Cosmological Fine-Tuning and the Inverse Gambler's Fallacy

Alex Schell (January 2014)

ABSTRACT: According to the fine-tuning argument, the anthropic fine-tuning of our universe confirms the existence of a multiverse. A prominent objection holds that, although a multiverse makes the existence of *some* life-permitting universe more likely than a single universe does, it does not make it any more likely that *our* universe should permit life. Hence, the multiverse hypothesis is not confirmed by the logically strongest description of our evidence. I examine recent replies to this objection and find them to share an implausible implication. I then develop a reply that avoids this implication, by formulating the fine-tuning argument as an instance of anthropic reasoning. This move is shown to be highly profitable, undermining several further objections to the fine-tuning argument.

1 Introduction

Several authors have argued that cosmological *fine-tuning*—the apparent fact that the life-permitting ranges of several physical constants and initial conditions of our universe are very narrow—is evidence for the existence of a vast and heterogeneous ensemble of universes, or *multiverse* (Leslie [1989]; Smart [1989]; Bostrom [2002a], ch. 2; van Inwagen [2002], ch. 9). This *fine-tuning argument* (FTA) runs roughly as follows:[1] fine-tuning suggests that it is unlikely that any given universe will instantiate life-permitting physical constants and initial conditions. Thus, if there is only one universe, it is unlikely to permit life. In contrast, a multiverse provides many independent trials, making it more likely that some universe permits life. Furthermore, since only life-permitting universes are observable, the multiverse hypothesis explains why we *observe* a life-permitting universe.

According to a prominent objection, the FTA trades on a confusion between the evidence that *this* universe permits life and the weaker statement that *some* universe permits life; the stronger statement is claimed to be no more likely on the multiverse hypothesis (Hacking [1987]; White [2000]). Hacking argues that the FTA is analogous to the 'inverse gambler's fallacy' committed by a gambler who enters a casino and, upon witnessing a throw of dice result in a

---

[1] A more familiar form of the fine-tuning argument is the design argument from fine-tuning, which I will not discuss here. For a detailed defense, see (Collins [2009]).

double-six, elevates his credence that the throw he witnessed was not the first one that night. Call this objection the 'inverse gambler's fallacy objection' (IGF).

In this article, I focus on White's ([2000]) formulation of IGF, which avoids some criticisms raised against Hacking's original formulation. I examine several replies to White in the literature and find that all implausibly imply that the FTA succeeds regardless of whether fine-tuning holds. By constructing a reply that steers clear of this implication, I show that White's argument nevertheless fails. Finally, I use the resources developed along the way to defuse several further objections to the FTA.

2 White's Formulation of IGF

White's argument employs the Bayesian account of confirmation, according to which a hypothesis is confirmed when its probability is raised by conditionalization on the evidence:

> Given background knowledge k, evidence e confirms a hypothesis h if and only if $Pr(h|k\&e) > Pr(h|k)$.

Equivalently, e confirms h if and only if the *likelihood ratio* $Pr(e|h\&k)/Pr(e|{\sim}h\&k)$ exceeds 1.

Let m denote the number of actually existing universes. Two hypotheses about the value of m are of interest: the single-universe hypothesis M_1 ($m = 1$) and the multiverse hypothesis M_v ($m = v$, for some $v >> 1$). Assume for simplicity that M_v and M_1 exhaust the possibilities.

White ([2000], p. 231) uses the following simplified model of fine-tuning. There is a large finite set $\{T_1, T_2, \ldots T_n\}$ of possible configurations of the relevant physical constants and initial conditions. Each universe independently instantiates a configuration sampled from a uniform distribution on $\{T_1, T_2, \ldots T_n\}$. T_1 is assumed to be the only life-permitting configuration.

Let E' stand for 'some universe instantiates T_1'. Increasing the number of independent universes makes the instantiation of any given configuration more likely, so E' is more likely on M_v than on M_1. That is, the likelihood ratio $Pr(E'|M_v)/Pr(E'|M_1)$ exceeds 1, and E' confirms M_v.[2] The probability of at least one success in a set of q independent Bernoulli trials, with probability p of success on each trial, is $1 - (1 - p)^q$, so we have $Pr(E'|M_v)/Pr(E'|M_1) = [1 - (1 - 1/n)^v]/(1/n)$. For large n and v, this ratio is quite high, so E' may substantially confirm M_v.

White points out that our relevant evidence is not exhausted by E': we know not only that *some* universe instantiates T_1 but also that *ours* does. (Let E stand for 'α instantiates T_1', where α rigidly designates our universe.) In accordance with the *principle of total evidence* (Good [1967]), we should then conditionalize on E rather than on the strictly weaker E'. White ([2000], p. 232) further argues that, because α instantiates its configuration independently of other

[2] I follow White in leaving the background knowledge k implicit. The choice of k needs to be quite sparse. In particular, k cannot include anything that entails the existence of life, as this would give us a likelihood equality. Sober ([2009]) argues that k must include the existence of observers, which would undermine likelihood formulations of the FTA. Kotzen ([2012]) disputes this, correctly in my view. See also (Monton [2006]; Pust [2007]; Collins [2009], pp. 241-4).

universes, the probability that α instantiates T_1 is just $1/n$, independently of whether M_v or M_1 is true. It follows that $Pr(E|M_v)/Pr(E|M_1) = 1$, and hence that E is neutral with respect to M_v (i.e., confirms neither M_v nor its negation M_1).

A final detail to be made explicit is the *epistemic process* by which the evidence E was obtained. Consider Hacking's ([1987]) gambler story: a gambler enters a casino and witnesses a pair of dice being thrown and coming up double-six. Thinking about whether this was the first throw that night or there were previous throws, the gambler reasons that the double-six outcome is more likely to occur if there are many throws rather than only one, and elevates his credence in many throws. This reasoning is clearly fallacious. Leslie ([1988]) tells a modified story: the gambler is asked to wait outside until a double-six comes up, whereupon he is let in to look at the result. In this case, the gambler does have evidence that the throw he observed is not the first one. Even though the gambler's evidence looks the same in the two scenarios, the crucial difference lies in the epistemic process by which he obtained the evidence.[3]

White points out that in Leslie's story, the inference to the existence of many throws is supported by a link between the *existence* and the *observation* of the outcome: if some double-six comes up at any time, the gambler observes one. There is no such link in Hacking's story, where the gambler observes the outcome of a throw specified independently of its outcome. White ([2000], pp. 237-8)

[3] A classic example of the relevance of the epistemic process is the Monty Hall problem, or the structurally identical 'three prisoners' story (Pearl [1988], pp. 58-62). See also (Bostrom [2002a], pp. 16-8; Bradley [2009]).

argues that our situation resembles Hacking's story: our identity is tied to the particular universe that we are in, so we could not have existed in any universe but *α*.[4] Thus, *α* is like the dice throw specified independently of its outcome: our observation that it instantiates T_1 is independent of whether or not there are other universes.

More formally, we can strengthen the evidence from E ('*α* instantiates T_1') to 'we observe that *α* instantiates T_1' (E_+), where 'observe' is factive and 'we' is a placeholder for your proper name.[5] Since E_+ entails E and so is equivalent to $E\&E_+$, we can expand the likelihood ratio of E_+ as follows:[6]

$$\frac{\Pr\left(E_+ \mid M_v\right)}{\Pr\left(E_+ \mid M_1\right)} = \frac{\Pr\left(E \mid M_v\right)}{\Pr\left(E \mid M_1\right)} \cdot \frac{\Pr\left(E_+ \mid M_v \,\&\, E\right)}{\Pr\left(E_+ \mid M_1 \,\&\, E\right)} \tag{1}$$

If we could not have existed in universes other than *α*, then, due to the causal isolation of *α* from other universes that might exist, the probability of E_+ conditional on E is independent of whether or not there are other universes. That is, E_+ and *m* are conditionally independent given E, or $Pr(E_+|M_v\&E)$ =

---

[4] White appeals to the necessity of material origins to support this view, but claims that his argument requires only that there be no 'mechanism ... linking the multiplicity of universes with our existence' (ibid., p. 238). What White needs is the claim that, conditional on E, our existence is equally likely on M_v and M_1. Barring the implausible assumption that we could have existed in other universes only if *α* had not permitted life, this commits White to assigning zero credence to our existence conditional on *α* not permitting life.

[5] That is, E_+ is the appropriate instance of the schema '<*your name*> observes that *α* instantiates T_1'.

[6] It follows from the definition of conditional probability that $Pr(a\&b|c) = Pr(a|c) \cdot Pr(b|c\&a)$.

Pr(E$_+$|M$_1$&E). By (1) it follows that Pr(E$_+$|M$_v$)/Pr(E$_+$|M$_1$) = Pr(E|M$_v$)/Pr(E|M$_1$). So we have Pr(E$_+$|M$_v$)/Pr(E$_+$|M$_1$) = 1.

Thus, White purports to show that the probabilistic inference from E' to the multiverse hypothesis is defeated by a further relevant component E$_+$ of our total evidence. In the next section, I will discuss the most popular reply to this argument.

3 The Symmetry Reply

Bostrom ([2002a]) points out that White's argument has the odd consequence that once it is known that *some* universe instantiates T_1, learning that α instantiates T_1 disconfirms M_v. This is because the likelihood ratio of E (which is equivalent to E'&E) can be expanded as follows:

$$\frac{\Pr(E \mid M_v)}{\Pr(E \mid M_1)} = \frac{\Pr(E' \mid M_v)}{\Pr(E' \mid M_1)} \cdot \frac{\Pr(E \mid M_v \,\&\, E')}{\Pr(E \mid M_1 \,\&\, E')}$$

Since Pr(E'|M$_v$)/Pr(E'|M$_1$) > 1, it follows that Pr(E|M$_v$)/Pr(E|M$_1$) = 1 only if Pr(E|M$_v$&E')/Pr(E|M$_1$&E') < 1. That is, White's conclusion requires that, given E', conditionalization on E disconfirms M_v, reversing the initial confirmation of M_v by E'. Conversely, conditionalization on a universe other than α instantiating T_1 must then further confirm M_v. But intuitively it seems that once E' is known, learning which particular universe instantiates T_1 should be neutral with respect to M_v, unless α is somehow distinguishably special and different from other universes, independently of its configuration (ibid., pp. 20-1).

White's model does in fact assign special status to our universe: in the model, the existence of α is ensured by both M_v and M_1, whereas every other possible universe can exist only if M_v is true.[7] Bradley ([2005]) and Juhl ([2005]) point out that White's argument only requires the weaker independence assumption that $Pr(\alpha \text{ exists}|M_v) = Pr(\alpha \text{ exists}|M_1)$. This can be shown as follows. Let A stand for 'α exists'. Since E is equivalent to A&E and $Pr(E|M_v\&A) = Pr(E|M_1\&A)$, it follows that $Pr(E|M_v) = Pr(E|M_1)$ if and only if $Pr(A|M_v) = Pr(A|M_1)$.

Both assumptions seem implausible. On the strong assumption that $Pr(A|M_v) = Pr(A|M_1) = 1$, α is the only possible universe whose existence is independent of the number of universes. On the weaker assumption that $Pr(A|M_v) = Pr(A|M_1)$, α still belongs to a select minority of possible universes whose existence is independent of *m*. Bradley ([2005], pp. 12-3) describes the former as 'chauvinism of the highest order' and takes the latter to involve assigning still too high a prior credence to A. It seems reasonable to ensure that our model treats no universe as physically special, or privileged in the actual world (Juhl [2005], pp. 345-6). This consideration motivates what I will call the *symmetry reply* to IGF, which holds that all possible universes have the same chance of actually existing (ibid., p. 346; Oppy [2006], pp. 219-20).

---

[7] For instance, White writes that 'the probability of [α instantiating T_1] is just $1/n$, regardless of how many other universes there are' ([2000], p. 232). Given White's other assumptions, this requires that M_v and M_1 each guarantee the existence of α. This explains the consequence noted by Bostrom: if α is the only universe and some universe instantiates T_1, then α instantiates T_1 ($Pr(E|M_1\&E') = 1$). If there are many universes and some instantiate T_1, α need not be one of them ($Pr(E|M_v\&E') < 1$) (cf. White [2000], pp. 247-8; Draper et al. [2007], pp. 292-3).

Accordingly, suppose that there are finitely many possible universes, of which a random subset of size *m* actually exist. Then for every possible universe, the chance that it exists is m/k, where k ($\geq m$) is the number of possible universes. Let 'SYM' denote the model obtained by adding this assumption to White's model. To find out how the modification affects White's argument, we expand the likelihood ratio of E_+ (equivalent to $A\&E\&E_+$) as follows:

$$\frac{\Pr(E_+ \mid M_v)}{\Pr(E_+ \mid M_1)} = \frac{\Pr(A \mid M_v)}{\Pr(A \mid M_1)} \cdot \frac{\Pr(E \mid M_v \,\&\, A)}{\Pr(E \mid M_1 \,\&\, A)} \cdot \frac{\Pr(E_+ \mid M_v \,\&\, E)}{\Pr(E_+ \mid M_1 \,\&\, E)} \qquad (2)$$

The first ratio is modified in SYM: $Pr(A|M_v)/Pr(A|M_1) = (v/k)/(1/k) = v$. As before, we have $Pr(E|M_v\&A) = Pr(E|M_1\&A) = 1/n$ and $Pr(E_+|M_v\&E) = Pr(E_+|M_1\&E)$. It follows by (2) that $Pr(E_+|M_v)/Pr(E_+|M_1) = v > 1$. This is contrary to White's conclusion, which is thus shown to depend on an implausible asymmetry in his model cosmology.

A majority of replies to White's argument are variations on the symmetry reply as formulated above.[8] The symmetry reply can also be given a purely

---

[8] Besides the replies due to Bradley ([2005]) and Juhl ([2005]), there is a more common variant of the symmetry reply concerned with the identity conditions of universes. White ([2000], pp. 231-2) takes configurations to be contingent properties of universes. Several philosophers have noted that on an alternative metaphysics on which configurations are essential (perhaps defining) properties of the universes instantiating them, E confirms M_v and White's argument fails (Holder [2002]; Manson and Thrush [2003]; Juhl [2005], pp. 343-4; Oppy [2006], p. 220). The metaphysical pictures considered by these authors have in common the suppositions that α necessarily instantiates T_1 and that the actual universes are a random subset of the set of possible universes. Thus, the probability of A is proportional to *m* and the likelihood ratio of E equals *v*. I classify this reply as a variant of the symmetry reply because all inferential work is done by the probability of

evidential formulation that does not assume that universes can be identified across M_v and M_1 or treated like randomly sampled possibilia. It suffices to note that, beyond its configuration, we have no information about our universe that could reverse the initial evidential impact of E' by being more expected on M_1 than on M_v. E' already entails that some *particular* universe permits life; the impact of E' is not reversed unless we know something special about our universe that indicates M_1 (Bostrom [2002a], pp. 20-1). Even if some universes are special like α is in White's model, we would have to know whether or not ours is one of them. We might have had such information—for instance, we might have found that our universe is the first in a temporal sequence of universes of uncertain length (Bradley [2005], p. 14)—but in fact we do not. Ontological symmetry is thus not a necessary component of the symmetry reply; mere evidential symmetry will do.

3.1 The promiscuity objection

In this section, I lay out an objection to the symmetry reply. What I will call the *promiscuity objection* faults the symmetry reply for having the counterintuitive implication that fine-tuning is an irrelevant premise of the FTA. To see why the symmetry reply has this consequence, consider the extreme case of no fine-tuning, where all members of $\{T_1, T_2, \ldots T_n\}$ are equally life-permitting. The existence of a life-permitting universe is now certain regardless of m, and hence neutral with respect to M_v. But the argument still goes through: the fact that α

A being proportional to m; the supposition that T_1 is an essential property of α plays no role. My objection to the symmetry reply in the next section applies equally to this variant.

permits life remains v times more likely on M_v than on M_1, but only because the *existence* of α remains more likely on M_v than on M_1 by the same factor. All inferential work is done by the existence of α; the probability that α permits life plays no role, as long as it takes the same value regardless of how many universes there are.[9]

The same goes for the evidential form of the symmetry reply. Even if we have no evidence about our universe that cancels the evidential impact of E', the magnitude of the impact of E' is independent of whether fine-tuning holds. Supposing again that all configurations permit life, it is certain that some universe *permits life*. But E', the fact that some universe *instantiates* T_1, is still more likely on M_v than on M_1 by the same factor $[1 – (1 – 1/n)^v]/(1/n)$.

Thus, the symmetry reply implies that the FTA has nothing to do with fine-tuning or life-permittingness in particular, instead requiring only that the universe be contingent or improbable in some way. According to the promiscuity objection, this is a problematic consequence. First, the FTA would seem to 'prove too much': mere contingency is trivial evidence, whereas the cosmological question on which it supposedly bears is quite nontrivial. Second, it cannot be that for every possible type of universe, observing a universe of that type confirms M_v (White [2000], p. 246). Unsurprisingly, philosophers writing about fine-tuning have tended to regard a multiverse inference based on mere contingency as

[9] Draper et al. ([2007], p. 296) make an analogous point in reply to Holder's ([2002]) formulation of the symmetry reply, which differs somewhat from mine (see footnote 8).

unreasonable and hold that the FTA requires a more substantive role for life-permittingness.[10]

I propose as a constraint on satisfactory replies to IGF the requirement that they avoid the promiscuity objection. If the FTA is salvageable, it should be sensitive to the degree of fine-tuning, which must play a substantive role distinct from that of contingency alone.

Though motivated by a seemingly genuine shortcoming of White's formulation of IGF, the symmetry reply does not meet this requirement. It is thus at best an incomplete account of where IGF goes wrong. In the following section, I will lay out a general strategy for improving on the symmetry reply in this respect.

4 Sampling Analogies

4.1 Biased sampling preliminaries

Suppose you have two exhaustive hypotheses about the contents of an urn: that it contains exactly one ball (H_1), and that it contains exactly ten (H_{10}). Prior to being placed in the urn, each ball was colored either black or white (based on the outcome of a fair coin toss) and given either a glossy or matte finish (based on

---

[10] While they may disagree about whether or not the FTA has more resources than mere improbability or contingency, friends and foes of the argument generally agree that a multiverse inference from mere contingency is unreasonable (Leslie [1989], ch. 5; Smith [1994]; White [2000], p. 246; Bostrom [2002a], pp. 23-31; Heller [2008]). Similar views are expressed in (Manson [2000], pp. 346-7; McGrew et al. [2001], p. 1032; Oppy [2006], p. 226). An exception is Juhl ([2005], p. 344), who accepts that fine-tuning is irrelevant to the inference.

another fair coin toss). You obtain a one-ball sample from the urn, which is white and glossy. How does this information bear on H_{10} and H_1?

One might think that H_{10} is confirmed, since the urn is more likely to contain a glossy white ball on H_{10} than on H_1, by a factor of $[1 – (1 – 0.25)^{10}]/0.25 \approx 3.8$. Yet this reasoning leaves out some evidence: you know not only that there is a glossy white ball in the urn (GW) but also the logically stronger fact that a glossy white ball was sampled from the urn (GW*). The reasoning would still lead to the right conclusion if the sampling process were maximally biased toward glossy white balls, selecting a glossy white ball just in case there are any in the urn (thus making GW and GW* coextensive).

If the ball is drawn randomly, GW* is neutral with respect to H_{10} and H_1. A glossy white ball, though more likely to exist on H_{10}, is equally likely to be randomly sampled on either hypothesis. The probability of obtaining a glossy white ball by random sampling equals the expected fraction of such balls in the urn, which is just 0.25 independently of H_{10} or H_1. Bradley ([2009]) discusses similar examples, drawing the lesson that the output of biased sampling (but not random sampling) carries information about the number of balls.[11]

[11] If this lesson seems odd at first, this may be because the urn example is quite unusual. Much familiar statistics concerns inferences from a sample to the distribution of properties in a population of large or irrelevant size (that is, estimation of *population parameters* from a sample). For this purpose, random samples are more straightforwardly informative than biased samples. In contrast, Bradley and I are interested in making inferences about the unknown size of a population with a known distribution.

Suppose you instead used *white-biased sampling*, a selection procedure that randomly selects a ball from the set of white balls in the urn (if there are any) or selects nothing (if there are none). GW* entails that there is a white ball in the urn (W), so we have

$$\frac{\Pr\left(GW^{*} \mid H_{10}\right)}{\Pr\left(GW^{*} \mid H_{1}\right)} = \frac{\Pr\left(W \mid H_{10}\right)}{\Pr\left(W \mid H_{1}\right)} \cdot \frac{\Pr\left(GW^{*} \mid H_{10} \,\&W\right)}{\Pr\left(GW^{*} \mid H_{1} \,\&W\right)} \tag{3}$$

(These probabilities are conditional on the use of white-biased sampling.) White-biased sampling selects a random white ball, so (conditional on W) the probability that the sampled ball is glossy equals the expected fraction of glossy balls in the set of white balls: $Pr(GW^*|H_{10}\&W) = Pr(GW^*|H_1\&W) = 0.5$. It follows by (3) that the likelihood ratio of GW* equals that of W, which is $[1 – (1 – 0.5)^{10}]/0.5 \approx 2.0$. The apparent lesson is that when a selection procedure combines bias and randomness, its biased component determines the likelihood ratio for a sample of one. (Note also that it would be misguided to object that the ball had the same chance of being glossy white whether or not it was the only ball in the urn.)

To generalize, consider a population consisting of an unknown positive number N of objects. Each object may or may not exhibit some arbitrary properties p and q, which are assigned to each object by a stochastic process that does not depend on N and has no memory (such that the properties of distinct objects are independent and identically distributed random variables, with a distribution independent of N). Let p-*biased sampling* be a selection procedure that randomly selects exactly one object from the set of objects exhibiting p (‘p-objects’) if there are any in the population, or else selects nothing. We define the following:

(K_p) *p*-biased sampling is used

(P) There is a *p*-object in the population

(PQ*) An object exhibiting *p* and *q* (a '*p*&*q*-object') is selected

Our conclusion is that, given *p*-biased sampling and assuming that the choice of this selection procedure is independent of the population contents, PQ* and the population size *N* are conditionally independent given P. It follows that for any hypotheses H_i and H_j about the value of *N* to which nonzero credence is assigned, we have

$$\frac{\Pr\left(PQ^*|H_i \& K_p\right)}{\Pr\left(PQ^*|H_j \& K_p\right)}=\frac{\Pr\left(P|H_i\right)}{\Pr\left(P|H_j\right)} \tag{4}$$

In other words, P captures all information in PQ* that is relevant to the population size. The extra information, that the selected *p*-object exhibits *q*, is neutral.

To show why this conclusion holds, we first expand the likelihood ratio of PQ* (equivalent to P&PQ*):

$$\frac{\Pr\left(PQ^*|H_i \& K_p\right)}{\Pr\left(PQ^*|H_j \& K_p\right)}=\frac{\Pr\left(P|H_i \& K_p\right)}{\Pr\left(P|H_j \& K_p\right)}\cdot\frac{\Pr\left(PQ^*|H_i \& K_p \& P\right)}{\Pr\left(PQ^*|H_j \& K_p \& P\right)}$$

As K_p is independent of the population contents, P and K_p are independent conditional on *N*: we have $Pr(P|H_i\&K_p) = Pr(P|H_i)$ and $Pr(P|H_j\&K_p) = Pr(P|H_j)$. Conditional on P, the probability that a *p*&*q*-object is selected equals the expected fraction of *p*&*q*-objects in the set of *p*-objects. This is just the probability that a given *p*-object exhibits *q*, which is independent of *N*, as each object is independently assigned its properties from a distribution independent of *N*.[12] That is, we have $Pr(PQ^*|H_i\&K_p\&P) = Pr(PQ^*|H_j\&K_p\&P)$. Thus we obtain (4).

We can weaken the assumption that the properties of distinct objects are independent. All we need for (4) is for *q* to be instantiated 'symmetrically' in *p*-objects, in the sense that the expected fraction of *p*&*q*-objects in the set of *p*-objects is independent of *N*. As long as the distribution of *q* in *p*-objects does not depend on *N*, this condition may also be satisfied if the *q*-states of distinct *p*-objects exhibit dependencies, for instance if they are like members of a sample drawn without replacement.[13]

4.2 Bradley's reply to White

This result provides an outline for constructing versions of the FTA that avoid the promiscuity objection. Here is one way to fill out this outline: suppose, fancifully, that (i) we observe some life-permitting universe just in case there are any, and furthermore that (ii) we are equally likely to observe any existing life-permitting

[12] This does not require that *p* and *q* be independent of each other, as color and finish are in the urn example. Nothing here hinges on the probability that a not-*p*-object exhibits *q*, so it does not matter how this compares to the probability that a *p*-object exhibits *q*.

[13] Let $\{X_1, X_2, \ldots X_n\}$ be the set of indicator variables for the *q*-state of each of *n* *p*-objects. Then *q* is symmetrically instantiated in *p*-objects if and only if the expected value of $\Sigma X_i/n$ is independent of *n*, or equivalently, ΣX_i is proportional to *n*. If the X_i are independent and identically distributed, their sum follows the binomial distribution, which satisfies this condition. If the X_i are like members of a sample drawn without replacement, their sum follows the hypergeometric distribution, also satisfying this condition.

universe. For the purposes of inferring the number of universes, we can then model our observations about the universe as the output of *life-biased sampling*: a selection procedure that selects a random life-permitting universe just in case there are any. Provided that the identity and precise configuration of a life-permitting universe are symmetrically instantiated properties, we expect that the degree to which the multiverse hypothesis is confirmed tracks the degree to which it raises the probability that a life-permitting universe exists.

Bradley ([2009], [2012]) takes a similar approach in his proposed reply to IGF, which combines the symmetry reply with premises about the modal metaphysics of our existence. In particular, Bradley argues against White's assumption that we could only have existed in α, and in favor of the view that we could have existed in any life-permitting universe. Our assumption (ii) is a natural extension of this view.[14] Furthermore, Bradley assumes a weakened form of (i). Though Bradley is not concerned with the promiscuity objection, it is instructive to look at the conditions in which his approach can avoid it.

Recall that in SYM we have a set of k possible universes, of which a random subset of size m actually exists, and life-permittingness (T_1) is instantiated independently with probability $1/n$ in each universe. Consider the more general model SYM# where T_1 is one of l life-permitting configurations, where $1 \leq l \leq n$. Let E" stand for 'some universe permits life' and E_+" for 'we observe a life-permitting universe'. We have

[14] For the moment, I ignore the possibility of configurations that are life-permitting to different degrees, or permit only different forms of life.

$$\frac{\Pr(E_+ \mid M_v)}{\Pr(E_+ \mid M_1)} = \frac{\Pr(E'' \mid M_v)}{\Pr(E'' \mid M_1)} \cdot \frac{\Pr(E_+'' \mid M_v \,\&\, E'')}{\Pr(E_+'' \mid M_1 \,\&\, E'')} \cdot \frac{\Pr(E_+ \mid M_v \,\&\, E_+'')}{\Pr(E_+ \mid M_1 \,\&\, E_+'')} \tag{5}$$

Bradley makes the simplifying assumption that (i') our observing a life-permitting universe is conditionally independent of *m* given that some universe permits life, that is, that $Pr(E_+''|M_v\&E'') = Pr(E_+''|M_1\&E'')$. Individually and jointly, the properties 'being α' and 'instantiating T_1' are symmetrically instantiated in life-permitting universes: the probability that a random life-permitting universe is α and instantiates T_1 is independent of *m*.[15] By (ii), it follows that $Pr(E_+|M_v\&E_+'') = Pr(E_+|M_1\&E_+'')$, by which (5) simplifies to $Pr(E_+|M_v)/Pr(E_+|M_1) = Pr(E''|M_v)/Pr(E''|M_1) = [1 - (1 - l/n)^v]/(l/n)$.

E_+ confirms M_v precisely to the degree to which E" is more likely on M_v than on M_1, which in turn depends on the fraction of life-permitting configurations l/n. The promiscuity objection is thus avoided. Once the instantiation of life-permittingness is conditionalized upon, the identity and precise configuration of the selected life-permitting universe carry no further

[15] Independently of *m*, we assign credence $1/l$ to a random life-permitting universe instantiating T_1, and credence $1/k$ to a random T_1-universe being α. (We assign credence $1/k$ to a *random universe* being α. As the identity and configuration of a universe are independent, we assign the same credence to a *random T_1-universe* being α.) It follows that we assign credence $1/lk$ to a random life-permitting universe instantiating T_1 and being α, independently of *m*. For closer analogy to the urn example, the identity of a universe may be thought of as a label assigned randomly without replacement to each actual universe. Then each universe, regardless of its configuration, has a chance $1/k$ of being assigned the label 'α'.

information about the value of *m*, as they are symmetrically instantiated properties subject to random sampling.

In arriving at this result, we relied on the simplifying assumption (i'). Given (ii) and an innocuous independence assumption, what is necessary for the conclusion is the following, slightly weaker assumption:[16]

(i") our existence and *m* are conditionally independent given E".

This assumption is implausible. E" does not guarantee the instantiation of all physical properties necessary for our existence (the existence of a certain type of nervous system, with a certain environment and history, etc.). Hence, our existence is plausibly more likely if many universes permit life than if only one does. Even if each universe were spatially infinite (in which case E' might guarantee our existence), other life-permitting configurations might be sufficiently different from ours to preclude our existence (hence E" would still not guarantee our existence).

Without (i"), the present approach is susceptible to the promiscuity objection. For suppose that our particular form of life can only exist in a small subset of all life-permitting configurations. With the appropriate analogue of (ii), the present approach now gives the verdict that the likelihood ratio of E_+ exceeds that of E". Even if life-permitting configurations were common, E_+ might still substantially confirm M_v, not because life is unlikely, but because human beings

[16] Within SYM#, (ii) and the assumption that E_+" and *m* are conditionally independent given our existence (cf. Bradley [2009], p. 66, [2012], p. 164) entail that the likelihood ratio of E_+ equals that of E" just in case (i") holds.

are unlikely. Proponents of the FTA have tended to disavow this anthropocentric style of argument (Leslie [1989], pp. 136-7; see also Carter [1983]).

None of this is a problem for Bradley, who is not concerned with the promiscuity objection and employs (i’) only as a conservative simplification ([2012], p. 166). But if we are to avoid the promiscuity objection, we are forced to look elsewhere.

5 Anthropic Reasoning with Indexical Information

In this section I build on the outline in section 4.1 to develop a version of the FTA that enables the symmetry reply to successfully avoid the promiscuity objection without relying on controversial assumptions about the modal metaphysics of our existence. This version of the FTA is also more faithful to the FTA as originally proposed. Early proponents such as Leslie ([1989]) stress that the FTA is an instance of *anthropic reasoning*, a methodology aiming to correct for the *observation selection effects* arising from the fact that our observations of the universe are restricted to ‘anthropic’ circumstances necessary for the existence of observers.[17] While often verbally mentioned in presentations of the FTA, this

[17] Cf. (Carter [1983]; Livio and Rees [2005]). Anthropic reasoning is often motivated by contrasting it with the ‘Copernican’ alternative of assuming that our observed circumstances are a representative sample of the set of circumstances in the world. Anthropic reasoning prescribes (roughly) that we consider our observed circumstances a representative sample of the set of circumstances in which observers can find themselves. This involves treating our observed circumstances as an ‘anthropically biased’ sample of the set of circumstances in the world (Carter [2006]).

dimension of the FTA is almost never incorporated into formalizations of the argument (with the sole exception of Bostrom [2002a], pp. 185-92).[18]

The anthropic character of the FTA is made explicit by adopting the *self-sampling assumption* (SSA), an explication of anthropic reasoning (Bostrom [2002a], ch. 3). I use SSA to take into account the indexical ('self-locating') statement 'I observe that α instantiates T_1' (E*), which is implied by our having access to the proposition E_+. I show that, in conjunction with the symmetry reply, this approach—which I call the *anthropic FTA*—can approximate the results obtained in the previous section and thus rebut IGF while steering clear of promiscuity. Furthermore, the anthropic FTA avoids a number of additional objections to the FTA.

5.1 The self-sampling assumption

Bostrom ([2002a], p. 57) defines SSA as follows:

(SSA) One should reason as if one were a random sample from the set of all observers in one's reference class.

Treating oneself as a random observer in this metaphorical sense amounts to having prior credences subject to the following constraint: given that the world is

[18] Here I commit myself to the perhaps controversial claim that what has come to be called the *anthropic objection* to the FTA (Sober [2004], [2009]; Pust [2007]) has little to do with anthropic reasoning. Yet regardless of whether the 'anthropic' label is more aptly applied to my forthcoming approach or to the objection by Sober and Pust, it should be uncontroversial that the two are radically different from each other. For instance, the anthropic version of the FTA will be shown to be in a unique position to *avoid* the anthropic objection.

such that a fraction *f* of observers in one's reference class have some property *p*, one assigns a prior credence *f* to the indexical statement 'I am an observer with property *p*' (Bostrom [2002b], p. 619).

SSA has implications for how indexical and non-indexical beliefs interact. Here is a thought experiment due to Bostrom that illustrates how SSA connects indexical evidence with non-indexical hypotheses:

> In an otherwise empty world there are three rooms. God tosses a fair coin and creates three observers as a result, placing them in different rooms. If the coin falls heads, He creates two observers with black beards and one with a white beard. If it falls tails, … He creates two whitebeards and one blackbeard. All observers are aware of these conditions … [and] know the color of their own beard. You find yourself in one of the rooms and you see that you have a black beard. What credence should you give to the hypothesis that the coin fell heads? (Bostrom [2002b], p. 619)

While both heads (H) and tails (T) are perfectly compatible with your non-indexical evidence ('there is a blackbeard'), you also know the indexical fact 'I am a blackbeard' (B), which SSA brings to bear on H and T. Assuming that all observers (but not God) are in your reference class, SSA prescribes that your conditional prior credences for B match the fraction of blackbeards: $Pr(B|H) = 2/3$ and $Pr(B|T) = 1/3$. Given a prior credence $Pr(H) = 0.5$, conditionalization on B yields the posterior credence $Pr(H|B) = 2/3$. H is confirmed because your observations are more typical on H, in the sense that a greater fraction of observers are making them (ibid., pp. 620-1).

More generally, SSA licenses us to treat indexical statements of the form 'I observe that *e*' as methodologically equivalent to 'a randomly chosen observer observes that *e*'. Note that this entails that there are observers in the first place. So, other things equal, SSA favors hypotheses proportionally to how likely they make the existence of observers. Conditional on the existence of observers, SSA favors hypotheses proportionally to how likely they make it that a randomly chosen observer makes your observations. In other words, SSA licenses us to model our observations as the output of an observer selection procedure that selects the observations of a random observer just in case there are observers.

5.2 The anthropic FTA

It is worth noting that without the symmetry reply, the anthropic FTA does not suffice to plausibly reply to White's formulation of IGF. This is because the application of SSA to White's model has the result that E* *disconfirms* M_v. We can intuitively see that matters would be somewhat complicated. Since the physical contents of distinct universes are independent and identically distributed, all observer-containing universes have the expected number of observers, and thus SSA prescribes that we expect to find ourselves in a random observer-containing universe. Finding ourselves in α, the only universe if M_1 is true, but one of many if M_v is true, would seem to support M_1. There is a tension between two elements of our evidence: M_v increases the probability that there are observers but decreases the expected fraction of observers finding themselves in α. The latter effect dominates, resulting in net disconfirmation of M_v.[19]

We can construct a more promising reply to White by combining the anthropic FTA with the symmetry reply. In what follows, I show this by evaluating the conclusions of the anthropic FTA in SYM#, beginning with a simplified, informal argument.

E* entails the existence of an observer-containing universe (R'), so we have

$$\frac{\Pr\left(E^*|\,M_v\right)}{\Pr\left(E^*|\,M_1\right)} = \frac{\Pr\left(R'|\,M_v\right)}{\Pr\left(R'|\,M_1\right)} \cdot \frac{\Pr\left(E^*|\,M_v \,\&\, R'\right)}{\Pr\left(E^*|\,M_1 \,\&\, R'\right)} \qquad (6)$$

Conditional on there being observers, any further SSA-licensed evidential impact of E* must stem from the expected observations of a random observer. In particular, we are interested in whether M_v and M_1 differ in the probabilities they assign to a random observer finding herself in the T_1-universe α. Two assumptions are sufficient to establish the absence of such a difference: first, that the expected number of observers in a universe is independent of the universe's identity (this follows from our assumption that the physical contents of distinct universes are independent and identically distributed random variables); second, that every life-permitting configuration has the same expected number of

[19] Consider the likelihood ratio of 'some universe contains observers'. If each universe has an independent chance r of containing observers, this ratio equals $[1 - (1 - r)^v]/r$, which is strictly smaller than v (it is strictly decreasing on $r \in (0,1]$ and its limit as $r \to 0$ is v). Given that some universe contains observers, then on M_1 we are certain to observe α, but on M_v we assign only credence $1/v$ to observing α. The net likelihood ratio is thus strictly smaller than 1.

observers (this a strong assumption that we will need to relax). Given these assumptions, the expectation of the identity and configuration of the universe inhabited by a random observer equals the expectation of the identity and configuration of a random life-permitting universe, which in section 4.2 was seen to be independent of *m* in SYM#. Hence, SSA prescribes that $Pr(E^*|M_v\&R') = Pr(E^*|M_1\&R')$, from which it follows by (6) that $Pr(E^*|M_v)/Pr(E^*|M_1) = Pr(R'|M_v)/Pr(R'|M_1) = [1 - (1 - r)^v]/r$, where *r* is the probability that a given universe contains observers. I establish this conclusion formally in the appendix.

The degree of confirmation of M_v is thus sensitive to the degree of fine-tuning, which plays a substantive role in virtue of influencing how likely a universe is to contain observers.[20] As in section 4.2, the identity and configuration of an observer-containing universe do not carry information about *m* because they are symmetrically instantiated properties subject to (metaphorical) random sampling.

We can now relax the assumption that all life-permitting configurations have the same expected number of observers.[21] If we drop this assumption, the above conclusion does not apply to E* but only to the weaker statement ‘I observe

[20] Note that, although (other things equal) the magnitude of confirmation of M_v increases with the degree of fine-tuning, fine-tuning is not necessary for confirmation of M_v. Fine-tuning plays a substantive role in the anthropic inference to M_v solely in virtue of making the existence of observers unlikely in any given universe. If observers were unlikely even in life-permitting universes ($r << l/n$), M_v is substantially confirmed even if l/n is high.

[21] [Acknowledgement removed for blind review]

that α permits life'. It remains an open question whether our observed configuration carries additional information about *m*.

Let λ denote the *anthropic index* of T_1: the ratio of the expected number of observers in a T_1-universe to that in a universe with another life-permitting configuration. Recall that in SYM# there are l life-permitting configurations including T_1. The probability of observing T_1 given that exactly p universes permit life is given by

$$g(p;\lambda,l)=\sum_{i=0}^{n}\binom{p}{i}(1/l)^{i}(1-1/l)^{p-i}\frac{\lambda i}{\lambda i+p-i}$$

We can put bounds on the additional evidential impact of observing T_1 (conditional on 'I observe that α permits life') by evaluating the likelihood ratio for the limiting case where $p\to\infty$ and $p = 1$, respectively:

$$\lim_{x\to\infty}\frac{g(x;\lambda,l)}{g(1;\lambda,l)}=\frac{l\lambda}{\lambda+l-1} \tag{7}$$

For $\lambda = 1$, this expression equals 1. Hence, conditional on 'I observe that α permits life', observing T_1 is neutral if T_1 has an average anthropic index. For $\lambda >$ 1, the ratio exceeds 1; observing T_1 further supports M_v. For $\lambda < 1$, the ratio is smaller than 1: observing T_1 disconfirms M_v, potentially more so than R' confirms M_v. (What is happening here is that at $p = 1$, the expectation of the observed configuration is determined solely by the physical probability distribution over life-permitting configurations, whereas at $p\to\infty$, the expectation is a full 'anthropic weighting' of the physical probabilities. The expectation of observing T_1 in the infinite limit is increased by the anthropic weighting if $\lambda > 1$ and decreased if $\lambda < 1$.)

Though these results add provisos to the FTA, they are not otherwise problematic. In particular, the result in the case of $\lambda > 1$ does not make the anthropic FTA vulnerable to the promiscuity objection. Rather, it is a direct consequence of paying attention to degrees of observer-friendliness in addition to the less discriminating predicate of life-permittingness. Consider the case in which all life-permitting configurations but T_1 are minimally anthropic, such that λ is very large. We should expect this to behave similarly to the case where T_1 is the only life-permitting configuration. Indeed, as λ increases, the likelihood ratio in (7) approaches *l*, such that the net likelihood ratio of E* approximates *n*, as expected if T_1 is the only life-permitting configuration.[22]

By analogy, we should expect the case of $\lambda \approx 0$ (T_1 being minimally anthropic) to have results similar to those of failing to observe any life-permitting universe, which would 'falsify' the hypothesis of a sufficiently large multiverse. In agreement with this intuition, the likelihood ratio in (7) approaches 0 as λ approaches 0. This points to an empirical possibility in which the FTA would fail on its own terms: if our configuration is minimally anthropic, the evidential impact of R' is outweighed by the impact of observing T_1, resulting in net disconfirmation of M_v even if life-permitting configurations are rare. The confirmation of M_v thus continues to be sensitively dependent on the degree of fine-tuning, where this is understood to relate to the observer-friendliness (a

[22] For a sufficiently large v, and assuming that $r \approx l/n$, the likelihood ratio of 'I observe that α permits life' approximates $1/(l/n) = n/l$. If T_1 is the only life-permitting configuration, the likelihood ratio of E* approximates $1/(1/n) = n$.

continuous feature) and not merely to life-permittingness (an all-or-nothing predicate).

The finer discrimination afforded by dropping the simplifying assumption is not only unproblematic, it is also a considerable virtue in at least two respects. First, it makes the FTA robust to the possibility of *freak observers*. These are observers produced by random fluctuations (of a thermal or quantum nature). In anthropic universes, the density of freak observers is astronomically lower than that of evolved observers; this need not be the case in non-anthropic universes, where freak observers may dominate in relative terms. One problem posed by freak observers is that even minimally anthropic universes dominated by freak observers are strictly speaking life-permitting. It seems that this detail should not matter much: a sufficiently low observer density should be similar to zero observer density. Yet the possibility of freak observers has a discontinuous effect on simple versions of the FTA that use the existence of a life-permitting universe (E") or of observers (R') as their evidence base. Consider again the case where all life-permitting configurations but T_1 are minimally anthropic, dominated by freak observers ($\lambda >> 1$). If we further suppose that universes are sufficiently large and that most configurations support freak observers (such that $r \approx l/n \approx 1$), E" and R' are each approximately neutral with respect to M_v. If we drop the observer density in non-T_1 configurations from any positive ε to 0, the likelihood ratio (of either E" or R') makes a discontinuous jump from approximately 1 to approximately *n*. As we have seen above, the anthropic FTA does not exhibit this counterintuitive discontinuity.[23]

Second, the anthropic FTA avoids an objection due to Pust ([2007]), who argues for a strong version of the Bayesian problem of old evidence. Pust holds that on any plausible account of epistemic probabilities, we cannot assign credence less than 1 to the existence of observers, even when considering hypothetical credences that 'subtract' old evidence from our background knowledge. This has the result that the existence of observers cannot play the role of positively relevant evidence in a Bayesian argument. Pust concludes that this 'strong problem of old evidence' undermines the FTA. Sober ([2004], [2009]) objects to the design argument from fine-tuning on similar grounds, and his objection also applies to the FTA. Though the objection—confusingly termed the *anthropic objection*—has been criticized (Weisberg [2005]; Monton [2006];

[23] Freak observers may nevertheless threaten the wider relevance of the FTA. Since fine-tuning is often taken to support the disjunctive conclusion that 'God is real and/or there exist vastly many, very varied universes' (Leslie [1989], p. 204), it is of interest how the multiverse hypothesis fares not only relative to the single-universe hypothesis, but also relative to the design hypothesis. Collins ([2009], pp. 265-71, [2013]) argues that the anthropic advantage enjoyed by anthropic configurations is insufficiently small to outweigh their relative physical improbability among nominally life-permitting configurations. As a result, the multiverse is dominated by freak observers and does not explain fine-tuning for evolved observers. I will not comment on this argument except to say that the present framework greatly clarifies what is at issue: the problem is that, if $1 < \lambda << l$, the probability of observing T_1 is, though higher on M_v than on M_1, very low on M_v, potentially sufficiently low to give the design hypothesis a decisive likelihood advantage over M_v.

Kotzen [2012]), the anthropic FTA can succeed even if these criticisms fail. The upshot of the anthropic objection is that, given what the implicit background knowledge *k* must contain, we have $Pr(\text{R'}|\text{M}_v)/Pr(\text{R'}|\text{M}_1) = 1$. Though this suffices to undermine simpler versions of the FTA, the anthropic FTA remains successful, provided that $\lambda > 1$.

These strengths come at the price of requiring more detail in the empirical fine-tuning premise of the FTA. Depending on whether or not we want to avoid the anthropic objection, we require either that $\lambda > 1$ or else that λ be at least large enough for our observation of T_1 to not completely reverse the evidential impact of R'. This is a previously unnoticed requirement of the FTA. Given that most life-permitting configurations are dominated by freak observers at extremely low densities, even the stronger burden could plausibly be met, though such a task is beyond the scope of this article.

Whether or not the anthropic FTA is ultimately successful, we have shown it to be substantially stronger than simpler formulations of the FTA. In conjunction with the symmetry reply, the anthropic FTA provides a satisfactory reply to IGF that avoids the promiscuity objection. It is furthermore robust in the face of freak observers and may avoid Sober's and Pust's anthropic objection. In the following section, I show that the anthropic FTA successfully avoids two further objections to the FTA.[24]

[24] Note that E*, the description of the evidence I have focused on, omits some information in E_+: the total evidence $\text{E*\&E}_+$ is equivalent to E* & 'I am <*your name*>'. Our conclusions will remain unaltered as long as 'I am <*your name*>' and *m* are conditionally independent given E*. This independence assumption holds true if, given that α is a certain way, what happens outside of α

6 What Anthropic Reasoning Can Explain

Several philosophers have raised worries about the FTA that stem from the fact that the multiverse hypothesis is *indiscriminate*, in the sense that it raises the probability of the existence of every type of universe, and not only of intuitively noteworthy universes such as life-permitting ones. One of these worries is the promiscuity objection to the FTA, discussed in (Smith [1994]; White [2000], p. 246; Heller [2008]). Having shown how the anthropic FTA avoids this objection, I will next consider two related recent objections and, in doing so, highlight some general features of anthropic reasoning as explicated by SSA.

The first objection aims to show that the FTA presupposes religious and axiological premises. Following authors such as Bradley ([2001]), Manson ([2003]) holds that the FTA requires the premise that fine-tuning is surprising or 'in need of an explanation' in a way that mere contingency is not. On a standard probabilistic analysis of surprise, this means that there is 'some initially implausible (but not wildly implausible) [explanation] K' that is confirmed by fine-tuning but not by mere contingency (Horwich [1982], pp. 101-2). Manson

---

does not affect what individuals are expected to exist in α (thus not affecting the SSA-prescribed prior credence assigned to 'I am <*your name*>'). This condition is met if we restrict ourselves to qualitative descriptions of individuals. If we admit numerical identity (and non-observational knowledge of it), the condition is still implied by White's view that our identity is tied to the particular universe we inhabit. Our conclusions would require modification only if we were to admit transcendent knowledge of identity facts and a metaphysical view on which the individuals expected to exist in α depend on whether there are other universes.

argues that due to its indiscriminateness, the multiverse hypothesis cannot play the role of K. The other candidate is the design hypothesis, which, in conjunction with an axiological premise about the value of life, can play the role of K and thereby support the premise that fine-tuning specifically is surprising. In the absence of other candidates for K, the FTA thus requires that the design hypothesis be taken seriously, for instance that it not be 'wildly implausible'.

With its strengthened evidence base, the anthropic FTA is not susceptible to this objection. While it is true that the *existence* of every type of universe can be made arbitrarily likely in a sufficiently large universe, this is not true of *observations* of universes: it is not true that every type of universe is likely to be observed by a random observer in a sufficiently large multiverse. We have seen that the multiverse hypothesis is confirmed by fine-tuning: it raises the probability that a random observer observes a life-permitting universe. But it is not confirmed by mere contingency in the absence of an observation selection effect: given some life-permitting universe (and assuming that all life-permitting universes are equally so), the multiverse hypothesis does not raise the probability that a random observer observes any particular type of life-permitting universe. This is for the same reason that the glossy finish of a white ball obtained by white-biased sampling is neutral with respect to the population size: the relevant property is symmetrically instantiated and subject to random sampling. Hence, the multiverse hypothesis can play the role of K in supporting the premise that fine-tuning specifically is surprising, and the design hypothesis need not be considered when setting up the FTA.[25]

The second objection takes the form of a skeptical challenge. It holds that if the FTA is valid and the multiverse hypothesis explains the fine-tuning of our universe, it likewise explains any other physically possible but locally improbable events, which are arbitrarily likely to obtain in a sufficiently large multiverse (Craig [2003]; Collins [2009], pp. 256-62; Plantinga [2011], pp. 213-4). Thus, if the FTA is valid, the multiverse hypothesis undercuts most scientific and ordinary probabilistic reasoning. For example, Plantinga claims that, if the multiverse hypothesis explains fine-tuning, it also explains the suspiciously good fortune of a poker player, undercutting the commonsense explanation that he is cheating. Thus, unless the proponent of the FTA can explain why the explanation works in one case but not the other, the FTA is suspect.

This objection also trades on an inadequately weak description of the relevant evidence. It is true that in a sufficiently large multiverse it is likely that some honest poker player is dealt a suspiciously convenient series of hands. The same is not true of the indexically strengthened evidence, taken into account by SSA: a random observer in the multiverse is more likely to be dealing with a cheat, as successful cheats are plausibly more common in the multiverse than are honest players blessed with comparably convenient cards. Hence the multiverse

[25] The careful reader will notice that the probabilistic analysis of Manson's surprisingness premise (with $K = M_v$) is precisely the *conclusion* of the anthropic FTA. Hence the surprisingness of fine-tuning plays no substantive role as a *premise* of the argument (cf. Bostrom [2002a], pp. 31-2).

hypothesis does not screen off the evidence from its commonsense explanation. More generally, in the absence of observation selection effects, differences between hypotheses in the conditional probability they assign to individual instances of an event of type *e* translate straightforwardly to differences in the fraction of observers observing an instance of *e*. Differences of the latter type allow SSA to bring an observation of *e* to bear on the competing hypotheses, maintaining the intuitive connection between theory and observation even in a very large multiverse.[26]

The lesson is that only properties correlated with the presence of observers, and thus subject to an observation selection effect, are candidates for anthropic explanation. Other properties are symmetrically instantiated and subject to random sampling, and therefore do not support (and are not explained by) ensemble hypotheses.[27] This insight alone has sufficed to dispel three objections to the FTA. Though it is by no means a new result (see Leslie [1989], pp. 123-4; Bostrom [2002a], pp. 189-90 for informal and formal statements, respectively), the scarcity of formal statements of the anthropic FTA has led to this lesson remaining underappreciated in the recent literature.

---

[26] See (Bostrom [2002b]) for a detailed exposition of this point in the context of a single, spatially infinite universe.

[27] These other properties include not only properties uncorrelated with the presence of observers, but also any properties *p* whose correlation with the presence of observers is screened off by a more basic property *p*' (where instantiating *p* entails instantiating *p*'). For example, if T_1 is one of *l* equally anthropic configurations, the multiverse hypothesis is supported by our observation of an anthropic configuration but is not further supported by our observation of T_1.

7 Conclusion

I started by examining White's formulation of IGF and several replies to it. The symmetry reply, a prominent and *prima facie* compelling reply to White's formulation of IGF, proved vulnerable to the promiscuity objection: all its extant formulations have the problematic consequence that the multiverse hypothesis is confirmed, not just by fine-tuning, but by contingency in any form. I proceeded to introduce the biased sampling framework, which I used to sketch a general route to circumvent the promiscuity objection. We found that Bradley's ([2009], [2012]) strengthening of the symmetry reply is capable of avoiding the promiscuity objection, though only with help from an implausible independence assumption.

I went on to show that the anthropic FTA, which uses SSA to take into account an indexical component of our evidence, enables the symmetry reply to avoid the promiscuity objection. The anthropic FTA has additional virtues. Unlike simpler versions of the FTA, it is robust to the possibility of freak observers, can avoid a prominent objection due to Sober ([2004], [2009]) and Pust ([2007]), and is immune to two objections to the FTA from the indiscriminateness of the multiverse hypothesis. I have also pointed out a previously unnoticed vulnerability: the success of the anthropic FTA requires that the observer density of our configuration not be too far below average among life-permitting configurations.

I conclude that, properly formulated, the FTA does not commit any fallacy of probabilistic reasoning and is remarkably resilient in the face of miscellaneous objections. The anthropic FTA thus represents a substantial improvement over most versions of the argument discussed in the recent literature. I do not, of course, claim to have shown that the anthropic FTA ultimately succeeds. Its prospects depend on matters beyond the scope of this article, such as on whether fine-tuning can be described as improbable in the required sense (McGrew et al. [2001]; Colyvan et al. [2005]; Monton [2006]; Collins [2009], pp. 249-52), or whether the observer density of our configuration is sufficiently large. However, while the FTA may ultimately succumb to these problems or others, none of the objections discussed in this article are fatal to the argument.

Acknowledgements

[removed for blind review]

Appendix

Let R stand for 'α contains observers'. As E* entails R, we can expand the likelihood ratio of E* as follows:

$$\frac{\Pr\left(E^*|M_v\right)}{\Pr\left(E^*|M_1\right)}=\frac{\Pr\left(R|M_v\right)}{\Pr\left(R|M_1\right)}\cdot\frac{\Pr\left(E^*|M_v\,\&\,R\right)}{\Pr\left(E^*|M_1\,\&\,R\right)}=$$
$$=\frac{r\left(v/k\right)}{r\left(1/k\right)}\cdot\frac{\Pr\left(E^*|M_v\,\&\,R\right)}{\Pr\left(E^*|M_1\,\&\,R\right)}=v\frac{\Pr\left(E^*|M_v\,\&\,R\right)}{\Pr\left(E^*|M_1\,\&\,R\right)}$$

Conditional on M_1&R, α is the only universe containing observers, so SSA prescribes assigning *Pr*(E*|M_1&R) = 1. Conditional on M_v&R, the expected fraction of observers inhabiting α depends on how many other universes contain observers. Let U_i stand for 'there are exactly *i* observer-containing universes other than α'. Conditional on M_v&R, E* entails the exclusive disjunction of U_0, U_1, … U_{v-1} and so is equivalent to the exclusive disjunction of E*&U_0, E*&U_1, … E*&U_{v-1}. By the sum rule of probability theory it follows that

$$\Pr\left(E^*|M_v\,\&\,R\right)=\sum_{i=0}^{v-1}\Pr\left(E^*\&U_i\,|\,M_v\,\&\,R\right)=$$
$$=\sum_{i=0}^{v-1}\Pr\left(U_i\,|\,M_v\,\&\,R\right)\cdot\Pr\left(E^*|M_v\,\&\,R\,\&\,U_i\right)$$

The values of *Pr*(U_i|M_v&R) are given by

$$\Pr\left(U_i\,|\,M_v\,\&\,R\right)=\binom{v-1}{i}r^i\left(1-r\right)^{v-1-i}$$

On the assumption that all life-permitting universes have the same expected number of observers, SSA prescribes that *Pr*(E*|M_v&R&U_i) = 1/(*i* + 1). We now have

$$\Pr\left(E^* \mid M_v \,\&\, R\right) = \sum_{i=0}^{v-1} \binom{v-1}{i} r^i \left(1-r\right)^{v-1-i} \frac{1}{i+1}$$

Let $j = i + 1$. We can now express the likelihood ratio of E* as follows:

$$\frac{\Pr\left(E^* \mid M_v\right)}{\Pr\left(E^* \mid M_1\right)} = v \sum_{j=1}^{v} \binom{v-1}{j-1} r^{j-1} \left(1-r\right)^{v-j} \frac{1}{j} =$$

$$= r^{-1} \sum_{j=1}^{v} \binom{v}{j} r^j \left(1-r\right)^{v-j} = \frac{1-\left(1-r\right)^v}{r}$$

References

Bostrom, N. [2002a]: *Anthropic Bias: Observation Selection Effects in Science and Philosophy*, New York: Routledge.

Bostrom, N. [2002b]: 'Self-Locating Belief in Big Worlds: Cosmology's Missing Link to Observation', *Journal of Philosophy*, **99**, pp. 607-23.

Bradley, D. J. [2005]: 'A Defence of the Fine-Tuning Argument for the Multiverse', <www.fitelson.org/few/few_05/bradley.pdf>

Bradley, D. J. [2009]: 'Multiple Universes and Observation Selection Effects', *American Philosophical Quarterly*, **46**, pp. 61-72.

Bradley, D. J. [2012]: 'Four Problems about Self-Locating Belief', *Philosophical Review*, **121**, pp. 149-77.

Bradley, M. C. [2001]: 'The Fine-Tuning Argument', *Religious Studies*, **37**, pp. 451-66.

Carter, B. [1983]: 'The Anthropic Principle and its Implications for Biological Evolution', *Philosophical Transactions of the Royal Society of London A*, **310**, pp. 347-63.

Carter, B. [2006]: 'Anthropic Principle in Cosmology', in J.-C. Pecker and J. Narlikar (*eds*), 2006, *Current Issues in Cosmology*, Cambridge: Cambridge University Press, pp. 173-9.

Collins, R. [2009]: 'The Teleological Argument: An Exploration of the Fine-Tuning of the Universe', in W. L. Craig and J. P. Moreland (*eds*), 2009, *The Blackwell Companion to Natural Theology*, Malden, MA: Wiley-Blackwell, pp. 202-81.

Collins, R. [2013]: 'Modern Cosmology and Anthropic Fine-Tuning: Three Approaches', in R. D. Holder and S. Mitton (*eds*), 2013, *Georges Lemaître: Life, Science and Legacy*, Heidelberg: Springer, pp. 173-92.

Colyvan, M., Garfield, J. L., and Priest, G. [2005]: 'Problems with the Argument from Fine Tuning', *Synthese*, **145**, pp. 325-38.

Craig, W. L. [2003]: 'Design and the Anthropic Fine-Tuning of the Universe', in N. A. Manson (*ed*), 2003, *God and Design: The Teleological Argument and Modern Science*, New York: Routledge, pp. 155-77.

Draper, K., Draper, P., and Pust, J. [2007]: 'Probabilistic Arguments for Multiple Universes', *Pacific Philosophical Quarterly*, **88**, pp. 288-307.

Good, I. J. [1967]: 'On the Principle of Total Evidence', *British Journal for the Philosophy of Science*, **17**, pp. 319-21.

Hacking, I. [1987]: 'The Inverse Gambler's Fallacy: The Argument from Design. The Anthropic Principle Applied to Wheeler Universes', *Mind*, **96**, pp. 331-40.

Heller, M. [2008]: 'Hudson Fine Tunes His Way to Hyperspace', *Philosophy and Phenomenological Research*, **76**, pp. 436-43.

Holder, R. D. [2002]: 'Fine-Tuning, Multiple Universes and Theism', *Noûs*, **36**, pp. 295-312.

Horwich, P. [1982]: *Probability and Evidence*, Cambridge: Cambridge University Press.

Juhl, C. F. [2005]: ‘Fine-Tuning, Many Worlds, and the ‘Inverse Gambler’s Fallacy’’, *Noûs*, **39**, pp. 337-47.

Kotzen, M. L. [2012]: ‘Selection Biases in Likelihood Arguments’, *British Journal for the Philosophy of Science*, **63**, pp. 825-39.

Leslie, J. A. [1988]: ‘No Inverse Gambler’s Fallacy in Cosmology’, *Mind*, **97**, pp. 269-72.

Leslie, J. A. [1989]: *Universes*, New York: Routledge.

Livio, M. and Rees, M. J. [2005]: ‘Anthropic Reasoning’, *Science*, **309**, pp. 1022-3.

Manson, N. A. [2000]: ‘There Is No Adequate Definition of “Fine-tuned for Life”’, *Inquiry*, **43**, pp. 341-52.

Manson, N. A. [2003]: ‘Cosmic Fine-Tuning, ‘Many Universe’ Theories and the Goodness of Life’, in W. B. Drees (*ed*), 2003, *Is Nature Ever Evil? Religion, Science and Value*, New York: Routledge, pp. 139-45.

Manson, N. A. and Thrush, M. J. [2003]: ‘Fine-Tuning, Multiple Universes, and the “This Universe” Objection’, *Pacific Philosophical Quarterly*, **84**, pp. 67-83.

McGrew, T. J., McGrew, L., and Vestrup, E. M. [2001]: ‘Probabilities and the Fine-Tuning Argument: a Sceptical View’, *Mind*, **110**, pp. 1027-38.

Monton, B. [2006]: ‘God, Fine-Tuning, and the Problem of Old Evidence’, *British Journal for the Philosophy of Science*, **57**, pp. 405-24.

Oppy, G. R. [2006]: *Arguing About Gods*, Cambridge: Cambridge University Press.

Pearl, J. [1988]: *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*, San Mateo, CA: Morgan Kaufmann.

Plantinga, A. [2011]: *Where the Conflict Really Lies: Religion, Science, and Naturalism*, New York: Oxford University Press.

Pust, J. [2007]: 'Cartesian Knowledge and Confirmation', *Journal of Philosophy*, **104**, pp. 269-89.

Smart, J. J. C. [1989]: *Our Place in the Universe: A Metaphysical Discussion*, Oxford: Blackwell.

Smith, Q. P. [1994]: 'Anthropic Explanations in Cosmology', *Australasian Journal of Philosophy*, **72**, pp. 371-82.

Sober, E. [2004]: 'The Design Argument', in W. E. Mann (*ed*), 2004, *Blackwell Guide to the Philosophy of Religion*, Oxford: Blackwell Publishers, pp. 117-47.

Sober, E. [2009]: 'Absence of Evidence and Evidence of Absence: Evidential Transitivity in Connection with Fossils, Fishing, Fine-Tuning, and Firing Squads', *Philosophical Studies*, **143**, pp. 63-90.

Van Inwagen, P. [2002]: *Metaphysics*, 2nd ed., Boulder, CO: Westview Press.

Weisberg, J. [2005]: 'Firing Squads and Fine-Tuning: Sober on the Design Argument', *British Journal for the Philosophy of Science*, **56**, pp.809-21.

White, R. [2000]: 'Fine-Tuning and Multiple Universes', *Noûs*, **34**, pp. 260-76. Reprinted with postscript in N. A. Manson (*ed*), 2003, *God and Design: The Teleological Argument and Modern Science*, New York: Routledge, pp. 229-50. (Page references are to the reprint.)