IS THERE A PLACE IN BAYESIAN EPISTEMOLOGY FOR THE PRINCIPLE OF INDIFFERENCE?

by

Dylan Barth

Submitted in partial fulfillment of the

requirements for Departmental Honors in

the Department of Philosophy

Texas Christian University

Fort Worth, Texas

May 7, 2018

IS THERE A PLACE IN BAYESIAN EPISTEMOLOGY FOR THE PRINCIPLE OF INDIFFERENCE?

Project Approved:

Supervising Professor: William Roche, Ph. D

Department of Philosophy

Susan Staples, Ph. D

Department of Mathematics

Blake Hestir, Ph. D

Department of Philosophy

ABSTRACT

Within the field of epistemology there exists a movement to formalize philosophical inquiry by appealing to probability calculus as a requirement for rationality concerning credences. Among a subset of formal epistemologists, the question of how to assign probability to propositions for which we have little evidence is of importance. In this paper I examine one of the principles that some philosophers claim bears on this issue, the principle of indifference, and issues surrounding its justification. I seek to evaluate the current literature and set forth a framework for future inquiry concerning the principle.

Dylan Barth

Is There a Place in Bayesian Epistemology for the Principle of Indifference?

1.0 Introduction

Essential to the contemporary scientific worldview is a deep understanding of probability and statistics for without these, the collection of data has little connection to meaningful discussions that underpin societal and technological development. However it is not only the scientific realm in which this understanding is useful, so too has an understanding of probability and statistics been informative to philosophers' pursuit of wisdom. Recently the school of thought Bayesian Epistemology has become very influential. The field of Bayesian Epistemology is comprised of philosophers who use a theorem from the probability calculus called Bayes' Theorem to aid in understanding rationality and inductive logic. Bayesian Epistemology has proven useful in shedding new light on issues that have been part of philosophical discourse for long periods of time. Within this emerging field, there is a very important debate on the foundations of our beliefs when there is little empirical evidence to appeal to.

There are two approaches that are taken when in this special case. One approach is of the Subjective Bayesians who hold two principles to be central. The first is that our level of belief should be probabilistically coherent, meaning that the laws of the probability calculus ought to be applied to our levels of belief. The second states that our level of belief ought to be updated in line with the use of Bayes' Theorem when new information is learned. Objective Bayesians maintain these two principles, but additionally claim that there are other constraints on prior probabilities. The inclusion of these extra constraints and which of them ought to be maintained is not without debate among Objective Bayesians.

One of these constraints that is often debated is the Principle of Indifference which informally stated dictates that given n possible events and no further information, one should assign a credence of 1/n to each possibility. In this paper I will discuss the issues surrounding the principle, and give an argument for how we ought to proceed with our discussion of it. Further I will lay out a framework for future work regarding the search for justification of the principle.

1.1 Usage of Bayes' Theorem

Bayes' Theorem is derived directly from the axiom of probability calculus that describes a conditional probability. The statement of Bayes' Theorem is as follows:

$$\Pr(H|E) = \frac{\Pr(E|H)\ \Pr(H)}{\Pr(E)}$$

Which is just to say that the probability of a hypothesis *H* given some evidence *E* is equivalent to the probability of some evidence *E* given that *H* is true multiplied by the quotient of the unconditional probabilities of *H* and *E* respectively. The probability that is defined on the left hand side of the equation is denoted a posterior probability, while the unconditional probabilities in the formula are denoted prior probabilities. This is because the former can only be calculated when some evidence is presented, and the latter is determined before that evidence is learned. A good way to see the uses of this theorem is through an example.

> <u>Example:</u> *What is the probability that a given person develops lung cancer given that she is a smoker?*

Of course there are extant studies that answer this question, but for the sake of the example these can be ignored. If we take frequency data to be indicative of future probabilities, then this question is easy to answer given some easy to measure proportions. The three things we need to solve this problem are the proportions of those who smoke, who have lung cancer, and who smoke amongst those who develop lung cancer. The resulting equation would look as such, substituting *S* in for the proposition "Joe is a smoker" and *L* for "Joe has lung cancer" for some random person Joe:

$$\Pr(L|S) = \frac{\Pr(S|L)\ \Pr\ (L)}{\Pr\ (S)}$$

This probability can be used in conjunction with a similar calculation of the probability that a person has lung cancer despite not being a smoker (we would need to calculate $Pr(L|{\sim}S)$ to make the comparison) to argue that those who abstain from smoking are less likely to have lung cancer.

Though this is a straightforward use of Bayes' Theorem to describe a rational set of credences, cases without frequency data are harder to fit into this framework. Further, if Bayes' Theorem is used to update credences, then there must have been at some point an initial credence in any given proposition. Whatever a particular credence is in the present includes information pertinent to its objective chance from the past. At some point there must bottom out a point where no pertinent information is known *a posteriori*. In other words, today's prior probability is yesterday's posterior probability. It is around this topic that Bayesian Epistemologists split into the two main camps, Subjective and Objective Bayesians. Subjective Bayesians, as previously stated, require only that a set of credences be probabilistically coherent in order to be rational in light of little to no evidence, and that credences be updated as information pours in. Objective Bayesians seek to put further constraints on initial credences, some even dictating that there exists a singular objectively rational credence in any given proposition when there is little to no evidence. This tension is one that will be the subject of this thesis, what is to be done about initial credences?

1.2 Formal Epistemology

This is a very useful place for Bayes' Theorem, however in this paper I am more concerned with philosophical issues. In Formal Epistemology, one's belief-states are described as analogous to probability states. This is a divergence from the classical way in which beliefs are discussed as a binary system. As opposed to being in one of two states – belief or disbelief in a proposition – beliefs fall on a scale from 0 to 1 inclusive. Intuitively this makes sense, as most of us believe the proposition that Kevin Bacon was the protagonist of the movie Footloose more than we believe a fair coin will come up heads when flipped. It also describes our beliefs through inductive logic well. Consider the following case.

> Example: *Your local meteorologist predicts that it will certainly rain tomorrow, and not having an umbrella handy, you wonder whether it will actually rain. The screen behind her indicates a chance of 90% that it will rain. You also note that she has been wrong about half the time when predicting rain the next day.*

Instead of asking if you believe whether it will rain or not, in this case it clearly makes sense to ask to *what degree* do you believe it will rain. You cannot be certain that it will or will not rain, and to attribute either belief to yourself would be mischaracterizing the relationship between yourself and the prediction of rain. One solution might be to say that you neither believe the claim nor its negation, but rather are in a state of agnosticism about whether or not it will rain tomorrow. However, if this route were to be explored, certainly most of our beliefs would fall into this third category. As is clear, your confidence in rain changes depending on the day even if it never reaches a state of certainty or near certainty on any given day. Dividing these beliefs into more fine-grained categories allows us to grasp more information about not *whether* a proposition is believed, but *how* it is believed.

Additionally *prima facie* the way in which a person would arrive to a conclusion given the scenario in the example is clear. A simple way to explain it is that a person would balance the probability that it will rain with the probability that the information came from a reliable source. A well-informed decision would take into account the fact that despite having the most technologically advanced equipment available, the meteorologist somehow manages to make the wrong prediction 50% of the time. Thus, balancing that fact with the probability given for rain tomorrow might reasonably lead to a confidence just shy of 50% that it will rain tomorrow. This “internal calculation” is paralleled closely by an actual calculation using the probability calculus. It seems, then, that using probability to describe beliefs also aids in understanding the use of inductive logic.

Lastly, using probability to guide beliefs provides a useful concrete guideline to forming them rationally. Consider the classical gambler's fallacy. After observing a roulette wheel produce a red number 10 times in a row, a misguided gambler bets all of his money on black. After all, what are the odds that the roulette wheel produces 11 red numbers in a row? A gambler using this line of reasoning might find knowledge of probability to be useful to him. What he's actually interested in is the probability that the next roll of the wheel produces black given that the last ten rolls have been red. Note the difference here, the first statement is concerned with 11 spins of the wheels taken all together, yet a more thorough statement would diverge the spins that have happened from those that have yet to happen. Since roulette wheel spins are independent from one another – meaning that the outcome of one spin won't affect the probability of the next spin – the objective chance that he wins the bet is 0.49 as opposed to the supposed very high credence that he has placed on the same proposition. Not only would knowledge of probability be useful to the gambler, it would seem to prescribe a belief that is in accordance with reality.

These are some instances that are hardly conclusive proof of the usefulness of probability calculus; yet they serve to illustrate the potential it holds in describing and prescribing beliefs in real world situations. More interesting, however, is the usage of principles such as the principle of indifference when very little evidence can be gathered. In the next section I describe the principle and some often discussed problems with it.

1.3 Principle of Indifference

The statement of the principle of indifference is an intuitive idea to most at first. The principle mirrors many grade school probability problems often including differently colored balls drawn from a bag or urn. The formal statement of the principle of indifference is as follows:

> *Principle of Indifference: Given n mutually exclusive and jointly exhaustive events and no further evidence to prefer any over the rest, then you ought to assign each of n events a credence of 1/n.*

This is likely an intuitively true claim. An example can be useful to illustrate. Suppose you have an urn filled with 7 different kinds of balls, each kind different from the rest by only the color of the ball. One of these colors is green. Knowing nothing else about the problem, to what degree might you believe that the color of a randomly chosen ball is not green? The natural use of the POI is to partition our probability space across the various colors of balls and conclude that we should assign a probability to the ball's being green of 1/7. Therefore, the probability that the ball is not green is 1-1/7=6/7. This example may seem to bolster the appeal of the POI, but in the next section I will discuss another example that will highlight a critical issue that Objective Bayesians face when trying to appeal to the veracity of the POI.

The first and most prominent objection to the use of the POI is that depending on how the space of possibilities is partitioned the use of the POI can lead to contradicting probabilities. Consider the following example:

> Example: *There is a factory that produces cubes that have edge lengths between 0 and 2 cm, what is the probability that a randomly chosen cube will have an edge length between 0 and 1 cm? Further, what is the probability that a randomly chosen cube with have a volume between 0 and 1 cm^3?*

Applying the POI, the probability the edge length is from 0 to 1 cm should be ½, as it covers that proportion of the possible outcomes. Similarly, the probability that the volume is from 0 to 1 cm^3 is 1/8 as the domain of the possible volumes stretches from 0^3 to 2^3 or 0 to 8, and the volumes from 0 to 1 cover 1/8 of the domain. However, the same cubes that have side lengths from 0 to 1 cm also have volume from 0 to 1 cm^3, it would seem that the conclusions here derived directly from the POI would lead to contradictory conclusions. That's an issue that many Formal Epistemologists have tried to tackle.

Furthermore the POI has another flaw that deserves discussion. Despite its seemingly useful nature, the POI is unjustified. Why ought we to believe the POI, is there an argument to be made for it? Many philosophers have undertaken this task and there have been attempts to prove the principle, to demonstrate its usefulness, to place constraints on how to use it, and more. In a further section I will discuss several attempts to do so as it will be informative to the heart of what I believe to be the divide over the POI. In the next section, I will discuss larger divides among Bayesian epistemologists.

2.0 Bayesian Divide

Formal Epistemology is useful in situations like the ones I have already discussed where abundant frequency data is available or objective chances are well established, however as noted before there are certainly situations in which it is not as clearly practical. Often an issue that philosophers address in this field is how to navigate scenarios in which one must assign a credence to a proposition with very little or no evidence. For example, suppose a rational agent wants to assign a degree of belief to the proposition that an unfair coin's next flip lands on tails. The only things known about the coin are that it has 2 possible outcomes and that they are not equally likely. How might one assign a credence to the proposition that the coin lands tails if frequency data is not available, nor is a well-established objective chance? Subjective Bayesians have an answer for this problem; any credence that is not ruled out by evidence is permissible. It might seem true that under this example that any level of credence that is not 0.5

is as valid as any other. This view inherits a central property from Bayes' Theorem. After several flips of the coin, whichever prior probability a person may have assigned other than 0 or 1[1], as evidence accrues the credences of agents following Bayes' Theorem will update their beliefs according to evidence and ultimately approach the true objective probability. For instance, let's suppose two radically different positions on the prior probability scale – 0.9 and 0.1. Assume after 100 flips the coin has landed on tails twice. After updating their credences that the next flip will return tails, the agents will have approached a state where their beliefs approximate the objective chances -- .155 and .022 respectively. After 1000 flips, if the coin has turned tails only 24 times, the rational agents' new credences will even closer match the objective chances.

Though the previous example touts a case for Subjective Bayesian constraints on initial credences, other cases are not so clearly analogous. There are many situations in every day life where we have substantially less evidence than the repeated coin-flipping example above. An easy way to find such examples is to ask close friends for their political opinions. Fairly shallow probing might reasonably reveal that even deeply held positions are based not on statistical inference, but rather intuition, imitation, or ideology. However, it

[1] 0 and 1 are "sticky" probabilities. If a proposition A has a prior of 0 or 1, its posterior cannot be changed by Bayesian updating. This makes some sense intuitively as the probabilities 0 and 1 are reserved for logical truths and falsehoods respectively.

cannot be expected of individuals to do their own rigorous research for every one of their beliefs, and even if it were it is often that we are not in a position to do so.

For instance, to what level do we believe that washing our hands regularly decreases the chances of catching an illness? It is common knowledge that this is the case, but is it the case that this “common knowledge” matches the empirical rigor of the repeated coin-flipping example? Or what of the proposition that it takes a photon approximately 8 minutes to travel from the Sun to Earth? How about the mathematical statement that two negative numbers multiplied results in a positive? What of historical facts, nutritional information, or weather prediction? These are all commonplace ideas that in their respective ways demonstrate that the kind of epistemic rigor necessary to justify even simple propositions is not as common or achievable as it may seem. The strategy of an “ends justify the means” approach that Subjective Bayesians take, then, is less useful than it appears to be once given less charitable examples. For many cases like ones listed above, objective constraints on prior probabilities would be useful in guiding belief when lacking such a rigorous amount of evidence.

One particularly interesting case will be discussed in the next section as an extended example of the kinds of problems that can be solved by the POI. It will also serve as a guide for how to apply the POI when guidelines are set later in the paper. Lastly this will serve as an example of how useful Objective Bayesianism can be in light of a longstanding epistemological issue.

3.0 A Case Study of Formal Epistemology

A case where Formal Epistemology is enlightening to classical problems in Epistemology is that of the debate surrounding the Transmissibility Argument. Discussion of this will highlight both the usefulness of Formal Epistemology in philosophy and also critical issues surrounding constraints on Subjective Bayesianism.

3.1 Transmissibility

The Transmissibility Argument appeals to claims about ordinary knowledge that seems to entail knowledge that we cannot know. The argument is as follows:

> **1)** If you can know that you have hands (*HANDS)* and you know that *HANDS* entails that you're not a Brain in a Vat (*not-BIV)*, then you can know *not-BIV*.
>
> **2)** You know *HANDS* entails *not-BIV*.
>
> **3)** You cannot know *not-BIV.*
>
> ---
>
> **4)** You cannot know *HANDS.*

The Brain in a Vat hypothesis, that all of ones sensations and experiences are completely simulated and that in reality the subject is a body-less brain in a vat being controlled by electrical impulses that give the mere impression of familiar reality. This argument relies on the transmission (hence the name) of knowledge across entailments. It seems obvious prima facie that if you can know something, and you know that this fact entails something else, then you can know that which is entailed. In other words, if you can know proposition P, and you know that P ==> Q, then you can also know Q. However, if this is true, then according to the argument above we cannot in fact know that we have hands (or any other ordinary knowledge for that matter). Note that this argument does not in fact rely on the particular skeptical hypothesis BIV; any equivalently skeptical proposition could be inserted into the argument and the conclusion would follow. I will call this property of the Transmissibility Argument the "reproductive property". Note that any solution that relies on the particulars of BIV to explain away the Transmissibility Argument is not sufficient to defeat the argument in a general sense due to the reproductive property.

3.2 Dogmatism

One approach to solving the Transmissibility Argument is appealing to Dogmatism. The particular solution that Dogmatism posits is that if you have a visual experience of having hands for the first time, and have no reason to suspect BIV, then you are justified in believing HANDS even if you were not already justified in believing not-BIV. Thus, you could easily avoid the skeptical solution to the Transmissibility Argument by being justified in believing HANDS, and using that to justify your belief in not-BIV. Therefore, premise 3 of the Transmissibility Argument would not be true and the argument would cease to be sound.

Dogmatism has the added benefit of addressing the reproductive property as well; for any skeptical hypothesis that could replace BIV an equivalent observation of the world around you, under Dogmatism, would allow you to justifiably believe the proposition that it opposes that serves as a foundation for believing that the skeptical hypothesis is false. In fact, even believing HANDS allows a subject to fend off most skeptical hypotheses, as implicit in the proposition is that the world around the subject is not an illusion. This is why knowing HANDS allows one to justifiably believe not-BIV.

3.3 White's Objection

Roger White provides an objection to Dogmatism's solution to the Transmissibility Argument based in probability theory[2]. First White posits that if

---

[2] Roger White, *Problems for Dogmatism*, (Philosophical Studies, 2006)

BIV were true, that would imply that you have a visual experience of having hands.

(a) BIV ==> VE(HANDS)

Therefore the probability of BIV increases when one has a visual experience of having hands[3].

(b) Pr(BIV|VE(HANDS)) > Pr(BIV)

If that is true, then the inverse is also true: that is the probability of not-BIV goes down when one has a visual experience of having hands[4].

(c) Pr(not-BIV|VE(HANDS)) < Pr(not-BIV)

Because HANDS entails not-BIV and not vice versa the probability of not-BIV will always be higher than HANDS[5]. Therefore, the probability of not-BIV given a visual experience of having hands is higher than the probability of HANDS given a visual experience of having hands.

(d) Pr(not-BIV|VE(HANDS)) > Pr(HANDS|VE(HANDS))

This particular result is an issue for Dogmatism; because one of the vital claims that it makes is that you can be justified in believing HANDS given a visual experience even if you're not already justified in believing not-BIV. It seems that

---

[3] It follows from Bayes' Theorem that, as long as $Pr(H)$ and $Pr(E)$ are not 0 or 1, if H entails E, then $Pr(H|E) > Pr(H)$.

[4] Because $Pr(\sim H|E) = 1 - Pr(H|E)$ is a theorem of probability calculus and $Pr(H|E)>Pr(H)$, then $Pr(\sim H|E)<Pr(\sim H)$.

[5] If H entails H* but not vice versa, then $Pr(H|E)<Pr(H^*|E)$

if you come to believe HANDS on the basis of VE(HANDS), then you should have had a high credence in ~BIV before your visual experience, which defeats the purpose of the argument for Dogmatism.

This argument, though a strong case against Dogmatism, may potentially lead to skepticism. If the probability of HANDS increases once we have a visual experience of having hands and the probability of not-BIV is always higher than HANDS (d), then the probability of HANDS given VE(HANDS) has a ceiling of the probability of not-BIV given VE(HANDS). Because the probability of not-BIV decreases once we have a visual experience of having hands (c), the prior probability of not-BIV (that is, before we have a visual experience of having hands) is integral in determining the probability of HANDS given VE(HANDS). One way to partition the space of possible solutions is to give all solutions equal probability and to split the solutions into two categories, BIV and not-BIV. Thus, the probability of each of these (before we learn anything about the world) is 1/(number of solutions) or 1/2. However, if the prior probability of not-BIV is 1/2, then the probability of HANDS given VE(HANDS) can never be higher than that. It may be the case that we are stuck with a skeptical solution after all.

3.4 Rebuttal

Though a daunting result of White's argument, it may not be the case that the prior probability of *not-BIV* is as low as one half. Partitioning the solution space into *BIV* and *not-BIV* may not capture the whole picture, after all it does not make sense to partition the rolled number on a fair dice as either 2 or not 2 and conclude that the probability is 1/6 that a 2 will appear. Part of the reason White is incorrect is that of all the possible hypotheses about the ontology of our world as we experience it, *BIV* is only one while *not-BIV* encompasses every other hypothesis possible. If we were to give every hypothesis equal probability, then each would still have a prior probability of $\frac{1}{number\ of\ solutions}$, but since *not-BIV* encompasses every solution outside of *BIV*[6], and together they span all possible solutions, the probability of *not-BIV* would be $1 - \frac{1}{number\ of\ solutions}$. Since the number of solutions is infinite (or at least pragmatically so) the probability of *not-BIV* is very close to one. Thus, we are justified in believing *not-BIV* a priori[7], and the probability of *HANDS* given *VE(HANDS)* has a reasonably high ceiling.

One worry a skeptical reader might raise is that implicit in this argument is the claim that we can know the probability of *not-BIV* prior to learning about the world around us, which could strike one as quite unintuitive. After all, we

[6] This includes some skeptical solutions, of course, so the problem isn't *really* solved at all.
[7] This is because BIV must be very specific to imply VE(HANDS).

would have very little evidence to go on by the very nature of the problem. One easy solution that is available to us is to appeal to the Subjective Bayesians, who require only that a set of prior probabilities be coherent. This may be defended by the fact that after enough evidence is given; even seemingly irrational sets of belief will all converge to one rational probability as has already been discussed. Therefore, a Subjective Bayesian would have no trouble wrestling with defining this prior probability. Objective Bayesians might appeal to epistemic virtues such as parsimony or explanatoriness to navigate this problem.

It may seem like we have a definitive solution to the problem of prior probabilities with respect to skeptical hypothesis, but alas there is yet another problem to overcome. The previous two sections of this paper have strayed from the heart of the issue; White's objection fails to address the reproductive property of the Transmissibility Argument. White argues against *BIV*, but not against all skeptical solutions. Propositions (a) through (d) could easily be reformulated to fit the reproductive property for a generalization of all skeptical hypotheses. This might look like replacing the proposition *BIV* with a proposition along the lines of "despite their appearances, your experiences are fictitious and do not represent the actual reality that you exist in" which can be said to generalize the skeptical hypothesis. This would entail, still, that one would indeed *have* these experiences that would seem to be of a non-skeptical reality (a). Due to this entailment, the probability of a generalized skeptical argument being true would increase given some seemingly real experiences (b). The contrapositive of this is still true; the probability of a generalized skeptical argument being *not* true decreases given these seemingly real experiences (c). Because (under Dogmatism) these experiences would entail that a generalized skeptical argument is *not* true, the probability that a generalized skeptical argument is *not* true is larger than the probability that these experiences are veridical[8] given that you're having them (d). So we've seen that the first four points of White's argument can address the reproductive property, but when addressing the worries about resurfacing skepticism, this virtue is lost.

Recall that the prior probability of *not-BIV* was integral in having high confidence in *HANDS* given *VE(HANDS)*. The parallel case when generalized is that the prior probability of a generalized skeptical argument being *not* true is integral in having high confidence in our experiences being real. Recall also that the solution to this was enumerating the solutions available to this problem and showing that the prior probability of *not-BIV* (if one uniformly distributes the probability, say if one is a proponent of the principle of indifference) was $1 - \frac{1}{number\ of\ solutions}$. But the crux of this issue was that *not-BIV* represented the set of all of the hypotheses available with the exception of one, *BIV.* If we try to

[8] "Real" here simply means that the experiences you are having are indicative of the reality that you exist in; that you actually have hands so to speak.

overcome the reproductive property with a generalization in this case, the argument is no longer the same. That is, the prior probability of a generalized skeptical solution being *not* true cannot be $1 - \frac{1}{number\ of\ solutions}$ because the probability of a generalized skeptical solution is not the same as the probability of a token example of a skeptical solution. In fact (assuming uniform distribution) the probability of a generalized skeptical solution is (assuming all token skeptical solutions are mutually exclusive) the sum of the probabilities of every individual token skeptical solution. Because there are a large number of skeptical solutions, we have the following principle:

(s) Pr(*Generalized Skeptical Solution*) >> Pr(*BIV*)

If (s) is true then the probability that our experiences are indicative of our reality is $1 - \frac{number\ of\ skeptical\ solutions}{number\ of\ solutions}$[9]. This latter ratio is seemingly incomprehensible; there is no obvious way to evaluate such a statement. We are left without a numerical answer and until one is found there is no strong argument against skepticism in this form. That is, unless the Subjective Bayesians are to be believed which would then allow for the latter ratio to be as small as necessary so long as it is coherent with our other co-related beliefs.

[9] This is only true if one accepts the principle of indifference. If that is not the case, the probability is still incomprehensible and we are left without a solution.

4.0 Attempts to Justify the Principle of Indifference

As mentioned at the end of section 2, there have been many attempts to justify the POI in various ways. These efforts have ranged in scope from attempting to prove the principle outright, to attempting to constrain it so that contradictions no longer occur. In this section I will describe a few of these papers in order to understand the heart of the POI and suggest a solution based on the shortcomings of the attempts preceding it.

4.1 Hawthorne et al.

Hawthorne et al. seek to prove that the POI follows logically from another principle that is widely accepted amongst Bayesian Epistemologists called the Principal Principle. The statement of this principle is as follows:

> *Principal Principle:* *$Pr(A|XE)=x$, where X is the proposition that the chance of A at time t is x and E is any admissible information at time t.*

This is rather intuitive, as it is not a particularly strong statement. It should be noted that the concept of admissible propositions refers to propositions that contain information about events occurring no later than time t. For instance, admissible information about the outcome of a coin flip at time t would include all previous flips of the coin occurring up until time t. While that may be pertinent information to assigning a probability, any information concerning only events before time t is admissible. For the same situation, the information that Catherine the Great was born in Prussia is also admissible (so long as the coin is flipped after her birth). If A is the proposition that the outcome of a coin flip at time t is heads, X is the proposition that the objective chance of A is ½ at time t, and E is the information that Catherine the Great was born in Prussia, then according to the PP the probability that the coin flip at time t will result in heads is ½. This should not be a particularly surprising result.

If Hawthorne et al. are right, it may entail some information that would put this long debate to rest. The authors use a string of deductions to get from the PP to the POI, and it should be noted that the proof is valid. For the sake of brevity, I will not go into detail through the entire proof. Instead I will be discussing one of their assumptions and raise potential concerns to the veracity of their claim. The crux of the argument rests on the conditionalization of a claim upon a biconditional statement. This just means that they consider the probability of a claim given that an "if and only if" statement concerning the claim is true. In particular they assume that[10]:

$$Pr(A \mid (A \Leftrightarrow F)XE) = Pr(A \mid FXE)^{11}$$

---

[10] Hawthorne et al., *The Principal Principle Implies the Principle of Indifference*, (The British Journal for Philosophy of Science, 2017)

[11] The biconditional in this statement is meant to represent a material biconditional.

where F is a contingent statement and is unrelated to A through E. This may seem intuitively true, as it seems that F is unrelated to A by definition, so there could not be any effect on A from F, and that learning a biconditional would similarly not provide any reason to believe or not to believe A.

However, this is a tempting assumption to make but it turns out to be false. The reason why is best illustrated through example[12].

> <u>*Example:*</u> *Your roommate is a gambler. He even gambles about whether to gamble! Whenever he goes to the track he flips a fair coin. If it comes up heads he bets on his favorite horse, Speedy. If it comes up tails he doesn't place any bets.*
>
> *Your roommate is at the races today. Your credence that his coin flip came up heads is ½. You don't know anything about horse racing (except that race outcomes aren't influenced by coin flips), but you know that there were 6 horses in Speedy's race. So you assign credence 1/6 that Speedy won the race.*
>
> *Your roommate comes home grinning. Only two things make him that smug: either Speedy won the race after your roommate bet on him, or Speedy lost but your roommate didn't bet. Now that you've seen the grin, how confident are you that the coin came up heads?*

---

[12] Hart and Titelbaum, *Intuitive Dilation?*, (Thought: A Journal of Philosophy, 2015)

At first glance, you might affirm that the credence you have that the coin landed on heads remains ½. While this is the intuitive answer, it turns out that the answer should be 1/6!

How can this be when you seemingly learned nothing about the outcome of the coin at all? The answer is a simple matter of mathematics. Formally, the probability that heads is flipped (H) given that your roommate is smiling (S) can be stated as the following using Bayes' Theorem:

$$Pr(H|S) = \frac{Pr(H)Pr(S|H)}{Pr(S)}$$

Using the Law of Total Probability, we can conditionalize S with H and ~H to get a probability that your roommate is smiling.

$$Pr(H|S) = \frac{Pr(H)Pr(S|H)}{Pr(S|H)Pr(H) + Pr(S|{\sim}H)Pr({\sim}H)}$$

Since his smile tells us that either Speedy won (Sp) or Speedy lost (~Sp), we can replace the proposition S with Sp&H ∨ ~Sp&~H, which is equivalent to the biconditional Sp ⇔ H.

$$Pr(H|S) = \frac{Pr(H)Pr(Sp\&H \lor {\sim}Sp\&{\sim}H\,|H)}{Pr(Sp\&H \lor {\sim}Sp\&{\sim}H\,|H)Pr(H) + Pr(Sp\&H \lor {\sim}Sp\&{\sim}H\,|{\sim}H)Pr({\sim}H)}$$

We know the prior probability of a fair coin flip to be 1/2, and that given he flipped heads, the probability that he is smiling is merely the probability that speedy wins which is 1/6. And given that he did not flip heads, the probability that he is smiling is merely the probability that speedy lost which is 5/6.

$$Pr(H|S) = \frac{\frac{1}{2} * \frac{1}{6}}{(\frac{1}{6} * \frac{1}{2}) + (\frac{5}{6} * \frac{1}{2})} = \frac{1}{6}$$

While this may be intuitive for mathematically inclined individuals, I would like to offer another more broad explanation. If we consider in general what it means to conditionalize on a biconditional, we can see that it surely is not the case that the claim Hawthorne et al. made is true. Firstly, consider the world in which we are given that $A \Leftrightarrow F$ and note that though this includes prima facie little information about A, it does contain information about the relationship between A and F. Notably, similar to the example, that either A is true and F is true or A is false and F is false. That means that all possible worlds where their truth-values differ are ruled out when given $A \Leftrightarrow F$. Now consider the comparison Hawthorne et al. make in their assumption; the world in which we are given that F is true. Note that this too contains any situation where A is true and F is true, however it does not allow for a situation where A is false and F is false, since we are given F. This disjoint is qualitatively why the assumption that Hawthorne et al. make is false.

Since the proof that Hawthorne et al. turned out to be based on an assumption that was false, their argument is not sound. Many others have tried to justify the POI with analytical proof, but with similar levels of success. There may yet be a proof of this sort yet, but Hawthorne et al. were a promising contender and came short. In the next section, I will review a different approach to justifying the POI outside of pure logical deduction.

4.2 Weisberg

Jonathan Weisberg attempts to resolve some of the tension surrounding constraints on prior credences by introducing a novel way to use what is called

Inference to the Best Explanation or IBE for short[13]. IBE is a subset of the method of reasoning known as induction. IBE leads users to conclusions by the light of an alleged epistemic virtue denoted explanitoriness. The general idea is that the better that a hypothesis explains some given evidence, the more you ought to believe it. More formally a statement similar to the following may be made.

> *Inference to the Best Explanation:* *Given some evidence E and a set of competing hypotheses* $H_1, H_2, \ldots, H_n$*, infer that the* H_i *which best explains E is true.*

An example may be illuminating in this situation. Imagine every Friday evening you and your friends spend your evening at the local Pub to drink some bourbon. Given this habit, you're well aware of how alcohol affects your perception of the world around you. One Friday a new bartender is working and after several drinks you feel completely sober and find absent the normal impairment you would experience after quite a few drinks. After asking your friends you find that they too are experiencing a lack of effect from their drinks. You entertain two hypotheses, the first of which is that your tolerance has spontaneously risen and the amount of alcohol needed to reach normal effects is more than it has been in the past. The second hypothesis is that the new bartender has shorted your group the alcohol in their cocktails. Following the light of IBE, you might conclude that the second hypothesis is preferable, as it

13 Jonathan Weisberg, *Locating IBE in the Bayesian Framework*, (Synthese, 2009)

not only explains your sobriety but also that of your friends. This is true whereas the first hypothesis would only explain your sobriety and not theirs. Therefore, using IBE you may conclude that the second hypothesis is the better one.

There are many issues with IBE that I will leave untouched in this paper, but suffice it to say that IBE is not without its critics[14]. Weisberg seeks to find a place for IBE in the constraints on prior probabilities in line with Objective Bayesians in his paper "Locating IBE in the Bayesian Framework." The most glaring issue with the Principle of Indifference is that it does not in itself offer guidance on how to partition the space of possible events. Weisberg suggests that IBE might be useful for narrowing this infinitely large scope by using explanatoriness to pick out the partitions that dovetail with IBE. His general proposed procedure for doing this consists of first identifying the potential explanations for a scenario that we are interested in, within each potential explanation use the Principle of Indifference to assign a uniform distribution of probability across the possible parameters, then weigh each partition by its relative explanatoriness. Now, this is a far cry from giving an objective constraint to prior probabilities, but it might be quite useful in some contexts. The following is a summary of the example Weisberg gives in his paper to demonstrate the usefulness of this procedure.

---

[14] Bas C van Frassen, *Laws and Symmetry*, (Oxford University Press, 1989) is an early example of such criticism.

Joel seeks to arrive at city A by train by 4:00pm and asks an employee of the train station when the train that departs at 3:00pm will arrive at city A. The employee replies that it will be somewhere between 3:50pm and 4:10pm. To what degree should Joel believe he would make it to city A on time? The Principle of Indifference may be invoked to give a credence spread uniformly across the times estimated by the employee of the train station, resulting in a credence of 1/2 that Joel will arrive on time. However, suppose that the given timeframe corresponds to train speeds between 20mph and 50mph, where a speed of 40mph will get Joel to city A by 4:00pm. Applying the Principle of Indifference to this partition, Joel should believe he would be on time with a credence of 2/3. Being that there's no reason to prefer one over the other, Weisberg suggests that we choose the more explanatory parameter to partition with the Principle of Indifference. In this case, speed explains arrival time not the other way around, so Joel is, under Weisberg's suggestion, justified in weighing the second partition more than the first and believing that he will arrive on time with a credence of 2/3.

As useful as this approach is in the example given above, in many situations this approach does not differ significantly from Subjective Bayesian approaches. After all, there are no extant objective rules to determine what it means to be a better explanation. At the very least, such rules are the subject of an open debate. Furthermore, this may miss the spirit of the POI as its use is when there exists very little evidence for a proposition, yet Weisberg's approach would require that the explanatory relationships between propositions be known *a priori*. It would be hasty to claim that in the extended example from section 3 has a clear parameterization that can be weighed by its explanatory virtue. In other words, does White's presentation of the problem carry with it more explanatory power or does the alternative that I present do so? With many *a priori* claims it would be difficult to know their explanatory value before appealing to evidence from the external world. Moreover there may be examples where explanatoriness fails to make a distinction between two hypotheses.

Furthermore, the ability to explain given evidence may be cast in Bayesian terms which would eliminate the need for such concerns altogether. Let us consider that either explanatoriness can or cannot be cast in terms of probabilities. If it is, the appeal to IBE is unnecessary as it already falls within Objective Bayesian constraints. This seems to be an intuitive interpretation as one might cast explanatoriness in the following way: A hypothesis H′ explains some evidence E better than another hypothesis H given some background information B only if $Pr(E|H' \& B) > Pr(E|H \& B)$. Obviously if this is how explanatoriness is described, the appeal to explanatoriness is not necessary. The preference for H′ over H is contained within the Bayesian framework already.

Any other appeal to probability calculus to describe explanatoriness would similarly be contained in the Bayesian framework.

If explanatoriness is not cast in terms of probability, then it is reducible to Subjective Bayesianism. It would be an individual's choice to determine which hypotheses are better at explaining the evidence, especially given no objective common ground to argue in over which hypothesis is the best at explaining some evidence. Furthermore, if explanatoriness does not change the probability of a proposition's being true, then for what reason should it be accepted that it tracks truth? If a proposition is better at explaining evidence, but is equally probable then it seems as if any divergence in credence from the objective probabilities would be strictly out of line with any of the Bayesian philosophies[15].

Lastly, Weisberg might suggest that his use of explanatoriness falls into neither of these categories neatly. After all, the point was to use explanatoriness to choose which partition to use the POI on; the problem that he is trying to solve is one that strictly falls out of the scope of our normal use of probability calculus. However, contrary to what Weisberg suggests, this would not in fact be an objective constraint on prior probabilities. Individual differences that comprise the notion of explanatoriness would still exist, especially when considering to what degree a hypothesis explains some given evidence. This undermines the

[15] William Roche, *The Perils of Parsimony*, (Unpublished)

last step in Weisberg's procedure for finding a partition over which to apply the POI. It seems as if, particularly without any objective constraints on what makes a hypothesis more or less explanatory, even this use of explanatoriness is reducible to Subjective Bayesianism. This is further bolstered by the fact that no justification for using explanatoriness to navigate the various partitions of probability is given outside of the fact that it would appease proponents of IBE without violating the rules of Bayesianism. This appears to be, at heart, an *ad hoc* addition to Objective Bayesianism.

That being said, there is something intuitive about what Weisberg says. It could be the case that we can use the epistemic virtue of explanatoriness to choose our partition for the POI in a different and more justified manner. On such a topic, Michael Huemer may have a resolution for the problems encountered so far.

4.3 Huemer

Huemer proposes in his paper "Explanationist Aid for the Theory of Inductive Logic" that in contrast to the approach that Weisberg takes, that we use not explanatoriness itself, but rather something that Huemer denotes explanatory priority[16]. This in short is the relationship between two propositions A and B such that A explains B only if A is explanatorily prior to B. The details of this are

[16] Michael Huemer, *Explanationist Aid for the Theory of Inductive Logic*, (British Journal for Philosophy of Science, 2009)

fleshed out in the next paragraph. A useful tool, Huemer claims, arises from viewing explanatoriness in light of explanatory priority. This allows us to define a scale on which things are more or less basic such that the guiding principle for the POI is no longer to spread ones credence over the partition that does the best job explaining a proposition, but rather the partition that is the most explanatorily basic. Huemer calls this the Explanatory Priority Proviso to the Principle of Indifference. More formally stated, suppose there are two alternative partitions of the possible events that are both mutually exclusive and jointly exhaustive $h_1,h_2,\ldots,h_i$ and $j_1,j_2,\ldots,j_k$. Further, suppose that each member of the h partition is explanatorily prior to each member of the j partition. The Explanatory Priority Proviso prescribes that the h partition be preferred over the j partition.

Some sufficient conditions Huemer lists in his paper about explanatory priority follow. Note that this is not an exhaustive list, but that these are merely examples of what he takes to be explanatory priority:

> <u>*Causal priority:*</u> *If A (partly) causes B, then the occurrence of A (that is, the fact that A occurs) is 'prior' to that of B in the order of explanation, meaning that A's occurrence is a candidate to figure in an explanation of B's occurrence, whereas B's occurrence is not fit to serve in an explanation of A's.*

This particular claim is a metaphysical one. In order to make a judgment about causal priority, we must understand first the causal relations between entities within the various hypotheses. For example, given two possible hypotheses that explain a chemical structure one appealing to the atomic level and the other to the subatomic level, we can appeal to causal priority to prefer the more causally basic hypothesis (because the existence of nuclear particles is *caused* by an arrangement of subatomic particles). In many situations, such as in the natural sciences, this information is readily available. However, this information is less readily available when comparing *a priori* claims.

> <u>*Temporal priority:*</u> *If A is a fact about events or states that are temporally prior to (exist before) the events or states that B concerns, then A is explanatorily prior to B. (A may still, of course, fail to satisfy some other requirement for explaining B.) For these purposes, an eternal or timeless fact may also be treated as prior to facts about what happens at particular times.*

This aspect of explanatory priority is one about the external world. The temporal aspect of explanation in short just requires that the partition over which the POI is applied be chosen to be temporally prior to events contained in competing partitions.

> <u>*The part–whole relation*</u>*: The existence, arrangement, and intrinsic features of the parts of an object are explanatorily prior to the existence and features of the whole.*

Though relatively straightforward, the part-whole relation describes the intuitive fact that the parts of a compound subject are explanatorily prior to the whole of it. As it relates to the POI, I earlier made a distinction of this sort to argue against White's use of the POI surrounding the BIV hypothesis. I argued that in essence, you should not lump all of the skeptical hypotheses into one proposition (not-BIV), but rather should include all parts of the whole in the partition.

> *<u>The in-virtue-of relation</u>: If B holds in virtue of A's holding, then A is explanatorily prior to B. The determinable-determinate relation may be a species of the in-virtue-of relation: if d is a determinate of D, then an object that has d will also have D in virtue of its having d. So a thing's having d will be explanatorily prior to its having D.*

For instance an object is scarlet in virtue of its being a shade of red. Therefore being some shade of red is explanatorily prior to being scarlet. This makes sense because in some sense being red explains being scarlet as being scarlet is a more specific color than the general shade of red.

> *<u>Supervenience:</u> At least some forms of supervenience are also instances of explanatory priority. For instance, the object on which I am seated is a chair in virtue of its parts having certain microphysical properties and relations, the properties and relations on which its chairhood supervenes. So the instantiation of those properties and relations is explanatorily prior to this object's being a chair.*

The difference between supervenience and the part-whole relationship is not immediately clear. Supervenience concerns sets of properties, while the part-whole relationship concerns objects. If the set of properties A supervenes on the set of properties B, there cannot be a difference with respect to A properties without a difference in B properties. Intuitively, then, some instances of supervenience would entail that B is more explanatorily prior to A, and thus our partition of probable events should be done at the level of B.

Now that we know the conditions Huemer sets out, let us now return to one of the examples concerning explanatoriness, that of Joel's credence that his train will arrive on time. The Explanatory Priority Proviso prescribes a partition not across time but rather across velocity, the same result as before. However, here it is important to note that there is a distinction to be made as an objection to this result might be that velocity is defined in terms of distance and time, therefore time is the more basic parameter on which we ought to partition our use of the POI. The distinction made by Huemer is that velocity is explanatorily more basic than time is because explanatory priority has to do with metaphysical basicness not conceptual basicness. The time at which Joel arrives is explained by the speed of the train, not the other way around. Similarly Joel could have gotten to his destination sooner if the train had gone faster, yet he could not have gone faster by getting there sooner.

Let us examine yet another of our examples, the original example given to demonstrate the flaws of the POI; that of the cube factory. On which partition should we place the POI, over the edge length or the volume? By similar reasoning to the previous example, the edge length is explanatorily prior to the volume of cubes. The volume of a cube is so because it has edges of a particular length, and if you were to construct a cube with a larger volume you would need to specify a longer edge length, not the other way around. Another way of explaining this is that a cube with volume 1 has volume 1 in virtue of its having side length 1, but a cube does not have side length 1 in virtue of its having a volume of 1.

This restriction on how one ought to partition the space of possible events has less utility than one may have hoped. While Huemer's approach does in fact inform some situations in which the POI is in need of clarification, an important use of the POI in philosophical inquiry is when assigning prior probabilities to *a priori* statements. There are significant worries to be had about determining explanatory priority *a priori*; though some examples may suggest an obvious hierarchy of explanatory priority it is often difficult to make such clear distinctions. For instance, recall the extended example in section 3. White proposes the partition of BIV and not-BIV to apply the POI onto, while I suggest an alternative that includes all skeptical hypotheses individually defined. Which of these is explanatorily prior to the other? The answer is not so clear given the guidelines given by Huemer. Even with the additional constraints placed on partitions, the POI is still vague when it comes to some *a priori* statements.

5.0 Future Direction

There is much debate over the Principle of Indifference to say the least. Several approaches exist to justify its use, whether by proving it through deductive proof as Hawthorne et al. did or narrowing the scope of its use in an attempt to dissolve contradictions that may arise as Huemer did. However, if Objective Bayesians are to find a justification of some sort, it would be useful to know which approach will be the most lucrative.

Firstly we must consider whether or not a proof of the kind that Hawthorne et al. attempted to carry out would be sufficient to justify the POI. Using a wide scope, the paper consists of an attempt to derive the POI from the probability calculus axioms and the principal principle. The approach did not pan out, but could a similar paper use the same strategy to attain its goal? Possibly, but one would have to be quite careful about the approach. Strictly speaking the POI is a normative statement while the axioms of probability calculus are descriptive statements. Any attempt to derive a normative statement from a set of purely descriptive statement needs to bridge what is classically referred to as the "is-ought" gap. This gap is described first by Hume, and places a divide between the set of statements concerning the way that reality is and the set of statements concerning the way individuals ought to act. Since the POI is a statement about how an individual ought to assign credences given little useful evidence, it cannot be shown to be true from a set of descriptive statements like the axioms of probability[17]. If some normative statement is used as a non-trivial part of the proof, there is no problem of bridging the is-ought gap as is true in the Hawthorne et al. paper.

Strictly speaking the POI cannot come from the axioms of probability, but could something similar be shown to be true without a normative component? A statement like this might mirror the POI in form, but assign a probability rather than a credence to a set of choices. This would, by the framework of the Objective Bayesians, be tantamount to a proof of the POI still. Note that this is not exactly the same approach as is contained in Hawthorne et al., but rather relies on facts about mathematics alone to justify the POI.

The effort to prove the POI may not matter in the end, though, as it turns out that there exist some statements within the framework of Objective Bayesianism that cannot be proven, but nevertheless are true. To find this out, it is necessary to explain Gödel's Incompleteness Theorem and show that it has implications within Objective Bayesianism.

---

[17] That is, unless the gap can be bridged, which is an open problem in metaphysics.

<u>5.1 Incompleteness and Undecidability</u>

Gödel's Incompleteness Theorem demonstrates the boundaries of any formal axiomatic system that contains arithmetic. It is formally defined as the following statement:

> *<u>Gödel's First Incompleteness Theorem:</u> Any consistent formal system F within which a certain amount of elementary arithmetic can be carried out contains statements that cannot be proven or disproven in F.*

This will not provide all the tools necessary, as this theorem does not provide that the statements which are unprovable will be true. Gödel later shows that some of these unprovable statements are true, but the proof for this is not important enough to discuss here. Without digging into the details of the proof, we must now show that Objective Bayesianism is dependent on the kind of system that is described in Gödel's Incompleteness Theorem.

The words "consistent formal system" refer to something very particular in the realm of mathematics. The requirement of consistency[18] is already assumed in the Objective Bayesian framework and is also present within probability calculus, so this requirement is quite clearly fulfilled. To be a formal system there are three main requirements: a finite set of symbols used to

[18] Consistency here is in the sense that a contradiction cannot come as a consequence of the axioms within that system.

construct formulas, rules by which these symbols must operate within the system, and a set of axioms that follows these rules. Formal systems are most frequently used in mathematics, however they are by no means limited to mathematical logic. Evidently the system described in Objective Bayesianism fits into this description (or at the very least depends on probability calculus which is certainly a formal system) as it has a finite set of symbols used to construct formulas – every example of the use of probability calculus shows this to be true. The rules by which these symbols operate are the rules of probability calculus in union with the extra-mathematical set of principles such as the Principal Principle. Undoubtedly the axioms of Objective Bayesianism do not defy the rules so set forth as well[19].

The next piece of the puzzle is determining whether or not Objective Bayesianism can carry out the kinds of arithmetic described by Gödel's theorem. The theorem in its primary statement is unclear, but as an elucidation we may state that the "amount of elementary arithmetic" that Gödel refers to is simply that the system may prove a sufficient collection of facts about the natural

[19] One might say that Objective Bayesianism contains only normative axioms and that it is dependent on the truth probability calculus, but does not contain its axioms specifically. In this case, Objective Bayesianism depends on a formal system rather than being one itself. As we will see, this disconnect comes out in the wash when all is said and done.

numbers. Since all of elementary arithmetic is a subset of probability calculus, this requirement is neatly contained.

Finally, the conclusion is that Objective Bayesianism will in fact contain unprovably true statements. However, there is an important fact about Gödel's incompleteness results that makes this conclusion slightly less impactful. The incompleteness only applies to the subset of the system that constitutes a formal system. So if Objective Bayesianism is only a formal system in virtue of its dependence on containing the axioms of probability calculus, that means that the only things that can be said about unprovable statements concern probability calculus itself. Nevertheless, as discussed previously, there likely is some descriptive statement that closely mirrors the POI in terms of probability not credence. At any rate some true conjectured rules that seem to have some wide bearing of evidence on them could in fact be impossible to prove. It very well might be the case that the POI is one of those rules, but unfortunately there currently does not exist a way to tell whether or not a statement is unprovable but true. Fortunately, however, unprovably true statements can be determined to be true or false using principles from outside the system they lie in. There is no formula for this strategy of proving statements, but there may be a successful route to take from areas like classical epistemology or an appeal to explanatory virtues, although such attempts likely would not fit into a Bayesian framework[20].

I believe that it is certainly within the realm of possibility that the POI falls inside the set of unprovably true propositions. It would explain the intuitiveness and failure to justify its use if it did. Furthermore, this gives a direction for future work on the POI; it justifies the approach that Huemer and Weisberg used of constraining our understanding of the correct way to partition when using the POI.

There are two worries I would like to address before concluding the paper. The first is that Gödel's work was based in self-referential statements, and the POI is not self-referential. For those readers who are vaguely familiar with the Gödel Incompleteness Theorem proof, this may be a salient issue. Luckily, we know that several long-standing conjectures within various subsets of mathematics have been proven to be "undecidable."[21] To be undecidable just means to be unprovably true or unprovably false. The important result is that among the proven undecidable statements, there exist plenty of non-self-referential statements.

---

[20] Elliot Sober, *The Principle of Parsimony,* (British Journal for the Philosophy of Science, 1981)

[21] van Heijenoort, From Frege to Gödel: A Source Book in Mathematical Logic 1879–1931 (Oxford University Press, 1967)

Secondly, I would like to address the proposal to prove the POI is undecidable. Undecidable statements are not guaranteed to be true or false once proven to be undecidable, which raises quite the difficult issue of “so what?” If the POI is proven to be undecidable, then we only know that it cannot be proven or disproven given the axioms of the system. This puts the philosophers interested in Objective Bayesianism in an interesting position, as they would know that no progress could be made on finding its truth value, yet would likely want to include or exclude it from their work. I suggest, as I will motivate in the next section, that this ultimately presents a bifurcation amongst epistemologists, those who will assume its veracity and those who will assume its falsehood.

5.2 An Example of Undecidable Statements

To understand not only the implications of proving that the POI is undecidable, but also why it benefits the Objective Bayesian to attempt such an endeavor, I find that it is easiest to explicate with an example from geometry.

Euclidean geometry is a ubiquitous way of understanding spatial relations between mathematical objects. Euclid himself derived much of what is commonly understood about geometry from just 5 axioms. The axioms are all very simple intuitive facts about geometry that can be reasonably accepted as axiomatic. That is, all except for Euclid’s 5th Postulate. The 5th axiom is unusually complicated and unintuitive. The statements of Euclid's 5th postulate is as follows:

> <u>*Euclid's 5th Postulate*</u>*: If a line segment intersects two straight lines forming two interior angles on the same side that sum to less than two right angles, then the two lines, if extended indefinitely, meet on that side on which the angles sum to less than two right angles.*

This just means that a line that crosses two parallel lines will create angles with those lines that when summed, equal to 180°. This is much less intuitive than most statements that we consider to be axiomatic; in fact it seems like the kind of statement that needs to be proven. However, as it turns out, Euclid's 5th Postulate is an undecidable statement within a system that assumes only the other 4 axioms. It cannot be proven or disproven from the other 4 axioms, and therefore its truth-value is unknown[22]. This discovery led mathematicians to explore formal systems where the postulate is false, eventually leading to the development of non-Euclidean geometries one of which was foundational for Einstein's theory of general relativity.

A reasonable suspicion may arise here, after all we know that the postulate is undecidable, yet it seems that mathematicians assign it a truth-value.

[22] Except, of course, in systems that assume its truth or falsehood. In which case, the truth-value is trivial. Interestingly, a system is consistent regardless of which truth-value is assigned, or even if none are.

That is the curious thing about undecidable statements, their truth-value must be assumed and cannot be proven otherwise by definition. In systems that assume its truth a Euclidean geometry is present, but in systems that assume its falsehood other properties arise. Undecidable statements are strange in this aspect, as they give theoreticians a choice about whether or not to include them in theoretical frameworks.

This harkens back to my statement concerning the bifurcation among epistemologists in the case that the POI is in fact undecidable. Systems of Objective Bayesianism could split into those systems that assume the POI and those that do not. Even then, some systems may be agnostic toward the POI. But what is to motivate an attempt to prove that the POI is undecidable if ultimately such a proof would leave proponents of the principle as proponents and those with a distaste for it with a distaste for it? There are a handful of reasons, not the least of which would be that the focus would shift to the development of partitioning guidelines, a much needed effort if the principle is to be implemented in good faith. However, in the next section, I will lay out some additional information about the POI that an attempt to prove its undecidability may grant.

5.3 A Framework for an Undecidability Proof

When contemplating the nature of how to prove that a statement is unprovable it is easy to become puzzled. After all, if the statement itself is unprovable, why should the statement that it is unprovable be provable? That seems to be at the very least an eyebrow raising issue, especially for those unfamiliar with advanced mathematics. In this section, I hope to elucidate this quandary by laying out the logical foundations of an undecidability proof.

Understanding the result of the example in the previous section is of the utmost importance, as I will be reverse-engineering a proof from it such that it will be more easy to understand not only how but *why* a proof of undecidability works. It is common for these types of statements to start out as intuitive conjectures, which attracts individuals who wish to prove its veracity. Given a conjecture, it is quite informative to carefully examine the information that a conjecture will provide if it is proven to be undecidable. From the last example, we know that once a conjecture is proven to be undecidable, valid logical systems where it is true or false can be constructed without loss of internal consistency. It is also clear that if the conjecture is truly undecidable, it cannot by definition be proven to be true or false unless its truth-value is assumed. Therefore, we may attempt to work backward and assume that the conjecture at hand is in fact true in some system containing only the remaining axioms of the system at hand. Once we assume this much, we must prove that the assumption does not lead to contradiction within the system, and therefore that the system is internally consistent. This step garners only that some system that assumes the truth of the conjecture is not self-contradictory, not that the conjecture is true – that would be circular reasoning since we assumed its truth for this part of the proof. A similar operation can be done but beginning from the assumption that the conjecture is false. If both of these steps are successful, the result is that we've proven that systems containing assumptions of any truth-value for the conjecture are consistent. How does this garner a status of undecidability? Undecidable statements are not provable or disprovable from the remaining axioms of a given system. Because we have shown that each system is consistent, we know that a *reductio ad absurdum* is impossible in any system containing only the remaining axioms for both the conjecture and its negation. This is a nice way to think about the result we've produced, but actually it is the case that this alone tells us that the conjecture is neither provable nor disprovable within the system. Therefore the conjecture, if all of the previous steps are successful, is undecidable.

But what if one of these steps fails? After all, the only way to be guaranteed that this approach will produce the expected result is to know the result beforehand. Let's assume that after assuming the conjecture is true, we find that the system is no longer consistent. This is just a *reductio ad absurdum* for the negation of the conjecture; the conjecture would be assumed true and it would be shown that a contradiction follows from this. Therefore, instead of a proof of the undecidability of the conjecture, a proof that the conjecture is false is produced! A proof that it is true would result from a similar event under the assumption that the conjecture is false. Therefore, even if the attempt to prove the undecidability of the POI fails[23], it fails in such a way that produces a result that is a boon to epistemologists still!

This approach differs importantly from the approach Hawthorne et al. take, as it is one of analytical proof, but not of derivation. Derivation proofs offer no consolation prize inherently if the efforts made do not pan out, while the method described above includes the possibility of failure accompanied by success in a different sense. Furthermore, I believe that this approach is in the spirit of Objective Bayesianism. It would require carefully made axioms, deductive proofs, and a formalization of the system as a whole. This type of detailed and purely logical approach dovetails well with the philosophy that requires that we have detailed logical approaches to all beliefs. This would open up other opportunities as well, for once a system has well defined axioms it is only a matter of time until many theorems are proven within the system, to say nothing of the final result of the attempt itself.

[23] "Fails" here is not meant to convey a failure of human ability to prove a proposition. Rather it is meant to convey that one of the steps needed to prove undecidability is shown to be impossible.

5.4 Conclusion

The Principle of Indifference is a controversial subject among Objective Bayesian philosophers due to its intuitiveness and the existence of what seem to be counter-examples to it. Many attempts have been made to justify the principle, both through derivation and through constraints on the act of partitioning, many if not all of which have failed in their attempts. This is a characteristic of undecidable statements that cannot be proven true or false. An attempt to prove that the principle is an undecidable statement would not only explain the difficulty with which philosophers have in justifying it, but also would provide deductive proof of its truth-value if such an attempt fails. It is because of these reasons that I believe there is a productive future in efforts to prove the principle is undecidable.