arguments regarding Sleeping Beauty.

In brief conclusion, there are three options open now, each of which is quite interesting. One option would be to bite the bullet and say that thirder arguments do indeed apply to Java, and that we should indeed alter our probabilities about ordinary events as we think stimulants and depressants may have entered our bloodstreams, and again each time we notice that we have entered into a new epistemic predicament (like hearing a clock strike 8:30, or imagining a pink elephant). This would be a striking – not to mention extremely implausible – practical upshot for thirder arguments that heretofore had seemed utterly academic and innocuous. Second, thirder might try to spell out their arguments more carefully, in such a way that they clearly apply to Sleeping Beauty but not Java. Such a revision and clarification of thirder arguments, if it could be done, would be a very useful step forward in these debates. The third, and I think best, option is to conclude that all these thirder arguments are irrevocably flawed, and that we should instead move to embrace a halfer position. This would be a striking and controversial conclusion, but one that I nevertheless think is correct.

References

- Arntzenius, F. 2002. Reflections on Sleeping Beauty. *Analysis* 62: 53-62.
- Dorr, C. 2002. Sleeping Beauty: in defence of Elga. *Analysis* 62: 292-96.
- Elga, A. 2000. Self-locating belief and the Sleeping Beauty problem *Analysis* 60(2): 143-147.
- Fisher, J. in prep. Disposition-Based Decision Theory.
- Fisher, J. & Wright, S. in prep. Sleeping Beauty and Two Theories of Probability.
- Gauthier, D. 1985. Maximization Constrained: The Rationality of Cooperation, in Campbell & Sowden (1985, pp. 75-93).
- Gauthier, D. 1986. *Morals by Agreement*. Oxford: Clarendon.
- Horgan, T. 2004. Sleeping Beauty awakened: new odds at the dawn of the new day. *Analysis* 64 (281), 10-21.
- Hitchcock, C. 2004. Beauty and the Bets. *Synthese* 139 (3), 405-420.
- Kierland, B., and Monton, B. 2005. Minimizing Inaccuracy for Self-Locating Beliefs. *Philosophy and Phenomenological Research*.
- Lewis, D. 2001. Sleeping Beauty: reply to Elga. *Analysis* 61: 171-76.
- Meacham, C. manuscript. Sleeping Beauty and the Dynamics of De Se Beliefs.