

Two Mistakes Regarding The Principal Principle

Christopher J. G. Meacham

Forthcoming in the British Journal for the Philosophy of Science

Abstract

This paper examines two mistakes regarding David Lewis' Principal Principle that have appeared in the recent literature. These particular mistakes are worth looking at for several reasons: the thoughts that lead to these mistakes are natural ones, the principles that result from these mistakes are untenable, and these mistakes have led to significant misconceptions regarding the role of admissibility and time. After correcting these mistakes, the paper discusses the correct roles of time and admissibility. With these results in hand, the paper concludes by showing that one way of formulating the chance-credence relation has a distinct advantage over its rivals.

1 Introduction

In "A Subjectivist's Guide to Objective Chance", David Lewis (1980) presents the canonical account of the relation between credence and chance, the Principal Principle. A number of different formulations of Lewis' principle can be found in the literature. Most of these differences are reasonable, harmless or both. But some of them are not. Some of them involve substantive mistakes.

Two mistakes are of particular interest. The first mistake replaces the reasonable initial credence function that appears in Lewis' original formulation with a subject's current credence function; the second mistake adds a time index to the subject's credences.¹ It's worth looking at these mistakes carefully for several reasons. First, the principles that result from these mistakes are problematic. The first mistake leads to inconsistencies, the second makes the Principal Principle too weak to apply to some cases. Second, these are natural mistakes. And getting clear on why they're mistakes will shed light on why certain natural ways of thinking about the relation between credence and chance are problematic. Third, and most importantly, these mistakes have led to misconceptions regarding the role of admissible evidence and the role of time in the chance-credence relation. And in the process of correcting these mistakes, the proper role of time and admissibility will become clear.

Examining these mistakes also has an unexpected benefit. In untangling the consequences of these mistakes, various merits and demerits of different formulations of the

¹For instances of the first, see Loewer (2001), Loewer (2004), Hajek (2002), Vranas (2002), Vranas (2004) and Lewis (1994). For instances of the second, see Vranas (2002) and Vranas (2004).

chance-credence relation will come to light. Once the dust settles, it will become clear that one formulation of the chance-credence relation is strictly better than the alternatives.

This paper will proceed as follows. In the next section I'll briefly sketch some background. In the third section I'll discuss the first mistake and its implications. In the fourth section I'll examine the role of admissible evidence. In the fifth section I'll discuss the second mistake, and its implications regarding the role of time. In the final section I'll explore the merits and demerits of different formulations of the chance-credence relation. With the lessons from the previous discussion in hand, I'll demonstrate that one way of formulating the chance-credence relation has a definite advantage over its rivals.

2 Background

2.1 The Chance Function

The core of Lewis' (1980) metaphysical account of chance can be expressed as a pair of claims about the form of the chance function:

The Function: Every possible chance assignment can be encoded by a function ch of two arguments, a *grounding argument* and an *object argument*. The grounding argument G picks out a chance distribution, $ch_G(\cdot)$. Given G , the object argument A picks out a value $x = ch_G(A)$.

The Arguments: The grounding argument of ch is a proposition: a conjunction of a complete chance theory, T , and a complete history up to a time at a world where that chance theory holds, H . The object argument of ch is also a proposition,² and $ch_G(\cdot)$ assigns a chance to every proposition to which an idealized credence function assigns a value.³

Lewis also provided an alternative characterization of the grounding argument as a time and world pair. But for his purposes, and ours, these two characterizations are interchangeable.

When conjoined with the Principal Principle, these two claims about the form of the chance function entail Lewis' other claims about chance, such as that $ch_G(\cdot)$ is a probability function, that "the past is no longer chancy" (if $H \Rightarrow A$, $ch_{TH}(A) = 1$), and so on.⁴ (An exception is Lewis' claim that determinism and chance are incompatible. Although Lewis

²In particular, a proposition about the way the world could be, or what Lewis (1979) calls a "*de dicto* proposition". This is given as a simplifying assumption in Lewis (1986), but in later work (such as Lewis (2004)) this is retained as a substantive assumption.

³This might seem unfaithful to Lewis, since Lewis explicitly declines to assume that a chance distribution assigns a well-defined value to every proposition. (See Lewis (1986), p.91.) But Lewis notes in the postscript that his reasons for declining stem from mathematical worries concerning whether a well-defined non-standard probability can be assigned to every proposition. And "plainly this reason for caution is no reason at all to think that the domains of chance distributions will be notably sparser than the domains of idealized credence functions." (Lewis (1986), p.132.) So the restriction to propositions that idealized credence functions are defined for keeps the characterization I provide faithful to Lewis.

⁴I use ' \Rightarrow ' to indicate strict implication, and ' \Leftrightarrow ' to indicate a strict biconditional. For ease of exposition, I will occasionally describe these relations as 'entails' and 'iff', respectively. (It should be clear from context what I mean.)

seems to have thought that this was required by his other commitments, it appears to be an independent assumption.⁵)

2.2 The Chance-Credence Relation

Let $up(\cdot)$ represent a rational subject's "ur-priors", her initial credence function before she's obtained any evidence. This is what Lewis calls a "reasonable initial credence function".⁶

Let $\langle ch(A) = x \rangle$ be the proposition that is true *iff* some TH obtains such that $ch_{TH}(A) = x$.⁷ More formally, $\langle ch(A) = x \rangle \Leftrightarrow T_1H_1 \vee T_2H_2 \vee \dots$, for all T_iH_i 's such that $ch_{T_iH_i}(A) = x$. Intuitively, $\langle ch(A) = x \rangle$ can be thought of as the proposition that at some time the chance of A is x . Note that $\langle ch(A) = x \rangle$ and $\langle ch(A) = y \rangle$ can both be true, since the chance of A can be x at some time and the chance of A can be y at some other time.

Similarly, let $\langle ch_t(A) = x \rangle$ be the proposition that is true *iff* some TH obtains such that $ch_{TH}(A) = x$, and H is a history up to time t , H^t . More formally, $\langle ch_t(A) = x \rangle \Leftrightarrow T_1H_1 \vee T_2H_2 \vee \dots$, for all T_iH_i 's such that $ch_{T_iH_i}(A) = x$ and $H_i = H_i^t$. Intuitively, $\langle ch_t(A) = x \rangle$ can be thought of as the proposition that at time t the chance of A is x .

Lewis presents two formulations of the Principal Principle. Lewis' first formulation is:

$$PP_1: up(A|\langle ch_t(A) = x \rangle E) = x, \text{ if } E \text{ is admissible with respect to } \langle ch_t(A) = x \rangle,^8 \quad (1)$$

assuming the terms in question are well-defined. (I'll take the "if well-defined" caveat to be implicit from now on, for both this and the other formulations of the chance-credence relation we'll look at.) When is evidence *admissible*? Lewis never provides a precise account—"I have no definition of admissibility to offer, but must be content to suggest sufficient (or almost sufficient) conditions"—but he suggests that information about the past and information about which chance theory obtains should both be admissible.⁹

Lewis' second formulation is:

$$PP_2: up(A|TH) = ch_{TH}(A). \quad (2)$$

Infinity caveats aside, Lewis takes these two formulations of the Principal Principle to be equivalent.¹⁰

While PP_1 makes use of the notion of admissibility, PP_2 does not. This makes Lewis' explicit ambivalence about what counts as admissible evidence puzzling. If PP_1 and PP_2

⁵Loewer (2001) argues that it is not required by Lewis' Humean commitments; Meacham (2005) and Hofer (2007) argue that it is not entailed by the Principal Principle and Lewis' other claims about chance.

⁶Lewis (1986), p.87. Lewis is not, of course, assuming that every rational subject must have the same ur-priors.

⁷In order to avoid Miller's Paradox, we should understand x , here and elsewhere, as a variable ranging over real numbers in the unit interval, not as (say) a schematic letter which we can replace with terms that designate these numbers (c.f. Lewis (1986)).

⁸Lewis (1986) originally assessed whether E was admissible relative to a time. But as Thau (1994) noted, and Lewis (1994) conceded, admissibility is relative to more than just time.

⁹Lewis (1986), p.92.

¹⁰Lewis (1986) notes that the derivations between PP_1 and PP_2 that he provides only go through for cases where $\langle ch_t(A) = x \rangle$ is equivalent to a *finite* disjunction of TH terms (see Lewis (1986), p.100). Lewis suggests that this lapse is unlikely to be important, and I'm inclined to agree. In the rest of this paper, I'll follow Lewis in putting this matter aside.

are equivalent, then providing a characterization of admissibility should be straightforward: call evidence “admissible” *iff* doing so is required to make PP_1 and PP_2 yield the same results. This puzzling feature of Lewis’ presentation has led some authors to believe that Lewis must have intended to add an admissibility clause to PP_2 as well as PP_1 .¹¹ We will return to examine the purpose of PP_1 ’s admissibility clause in section 4.

2.3 Assumptions

Following Lewis, I will assume that it’s rational to be *Bayesian*. Letting $cr_E(\cdot)$ represent the credence function of an agent whose total evidence is E , we can characterize Bayesian agents as agents with probabilistic credences who satisfy the following constraint:¹²

$$cr_E(\cdot) = up(\cdot|E), \text{ if defined.} \quad (3)$$

In what follows I will restrict my attention to Bayesian agents who satisfy the usual idealizations (logical omniscience, etc.). For convenience, I will also follow Lewis in assuming that the propositions under consideration can be identified with sets of possibilities; in particular, sets of possible worlds.

For the purposes of this paper, I will turn a blind eye to the question of whether Lewis’ Humean account of chance can be reconciled with the Principal Principle. I will also sidestep the question of whether the “New Principle” proposed by Lewis (1994) and Hall (1994) should replace Lewis’ original principle. These issues, although interesting, are orthogonal to the matters we’ll be concerned with. So I’ll circumvent them in what follows.

3 Assessing the First Mistake

3.1 The First Mistake

Recall Lewis’ first formulation of the Principal Principle:

$$PP_1: up(A|\langle ch_t(A) = x \rangle E) = x, \text{ if } E \text{ is admissible with respect to } \langle ch_t(A) = x \rangle.$$

The first mistake replaces the “reasonable initial credence function” $up(\cdot)$ with the current credence function, $cr(\cdot)$:

$$PP_1^*: cr(A|\langle ch_t(A) = x \rangle E) = x, \text{ if } E \text{ is admissible with respect to } \langle ch_t(A) = x \rangle. \quad (4)$$

3.2 Motivating the First Mistake

There are several reasons why it’s natural to make the first mistake.

1. Two reasons stem from Lewis himself. First, in his original paper he uses “ $c(\cdot)$ ” to represent the “reasonable initial credence function” $up(\cdot)$. When expressed in this notation, it’s easy to confuse PP_1 with PP_1^* . Second, and more importantly, Lewis himself

¹¹For example, see Vranas (2004) and Meacham (2005).

¹²Where the subject’s *total evidence* is understood here to be the conjunction of all of the evidence the agent has received so far.

endorses PP_1^* in one of his later papers.¹³ Since it's reasonable to take Lewis at his word, it's natural to follow him in making this mistake.

2. Lewis' reason for endorsing PP_1^* in his later paper stems from another source. Given Bayesianism, we can reformulate PP_1 as:

$$PP_1: cr_{\langle ch_t(A)=x \rangle E}(A) = x, \text{ if } E \text{ is admissible with respect to } \langle ch_t(A) = x \rangle. \quad (5)$$

Since a subject's credence in her evidence is 1, we can equate $cr_E(A)$ with $cr(A|E)$ as long as we explicitly state in the latter case that E is the subject's total evidence. So we can re-express PP_1 as:

$$PP_1: cr(A|\langle ch_t(A) = x \rangle E) = x, \quad \text{if } E \text{ is admissible with respect to } \langle ch_t(A) = x \rangle, \quad (6)$$

where $\langle ch_t(A) = x \rangle E$ is the subject's total evidence.

So far, so good. But we must be careful. It's easy to lose track of the total evidence clause and express this as:

$$PP_1^*: cr(A|\langle ch_t(A) = x \rangle) = x, \text{ if } E \text{ is admissible with respect to } \langle ch_t(A) = x \rangle.$$

Indeed, since PP_1^* looks reasonable at first glance, one may see little reason to keep the total evidence clause, even if one realizes that the resulting rule is not the same as PP_1 . This appears to be Lewis' reason for adopting PP_1^* in his later paper. Although he explicitly recognizes the distinction between PP_1^* and the formulation of PP_1 expressed by (5), he states that the latter is merely a special case of the former, and presents PP_1^* as the full-fledged Principal Principle.¹⁴

3.3 Why the First Mistake is Problematic

PP_1^* is problematic. Unlike PP_2 , PP_1^* does not take one's evidence into account. As a result, it leads to inconsistencies.

To see this, pick a T , H and A such that $ch_{TH}(A) = x \neq 1$, and consider a subject whose total evidence is THA . Then it follows from PP_1^* that:

$$1 = cr_{THA}(A) = cr_{THA|\langle ch_t(A)=x \rangle}(A) = cr_{THA|\langle ch_t(A)=x \rangle}(A|\langle ch_t(A) = x \rangle) = x \neq 1. \quad (7)$$

What went wrong? The problem is that PP_1^* doesn't take the subject's evidence into account. If we formulate the principle in terms of ur-priors, as Lewis originally did, then we don't need to worry about this, since the subject has no evidence. Likewise, if we formulate the principle in terms of a subject's current credence, and we explicitly take her current evidence into account, as (5) and (6) do, we avoid these problems. But if we formulate the chance-credence relation in terms of a subject's current credence, and fail to take her current evidence into account, then we'll run into contradictions.¹⁵

¹³See Lewis (1994), p.237-238. The admissibility clause Lewis provides is slightly different, but this has no bearing on the issues we're concerned with.

¹⁴See Lewis (1994), p.237-238.

¹⁵One might hope that realistic agents won't get evidence of the kind that will lead to contradictions. Unfortunately, realistic agents will get such evidence. To get such evidence, one merely needs to have observed the outcome of a chance event. Alternatively, one might try to avoid these worries by restricting PP_1^* so that it only applies to subjects at the time picked out by the chance proposition $\langle ch_t(A) = x \rangle$ in question. Then realistic agents won't have 'future' evidence of the kind that will lead to contradictions. This modification of PP_1^* helps to circumvent contradictions, but it leads to problems of a different nature (see section 5).

3.4 Consequences of the First Mistake

The first mistake skews our assessment of the role of admissibility. Once we make this mistake, it's easy to be led astray about the purpose of admissibility.

For example, Vranas (2004) argues that any tenable formulation of the chance-credence relation requires an admissibility clause:

“Assume you know (for sure) that $ch(A)$ is 50%. Does it follow that your credence in A should be 50%?... Given [the chance-credence principle without an admissibility clause] it does, but in fact it doesn't, because you may also know some inadmissible proposition”, such as the proposition that A occurred.¹⁶

With respect to PP_1^* , this argument makes sense. Stripping PP_1^* of the admissibility clause gives us:

$$PP_1^* -: cr(A | \langle ch_t(A) = x \rangle) = x. \quad (8)$$

And this principle leads to inconsistencies, for the reasons Vranas presents. If we apply this principle to an agent whose total evidence is A , with respect to the proposition $\langle ch_t(A) = 1/2 \rangle$, we get:

$$1 = \frac{cr_A(\langle ch_t(A) = x \rangle)}{cr_A(\langle ch_t(A) = x \rangle)} = \frac{cr_A(A | \langle ch_t(A) = x \rangle)}{cr_A(\langle ch_t(A) = x \rangle)} = cr_A(A | \langle ch_t(A) = x \rangle) = 1/2 \neq 1. \quad (9)$$

Vranas concludes that the purpose of the admissibility clause is to protect us from these inconsistencies. And if we've made the first mistake, it's natural to follow Vranas in thinking that this is why we need admissibility.

As we saw in section 3.3, adding the admissibility clause back to PP_1^* doesn't get us out of these inconsistencies, so this can't be why admissibility is needed. But put that aside. We won't even be tempted to follow this line of thought if we start with PP_1 instead of PP_1^* . If we strip PP_1 of the admissibility clause we get:

$$PP_1 -: up(A | \langle ch_t(A) = x \rangle) = x, \quad (10)$$

or equivalently:

$$PP_1 -: cr_{\langle ch_t(A) = x \rangle}(A) = x. \quad (11)$$

And these principles are not subject to the kind of worry Vranas raises. These principles only entail that a subject's credence in A should be $1/2$ if her *total* evidence is $\langle ch(A) = 1/2 \rangle$. If the subject is in possession of further evidence, such as A , then PP_1 - no longer requires her credence in A to be $1/2$. So Vranas' worry won't apply.¹⁷

PP_1 's admissibility clause isn't needed to escape Vranas' worry. But if this isn't what we need admissibility for, what *do* we need it for? We'll turn to this question next.

4 The Role of Admissibility

Why does PP_1 need an admissibility clause? One popular answer is that admissibility is needed to handle crystal ball cases. Another is that admissibility is needed in order to

¹⁶Vranas (2004), p.9.

¹⁷The first mistake is also what leads Vranas to criticize Hall's claim that the "New Principle" doesn't require an admissibility clause (see Vranas (2004), section 5). If we understand the New Principle correctly, in terms of ur-priors, then these criticisms won't apply.

make a chance-credence principle strong enough to be useful. However, nothing about these answers hangs on the particular form of PP_1 . If they succeed in showing why PP_1 needs an admissibility clause, then they will have shown why PP_2 needs an admissibility clause as well. So if either of these answers are correct, PP_2 should be modified and equipped with an admissibility clause.

The correct answer, on the other hand, will demonstrate why only PP_1 needs an admissibility clause. Let's look at each of these answers in turn.

4.1 Crystal Balls

One reason given for why admissibility is needed is that it's required in order to deal with crystal ball cases.¹⁸ And to handle such cases, both PP_1 and PP_2 require an admissibility clause.

As Hall (1994) points out, this answer is mistaken: no admissibility clause is needed to deal with crystal ball cases. To see this, let's look at how PP_2 handles a typical crystal ball case:

Suppose that there are crystal balls—devices which gave infallible predictions with respect to the outcomes of future events. And suppose that a crystal ball tells us at t_0 that the outcome of some chance event at t_1 will be A . Finally, suppose that the chance at t_0 of this outcome, A , is $1/2$. If our total evidence consists of the chance theory at this world and the history up to t_0 , TH , what should our credence in A be?

The correct answer seems to be 1 . But PP_2 appears to give the wrong answer:

$$cr_{TH}(A) = up(A|TH) = ch_{TH}(A) = 1/2, \quad (12)$$

unless we add an admissibility clause.

The way in which this argument fails depends on the status of the crystal ball's infallibility. If the crystal ball is actually infallible, then the case described above is incoherent. Since the crystal ball shows us at t_0 that A will come about, $TH \Rightarrow A$, and thus $1/2 = ch_{TH}(A) = up(A|TH) = up(ATH)/up(TH) = up(TH)/up(TH) = 1$, which is a *reductio* of the case described.

On the other hand, if we just believe (mistakenly) that the crystal ball is infallible, then PP_2 must declare our beliefs irrational. Let I be the proposition that the crystal ball is infallible. Since we believe that the crystal ball is infallible, our priors must be such that $up(I|TH) = 1$. Since $ITH \Rightarrow A$, it follows that:

$$1/2 = ch_{TH}(A) = up(A|TH) \geq up(AI|TH) = up(I|TH) = 1 \neq 1/2. \quad (13)$$

So PP_2 must declare our priors irrational. But since this is intuitively the right answer—you have no more reason to disbelieve the chances in this case than you do in any other—this isn't a mark against PP_2 .

¹⁸For example, see Strevens (1999), Vranas (2004).

4.2 Usefulness

Another answer is that the admissibility clause is needed in order to make the chance-credence relation strong enough to be interesting, and that Lewis' failure to add an admissibility clause to PP_2 as well as PP_1 was a mistake.

This answer stems from something like the following train of thought. Consider how these principles fare without admissibility. PP_2 tells us what our credences should be if our total evidence consists of some TH . And PP_1 -, the principle we get once we strip PP_1 of the admissibility clause, tells us what our credences should be when our total evidence consists of some $\langle ch_t(A) = x \rangle$. But this is of little use to people like us, who are never in either of these evidential situations—situations where our total evidence is all and only some TH , or all and only some proposition about the chance of A . By contrast, a principle like PP_1 is helpful:

$$PP_1: cr_{\langle ch_t(A)=x \rangle E}(A) = x, \text{ if } E \text{ is admissible with respect to } \langle ch_t(A) = x \rangle,$$

since we are often in situations where our total evidence can be expressed as $\langle ch_t(A) = x \rangle E$, for some E .

This answer fails to appreciate the strength of PP_1 - and PP_2 when coupled with Bayesianism.¹⁹ Consider PP_2 . It's true that PP_2 only applies directly to our credences when our total evidence is equal to some TH . But PP_2 doesn't need to apply directly in order to bear on our credences. Through PP_2 , the chances place constraints on our ur-priors. And the Bayesian rule generates our credences from our ur-priors and our evidence. So PP_2 allows the chances to indirectly constrain our credences, even when our total evidence doesn't equal the grounds of a chance distribution.

For a simple example, consider a subject whose total evidence is $B = T_1H_1 \vee T_2H_2$, where T_1H_1 and T_2H_2 are mutually exclusive. (If they weren't mutually exclusive, then one would entail the other, and we could represent the subject's evidence as just one of the two disjuncts.) Suppose that $ch_{T_1H_1}(A) = 1/2$ and $ch_{T_2H_2}(A) = 1/3$. In this case, PP_2 won't fix the subject's credence in A . But it will still place the appropriate constraints. In particular, it will require her credence in A to satisfy:²⁰

$$cr_B(A) = \frac{cr_B(T_1H_1)}{2} + \frac{cr_B(T_2H_2)}{3}. \quad (14)$$

4.3 The Strength of PP_1

Neither of these two answers for why we need admissibility succeeds. So what do we need the admissibility clause for? And why does only PP_1 need one?

To see why PP_1 requires an admissibility clause, consider the chance-credence relation that results once we strip it of this clause, PP_1 -:

$$PP_1-: up(A|\langle ch_t(A) = x \rangle) = x.$$

¹⁹Meacham (2005) provides a more detailed defense of the claim that PP_2 does not require an admissibility clause.

²⁰In detail: $cr_B(A) = up(A|B) = up(AB)/up(B) = up(AT_1H_1 \vee AT_2H_2)/up(B) = up(AT_1H_1)/up(B) + up(AT_2H_2)/up(B) = up(AT_1H_1)up(T_1H_1)/up(B)up(T_1H_1) + up(AT_2H_2)up(T_2H_2)/up(B)up(T_2H_2) = up(A|T_1H_1)up(T_1H_1)/up(B) + up(A|T_2H_2)up(T_2H_2)/up(B) = ch_{T_1H_1}(A)up(T_1H_1)/up(B) + ch_{T_2H_2}(A)up(T_2H_2)/up(B) = ch_{T_1H_1}(A)up(T_1H_1B)/up(B) + ch_{T_2H_2}(A)up(T_2H_2B)/up(B) = ch_{T_1H_1}(A)up(T_1H_1|B) + ch_{T_2H_2}(A)up(T_2H_2|B) = ch_{T_1H_1}(A)cr_B(T_1H_1) + ch_{T_2H_2}(A)cr_B(T_2H_2) = 1/2 \cdot cr_B(T_1H_1) + 1/3 \cdot cr_B(T_2H_2).$

Given PP_2 , we can derive PP_{1-} . (The proof is provided in Appendix A.) But given PP_{1-} , we cannot derive PP_2 . So PP_2 is strictly stronger than PP_{1-} . Furthermore, PP_{1-} 's weakness is problematic. There are intuitive relations between credence and chance that PP_{1-} cannot capture.

For example, consider the following case:

Consider a chance theory T_1 such that $ch_{T_1 H_1^t}(A) = 1/2$. Recall that $\langle ch_t(A) = 1/2 \rangle$ is equivalent to a disjunction of all of the $T_i H_i^t$'s for which $ch_{T_i H_i^t}(A) = 1/2$. For simplicity, let us assume there are just two $T_i H_i^t$'s in the disjunction associated with $\langle ch_t(A) = 1/2 \rangle$, $T_1 H_1^t$ and $T_2 H_2^t$. So $\langle ch_t(A) = 1/2 \rangle \Leftrightarrow T_1 H_1^t \vee T_2 H_2^t$.

Now consider a subject whose total evidence is $T_1 H_1^t$. What should her credence in A be?

Since the subject knows only that $T_1 H_1^t$ obtains, and $T_1 H_1^t$ entails that the chance at t of A is $1/2$, the subject's credence in A should be $1/2$. But PP_{1-} is too weak to impose this constraint.

We can see this explicitly. Here is an ur-prior function which (i) entails that the subject's credence in A will be 0 instead of $1/2$, and yet (ii) satisfies the constraints that PP_{1-} imposes:

$$\begin{aligned} up(T_1 H_1^t) &= a & up(AT_1 H_1^t) &= 0 & up(T_1 H_1^t \wedge T_2 H_2^t) &= 0 \\ up(T_2 H_2^t) &= a & up(AT_2 H_2^t) &= a \end{aligned}$$

Given these ur-priors, (i) the subject's credence in A will be 0:

$$cr_{T_1 H_1^t}(A) = up(A|T_1 H_1^t) = \frac{up(AT_1 H_1^t)}{up(T_1 H_1^t)} = \frac{0}{a} = 0. \quad (15)$$

But (ii) this ur-prior function is compatible with PP_{1-} , since:

$$\begin{aligned} up(A|\langle ch_t(A) = 1/2 \rangle) &= up(A|H_1^t T_1 \vee H_2^t T_2) &= \frac{up(AT_1 H_1^t \vee AT_2 H_2^t)}{up(T_1 H_1^t \vee T_2 H_2^t)} & (16) \\ &= \frac{up(AT_1 H_1^t) + up(AT_2 H_2^t)}{up(T_1 H_1^t) + up(T_2 H_2^t)} \\ &= \frac{0 + a}{a + a} = 1/2. \end{aligned}$$

Why isn't PP_{1-} strong enough to yield the desired result? PP_{1-} places a constraint on one's prior in A conditional on $T_1 H_1^t \vee T_2 H_2^t \vee \dots$, the disjunction associated with $\langle ch_t(A) = x \rangle$. But in order to get the result we want, we need to impose a more fine-grained constraint on one's priors. We need a principle which, like PP_2 , constrains one's prior in A conditional on each of the *individual* $T_i H_i^t$'s associated with $\langle ch_t(A) = x \rangle$. PP_{1-} constrains one's conditional priors with respect to the disjunction, but we also need to constrain one's conditional priors with respect to each of the individual disjuncts.

This is why PP_1 requires an admissibility clause. The admissibility clause makes PP_1 strong enough to impose the constraints we want it to impose. Consider the case above, for example, where the subject's total evidence is $T_1 H_1^t$ and $ch_{T_1 H_1^t}(A) = 1/2$. If we apply PP_1 instead of PP_{1-} , we get the answer we want:

$$cr_{T_1 H_1^t}(A) = up(A|T_1 H_1^t) = up(A|T_1 H_1^t \langle ch_t(A) = 1/2 \rangle) = 1/2, \quad (17)$$

where the last step employs the fact that the TH 's associated with $\langle ch_t(A) = x \rangle$ are admissible relative to it. So with the admissibility clause in place, PP_1 is strong enough to yield the desired results.

But when, exactly, is evidence admissible? We can use the requirement that PP_1 and PP_2 should be equivalent to determine the required notion of admissibility:

Admissibility: E is admissible relative to $\langle ch_t(A) = x \rangle$ iff $\langle ch_t(A) = x \rangle E$ can be expressed as a disjunction of some subset of the TH 's associated with $\langle ch_t(A) = x \rangle$.²¹

I.e., E is admissible relative to $\langle ch_t(A) = x \rangle$ if E intersects $\langle ch_t(A) = x \rangle$ 'cleanly'—the outline of their intersection traces the boundaries of TH 's. Given this characterization of admissibility, PP_1 can be derived from PP_2 and vice versa, in the way Lewis (1986) describes. (The proofs are reproduced in Appendix A.)

This characterization of admissibility entails that information about the past is admissible, as is information about which chance theory obtains. For E to be inadmissible, it needs to exclude some, but not all, of the possibilities in some TH associated with $\langle ch_t(A) = x \rangle$. Consider one of these TH 's. An inadmissible E needs to include some further information not entailed by TH . And TH entails everything about the past and the chance theory. So the inadmissible part of E can't be about the past or the chance theory; it must be about some other aspect of the world, such as the future.

4.4 Lewis and Admissibility

This leaves us with a final question: why is Lewis so ambivalent about admissibility? Given that his claim about the equivalence of PP_1 and PP_2 entails a precise characterization of admissibility, why does he claim that "I have no definition of admissibility to offer"?²² Lewis' ambivalence is even more puzzling in light of two further observations: Lewis employs the criterion for admissibility given above as a necessary condition for admissibility in his derivation of PP_1 from PP_2 , and the sufficient condition for admissibility Lewis offers—being about the past or the chance theory—entails that the above criterion is sufficient as well.²³ So why doesn't he take this criterion to be a necessary and sufficient condition for admissibility, as I've suggested?

Lewis refrains from adopting this characterization of admissibility for two reasons. The first concerns worlds with temporal abnormalities.²⁴ Suppose there are worlds where time can loop back on itself, such as the Gödel universe models of general relativity which admit closed timelike curves. At worlds like these, where the same event can be in both one's future and one's past, it's hard to know what to count as "information about the future" or "a history up to time". Or suppose there are possible worlds without a well-defined notion of global time, such as the Anti-De Sitter models of general relativity which don't admit Cauchy surfaces. At worlds like these, it's hard to know how to make sense of even simple notions like the "chance at t ". In either case, it's hard to make sense of the characterization of admissibility given above, which relies on these kinds of temporal

²¹This characterization of admissibility has the consequence that E is admissible relative to $\langle ch_t(A) = x \rangle$ in the trivial case where E and $\langle ch_t(A) = x \rangle$ are mutually exclusive. But one could also define things the other way. Nothing of substance hangs on this choice.

²²Lewis (1986), p.92.

²³See Lewis (1986), p.94-96 and 99-100.

²⁴See Lewis (1986), p.94.

notions.²⁵ So we can only take this characterization of admissibility to be adequate if we restrict ourselves to worlds without such temporal oddities.²⁶

Second, Lewis worries that information about the chances themselves will be inadmissible if one adopted a Humean account of chance.²⁷ If so, then the characterization of admissibility given above will be inadequate. This is an interesting topic, but since I'm side-stepping issues particular to Humeanism, I won't discuss it further here.

So Lewis has good reasons for declining to adopt the characterization of admissibility suggested above. If we restrict ourselves to worlds without temporal anomalies, and put aside issues regarding Humeanism, then this characterization will suffice. But once we relax these restrictions, this characterization may no longer be adequate.

5 Assessing the Second Mistake

5.1 The Second Mistake

Recall Lewis' second formulation of the Principal Principle:

$$\text{PP}_1: \text{up}(A|\langle ch_t(A) = x \rangle E) = x, \text{ if } E \text{ is admissible with respect to } \langle ch_t(A) = x \rangle.$$

The second mistake is to formulate the principle in terms of time-indexed credences.

Let $cr_{t,E}(\cdot)$ be the credence at t of a subject whose total evidence is E . (When cr has only one subscript, it should be clear from context which is intended.) In the literature, the second mistake generally appears together with the first mistake.²⁸

$$cr_t(A|\langle ch_t(A) = x \rangle E) = x, \text{ if } E \text{ is admissible with respect to } \langle ch_t(A) = x \rangle. \quad (18)$$

That said, the second mistake does not require the first. We can isolate the second mistake by using Bayesianism to formulate PP_1 in terms of a subject's current credences:

$$\text{PP}_1: cr_{\langle ch_t(A) = x \rangle E}(A) = x, \text{ if } E \text{ is admissible with respect to } \langle ch_t(A) = x \rangle,$$

and then adding a time-index to the credence function:

$$\text{PP}_1^{**}: cr_{t;\langle ch_t(A) = x \rangle E}(A) = x, \text{ if } E \text{ is admissible with respect to } \langle ch_t(A) = x \rangle. \quad (19)$$

5.2 Motivating the Second Mistake

Here are two natural trains of thought that lead to this mistake.²⁹

²⁵I thank Phillip Bricker for suggesting this way of understanding Lewis.

²⁶Although Lewis highlighted these worries in his discussion of admissibility, these possibilities also make trouble for Lewis' account of the grounding argument of the chance function. So the impact of these possibilities cuts deeper than just the characterization of admissibility: the intelligibility of Lewis' account of the grounding argument also depends on this restriction of the possibilities under consideration.

²⁷See Lewis (1986), p.130.

²⁸I.e., see Vranas (2002), Vranas (2004).

²⁹Yet another motivation is to time index things in order to circumvent the kind of contradictions brought about by the first mistake. But since I am trying to treat these mistakes separately, I will put this motivation aside.

1. Consider a subject whose total evidence consists of what the chance of A will be at several different times, t_1 - t_n . And suppose that the chance of this outcome at each of these times is different. Which of these chances should her credence in A line up with?

A natural thought is that we should want our credences to line up with whatever the *current* chance is. So we should add a time index to the credence function on the left hand side of the Principal Principle, and PP_1 should really be:

$$PP_1^{**}: cr_{t; \langle ch_t(A)=x \rangle E}(A) = x, \text{ if } E \text{ is admissible with respect to } \langle ch_t(A) = x \rangle.$$

2. In Lewis' formulation of PP_1 , the chance propositions $\langle ch_t(A) = x \rangle$ are time indexed. But what role is the time index playing? Suppose, for example, that we replaced $\langle ch_t(A) = x \rangle$ (the proposition that at t the chance of A is x) with $\langle ch(A) = x \rangle$ (the proposition that at some time the chance of A is x). Doing so yields:

$$PP_3: cr_{\langle ch(A)=x \rangle E}(A) = x, \text{ if } E \text{ is admissible with respect to } \langle ch(A) = x \rangle, \quad (20)$$

where E is admissible with respect to $\langle ch(A) = x \rangle$ iff $E \langle ch(A) = x \rangle$ can be expressed as a disjunction of a subset of the TH 's associated with $\langle ch(A) = x \rangle$.

Is there anything wrong with PP_3 ? Here's a natural gut-reaction response:

PP_3 can't be right. Consider an agent who knows nothing except a single fact about chance—that at some time the chance of A will be x . PP_3 says that this agent's credence in A ought to be x . But that's crazy! What does it matter if at *some* time the chance of A is x ? Surely we only have reason to set our credences equal to the chances if they're the *current* chances.

With this thought in mind, it seems we should add a time index to the credence function on the left hand side of (5.1), and PP_1 should really be:

$$PP_1^{**}: cr_{t; \langle ch_t(A)=x \rangle E}(A) = x, \text{ if } E \text{ is admissible with respect to } \langle ch_t(A) = x \rangle.$$

5.3 Why the Second Mistake is Problematic

Unlike PP_1^* , PP_1^{**} is not inconsistent. The problem with PP_1^{**} is that it won't allow us to make all of the inferences we'd like to make.

Say you're told about a coin toss in 1900 A.D., before you were born. You know the chance right before the coin toss was $\langle ch_{1900}(h) = 1/2 \rangle$. And let's stipulate that you don't have any evidence that's inadmissible relative to $\langle ch_{1900}(h) = 1/2 \rangle$.³⁰ What should your credence be that the coin came up heads?

³⁰This is not a realistic case, since someone like us *would* have inadmissible evidence, information that "cuts across" the TH 's associated with $\langle ch_{1900}(h) = 1/2 \rangle$. But the worry described below still arises, albeit less directly. I.e., consider someone who is told about this coin toss, and who has the kind of inadmissible evidence that we might have. What should her credence be in h ? As Lewis (1986) notes in his discussion of a similar case (p.115-116), it depends. PP_1 will entail that her prior in h conditional on the admissible part of her evidence should be $1/2$. If we further stipulate that the inadmissible part of her evidence is independent of the outcome of the coin toss, then it will follow that her credence should remain $1/2$. On the other hand, if no such stipulation is made, then it won't follow that her credence in heads should remain $1/2$. This is all as it should be. But PP_1^{**} cannot yield these results, since PP_1^{**} will not entail that her prior in h conditional on her admissible evidence should be $1/2$.

It seems your credence should be $1/2$. And this is what PP_1 tells us to believe: you know the chance of the coin coming up heads was $1/2$, and you don't have any evidence that's inadmissible relative to that chance. But PP_1^{**} won't deliver this result. You don't know the *current* chance of the coin coming up heads—you don't know whether it's 0 or 1—so PP_1^{**} won't apply.

It might seem like Bayesianism can rescue PP_1^{**} here. If your credence in heads was appropriately constrained at the time of the coin toss, and you updated using the Bayesian rule, won't you end up with correct credences later on? If so, doesn't this get PP_1^{**} out of the problem? Yes, you will end up with the right credence, but no, this won't get PP_1^{**} out of the problem. In the case described, the coin is flipped before you were born. So your credence in heads can't have been appropriately constrained at the time of the coin toss. And once that time has passed, it's too late.

It's natural at this point to think that a shift to a constraint on ur-priors is called for. If we formulate the chance-credence principle as an ur-prior constraint, we don't need to worry about when you were born. But this move isn't available to the proponent of PP_1^{**} . Because the credences in PP_1^{**} are time indexed, there's no easy way to reformulate it in terms of ur-priors.

5.4 The Role of Time

PP_1^{**} incorporates time in two ways: it tracks the time at which a subject has the credence, and it tracks the time at which the chance obtains. As we've just seen, incorporating time in this first way is undesirable. We don't want the time at which the credence is held to matter.

What about incorporating time in the second way? Does it matter whether we track the time at which the chances obtain? No. We can see this in the following way. If we remove the time index from PP_1 , we get PP_3 :

$$PP_3: cr_{\langle ch(A)=x \rangle E}(A) = x, \text{ if } E \text{ is admissible with respect to } \langle ch(A) = x \rangle.$$

And PP_1 and PP_3 are equivalent: given PP_3 you can derive PP_1 and PP_2 , and vice versa. (The proofs are provided in Appendix B.) Thus PP_3 adequately captures the relation between credence and chance that Lewis envisioned, even though it doesn't track the time at which the chance obtains. So nothing is lost by discarding the time index that appears in PP_1 .

To get a better feel for why this is the case, consider again a subject who knows nothing except a single fact about chance, $\langle ch_t(A) = x \rangle$. Regardless of what the time t is, PP_1 will require that her credence in A be x . And since $\langle ch(A) = x \rangle$ is just the disjunction of $\langle ch_t(A) = x \rangle$'s for all t 's, PP_1 will impose the same constraint if her only evidence was that the chance of A was x at *some* time. So in this case the time at which the chance obtains is irrelevant.

In more complicated cases the time at which the chance proposition obtains can play a useful role: it can be used as an easy way to gauge the admissibility of different kinds of evidence. But what is and isn't admissible ultimately stems from structural features of the evidence and the chance proposition. And while interesting facts about the relation between admissibility and the time fall out as a consequence of these structural features (as we saw in section 4.3), time itself isn't an integral part of this structure. This is why we

can use the same characterization of admissibility for PP_1 and PP_3 , even though the latter doesn't keep track of the time at which the chance obtains.

We can index chances to times because of Lewis' choices with respect to the grounding argument. But while this plays an important part in Lewis' metaphysical account of chance, it is not a substantive part of Lewis' epistemic account of chance. With respect to the chance-credence relation itself, the role of time is dispensable.

6 Assessing The Chance-Credence Relation

We've seen a number of attempts to formulate the relation between credence and chance. Which of these formulations should we prefer?

We can rule out some of these formulations for reasons we've already seen. Both PP_1^* and PP_1^{*-} are inconsistent. And neither PP_1^- nor PP_1^{**} is strong enough to impose the desired constraints on our credences.

What about the three formulations of the Principal Principle, PP_1 , PP_2 and PP_3 ? The details of these assessments will depend on the assumptions we make about the grounding argument of chance distributions. So we'll first assess the different formulations of the chance-credence relation given Lewis' account of the grounding argument. Then we'll relax this assumption, and reassess the merits of these formulations once we allow for more general kinds of grounding arguments.

6.1 Take I: Lewis' Grounding Argument

First, let's assess the different formulations of the chance-credence relation given Lewis' account of the grounding argument.

Let's start by considering the two formulations Lewis offers, PP_1 and PP_2 . As we've seen, these two formulations are equivalent: PP_1 entails PP_2 , and vice versa. But this doesn't mean that the two formulations are equally good. PP_2 , unlike PP_1 , does not have an admissibility clause. And this makes PP_2 much more convenient to work with. We can see this in Lewis' (1986) paper: after introducing the two formulations of the Principal Principle, he works exclusively with PP_2 .

What about PP_3 ? As we've seen, PP_3 is also equivalent to PP_2 . But, like PP_1 , PP_3 is burdened with an admissibility clause, and this makes PP_3 harder to use.

What if we removed the admissibility clause from PP_3 ? This would yield PP_{3-} :

$$PP_{3-}: up(A | \langle ch(A) = x \rangle) = x. \quad (21)$$

But PP_{3-} suffers from the same problems as PP_1^- . PP_{3-} is too weak to capture all of the relations between credence and chance that we want. As a result, PP_{3-} is not a satisfactory formulation of the chance-credence relation.

So given Lewis' account of the grounding argument, PP_2 is the most convenient formulation of the Principal Principle.

6.2 Take II: General Grounding Arguments

Now let's assess the merits of these formulations once we allow for more general kinds of grounding arguments.

The grounding arguments Lewis employs have two distinctive features. First, they are *time-indexed*: every grounding argument is naturally associated with some time t . In particular, TH is associated with the time which H is a history up to. Second, there are *no partial overlaps*: two grounding arguments overlap *iff* one is a subset of the other. Consider two TH 's. If the T 's differ or the H 's are incompatible, then they'll be mutually exclusive. If the T 's match and the H 's are compatible (just of different lengths), then the TH corresponding to the longer history will be a subset of the other TH .

These features fit well with many of our intuitions regarding chance. But in order to accommodate statistical mechanical chances and the like, we might want to allow for grounding arguments which have neither of these features.³¹ So let's consider more general kinds of grounding arguments, and then reassess the different formulations of the chance-credence relation.

1. What happens if we allow grounding arguments which aren't time indexed? Allowing such arguments creates problems for PP_1 , because there's no natural way to generalize PP_1 in order to accommodate them. PP_2 and PP_3 , on the other hand, have little trouble accommodating arguments that aren't time indexed. Let's look at each of these principles in turn.

First, consider PP_1 . To generalize PP_1 , we need to generalize our characterization of $\langle ch_t(A) = x \rangle$. The natural way to do this is to replace the TH 's with some more general grounding argument G , and to characterize $\langle ch_t(A) = x \rangle$ as:

$$\langle ch_t(A) = x \rangle \Leftrightarrow G_1 \vee G_2 \vee \dots, \text{ for all } G_i\text{'s such that } ch_{G_i}(A) = x \text{ and } G_i = G_i^t. \quad (22)$$

But this definition only makes sense if we're able to associate the grounding argument with a time. And once we allow grounding arguments which aren't time indexed, we no longer have a natural way to do this. Since the intelligibility of PP_1 hangs on the intelligibility of $\langle ch_t(A) = x \rangle$, our failure to generalize $\langle ch_t(A) = x \rangle$ means we can't generalize PP_1 either.

Next, consider PP_3 . To generalize PP_3 , we need to generalize our characterization of $\langle ch(A) = x \rangle$. But time doesn't play a special role in the characterization of $\langle ch(A) = x \rangle$, so this generalization is unproblematic. We can simply replace the TH 's with G 's, and characterize $\langle ch(A) = x \rangle$ as:

$$\langle ch(A) = x \rangle \Leftrightarrow G_1 \vee G_2 \vee \dots, \text{ for all } G_i\text{'s such that } ch_{G_i}(A) = x. \quad (23)$$

Of course, we can no longer think of $\langle ch(A) = x \rangle$ as the proposition that at some time the chance of A is x , since chances are no longer time-indexed. Instead, $\langle ch(A) = x \rangle$ should just be thought of as the proposition that some grounding argument obtains for which the chance of A is x .

With these adjustments in hand, we can formulate a generalized version of PP_3 :

$$GPP_3: up(A | \langle ch(A) = x \rangle E) = x, \text{ if } E \text{ is admissible with respect to } \langle ch(A) = x \rangle, \quad (24)$$

or, in terms of credences:

$$GPP_3: cr_{\langle ch(A) = x \rangle E}(A) = x, \text{ if } E \text{ is admissible with respect to } \langle ch(A) = x \rangle. \quad (25)$$

³¹For proposals along these lines, and the reason such features need to be abandoned in order to accommodate statistical mechanics, see Arntzenius (1995), Loewer (2001), Meacham (2005), Hoefer (2007) and Nelson (2008).

Finally, consider PP_2 . Formulating a generalized version of PP_2 is straightforward. We simply replace TH with G , and PP_2 becomes:

$$GPP_2: up(A|G) = ch_G(A), \quad (26)$$

or, in terms of credences:

$$GPP_2: cr_G(A) = ch_G(A). \quad (27)$$

2. What happens if we also allow grounding arguments which can partially overlap? The effects of this change are more subtle, though no less important.

Once we allow such arguments, GPP_2 and GPP_3 are no longer equivalent. The derivations between PP_2 and PP_3 rely crucially on the assumption that grounding arguments do not partially overlap.³² Once we allow grounding arguments that can partially overlap, the analogous derivations won't hold between GPP_2 and GPP_3 .

Since these principles are no longer equivalent, we have a substantive choice to make. Which of these two principles— GPP_2 or GPP_3 —should we prefer?

A little consideration reveals that GPP_2 is preferable to GPP_3 . The reason is that only GPP_2 can accommodate theories with partially overlapping arguments of the kind that we're interested in, like statistical mechanics. If we try to apply GPP_3 to such theories, we'll run into contradictions.

These inconsistencies arise for statistical mechanics, but we can see how these problems arise using a simpler theory. Consider a chance theory T which holds at only three worlds, w_1 , w_2 and w_3 . Let every subset of $\{w_1, w_2, w_3\}$ correspond to the grounds of some chance distribution $ch_G(\cdot)$, and let $ch_G(\cdot)$ assign an equal chance to each of the worlds in G . So the distribution grounded by $G = w_1 \vee w_2 \vee w_3$ will assign a chance of $1/3$ to each of the three worlds, the distribution grounded by $G = w_1 \vee w_2$ will assign a chance of $1/2$ to w_1 and w_2 , and so on.³³

Now consider a subject whose total evidence is $w_1 \vee w_2 \vee w_3$. What should her credence in w_1 be?

GPP_2 delivers the correct result:

$$cr_{w_1 \vee w_2 \vee w_3}(w_1) = ch_{w_1 \vee w_2 \vee w_3}(w_1) = 1/3. \quad (28)$$

What about GPP_3 ? First consider the proposition $\langle ch(w_1) = 1/3 \rangle$. $\langle ch(w_1) = 1/3 \rangle$ corresponds to the disjunction of all of the grounding arguments G such that $ch_G(w_1) = 1/3$. In this case, there is only one such grounding argument: $G_a = w_1 \vee w_2 \vee w_3$. Thus $\langle ch(w_1) = 1/3 \rangle = G_a = w_1 \vee w_2 \vee w_3$. So GPP_3 entails:

$$cr_{w_1 \vee w_2 \vee w_3}(w_1) = cr_{\langle ch(w_1) = 1/3 \rangle}(w_1) = 1/3. \quad (29)$$

Now consider the proposition $\langle ch(w_1) = 1/2 \rangle$. $\langle ch(w_1) = 1/2 \rangle$ corresponds to the disjunction of all of the grounding arguments G such that $ch_G(w_1) = 1/2$. In this case,

³²In particular, the proofs depend on being able to replace a disjunction of grounding arguments with a disjunction of a subset of these arguments that are mutually exclusive. (See Appendix B.) If we assume there are no partial overlaps, we can always do this: just eliminate the grounding arguments which are proper subsets of other grounding arguments. But if we allow for partially overlapping grounding arguments, we can't always do this, and the proofs no longer work.

³³Recall that many different G 's can be true at the same world. (Given Lewis' grounding argument, for example, there will be infinitely many different TH's that are true at a world, one for each time.)

there are two such grounding arguments: $G_b = w_1 \vee w_2$ and $G_c = w_1 \vee w_3$. Thus $\langle ch(w_1) = 1/2 \rangle = G_b \vee G_c = (w_1 \vee w_2) \vee (w_1 \vee w_3) = w_1 \vee w_2 \vee w_3$. So GPP₃ entails:

$$cr_{w_1 \vee w_2 \vee w_3}(w_1) = cr_{\langle ch(w_1) = 1/2 \rangle}(w_1) = 1/2. \quad (30)$$

So GPP₃ makes inconsistent demands: it requires the subject's credence in w_1 to be both $1/3$ and $1/2$.

The source of GPP₃'s problem is that there are chance theories, like statistical mechanics or the simple theory just described, for which propositions asserting different chances are equivalent; i.e., correspond to the same set of worlds.³⁴ And since GPP₃ constrains credences via these propositions, this leads it to make inconsistent demands: a subject whose total evidence corresponds to these propositions is required to adopt inconsistent credences. GPP₂ avoids these difficulties because it doesn't constrain credences through propositions about chance. Instead, it employs the grounding argument directly.

6.3 The Winner

We've seen a number of formulations of the chance-credence principle. But one formulation:

$$\text{GPP}_2: cr_G(A) = ch_G(A),$$

has several advantages over the others.³⁵ It has no admissibility clause. It is strong enough to impose the desired constraints on our credences. It doesn't fail to apply in some cases. And unlike the other formulations of the chance-credence relation that we've looked at, it is intelligible and consistent regardless of what we take the grounding arguments to be.³⁶

References

- Arntzenius, Frank. 1995. "Chance and the Principal Principle: Things Ain't What They Used To Be." Unpublished Manuscript.
- Hajek, Alan. 2002. "Interpretations of Probability." The Stanford Encyclopedia of Philosophy.
- Hall, Ned. 1994. "Correcting the Guide to Objective Chance." *Mind* 103:505–517.

³⁴For those familiar with statistical mechanics, we can construct a statistical mechanical analog to this case in the following way. Take the grounding argument of a statistical mechanical chance distribution to be conjunction of the chance theory and the relevant macrostate. Consider every statistical mechanical macrostate TM_i such that $ch_{TM_i}(A) = 1/2$, for some A . This will consist of every union of A with another region of the phase space that has the same Liouville measure as A . So the disjunction of all TM_i 's will encompass the entire state space, S . Similar reasoning shows that the disjunction of all TM_j 's such that $ch_{TM_j}(A) = 1/3$ will also encompass all of S . It follows from GPP₃ that both $cr_S(A) = cr_{\langle ch(A) = 1/2 \rangle}(A) = 1/2$ and that $cr_S(A) = cr_{\langle ch(A) = 1/3 \rangle}(A) = 1/3$. Contradiction.

³⁵Notational variations aside, this formulation of the chance-credence principle is the identical to the formulations endorsed by Arntzenius (1995) (the "Scientific Principle"), Meacham (2005) (the "Basic Principle") and Nelson (2008).

³⁶I would like to thank Frank Arntzenius, Ned Hall, Barry Loewer, Tim Maudlin and Michael Strevens for helpful discussions on these topics. I would also like to thank Phillip Bricker, Maya Eddon and an anonymous referee for discussion and comments on this paper.

- Hofer, Carl. 2007. "The Third Way on Objective Probability: A Skeptic's Guide to Objective Chance." *Mind* 116:549–596.
- Lewis, David. 1979. "Attitudes *De Dicto* and *De Se*." *The Philosophical Review* 88:513–543.
- Lewis, David. 1986. A Subjectivist's Guide to Objective Chance. In *Philosophical Papers*, Vol. 2. Oxford University Press.
- Lewis, David. 1994. "Humean Supervenience Debugged." *Mind* 103:473–490.
- Lewis, David. 2004. "How Many Lives Has Schrodinger's Cat?" *Australasian Journal of Philosophy* 82:3–22.
- Loewer, Barry. 2001. "Determinism and Chance." *Studies in the History of Modern Physics* 32:609–620.
- Loewer, Barry. 2004. "David Lewis' Theory of Objective Chance." *Philosophy of Science* 71:1115–1125.
- Meacham, Christopher J. G. 2005. "Three Proposals Regarding a Theory of Chance." *Philosophical Perspectives* 19:281–307.
- Nelson, Kevin. 2008. "On Background: Using Two-Argument Chance." Forthcoming in *Synthese*.
- Strevens, Michael. 1999. "Objective Probability as a Guide to the World." *Philosophical Studies* 95:243–275.
- Thau, Michael. 1994. "Undermining and Admissibility." *Mind* 103:491–503.
- Vranas, Peter. 2002. "Who's Afraid of Undermining? Why the Principal Principle might not contradict Humean Supervenience." *Erkenntnis* 57:151–174.
- Vranas, Peter. 2004. "Have your cake and eat it too: The Old Principal Principle reconciled with the New." *Philosophy and Phenomenological Research* 69:368–382.

Appendix: Proofs

These derivations employ the following characterization of admissibility: *E* is admissible relative to $\langle ch_t(A) = x \rangle$ iff $\langle ch_t(A) = x \rangle E$ can be expressed as a disjunction of a subset of the *TH*'s associated with $\langle ch_t(A) = x \rangle$. Likewise, *E* is admissible relative to $\langle ch(A) = x \rangle$ iff $\langle ch(A) = x \rangle E$ can be expressed as a disjunction of a subset of the *TH*'s associated with $\langle ch(A) = x \rangle$. For simplicity, I will assume in the following derivations that there are only finitely many terms. (The proofs in Appendix A are recreated from Lewis (1986), p.96-100.)

Appendix A: Proofs Regarding the Relation Between PP₂ and PP₁/PP₁-

PP₂ ⇒ PP₁:

If E is admissible relative to $\langle ch_t(A) = x \rangle$, then:

$$\begin{aligned}
 up(A|\langle ch_t(A) = x \rangle E) &= up(A|T_1H_1^t \vee \dots \vee T_nH_n^t) & (31) \\
 &= \frac{up(A \wedge (T_1H_1^t \vee \dots \vee T_nH_n^t))}{up(T_1H_1^t \vee \dots \vee T_nH_n^t)} \\
 &= \frac{\sum_i up(AT_iH_i^t)}{\sum_j up(T_jH_j^t)} \\
 &= \frac{\sum_i up(T_iH_i^t) \cdot up(A|T_iH_i^t)}{\sum_j up(T_jH_j^t)} \\
 &= \frac{\sum_i up(T_iH_i^t) \cdot ch_{T_iH_i^t}(A)}{\sum_j up(T_jH_j^t)} \\
 &= x \cdot \frac{\sum_i up(T_iH_i^t)}{\sum_j up(T_jH_j^t)} \\
 &= x.
 \end{aligned}$$

PP₁ ⇒ PP₂:

Suppose $ch_{TH^t}(A) = x$. Then $TH^t \Rightarrow \langle ch_t(A) = x \rangle$, and thus:

$$\begin{aligned}
 up(A|TH^t) &= up(A|\langle ch_t(A) = x \rangle TH^t) & (32) \\
 &= x \\
 &= ch_{TH^t}(A),
 \end{aligned}$$

where the second step employs PP₁ and the fact that TH^t is admissible relative to $\langle ch_t(A) = x \rangle$. (This follows from the characterization of admissibility given above.)

PP₂ ⇒ PP₁-:

Once we replace E with a tautology, the derivation is identical to that given for PP₂ ⇒ PP₁.

Appendix B: Proofs Regarding PP₃/PP₃-

PP₂ ⇒ PP₃:

If E is admissible relative to $\langle ch(A) = x \rangle$, then:

$$\begin{aligned}
 up(A|\langle ch(A) = x \rangle E) &= up(A|E \wedge (\langle ch_{t_1}(A) = x \rangle \vee \dots \vee \langle ch_{t_n}(A) = x \rangle)) & (33) \\
 &= up(A|(E \wedge \langle ch_{t_1}(A) = x \rangle) \vee \dots \vee (E \wedge \langle ch_{t_n}(A) = x \rangle)) \\
 &= up(A|T_1H_1 \vee \dots \vee T_mH_m) \\
 &= up(A|T_1H_1 \vee \dots \vee T_rH_r)
 \end{aligned}$$

$$\begin{aligned}
&= \frac{up(AT_1H_1 \vee \dots \vee AT_rH_r)}{up(T_1H_1 \vee \dots \vee T_rH_r)} \\
&= \frac{\sum_i up(AT_iH_i)}{\sum_j up(T_jH_j)} \\
&= \frac{\sum_i up(A|T_iH_i) \cdot up(T_iH_i)}{\sum_j up(T_jH_j)} \\
&= \frac{\sum_i ch_{T_iH_i}(A) \cdot up(T_iH_i)}{\sum_j up(T_jH_j)} \\
&= \frac{\sum_i x \cdot up(T_iH_i)}{\sum_j up(T_jH_j)} \\
&= x.
\end{aligned}$$

The fourth step replaces the disjunction $T_1H_1 \vee \dots \vee T_mH_m$, where the terms need not be mutually exclusive, with a disjunction of a subset of these terms, $T_1H_1 \vee \dots \vee T_rH_r$, which are mutually exclusive. We can do this because if any two elements T_iH_i and T_jH_j of the disjunction aren't mutually exclusive, then one will be a subset of the other, and we can discard it.

PP₃ ⇒ PP₂:

Suppose $ch_{TH}(A) = x$. Then $TH \Rightarrow \langle ch(A) = x \rangle$, and thus:

$$\begin{aligned}
up(A|TH) &= up(A|\langle ch(A) = x \rangle TH) \\
&= x \\
&= ch_{TH}(A),
\end{aligned} \tag{34}$$

where the second step employs PP₃ and the fact that TH is admissible relative to $\langle ch(A) = x \rangle$. (This follows from the characterization of admissibility given above.)

PP₂ ⇒ PP₃•:

Once we replace E with a tautology, the derivation is identical to that given for PP₂ ⇒ PP₃.