Perspective Reasoning and the Solution to the Sleeping Beauty Problem

Xianda Gao[1]
November 2018

This paper proposes a new explanation for the paradoxes related to anthropic reasoning. Solutions to the Sleeping Beauty Problem and the Doomsday argument are discussed in detail. The main argument can be summarized as follows:

1. Our thoughts, reasonings and narratives inherently comes from a certain perspective. With each perspective there is a center, or using the term broadly, a self.
2. The natural first-person perspective is most primitive. However we can also think and express from others' perspectives with a theory of mind.
3. A perspective's center could be unrelated to the topic of discussion so its *de se* thoughts need not to be considered, e.g. the perspective of an outside observer. Let's call these the third-person perspective.
4. First-person reasoning allows primitive self identification as I am inherently unique as the center. Whereas from third-person perspective I am not fundamentally special comparing to others so a reference class of observers including me can be defined.
5. It is my contention that reasonings from different perspectives should not mix. Otherwise it could lead to paradoxes even independent of anthropic reasoning.
6. The paradoxes surrounding anthropic reasoning are caused by the aforementioned perspective mix. Regarding the sleeping beauty problem the correct answer should be double halving. Lewisian halving and thirding uses unique reasonings from both first and third-person perspectives.
7. Indexical probabilities such as "the probability that this is the first awakening" or "the probability of me being one of the first 100 billion human beings" also mixes first- and third-person reasonings. Therefore invalid.
8. Readers against perspectivism may disagree with point 1 and suggest we could reason in objective terms without the limit of perspectives. My argument is compatible with this belief. Objective reasoning would be analytically identical to the third-person perspective. My argument would become that objective reasoning and perspective reasonings should not mix. In the following I would continue to use "third-person perspective" but readers can switch that to "objective reasoning" if they wish so.

---

[1] Comments are welcome at gaoxianda@gmail.com

"Everyone is presented to himself in a particular and primitive way, in which he is presented to no one else" Gottlob Frege, (1918)

1.Two Ways of Reasoning

Consider this thought experiment:

Imagine waking up alone in an empty room. The following message is displaying on the wall: An advanced alien race has made a trillion clones of a certain earthling. You are one of the clones. Their process is highly accurate so it's humanly impossible to find any variation among the copies. They put each clone in an identical room as the one you are in right now. The question is, do you know which one is you?

At first look it seems to be a rather strange question. It is extraordinarily trivial because I can inherently tell who I am. I can simply point to myself and say "this is me" and call it a day. So the answer must be *yes* I do know which is me. However there appears to be more to the problem. Because I am having a hard time trying to differentiating me from the other clones. They could also point to themselves and say "this is me" and also be correct. So in what way am I unique? This is beyond my knowledge because it is humanly impossible to find any variations among the copies. In this sense the answer can only be *no* because I could not definitively describe which clone is me. This contradiction highlights the difference between first- and third-person reasoning.

From first-person perspective the question is essentially asking if I'm able to identify myself. Let us put aside the philosophical debates about the nature of self for a moment. Here I'm only using self in the context of everyday language, as a specification of an individual. In this sense the answer is of course yes. From first-person perspective everything I know, all of my mind's content, my feelings, thoughts, ideas and desires are ultimately based on experience or perception. As David Hume put it "I never catch myself at any time without a perception, and never can observe anything but the perception." It is how I know the world. However things in the world does not have equal effect to the perception. Some part of the world is more directly related to it while some part more distant. When there is a paper cut on my finger it immediately registers pain. However a debilitating wound of a frontline soldier can only affect the perception though a series of event (such as a news report) which eventually lead to some sort of sensory input. The part of the world that is more immediate to the perception, things most intimate to the experience, can be regarded as the internal self. The rest of the world which is more remote can be seem as the external environment. This suggests the word self and environment are just relevant concepts with no clear cut difference. I think that is indeed the case. The convention of drawing the line at body's physical boundary is nothing but an arbitrary decision motivated by conveniences, but I digress. Regardless of what is a reasonable boarder of self, in the world described by the thought experiment, there is obviously something most directly related to the

perception - someone specific. Following the question's cue I can refer that specific clone as myself. Hence I point and say "This is me".

It seems reasonable to assume things exist in reality independent of my perception of them. So even though I cannot observe them, others have their own minds. Their thoughts, intentions, feelings and experiences are just as real as mine. If that's the case then with a theory of mind it is perfectly valid to ask how could other minds identify me. Most interestingly what makes me unique in the eyes of someone impartial towards any of the clones. This is the question from third-person perspective[2]. Here the immediacy to my perception is no longer useful since it is imperceivable to others. Even if an outsider can observe all minds I am no more intimate to its perception than the rest of the clones. Therefore the differentiation must be based on something independent of one's perspective: something objective. However because the cloning process is highly accurate I do not know any objective difference that would set me apart from the rest. So the information I have is not enough for an outsider to identify me. In this sense the answer is no, I do not know which one is me. Of course I can try to answer the question by performing actions and hoping objective differences arise among the clones.[3] For example I can hold up six fingers and then identify me as "the clone who's holding up six fingers". But there is always the possibility that somebody else among the trillion clones is doing the same. So the identification is never definitive. If each room were assigned an unique number I can definitively identify me by it when that number is revealed to me. Nevertheless this is obviously very different from the first-person reasoning where I simply knows "this is me" without needing any information.

The above example shows first- and third-person perspectives can potentially give opposite answers. It is crucial to acknowledge they are two completely different reasonings processes that are not interchangeable. The example also shows reasonings from different perspectives should be kept apart and not mixed. At least in this example that is quite intuitive. When asked "do you know which is you?" I imagine majority of people would want some clarification and ask what exact is it referring to. Can I just point and say "this is me"? Or do I need to provide an objective description? After all an answer of "yes AND no" is rather paradoxical. However in everyday life such distinction of perspectives are rarely needed. So we are apt to confound them and often arbitrarily switch perspectives in our logic. It shall be my contention that this perspective inconsistency is the root cause of many paradoxical arguments. These include the Doomsday Argument and the Sleeping Beauty Problem. Before I get into the particulars of each case I want to further analyze the properties and differences between the perspectives.

---

[2] Here on I am going to use third-person perspective to refer reasoning as if I am an outside observer, i.e. not as someone who is perceptually close to any of the clones. This way all clones are in equal positions. Perhaps due to this impartiality this perspective is very often used in logic. It should also be pointed out that this observer does not have to be any specific being who actually exist. An completely imaginary observer is suffice for the purpose of third-person reasoning.

[3] Let's ignore the debate of existence of freewill or if it possible for (near) identical clones in (near) identical environments to make different decisions. They are unrelated to the argument. For the sake of the narrative let's assume yes.

In the thought experiment when I identify myself in first-person all I can say is "this is me". It is a non-qualitative, non-descriptive reference. That is because qualities or features had no part in the reasoning. First-person self is subjective and primitive. Imagine yourself to be taller, shorter, being the opposite sex, born in a different time or some type of alien life-form all together. None of these changes would make it not you, as long as you are experiencing the world from its perspective making it the most immediate thing to the perception. This lack of quality and attributes suggest first-person self as the center of perception and consciousness is technically an immaterial concept. When we use the term self in everyday language to refer a person it is just an approximation of said concept. It also means the first-person identity is purely based on haecceity: the *thisness* of a thing. Compare that to the third-person identity which is based on quiddity, the *whatness* of a thing. Back in the thought experiment from third-person perspective I am defined by identifiable features, such as the room number. It is a qualitative and descriptive reference. Trying to mix first-person identity to a third-person identity would result in perspectively inconsistent problems. For example "Why am I(first-person) me(third person)?", "Why am I not Bill Gates?", or "How come I'm experiencing the world from this human's perspective instead of some alien's?". Logic reasoning could not provide any answer to such questions.

In the previous thought experiment, how similar the clones are is irrelevant in first-person reasoning. In fact I don't even have to acknowledge other clones' existence to say "this is me". Because from my perspective I am inherently special. Everything beside myself, including the other clones, are just part of the external environment. It does not matter if the universe is infinite or there are multiverses with countless observers similar to me. It is clear to me that this is me and not any of those. Plainly put, from first-person perspective I am one of a kind, a peerless entity. From third-person perspective however, I am no longer in such privileged position. There is no inherent property that makes me stand out from others. Distinctive features have to be used to specify me from a reference class of observers. Therefore this reference class only exists in third-person perspective. For these reasons I shouldn't regard my first-person self as a random sample from a reference class.

Another point worth mentioning is that first-person perspective is the logic basis of self-interest. Pleasure or pain, gain or loss, I value them most if they are of my own. Because that way they are most imminent to the perception with greatest effect on the experience. This primitive self-interest cannot be justified from third-person perspective as it is neutral towards all entities in the reference class. There is no reason to put one clone's interest over another just because their room numbers are different. This highlights the importance of first-person reasoning since the concept of rational agents, which is widely used in decision theory and economics, is based on this primitive self-interest.

However first-person reasoning is not without its limitations. One of its shortcoming is regarding the counterfactual scenario of me not existing. It is impossible to discuss such ideas from the first-person perspective because it is ontologically self-contradictory to reason from a perspective while also declaring that perspective's center do not exist. Clearly none can be

identified by immediacy to perception if there is no perception to begin with. If this is not obvious enough we can look at its contraposition. To see my own nonexistence cannot coexist with first-person reasoning I only need to realize that first-person reasoning will always affirm self existence. This was done by Rene Descartes with the famous cogito ergo sum, usually translated as "I think therefore I am". This idea is also captured by the intuitive notion that "I would always find myself exist". From first-person perspective there simply is no counterfactual for my existence. My nonexistence can only be meaningfully reasoned from a third-person perspective where I am specified by features instead of immediacy to perception. Consequently any discussion about the probability of my current existence must also be done from a third-person perspective.

Another perhaps more important limit of the first-person identify its *incommunicability*.[4] When I identify myself from first-person perspective the only criteria was the immediacy to perception. But because my perception is imperceivable to others this subjective identity is only meaningful to me. To effectively communicate I have to use third-person perspective and describe me by objective features. Perhaps the most common practice is the use of names. From first-person perspective assigning myself a name is redundant. I would never confuse myself with anybody else. However names make communication immensely easier by acting as an observable feature by all. Trying to directly communicate first-person self is futile. In everyday life when I say "this is me" to someone it may seem to be meaningful only because the listener uses contextual information to define me in third-person. The listener may be defining me as the guy who's talking, the one with a specific look or all observable traits combined. Getting rid of these contextual information and the sentence would be incomprehensible. In the thought experiment if you are trying to leave a message for an outsider to pinpoint you among all the clones then the least helpful thing would be writing down "this is me".

This incommunicability of first-person self identification has an interesting implication. It suggests in some rare cases two rational people sharing all their information could rightfully disagree with each other. Even if they are free to communicate. Each of them would think the other person's reasoning is correct although his result is different i.e. they would agree to disagree. Some readers might notice this seems to contradict Aumann's Agreement Theorem. I would argue this is justified. For this disagreement to occur one person's observable features shouldn't related to the matter at question other than he belongs to a certain reference class. So from third-person perspective he is just an unspecific individual from said reference class. While from first-person perspective he would naturally see himself as a distinct entity. Because this specification is based on the primitive first-person identification it is incommunicable. Thus they would have different interpretations for the same situation, leading to a disagreement. Such cases rarely occurs in everyday life. However in thought experiments involving the duplication of mental states it can happen. I will explain in more detail with examples in section 4.

---

[4] Another way to put it is "the *de se* cannot be reduced to the *de dicto*. There is something special about learning who oneself is that cannot be captured in learning about what features one process" (Liao, 2012)

2. Doomsday Argument

The Doomsday Argument (Carter and McCrea, 1983) is a probabilistic argument predicting the number of future humans yet to be born based on one's own birth rank. The argument can be summarized as follows. Let *N* be the total number of humans who were or will be born. Let *n* be my own birth rank among all humans. Without any other information I shall assume I am equally likely to be at any position from 1 to N. After learning my own birth ranking, say n is 100 billion, I can construct a posterior distribution of N, which suggests the total number of humans would more likely to be smaller than bigger. E.g. probability of N being 200 billion is greater than the probability of N being 200 trillion. Ceteris paribus, this suggest we should have a more pessimistic outlook about our future as a specie, i.e. the doom is closer than we normally think.

The argument pivots on the Self Sampling Assumption (SSA). The term is coined by Nick Bostrom (2002) which suggests all other things equal, I should reason as if I am randomly selected from the set of all *actually existent* observers (past, present and future) in my reference class. An alternative principle which rebuttals the doomsday argument is the Self Indication Assumption (SIA). It states that all other things equal, I should reason as if I am randomly selected from the set of all *possible* observers in my reference class.[5] To illustrate the difference between the two the following thought experiment is often used.

Imagine there are a thousand rooms numbered from 1 to 1000. A fair coin is tossed by the advanced alien. If it lands on heads a clone of an earthing would be created in each room from 1 to 1000. If it lands on tails a clone would be created in each room from 1 to 10 only. Suppose you wake up in one of the rooms and the experiment procedure was displaying on the wall in front of you. How should you guess which way the coin fell? Now suppose the room number is revealed to you. It turns out to be 7. How should you guess the coin now?[6]

According to Self Sampling Assumption when I wake up in the room I have no relevant information about the coin toss. So I ought to guess the fair coin landed with an even chance. Here I can calculate the conditional probability of my room number by treating myself as a randomly selected individual among all existing clones. For example, conditional on heads the probability that my room number being between 1 and 10 is 1% since only 10 out of the 1000 clones are in those range. While conditional on tails the probability that my room number being between 1 and 10 is 100% since all clones are in the first 10 rooms. After knowing my room number is 7 I have to adjust my answer about the coin. The probability of tails is now much higher since it is more likely to give the observed room number. With Baye's theorem it can be calculated to be around 99%. It is easy to see why such reasoning would lead to the doomsday argument when transferred to the future of human population. Under Self Sampling Assumption

---

[5] Nick Bostrom used the term in a slightly different way according to which this school of thought should be called SSA+SIA. However majority of the literature simply refer it as SIA. The argument is first proposed by Dieks (1992) and also used by Olum (2002) as a counter to the Doomsday argument.

[6] This example is used by Nick Bostrom (2002) with slight modifications.

a hypothesis with less people existing would more likely lead to my observed birth rank. Just like being tails, doom would more likely to be happening sooner rather than later.

The above answers changes drastically if Self Indication Assumption is used instead. According to which when I wake up in my room I do have relevant information about the coin toss. The information that I exist. Treating myself as randomly selected from the 1000 *possible* clones that might be created, I actually exist is an evidence favouring heads. Because if the coin landed heads all 1000 potential clones would be created. My conditional probability of existence would be 1. But if the coin landed tails only 10 out of the 1000 potential clones would be created. Meaning the probability of my existence would be only 1%. Using Baye's theorem the probability of heads can be calculated to be about 99%. After learning my room number is 7 this probability is reduced. Because I, the (possible) clone from room 7, is guaranteed to be created regardless of the coin toss result. So the the probability of heads is back at 50%. Transferring this logic to the real world, the doomsday argument based on my birth rank is exactly cancelled out by the fact that I exist. Perhaps due to this elegancy it is the leading counter to the doomsday argument. However the Self Indication Assumption has controversies of its own. It suggests I should strongly favour hypotheses with great number of observers simply because I exist. This could lead to absurd arguments such as the case of the Presumptuous Philosopher[7] (Bostrom, 2002). It also suggests simply because I exist the Great Filter is more likely to be ahead of us rather than behind. Meaning human's outlook as a specie should still be more pessimistic but due to a different reason from the Doomsday Argument (Grace, 2010).

Neither Self Sampling Assumption nor Self Indication Assumption stated how was "I" specified. It was treated as if it is obvious enough to be left undefined. But as explained in the first section, first-person self, i.e. primitive identification based on subjective proximity to perception, and third-person identity, i.e. identity based on distinct features among a reference class of observers, are very different concepts. Each identity must only be used in their own perspective without interchanging. Both Self Sampling Assumption and Self Indication Assumption confound the two by arbitrarily switching perspectives in their argument. This mix of different perspectives is why both of them are mistaken.

Just as the first thought experiment, here I can easily specify me in first-person by immediacy to perception. However I cannot specify me in third-person by objective differences relevant to the coin toss. At the time of awakening, from first person perspective I know specifically I, myself, exist. This is not surprising since self existence is a guaranteed observation from first-person perspective. On the other hand from third-person perspective since I cannot specify me the only information available is that an unspecified clone exists, i.e. there exist at least one clone. Again this is not surprising because no matter how the coin lands more than one clone would have been created. There is no new information no matter which perspective we choose to reason from. The

---

[7] In this thought experiment the universe either has a trillion observers or a trillion trillion trillion observers. All observations are indifferent towards both theories. Physicists are about to run a simple experiment to determined which theory is correct. But a philosopher claim such experiment is unnecessary and according to SIA the bigger universe must be correct with a probability of a trillion trillion to one.

odds ought to stay even. When the room number is revealed to be 7, in third-person this simply means a clone was created in room 7. That was certain to happen regardless of heads or tails. From first-person perspective I just learnt the room most immediate to perception is numbered 7, which means my existence does not depend on the coin toss result. So knowing that gives me no evidence about the coin toss. Again from both perspectives there is no new information and the probability stays at 50%. In short, whichever perspective we choose to reason from, as long as its consistent, the odd of coin toss stays at half. Transferring this logic to the real world means neither the fact that I exist nor my birth rank would have any effect on the outlook of human's future.

It may seem natural to ask "what is the probability that I am a clone from room 1 to 10". This question is actually ill-defined. In the question "I" am only specified in first-person by subjective proximity to perception. While "a clone from room 1 to 10" is a third-person identity based on objective difference among a reference class. Therefore the question cannot be fully interpreted from either perspectives and one must switch perspectives in order to comprehend it. So this seemingly innocuous question itself is perceptively inconsistent. The probability of me being a clone from room 1 to 10 is actually invalid therefore do not exist.[8]

Some may argue that I can identify myself in third-person by objective details observed after waking up. This third-person identity could then be used to specify me among all possible clones. This would result in a probability shift. For example, imagine in each room there is a deck of cards. After waking up I can shuffle it to get a random sequence. It can be stated that a clone with this sequence exists. If more clones were created then this statement would more likely to be true since that specific sequence would more likely to be produced by one of the clones. So this should be considered an evidence favouring the scenario with more clones.[9] By this logic the probability of heads should be much higher as suggested by Self-Indication Assumption. However this argument is flawed. That particular sequence was used in that statement only because I shuffled it after waking up. Of course a clone with such an ad hoc sequence exists. Since it is a guaranteed observation no probability update is warranted. Compare this to a different process. Here a particular sequence is first specified before checking any clone's deck. If the same sequences was recreated among all the clones then it is an evidence favouring heads. In this process the recreation of the specified sequence is not guaranteed. More clones leads to a higher chance of success. Therefore a bayesian update should be rightly carried out. The previous argument however made no such pre-commitment to any sequence. It essentially substitute the experiment process without pre-selection with a different process with pre-selection and performed calculation base on that different process. Of course its calculation

_____

[8] Since none of the major school of thoughts in anthropic reasonings endorse this idea. I expect there are going to be many objections regarding this invalidity. In Section 5 I will discuss in more detail of why this invalidity is not as obscure as it may seem.

[9] This example is only focusing on the deck of card for the ease of expression. All details observed after waking up can be used. The argument of using all details observed in probability calculation is called Full Non-indexical Conditioning. (Neal, 2006)

is wrong.[10] Consequently any details I observed after waking up, unless otherwise related to the coin toss, should not change the probability due to the lack of pre-commitment. [11]

Self-Sampling Assumption and Self-Indication Assumption both merge the two perspectives together. Doing so means defining the first-person self by features to enable the switch to third-person perspective. However as discussed earlier first-person self is only defined by subjective proximity to the perception. That's where their mistake lays. It does not matter which exact imaginary selection process was used. As a result they both have a subtle self-contradiction. At one hand all clones are considered equals. On the other hand one of the clones, myself, is inherently special such that no information is needed to specify it. The first part of Self-Sampling Assumption's argument, which states that I have no new information by simply finding myself exist, can be fully contained in first-person reasoning. So that part is perspectively consistent. Therefore it answers correctly that the probability of heads is 1/2. However by calculating the conditional probability of my room number the perspective is switched and first-person self is redefined as a result of a random selection among all existing clones. Consequently it incorrectly increases the probability to 99%. For Self-Indication Assumption the perspective switch happened right from the start. It specifies me in third-person by redefining my first-person identity as a result of a random selection among all possible clones. Using that specification a Bayesian update is performed. Therefore its answer of the probability being 99% is wrong. The reason that it gives the correct answer of 1/2 after knowing the room number is because it performed another perspectively inconsistent calculation which happens to cancel out the first mistake. Due to these reasons neither the Doomsday Argument nor the Presumptuous Philosopher is correct.

3. Sleeping Beauty Problem

Another perhaps more famous paradox caused by perspective inconsistency is the sleeping beauty problem (Elga, 2000). It describes an experiment where you are going to be briefly woke up either once or twice depending the result of a random coin toss. If the coin landed on heads you would be woken up only once on day 1. If tails you would be woken up twice, on day 1 as well as on day 2. At the end of each day your memory of the awakening would be erased. Now supposed you are awakened in the experiment, how confident should you be that the coin landed on heads? How should you change your mind after learning this is the first awakening?

This problem gained a lot of attention over the last decade. Though vigorously debated by many highly capable individuals no agreement has been reached regarding the correct answer. There are three main positions (Titelbaum, 2013). Thirders think the probability of heads is 1/3 at wake up and then 1/2 after learning it is the first awakening(e.g. Elga, 2000; Dorr,2002). Lewisian halfers think the probability of heads is 1/2 at the time of awakening and then 2/3 after learning

---

[10] Arguments confounding the selection process leads to paradoxical conclusions in problems independent of anthropic reasoning, e.g. Boy or Girl paradox.

[11] Another systematic counter for this argument is available in section 3.

this is the first awakening (e.g. Lewis. D, 2001). Double halfers think the probability of heads is 1/2 at the time of awakening and stays at 1/2 after learning it is the first awakening(e.g. Halpern, 2004; Bostrom, 2007; Meacham,2008). Currently the thirder view is the dominating position (Winkler, 2017). I believe double halving is the correct answer. However to my knowledge my reasoning is different from all previous published arguments for this position.

Lewisian halfer's reason can be summarized as follows. After waking up I have no new information about the coin toss. Therefore the probability ought to stay at 1/2. By treating this awakening as randomly selected from all actual awakenings I can calculate the conditional probability of this being the first. If the coin landed heads then this must be the first and only awakening. Whereas if the coin landed tails the probability of this being the first awakening is only half. Therefore knowing this is in fact the first awakening is an evidence favouring heads. A bayesian update can be carried out. As a result the probability of heads shall increase to 2/3. It is essentially the Self-Sampling Assumption applied to this problem.[12]

Thirder's reasoning agrees with the Self-Indication Assumption instead[13]. By its logic after waking up I do have new information[14]. The fact that I am awake *today*. According to Self-Indication Assumption today can be thought as a randomly selected day from day 1 or day 2. So if the coin landed tails then the probability of me being awake today would be 1. Whereas if the coin landed heads, the probability of me being awake today would be 1/2. The fact that I am awake today is evidence favouring tails. The probability for heads should decrease to 1/3. After learning that today is day 1, i.e. the day that I would be awake regardless of the coin toss result, the probability of heads would increase back to 1/2.

Similarities between the sleeping beauty problem and the alien's coin toss from previous section is obvious. The only notable difference is how the duplicated instances are actually created. In the previous problem it is achieved by cloning. Here it is done by memory wipe. This should not result in any change in the way we solve the problem.[15] The memory wipe can be regarded as an information barrier. It prevents any information to pass though to the following day. As a result my experience and perception would not be in anyway affected by what happens on the other day. Therefore from first-person perspective the person from the other awakening is simply not

---

[12] David Lewis did not specifically say an observer should treat himself as randomly selected among all actually existing observers. However his argument does agree with Self-Sampling Assumption's result. (Lewis, 2001)

[13] This is actually an oversimplification consider all the different arguments supporting thirding. But by and large thirders accepts the consequences of Self-Indication Assumption in general.

[14] Some thirder (e.g. Monton, 2002) including Elga (2000) do not think there is new information upon awakening. However many thirders with this position now argues finding oneself awake in the experiment is new information indeed (e.g. Weintraub, 2004; Horgan, 2004; Winkler, 2017)

[15] Many authors from different camps have used examples with physical duplication in their discussion. It is generally accepted that the method of duplication does not affect the underlying problem. However given then amount of literature on this problem there are disagreements about this equivalency as expected. Readers interested in such arguments can check out Kierland and Monton (2005).

me[16]. As a result the exact same first-person reasoning from the previous question can be applied here. Of course from third-person perspective I am usually regarded as the same individual in the two awakenings. However from third-person perspective the identity was never relevant to the question anyway. What is important is the indexical position of the awakening, i.e. whether it is day 1 or day 2. Due to the experiment setup no information relevant to the coin toss is available to specify today from the third-person perspective. Just as in the alien's coin toss problem the third-person identity could not be specified due to high accuracy cloning. Therefore the same third-person reasoning can be used as well. As a result the following solution would be very similar in format compare to the previous section.

One thing halfers and thirders have in common is that neither mentioned how is *today* or *this awakening* specified. The concept of *now* or *this moment* are meaningful due to their immediacy to perception i.e. their *thisness* (haecceity) . Therefore *today* or *this awakening* are only specified from first-person perspective. However due to the experiment setup no objective difference relevant to the coin toss is available to distinguish a day. So they cannot be specified in third-person.[17] As a result at the time of waking up I can reason in two ways. From first-person perspective I know I am awake today. Finding myself exist or conscious is a guaranteed observation from first-person perspective. Therefore it offers no new evidence. From third-person perspective the only information available is that beauty is awake on an unspecific day, or there is at least one awakening. This is not new information because regardless of the coin toss result there is at least one guaranteed awakening. The probability of the coin toss ought to stay even. After learning that this is the first awakening it can still be reasoned in two ways. From third-person perspective it is already known that there would be an awakening on day 1. So no new information is gained about the coin toss. From first-person perspective I learnt that my awakening does not depend on the coin toss.[18] However this doesn't tell me anything about how the coin fell. So the probability should still remains at 1/2. No matter which perspective was chosen, as long as it is consistent, the answer would be 1/2 at awakening and remains at 1/2 after learning today is day 1.

---

[16] Since first-person self is defined by subjective immediacy to perception it inherently identifies self at the present moment. This means the center of first-person perspective is a primitively identified person-stage instead of a persistent person. The notion of subjective persistent person has to be a logical posterior of this person-stage. Liao (2012) proposed the same idea while discussing centered worlds. Therefore another way to put it would be my subjective persistent self does not include the other awakening.

[17] It may seem words such as *today* or *now* are communicable so it should also be meaningful from third-person perspective. However this sense of communicability is misguided. It seems to be understandable if the message sender and receiver are temporally close or the time delay of communication is known such that each person's own sense of *now* can be used to approximate the another's, i.e. one's *de se* thought matters for the comprehension of the moment. Imagine finding a message in a bottle on the beach. If it says "a storm is brewing now…". It would give you no indication of what time it is referring to. Even though it is obvious the message sender clearly wrote *now* and knew exactly when he meant by it. This shows concepts such as *now* or *this moment* are incommunicable just as first-person self identity.

[18] Remember from first-person perspective I only regard whoever awake today as myself. Beauty awake on the other day, if there is another awakening, is primitively not me.

Same as before, the questions of "what is the probability that today is day 1?" or "what is the probability that this is the first awakening?" are invalid questions which have no answer. It is because these questions cannot be fully interpreted from either perspectives. From first-person perspective "today" or "this awakening" can be specified base on its immediacy to perception. However from first-person perspective this is my one and only awakening. On the other hand the sequence of the awakenings, and day 1 or 2, is meaningful from a third-person perspective. However from third-person perspective "today" or "this awakening" cannot be meaningfully specified. So these seemingly innocuous questions are actually incomprehensible either way thus perspectively inconsistent. Consequently they have no valid answer.[19]

As expected there are also arguments that tries to specify *this awakening* or *today* in third-person by details observed after waking up. In my opinion the best formulated one is the Technicolor Beauty (Titelbaum, 2008) where an insignificant detail was added to the original problem[20]. The experimenter would randomly choose one of the two days and put a piece of red paper on the wall, on the other day he would put out a piece of blue paper. The argument can be summarized as follows. When I wake up I would see a piece of paper. Say that paper is blue. Then it can be stated that beauty is awake on the blue day. Which is new information since I could not know that just from the experiment setup. Because the probability of beauty being awake on the blue day is one conditional on tails comparing to half conditional on heads this is an evidence favouring tails. As a result the probability of heads could be calculated by bayesian updating to be 1/3. If I see the red paper the same argument can still be applied due to symmetry. Therefore even before I open my eyes and see the colour of the paper I should already conclude that the probability of heads must be 1/3. Instead of repeating my previous argument here I am going to use a counter example. Imagine you are on an exotic island where all families have two children. The island is having their traditional festival of boys' day. On this day it is their custom for each family with a boy to raise a flag next to their door. Tradition also dictates in case someone knocks on the door then only a boy can answer. You notice about 3/4 of the families has raised a flag as expected. It should be obvious that if you randomly choose a family with a flag then the probability of that family having two boys is 1/3. You also know if you knock on the door a boy would come to answer it so by seeing him there is no new information. But not so fast. When you see the boy you can ask him "are you the older or the younger child?". Say he is the older one. Then it can be stated that the older child of the family is a boy. This is new information since I could not know that just by seeing the flag. If both children are boys then the conditional probability of the older child being a boy is one. If only one child is a boy then the conditional probability of the older child being a boy is only half. Therefore this new evidence favours the former. As a result the probability of this family having 2 boys can be calculated by bayesian updating to be 1/2. If the child is the younger kid the same reasoning can still be applied due to symmetry. Therefore even before asking I should already conclude the randomly chosen family's

---

[19] Once again I expect many objections about this invalidity since both halfers and thirders readily calculate these probabilities. More discussion of this is presented in Section 5.

[20] Similar arguments have also been proposed by Rosenthal's (2008) Sleeping Peon Problem. As mentioned before these arguments are closely related to Full Non-Indexical Conditioning (Neal, 2006)

probability of having 2 boys is 1/2 instead of 1/3. This is absurd. Given their resemblance it is quite clear technicolor beauty's argument is not valid.

Imagine before knocking on the family's door you make a pre-commitment among one of the two children. Say you randomly decide to see the older child specifically. If the older child comes out and it turns out to be a boy then it would be correct to increase the probability of two boys from 1/3 to 1/2. In this process it is entirely possible that you would see a girl. Therefore seeing a boy is indeed evidence favouring the family having more boys. But that's not the process of the exotic island problem where no such pre-commitment is made. In its process you would always meet a boy at the door. When asked if he is the older or younger child ,whatever the boy's answer is, obviously it is guaranteed to be a boy. You are looking right at him. For this process the probability of two boys must remain at 1/3. The mistake happened because the argument used a different process in its calculation even though that process also lead to the observation of "the older child is a boy".[21] With this in mind the mistake of technicolor beauty's argument is apparent. In the experiment there is no pre-commitment made to any color. That's why no matter which color beauty sees the same argument can be applied due to symmetry. However its calculation is based on a process with pre-commitment. For example the experimenter could randomly decide to check the room on the blue day specifically. If he saw beauty awake on that day then the probability of tails should be increased to 2/3. Obviously that is a completely different process from beauty simply opening her eyes and check the paper. Even though both process can lead to the observation of "beauty is awake on the blue day".[22] As a result technicolor beauty's calculation is invalid the same way as the exotic island problem does. Some may argue checking the color on *today* specifically is a pre-commitment among the two days thus its calculation is valid. There are two problems for this argument. First it completely negates any relevance of the color or any other details observed after waking up. Same calculation can be carried out base on the simple fact that beauty is awake *today*. Secondly the argument essentially becomes Self-Indication Assumption itself. Since it is well established that Self-Indication Assumption would lead to thirding it means this technicolor example makes no contribution to thirders' case.

Many authors have used arguments involving operationalization to present their case. It often involves repeating the experiment and counting the relative frequency in certain way (e.g. Groisman, 2008) . Sometimes it involves a monetary reward (e.g. Arntzenius, 2002). Others use elaborate Dutch Book arguments by setting up various kind of bets to show what the correct probability should be (e.g. Hitchcock, 2004; Draper & Pust, 2007). These argument are ineffective at convincing the opposition because people with different opinions would simply

---

[21] The possible outcomes of the process with pre-commitment are : the older child is a boy vs the older child is a girl. The possible outcomes of the exotic island problem without pre-commitment are: the older child is a boy vs the younger child is a boy. Therefore they convey different information even with the same observation.

[22] Like the previous example. The possible outcomes of the process with pre-commitment are: beauty is awake on the blue day vs beauty is asleep on the blue day. Whereas the possible outcomes of beauty simply opening her eyes without pre-commitment are : beauty is awake on the blue day vs beauty is awake on the red day. Therefore they convey different information even with the same observation.

disagree on the bookkeeping rule and/or beauty's objective. Nonetheless it is still interesting to see what should the operationalization looks like when perspective reasoning is considered. Remember due to the memory wipe whatever happens on the other day has no effect on the perception so from first-person perspective it is not myself. Hence from beauty's first-person perspective any repetition of the experiment has to happen without passing though a memory wipe. Since a memory wipe would happen at the end of the day the next repetition will need to fit within the time constrain. So the next experiment must be shorter, e.g. the first awakening in the morning and the second possible awakening in the afternoon. Further repetitions would be even shorter still to fit in the time slot the previous experiment has left. From an outside observer's third-person perspective every awakening should be treated the same. As a result the repetitions would not be chronologically sequential but nested with multiple levels instead, i.e. using memory wipes, the first level separate the duration of the experiment into two halves, the next level separate each resulting segment into two halves again, so on and so forth. It would be in a form similar to a supertask. Even if the experiment is repeated infinite number of times it would still have to happen within two days.[23] This may seems a little tedious. It is because in this problem the duplicated instances are created by a memory wipe at a predetermined time. If I am physically duplicated, e.g. by memory retained cloning, then the process would be much clearer. After going through a random cloning procedure, a repetition from my first-person perspective obviously means going through the random cloning procedure again. It would in no way involve killing me off which is the equivalent of letting the memory wipe happen in sleeping beauty's case. Other clones, even though complete irrelevant from my perspective, should go through the same procedure. Therefore from the perspective of an outside observer repetitions means the clones would keep on multiplying with compound duplication[24] just as the nested mode in sleeping beauty problem does. From first-person perspective a subject should only be interested in coin tosses that directly involves her subjective persistent self. After all if I see myself as inherently special then coin tosses involving me must also be uniquely different from others so there is no point considering other's coin tosses. If the experiment is repeated near infinite number of times then beauty would personally count the coin lands heads about half of the time, meaning the frequentist and bayesian approach gives the same answer. Obviously from an observer's third-persons perspective all the coin tosses for all the clones need to be considered. This is uncomplicated since outside observers are never involved in memory wipes or cloning. Clearly about half of all these tosses are landing on heads. Meaning frequentist and bayesian probability gives the same answer from third-person perspective as well.

As mention above, another controversy closely related to the frequentist argument is the use of bets. After each awakening beauty can be offered a bet which she can decide the fair price to

---

[23] Two days from the outside observer's third-person perspective to be precise.

[24] Borrowing from the concept of compound interest. It simply means resulting duplicates from previous experiments would experience further duplications in the future repetitions.

enter.[25] That fair price should in theory reflect the correct probability. However in sleeping beauty's case it turns out to be more complicated than that. With different motives beauty's decision actually reflects different probabilities. For example, if beauty is trying to maximize the total money belonging to all copies she should make her decision as if the probability of heads is 1/3. Whereas if she is trying to maximize the average money of each copy then she should make her decision as if the probability is 1/2. Even if beauty is simply selfishly trying to maximize her own personal gain the decision is still ambivalent depending on beauty's position on issues such as counterfactual identity and self non-existence (Armstrong, 2011). With these in mind it would be interesting to see how perspective reasoning links beauty's decision to probability. First of all for beauty's decision to properly reflect probability the bets' cost and rewards shall follow her first-person perspective. This means when duplication occurs the subject's money/debt shall duplicate with her. If the duplication is achieved by memory wipe as in sleeping beauty's case then whenever a period's memory is lost the accrued gain or loss of that corresponding period shall be expunged as well. This would ensure indistinguishable copies are in identical financial states. Furthermore this also makes beauty's cumulative gain or loss to be in agreement with her first-person experience if the experiment is repeated. Secondly for her decision to reflect probability the proper motive for beauty should be simple self-interest. While various forms of altruism towards copies are also valid objectives they are justified by intricate intentions such as ethics or evolutionary psychology. These considerations adds an another layer of complexity which could cause beauty's decision to deviate from probabilities. Whereas self-interest is primitive so the link between probabilities and decision is direct. It might be helpful to think as if each duplicate has her own account which is settled independently. With the above two points beauty's rational decision can be readily derived.[26] The fair price for beauty to enter a bet that pays 1 dollar in case of heads shall be 50 cents. This can be easily verified using the previously mentioned nested repetition model. Some may detest the idea of duplicating money and suggest gain or loss from the bets should remain unchanged. Others might think all clones from the experiment ought to be treated equally when making decisions. These ideas seems reasonable because they are justified, but only from a third-person perspective. To be consistent with the third-person perspective such ideas shall be in effect if the decision maker, i.e. the person entering the bets, is an outside observer. Obviously for such a decision maker it makes no sense to duplicate his money nor give preferential treatment to any specific clone. It is also trivial for the observer, i.e. someone who does not experience memory wipe or duplication, that the fair price for a bet on coin toss should be half the payout. This shows once again the probability of heads should be half from both first- and third-person perspectives.

---

[25] Halfers may disagree and argue only one bet should be offered per coin toss, e.g. a bet is offered on Monday only whereas the bet on Tuesday is just a placeholder or dummy with no actual cash exchange. However this is argument is not convincing. Firstly it treats two otherwise similar awakenings differently without a clear reason. Secondly it does not translate very well to cases with physical duplication. For example the duplication process could be the original person is split into two halves and each half can be connected with a artificially constructed counterpart to form a new body. In this process the resulting copies are symmetrical and it is not clear which one should be offered the dummy.

[26] Complex issues such as self-nonexistence and counterfactual identity are no longer relevant if perspective reasoning is used. E.g. from first-person perspective self nonexistence is an irrelevant concept. People in counterfactual worlds are isolated from perception so from first-person perspective none would be regarded as self.

4. Perspective Disagreements

As mentioned earlier perspective reasoning suggests in some cases two people in direct communication should disagree with each other. It has been discovered that any halfer position would lead to such a disagreement (Pittard, 2015). This deserves our attention because it seems counterintuitive. It shall be pointed out that Self-Indication Assumption, or thirder's reasoning in regards to the Sleeping Beauty Problem, would not result in such disagreements. This has been used as an evidence favouring Self-Indication Assumption(Grace, 2011). However here I would argue disagreements of this kind are valid. They should indeed have different probabilities to the same proposition.

Let's consider the following experiment. An alien has abducted you and one of your friends. You are put to sleep. The alien then tosses a fair coin. If it lands on heads it won't do anything to you. If it lands on tails it will clone you and put the clone into another identical room. The clone process is highly accurate so that the memory is retained. As a result the clone, as well as the original, can not tell if he is old or new. Meanwhile your friend never goes though any cloning process. After waking you up the alien let your friend choose one of the two rooms to enter. Suppose your friend has chosen your room. As a result you guys meet each other inside. How should you reason about the probability of the coin toss? How should your friend reason it?

For my friend the question is non-anthropic thus very simple. If the coin landed heads then 1 out of the 2 rooms would be empty. If the coin landed tails then both rooms would be occupied. Because the room that she randomly chose is occupied she now has new evidence favouring tails. As a result the probability of heads can be calculate by a simply bayesian update to be 1/3.

For me the question is not too complicated either.[27] After waking up I have no new evidence about the fair coin toss. So I ought to believe the probability of heads is 1/2. Because my friend is randomly choosing between two rooms, regardless of the coin toss result the probability of my room being chosen is always half. Therefore seeing my friend gives me no new information about the coin toss either. This means I should keep believing that the probability of heads to be 1/2.

Here the disagreement is apparent. Even though the two of us appear to have the same information about the coin toss we assign different probability to the same proposition. To make the matter more interesting nothing I could say would change her mind and vice versa. We can communicate however we like but nobody is going to revise their answer. This may seem strange but it is completely justified. The cause of this disagreement is our different interpretations of who is exactly in this meeting. Remember according to my friends' reasoning the evidence that causes the probability update is "the chosen room is occupied." The occupant, in case there are duplicates, could not be specified from her perspective. In other words, as long as there is

---

[27] Assuming I take the position of this paper.

someone in the room she will reason as such. This is expected since the cloning procedure is highly accurate so there is no objective feature relevant to the coin toss to differentiate each duplicate. However from my first-person perspective I can inherently specify the one whom she is in meeting with. It is me, myself. But this specification is only valid from first-perspective. Remember first-person identity is based on the immediacy to perception. Because perception is imperceivable by others this means my specification is incommunicable. I can keep telling her "this is me" and it would not mean anything to her. As a result the two of us would keep our own answers and remain in disagreement.

The above paragraph explained why the disagreement is valid in the sense of subjective probability. This disagreement is also valid with a frequentist interpretation, which in my opinion is also easier to understand. The experiment can be repeated many times and the relative frequency can be used to show the correct probability. From my perspective repeating the experiment simply involves me going back to sleep, and wake up again after a coin toss and the potential cloning process. Of course after waking up I may not be the same physical human being just as the case of the first experiment. But this does not matter because in first-person perspective I am defined primitively base on subjective identity instead of objective features or qualities. So I would always regard the one falling asleep on the previous night as part of my subjective persistent self since it was the center of my perception and my current subjective experience is a continuation of it. To make the procedure easier suppose I can check the previous coin toss result before falling back to sleep again. So each iteration can be summarize as to go to sleep, wake up, and check the coin. Imagine repeating this iteration 1000 times by my count. I would have experienced about 500 heads and tails each. Furthermore if my friend is involved then I would see her about 500 times with about equal number of occurrences after heads or tails. However for these 1000 coin tosses my friend would see an occupied room about 750 times. The extra 250 times would be due to seeing the other duplicate instead of me after tails. It is easy to see our relative frequency of heads with a meeting are indeed different, half for me, a third for her. Of course my friend should be involved in far more repetitions than I do since every duplication of me are indifferent from her perspective so she shall be involved with their repetitions as well. However her relative frequency would remain unchanged by the higher number of iterations.

One more thing to point out is that this disagreement is not because I am better at logical reasoning than my friend or the other way around. Neither is because anyone of us is in a epistemically more advantageous position. In fact I completely agree with her reasoning just as she agrees with mine. It is simply due to the difference in perspectives that caused my specification to be incommunicable and meaningless to her. In another word, we agree to disagree.

5. Invalid Probabilities

One topic of interest in anthropic reasonings is the probability distribution of oneself's indexical positions. For the Sleeping Beauty Problem this includes questions such as "what is the

probability that *this* is the first awakening?" or "what is the probability that *today* is day 1?". For the Doomsday argument this includes question of "what is the probability that *I* am among the first 100 billion human ever existed?". Concepts such as a self-evident *I, today* and *this present awakening* are only specified in first-person perspective whereas the notions of *all human beings*, different *days* and *awakenings* belonging in the same reference class is only valid in third-person reasoning. Hence these questions are perspectively inconsistent thus invalid, meaning such probabilities do not exist. To my knowledge this idea has not been proposed by any published paper. This section intends to discuss this idea in detail.

To remove other variables consider this uncomplicated experiment: you are taking part in a cloning experiment. At the end of the day, while sleeping, you would be cloned. The process is highly accurate such that it is humanly impossible to tell the resulting duplicates apart. All features among the two, including memories, would be indistinguishable. Note this cloning process is not contingent on any random event, it is guaranteed to happen. Also for the sake of differentiation the experimenters would randomly choose one of the two duplicates and put a piece of red paper in his room. A piece of blue paper would be put into the other room. Now suppose you wake up the following morning, it may seems natural to ask: "What is the probability that I am the original?"

This uncomplicated experiment can be seen as a prototypical problem about probability of indexical positions. In this experiment the cloning process is a certainty. Thus the distinction of *actually* existing observers versus *potentially* existing observers is otiose. Supporters of Self-Sampling Assumption and Self-Indication Assumption would give the same answer. It is also analogous to the Sleeping Beauty Problem when beauty is informed that the coin landed tails. When asked about which day it is Elga (2000) suggested that "even a highly restricted principle of indifference" would assign equal probability of today being day 1 or day 2. This has been accepted by halfers and thirders alike. Virtually all school of thoughts would give the answer as 1/2 based on similar reasons. The generality of the problem is apparent. Therefore arguments regarding this experiment need not to target any halfers or thirders specifically.

To show the question is perspectively inconsistent it is necessary to show that neither third- or first-person reasoning can meaningfully interpret the problem and provide an answer. Let's consider third-person perspective first. An integral part of the question is pin pointing which exact duplicate's indexical position is being asked. This is crucial since the truth value of a proposition about someone being the original depends on which particular duplicate it is referring to. This specification however proves to be an insurmountable challenge. As mentioned above taking a third-person perspective is to reason as an outside observer whose *de se* thoughts are unrelated to the topic of interest. From this perspective the duplicates are in equal positions. Consequently there is no inherently identifiable *me* among the the two. Therefore the question is incomprehensible. The next logical step would be to specify a duplicate by some objective features meaningful in third-person. For example I can check the color of my paper. Say I see

red. Then it can be asked "What is the probability that the red duplicate is the original?[28]". While this is a valid question, it is not equivalent to the original one[29]. First of all, the indexical question of "What is the probability that I am the original?" is defined the moment I become conscious[30]. However the new question is only defined after seeing the color of my paper as red. Secondly the experiment can be repeated such that each duplicate would undergo further fission. After waking up from another experiment the indexical question can be asked again without any modification. However due to the random selection process it is entirely possible for me to see a blue paper in a different iteration. Here the supposed equivalency between *me* and *the red duplicate* obviously breaks down. In order to save this equivalency it might be tempting to modify the new question as "what is the probability that the whatever-colour-I-see duplicate is the original?". But then the problem of defining *I* in third-person returns. This shows the indexical question is incomprehensible from third-person perspective and there is no equivalent question to substitute it.

A foreseeable counter against the above argument is that specifying a particular clone does not matter at all. After all regardless of which clone the clone is referring to, e.g. be it the red or the blue, the probability of him being the original is the same: 1/2. Therefore the answer to "the probability of me being the original" must also be 1/2. Discussion about how *I* am specified in third-person is nothing but pointless nitpicking. This argument is simply jumping the gun. Calculating the probability about a particular duplicate is part of a question's solution. Whereas specifying one of the duplicates is part of the question itself. A question is logically prior to its solution. The fact that different duplicates give the same numerical answer cannot be used to nullify the importance of specification. I imagine even the validity supporters would consider "the probability of me being the original" and "the probability of my counterpart being the original" to be completely different questions. Even though by whose logic their value would be the same, i.e. 1/2. This again shows specification is a critical part of the indexical probability which cannot be overlooked just because it does not affect the numerical answer.

We can also try to solve the problem from the first-person perspective. As previously discussed self is the center of first-person perspective. This mean I am uniquely different from anybody else which is identified by subjective immediacy to perception alone. This identity is primitive such that no other information is needed. Comparing to third-person reasoning first-person perspective have the unique advantage of knowing which is *me* inherently. So it does not have the problem of specification. However treating the two duplicates as equals and differentiating them basing on features such as the clone vs the original is third-person reasoning. Consequently first-person perspective could not interpret the two possibilities of the indexical question. This means first-person perspective could not comprehend the question either. There are arguments

---

[28] Same as before here I am only focusing on the color for the easy of expression. Any details observed after waking up can be used for specification.

[29] It is worth noting that the new question is not anthropic related. Also because the color is determined by a random selection among all duplicates it would not lead to paradoxes similar to the Doomsday Argument.

[30] That is, aside from the invalidity proposed here.

successfully assigning a value to the probability. However they inevitably switched to third-person reasoning at some point. A very commonly used argument is as follows: since half of all observers in question are original then the probability of me being the original must also be 1/2.[31] Sometimes this logic is presented in a more visual way: imagine all duplicates guesses he is the original then only half of them (one out of two) would be correct, suggesting the probability of me being the original must be half as well. Essentially this type of reasoning takes the average of all observers in question and treat that value as my indexical probability. While it is prima facie reasonable there is an underlying assumption involved. It assumes *I* am an ordinary member of all observers who belongs to the same reference class. This is in direct conflict with first-person reasoning where *I* am inherently unique. The notion that all duplicates are in the same reference class is only valid from third-person perspective.[32] Therefore the average-among-all arguments are not perspectively consistent.

Another foreseeable argument is that indexical possibilities are just not special at all. Before opening my eyes I can easily tell the probability of me seeing a piece of red paper is 1/2. Likewise the probability of me being the original should also be 1/2. This argument ignores the differences between seeing a piece of paper and being a certain individual. Which paper being most immediate to my perception is determined by a random experiment. Red and blue are just the two potential outcomes. These possibilities are comprehensible from the first-person perspective. Whereas the original and clone are third-person identities. There is no random experiment determining how *I* become one of the two. Of course one can imagine a experiment such that after the cloning procedure *I* get to be randomly incarnated into one of the bodies. With this experiment the indexical probability can be assigned just as the probability of seeing a certain colored paper. One obvious counter is that the experiment is entirely imaginary. However considering widely discussed Self-Sampling Assumption and Self-Indication Assumption are also based on imaginary processes this may not be seen as a serious shortfall by some. The definitive problem of this argument is its subtle switch to third-person perspective. Recall first-person self is primitively defined by subjective immediacy to perception. However the imaginary incarnation procedure happens before I am awake and conscious. It prevents first-person self identification.[33] Meaning the *I* in the argument has to be defined by some feature instead. Defining observers by features is third-person reasoning. Interestingly due to the nature of the imaginary experiment this feature can not be dependent on the physical body, meaning it must be something similar to the concept of a soul. Then which particular soul is the question referring to

---

[31] Nick Bostrom(2003) used this idea and called it a Bland Indifference Principle. In which he also adds the condition that there is have no evidence of me being more or less likely to be the original. But this condition already assumes the likelihood of me being the original as a valid notion so it is not presented here in this discussion.

[32] In another word endorsing the idea that I am just an ordinary duplicate means I am not inherently special so it wouldn't be possible to directly specify myself in the indexical probability.

[33] Of course one can argue during the imagine experiment I am conscious just that all memory of it is lost once I get embodied. But then from first-person perspective the person before incarnation is no longer perceptually immediate. So my subjective persistent self would not include any person-stage in the imaginary experiment. To have a unified identity during and after embodiment third-person definition has to be used.

would be critical. This brings back the specification problem mentioned earlier. Obviously the incarnation argument mixes two different perspectives as well.

The above shows indexical probabilities necessarily mixes reasoning from both perspectives, therefore invalid. Even without considering perspective consistency there are other problems with such probabilities. The indifference principle accepted by halfers and thirders alike is high road to skepticism (Schwarz, 2011). For example, provided one knows there are many brains in vats that emulate his own subjective experience then he ought to be very confident that he is one of those brains (Elga, 2004). Bostrom (2003) suggests someone who is confident that human would eventually be able to run ancestor-simulations must also believe we are living in a simulation now[34]. These arguments are very hard to take. Another problem is regarding the logic of assigning indexical probabilities. Marcoci (2018) showed that Elga's (2004) "highly restricted principle of indifference" would lead to the decision of not switching in the Monty Hall Problem. With these hard to explain problems and the paradoxes about anthropic reasoning it is safe to say the indexical probabilities' invalidity is not as absurd as it may first seem. Questions regarding these probabilities are perspectively inconsistent fallacies with no answer.

6. Discussions

In summary, my position is that reasonings from different perspectives should not mix such that arguments should be perspectively consistent. From this starting point the answer to the sleeping beauty problem can be shown to be double halving. Once the importance of perspective consistency is recognized various controversies surrounding the problem can be, in my opinion, thoroughly explained.

Comparing to other arguments supporting the double halving position perspective reasoning have several advantages. Earlier arguments all readily assign probabilities to perspectively inconsistent events such as "today being day 1". This cause the probability of heads to decrease after learning it is day 1 under extant Bayesian updating rules. Various new updating rules about self-locating information have been suggested to keep the probability unchanged at 1/2. However none seems to be satisfactory. Halpern's (2004) method, Meacham's (2008) compartmentalized conditonalization and Briggs' (2010) halfer rule all suggest beauty's probability should be 1/2 if the experience from the two awakening are strictly identical. However if there are even tiny differences in objective details then the answer should change (Titelbaum, 2013). e.g. in the case of Technicolor Beauty these methods suggest the probability should be 1/3 after seeing the

---

[34] Bostrom (2003) also suggested a third possibility that future humans could be technologically able to run such simulations but decides not to. Personally this alternative seems rather weak since the decision has to hold for every single individual over entire civilization history after such technology is available. Because even a few violators could run a great number of simulations and decisively alter the probability.

paper[35]. Such positions are very questionable since the paper's color is strictly independent of the coin toss. Furthermore older argument suffers the same draw back as the Lewisian halfers does. They would predict a fair coin toss yet to happen will land on heads with a probability greater than 1/2 (Titelbaum, 2012). To show this the original coin toss is moved to day 1 after the first awakening[36]. Another coin tosses is added on day 2 later than the potential second awakening. The second coin toss is completely inconsequential, i.e. nothing depends on how the coin lands. Then it can be asked "what is the probability that *today's* coin will land on heads?". With perspective reasoning this question is trivial. The question is only meaningful from first-first-person perspective since it uses the concept of *today*. It is obvious that the probability of heads is 1/2 for this specific coin toss yet to happen. However the earlier arguments would reach a different solution. Earlier theories would treat the probability of "today's coin landing on heads" as the sum of the probability that "today is day 1 and the first coin is landing on heads" and "today is day 2 and the second coin is landing on heads". Halfer position dictates the probability of "today is day 1 and the first coin is landing heads" should be 1/2. Therefore as long as the probability of "today is day 2 and the second coin is landing on heads" is not 0, which is hard to argue otherwise[37], their answer must be greater than 1/2 (Titelbaum, 2012). More importantly earlier double halving arguments do not provide a frequentist model nor any justification for the perspective disagreements discusses in previous sections. These should be enough to show that perspective consistency is the superior explanation for double halfer position.

Readers familiar with the David Lewis' (1979) centered and uncentered worlds may notice there are connections between these concepts and perspective reasonings. While trying to find a universal object for attitude David Lewis showed that uncentered propositions can be subsumed under centered propositions[38]. Perhaps partially due to this conclusion uncentered probabilities is often regarded as a type of centered probabilities. The discussion of sleeping beauty problem, and anthropic reasonings in general, can be seen as the debate of how to incorporate centered evidence while calculating the probability of an uncentered event. Obviously first-person perspective is centered reasoning. On the other hand, third-person perspective is to think from the viewpoint of someone whose *de se* thought is irrelevant which is analytically speaking

---

[35] The counter argument demonstrated in section 3 would not apply here. Because it actually requires differentiating first- and third-person perspectives to work. Since the process with pre-commitment is done by an outside observer while the process without pre-commitment is done by beauty herself. Without a clear distinction between the perspectives their difference becomes elusive. Furthermore without the justification for perspective disagreement it would be impossible to argue why beauty's probability can be different from the outsider's 1/3 since they could be in direct communication.

[36] Elga (2000) suggested this change in the original paper and argued it would still be the same experiment. To my knowledge no one has challenged it.

[37] There are rare arguments that suggests when beauty wakes up she should be absolutely certain it is day 1 even before being told so (Hawley, 2011). Traditionally they are regarded as quite implausible since it assign zero credence to a possibility (Titelbaum, 2013). However under perspective reasoning it is easy to see there are merits in those arguments since from beauty's first-person perspective this awakening is definitively her first.

[38] In his original paper (David Lewis, 1979) he used the term property for centered propositions and reserved the term proposition to indicate uncentered propositions only.

uncentered reasoning[39]. Therefore the perspective consistency argument proposed in this paper can also be seen as the argument against mixing centered and uncentered reasonings. Coincidentally the argument is not new if presented this way. Peter Lewis (2010) and Luna (2018) both claimed it is impossible to merge centered and uncentered probabilities and any attempt in mixing worlds of different kinds in the same logical framework would lead one astray. However they both concluded the centered and uncentered probabilities are different, i.e. the probability of "the coin landed on heads" is 1/2 but the probability of "this is a heads awakening" is 1/3. This is very hard to swallow because they have the same truth value, i.e. one proposition is true if and only if the other is true. Interestingly enough the root cause of this supposed probability discrepancy is precisely the mix of centered and uncentered reasoning. Luna (2018) used a frequentist approach to calculate the centered probability. As discussed in section 3, according to first-person perceptive centered reasoning the repetitions shall take a nested structure. Only awakenings beauty personally experienced[40] are relevant. As a result the relative frequency as the number of repetitions approaches infinity shall be 1/2. However Luna performed an uncentered calculation of relative frequency of awakenings. Here the repetition structure is irrelevant and all awakenings from different duplicates are used in the calculation. Consequently the probability was mistakenly deemed as 1/3. Peter Lewis (2010) used Dutch Book arguments to show the centered probability to be 1/3. In these arguments bets from different awakening are pooled together and settled in one account, which is valid only for outsider observers. For beauty's first-person centered reasoning the other duplicate, or beauty from another day, is primitively not herself. Therefore each one shall have a duplicated account that settles independently such that bets from two awakenings are never pooled together. This way it is easy to see beauty would avoid getting Dutch booked only if she keep her credence unchanged at 1/2 after waking up. These demonstrates properly separating centered and uncentered reasoning would not result in the counterintuitive conclusion of probability discrepancy. The correct probability should be 1/2 for both cases. Once again perspective consistency is the key in resolving the paradox.

In conclusion perspective reasoning is crucial in anthropic related problems. With it in mind the correct answer to the Sleeping Beauty Problem can be shown to be double halving. It also disproves the Doomsday Argument and the Presumptuous Philosopher problem. In my opinion perspective reasoning definitely deserves more attention as it could be the key to other paradoxical arguments.

References

Armstrong, Stuart. "Anthropic Decision Theory for Self-Locating Beliefs". arXiv preprint arXiv: 1110.6437, 2011 - arxiv.org

---

[39] For readers treating third-person perspective as objective thinking this equivalency to uncentered reasoning is even more direct.

[40] In other words, tosses happened to her subjective persistent self.

Arntzenius, Frank. "Reflections on Sleeping Beauty." *Analysis*, vol. 62, no. 273, 2002, pp. 53–62., doi:10.1111/1467-8284.00330.

Bostrom, Nick. "Anthropic Bias: Observation Selection Effects in Science and Philosophy", 2002, New York: Routledge

Bostrom, Nick. "Sleeping beauty and self-location: A hybrid model". *Synthese*, 157(1), 59–78. 2007

Bostrom, Nick. "Are We Living in a Computer Simulation?" *The Philosophical Quarterly*, vol. 53, no. 211, 2003, pp. 243–255., doi:10.1111/1467-9213.00309.

Brandon Carter and W. H. McCrea. "The anthropic principle and its implications for biological evolution." *Philosophical Transactions of the Royal Society of London*, 310(l512): 347363, 1983. do:10.1098/rsta.1983.0096.

Dieks, Dennis. "Doomsday--Or: The Dangers of Statistics." *The Philosophical Quarterly*, vol. 42, no. 166, 1992, p. 78., doi:10.2307/2220450.

Dorr, C. "Sleeping Beauty: In defence of Elga". *Analysis*, 62(4), 292–296. 2002

Draper, Kai, and Joel Pust. "Diachronic Dutch Books and Sleeping Beauty." *Synthese*, vol. 164, no. 2, 2007, pp. 281–287., doi:10.1007/s11229-007-9226-1.

Elga, Adam. "Self-locating belief and the Sleeping Beauty problem". *Analysis* 60: 143–47. 2000

Elga, Adam. "Defeating Dr. Evil with Self-Locating Belief." *Philosophy and Phenomenological Research*, vol. 69, no. 2, 2004, pp. 383–396., doi:10.1111/j.1933-1592.2004.tb00400.x.

Frege, Gottlob. "The Thought: A Logical Inquiry" Original paper (1918-1919) *Mind, New Series*, Vol. 65, No. 259. (Jul., 1956), pp. 289-311.

Grace, Katja. "Anthropic Reasoning in the Great Filter", 2010. Retrieved from http://www.academia.edu

Grace, Katja. "Agreement on Anthropics." *Meteuphoric*, 9 Mar. 2011, meteuphoric.com/2011/03/09/agreement-on-anthropics/.

Groisman, Berry. "The end of Sleeping Beauty's nightmare". *The British Journal for the Philosophy of Science*, 2008

Halpern, J. "Sleeping Beauty reconsidered: Conditioning and reflection in asynchronous systems". In Proceedings of the Twentieth conference on uncertainty in AI, pp. 226–234. 2004

Hitchcock, Christopher. "Beauty and the Bets." *Synthese*, vol. 139, no. 3, 2004, pp. 405–420., doi:10.1023/b:synt.0000024889.29125.c0.

Horgan, T. "Sleeping Beauty Awakened: New Odds at the Dawn of the New Day." Analysis, vol. 64, no. 1, 2004, pp. 10–21., doi:10.1093/analys/64.1.10.

Kierland, Brian, and Bradley Monton. "Minimizing Inaccuracy for Self-Locating Beliefs." *Philosophy and Phenomenological Research*, vol. 70, no. 2, 2005, pp. 384–395., doi:10.1111/j.1933-1592.2005.tb00533.x.

Lewis, David. "Sleeping Beauty: Reply to Elga". *Analysis*, 61(3), 171–176. 2001

Liao, Shen-Yi. "What Are Centered Worlds?" *The Philosophical Quarterly*, vol. 62, no. 247, 2012, pp. 294–316., doi:10.1111/j.1467-9213.2011.00042.x.

Marcoci, Alexandru) "Monty Hall saves Dr. Evil: on Elga's restricted principle of indifference". *Erkenntnis*. 2018. ISSN 0165-0106 DOI: 10.1007/s10670-018-0018-4

Meacham, Christopher J.G. "Sleeping Beauty and the Dynamics of De Se Beliefs." *Philosophical Studies*, 138: 245-269. 2008.

Monton, B. "Sleeping Beauty and the Forgetful Bayesian." Analysis, vol. 62, no. 1, 2002, pp. 47–53., doi:10.1093/analys/62.1.47.

Neal, R. M. "Puzzles of anthropic reasoning resolved using full non-indexical conditioning". *Technical Report No. 0607*, Dept. of Statistics, University of Toronto. 2006

Olum, Ken. "The Doomsday Argument and the Number of Possible Observers". *Philosophical Quarterly*. 52 (207): 164–184. arXiv:gr-qc/0009081, 2002. doi:10.1111/1467-9213.00260.

Pittard, John. "When Beauties Disagree: Why Halfers Should Affirm Robust Perspectivalism." *Oxford Studies in Epistemology* Volume 5, 2015, pp. 195–225., doi:10.1093/acprof:oso/9780198722762.003.0007.

Rachael Briggs. "Putting a value on Beauty". In Tamar Szabó Gendler and John Hawthorne, editors, Oxford Studies in Epistemology: Volume 3, pages 3–34. Oxford University Press, 2010.

Rosenthal, Jeffrey S. "A Mathematical Analysis of the Sleeping Beauty Problem." *The Mathematical Intelligencer*, vol. 31, no. 3, 2009, pp. 32–37., doi:10.1007/s00283-009-9060-z.

Schwarz, Wolfgang. "Changing Minds in a Changing World." *Philosophical Studies*, vol. 159, no. 2, 2011, pp. 219–239., doi:10.1007/s11098-011-9699-0.

Titelbaum, Michael G. "Ten Reasons to Care About the Sleeping Beauty Problem." *Philosophy Compass*, vol. 8, no. 11, 2013, pp. 1003–1017., doi:10.1111/phc3.12080.

Titelbaum, Michael G. "The Relevance of Self-Locating Beliefs." *Philosophical Review*, vol. 117, no. 4, 2008, pp. 555–606., doi:10.1215/00318108-2008-016.

Titelbaum, Michael G. "An Embarrassment for Double-Halfers." *Thought: A Journal of Philosophy*, vol. 1, no. 2, 2012, pp. 146–151., doi:10.1002/tht3.21.

Weintraub, R. "Sleeping Beauty: a Simple Solution." *Analysis*, vol. 64, no. 1, 2004, pp. 8–10., doi:10.1093/analys/64.1.8.

Winkler, Peter. "The Sleeping Beauty Controversy." *The American Mathematical Monthly*, vol. 124, no. 7, 2017, p. 579., doi:10.4169/amer.math.monthly.124.7.579