§1 Introduction

David Lewis's classic 1980 paper "A Subjectivist's Guide to Objective Chance" helped set the terms for one of the most vibrant debates in philosophy of science and metaphysics of the last 35 years. Lewis saw, as had others before him (e.g. Carnap 1945), that probability theory, as a branch of mathematics, has distinct applications to the epistemic and physical domains: very roughly, we must distinguish probability as a way of representing *rational credence* from probability as a way of describing certain aspects of the *physical world*. We therefore require not just more than one answer to the question "what is probability?", but an account of how these answers relate to each other. Now, we'll see below that some of the terms that Lewis set for the debate that followed him were confused; but the debate itself has been enormously philosophically productive. And one absolutely central point that has emerged quite clearly from it concerns the significance of deep metaphysical questions about the nature of the physical modalities (laws of nature, causation, chance) for our understanding of physical probability, and its connection to rational credence. That is what this essay explores. After laying out a useful formal framework (§2), and saying just what those deep metaphysical questions center on (§3), we'll explore in detail just how different "chance" (as we'll henceforth label the kind of probability that attaches to the physical world) looks, depending on the correct answer to those questions (§§4-6). I'll close (§7) by listing some particularly interesting open questions that, I hope, will continue to drive this fascinating debate forward.

§2 A formal framework

Lewis took for granted a certain understanding of, as it were, the *structure* of chance. On this understanding, the canonical expression of a fact about objective chances has this form:

The objective chance, at time t, that proposition P is true, is x.

Note three happy-making features of this understanding. It is admirably clear about what chances attach to; it steers us away from the confused thought that there is such a thing as *the* chance of P; but it does so without steering us toward the confused frequentist idea that chances are had only relative to a choice of reference class. On the other hand, it builds in a different confused thought, which is that chances are, in the first instance, defined *at times*. What of worlds such as our own, whose relativistic spacetime structures serve up no objective basis for talk of "times"? And even in worlds with Newtonian or neo-Newtonian spacetimes, we *might* want to allow that chances can attach to their initial conditions; while these might be chances *about* a time (namely, the first time), they seem not be chances *at* that, or any other, time. Our most basic way of talking about chance should not rule out this possibility.

Here is a better – because more flexible – framework. Take each possible world w to have associated with it a single "ur-chance" function $urch_w$, defined over a set S of possible worlds (which we will take to include w).[1] For a given proposition P, $urch_w(P)$ is the chance according to w that P is true, and is simply the sum of the probabilities $urch_w$ attaches to worlds in S that make P true.[2] Chance is thus modeled as a contingent feature of a world; but it is, in the first instance, a feature *of that world*, not of a *time* within it.

As it should, the framework readily allows for chances at times, in two different ways. Given some time t, let S_{tw} be the proposition true in a world v iff the state of v at t is exactly the same as the state of w at t; let H_{tw} be the proposition true in a world v iff the history of v up through t is exactly the same as the history of w up through t. Then either $urch_w(\bullet \mid S_{tw})$ or $urch_w(\bullet \mid H_{tw})$ can serve as "the" chances at t (in w). (Which one should we choose? We can, and should, let the answer depend on what particular use we have for a notion of chances-at-times.)

The framework also allows for chances over initial conditions – even when there is no first moment of time – as follows: Say that two worlds w and v "start out the same" iff there is some moment of time t such that they are exactly qualitatively alike up to t.[3] *Starting out the same* is an equivalence relation; so it partitions S.[4] We can thus take the chances over initial conditions to be given by the values $urch_w$ attaches to the elements of this partition.

---

[1] When the set S is non-denumerable, we do the usual thing: define $urch_w$ over an algebra of subsets of S. Also, we will see below some reason to think that worlds may have *multiple* ur-chance functions associated with them.

[2] With, of course, the usual axioms in place: these probabilities are non-negative, and sum to 1.

[3] In a relativistic setting: iff there is some Cauchy surface C in w and C* in v such that the portion of w to the past of C is exactly qualitatively the same as the portion of v to the past of C*.

[4] In a relativistic setting, we need to work a bit harder to get this relation to be an equivalence relation. One approach: assume that there is some way of partitioning each world in S into a set of non-overlapping but exhaustive Cauchy surfaces. In the definition of "starts out the same" given in the last footnote, restrict the choice of Cauchy surfaces to *these* ones.

Finally, in a relativistic setting we can, by the method of conditionalizing on a suitably chosen proposition, define useful notions of spatio-temporally "localized" chances. For example, given a region of spacetime R, let the proposition P exactly describe the state of the past light-cone of R; then we might identify the "chances-at-R" with $urch_w(\bullet \mid P)$.

I'm going to assume this framework, henceforth. But – full disclosure – there's at least one reasonable misgiving you might have, and that concerns the oft-endorsed thesis that "what's past is no longer chancy". Lewis, in taking chances to be defined in the first instance at times, could treat this thesis as *substantive* (as he does in Lewis 1980): For any time t, and any proposition H entirely about matters prior to t, the chance-at-t of H is 1. But on one way I've offered of defining chances-at-times, the thesis is straightforwardly false (or true, only on remarkable assumptions about the laws that give rise to the ur-chances), whereas on the other, it's perfectly trivial. I've yet to decide whether this is a feature or a bug. The case for "bug" is clear enough: Other things equal, don't adopt a framework that by itself makes it impossible to see some interesting thesis as both non-trivial and true. The case for "feature" is more subtle, and I'll just indicate the main idea: Perhaps we are tempted to think that it's an interesting question whether what's past is no longer chancy *only because* we remain in the grip of a bad metaphysical picture of time, according to which the past differs in some profound metaphysical manner from the future – say, by being "fixed" or "settled" or "immune from causal influence". But we should have learned enough physics by now to have given up any such picture. Which is not – of course! – to say that nothing profound or interesting distinguishes past from future; it's just that the distinction will be found in physics, not metaphysics.[5] And (punch line) physics has no use for the claim that what's past has, from the standpoint of the present, chance 1.[6]

One final point. The framework is quite flexible – much more flexible than you might need, depending on your views about the nature and source of objective chances. For example, if you think about chances as, in the first instance, characterizing the lawful transition from earlier to later states (see for example Maudlin 2007), you'll consider it incoherent to attach chances to initial conditions, and you'll likely have no interest in "backwards looking" chances (chances defined *at* times, concerning matters *prior* to those times). That doesn't prevent you from using the framework to represent

[5] Perhaps in the way described in Albert 2000, for example. For some discussion, see §4.8 and §5.8, below.

[6] And that might, in turn, be because physics has no need to assign t-chances to events prior to t; see the next paragraph.

what you think is real, concerning chances. It's just that surplus mathematical structure will come along for the ride, structure to which you attach no physical significance.

§3 The Great Divide

A philosophical account of the physical modalities faces a basic choice point. On the first, "Humean" branch, laws, causes, and chances are nothing more than certain kinds of *patterns in the non-modal phenomena*. Or – perhaps better – the *rationale* for having concepts of these modalities is simply to allow for the more effective organization of information about the non-modal facts. On the second, anti-Humean branch, the physical modalities have some sort of objective reality above and beyond that of the non-modal facts, and consist in some kind of *fundamental constraints* on how these facts unfold.

For more precision, we need an account of what "non-modal" comes to, and what "nothing more than" comes to. As to the first, we might begin by focusing on the most fundamental physical magnitudes instantiated at our world, and see these as "non-modal" in the following sense: the instantiation of one or more of these magnitudes by some particulars places no metaphysical constraints on their instantiation by other particulars. Slightly more precisely, we might capture the non-modal character of these magnitudes via the thesis that the *metaphysical* possibilities concerning their instantiation coincide with the *purely combinatorial* possibilities. (For discussion, see Hall 2010.) As to "nothing more than", a good start is contained in the well-known thesis of "Humean supervenience" (see, e.g., Lewis 1994): No two possible worlds differ in their physical modalities without differing, somehow, in their non-modal facts (or, given the account of the latter, in the instantiation within them of fundamental physical magnitudes).[7] Thus, an anti-Humean account can earn that status either by denying that fundamental physical magnitudes are non-modal,[8] or by positing some additional ingredient to reality from which the physical modalities derive.[9]

Like Jenann Ismael (this volume), I think that a deeper issue lurks behind this divide, one concerning the nature of explanation and understanding.

---

[7] To properly capture the idea that facts about physical modalities *reduce to* non-modal facts, we might want a stronger notion than supervenience; for example, perhaps the right account is that modal facts are *grounded in* non-modal facts. See, e.g., Rosen 2010.

[8] For example, perhaps part of *what it is* for something to have mass is for it to attract other things with mass (see e.g. Ellis 2002).

[9] For example, perhaps laws consist in higher-order relations of necessitation between universals (see e.g. Armstrong 1983).

Begin with a truism: Scientific inquiry aims to provide us with *understanding* of the world. Here is a fundamental philosophical question about such understanding. Does it consist in the possession of a *special kind of information*, or does it rather consist in having one's information *organized in a special sort of way*? ("Some of both" is also an option!)

To fix ideas, it will help to have some examples illustrative of each position. For the "special kind of information" position, examples seem awfully easy to come by. The window broke. Why? What *explains* the breaking? What do we need to know, in order to *understand* how this came about? Just this: Suzy threw a rock at it, which struck the window with sufficient force to break it. Of course there's plenty more we could add, in order to enrich our understanding; but it is quite striking that understanding seems to *begin*, at least, with knowledge of the breaking's *causes*. Which is to say, knowledge of a special kind of information about the breaking.

A second example. As a planet orbits the sun, the line joining it to the sun sweeps out (to a very close approximation) equal areas in equal times; this is Kepler's second law. What explains this regularity? Newton's laws of motion, together with the fact that the gravitational force of the sun on a given planet dominates all other forces on that planet, *and* is a central force (i.e., a force acting on a line joining the sun with the planet).[10]

Now for a very different example. Consider the following initial segment of an infinite sequence of natural numbers:

1,1,1,2,3,2,1,3,5,4,2,5,7,8,3,7,9,16,5,11,11,32,8,13,13,64,13,17,…

Perhaps you've figured out the rule that generates the sequence. Perhaps, on the other hand, you find it confusing. You don't understand it. You don't know why it has the form it does. If so, the following way of reorganizing the initial segment will make things crystal clear:

[10] Note one difference between the two examples. In explaining Kepler's second law, we pick out a range of facts some of which appear to have a *special metaphysical status*: they are laws of nature. (And yes, we should remember that for many philosophers, these appearances will be misleading.) By contrast, the fact that Suzy threw a rock at the window is hardly a metaphysically special sort of fact; what matters, rather, is that it bears (apparently!) a metaphysically special relationship (viz., *being a cause of*) to the event to be explained. In short, on a "special sort of information" view, explanatory information might itself be special, or might instead be ordinary information picked out because of its special relationship to the explanandum.

1,	1,	1,	2,
3,	2,	1,	3,
5,	4,	2,	5,
7,	8,	3,	7,
9,	16,	5,	11,
11,	32,	8,	13,
13,	64,	13,	17,…

Looking down the columns, we see that the sequence is just an interleaving of the odd numbers, powers of 2, fibonacci numbers, and prime numbers. Once you see this, you understand the sequence. But not by acquiring a special sort of information about it. (The sequence is, after all, not the sort of thing that has "causes", or that "metaphysically depends" on anything else.) To me, examples like this evoke in its purest form the idea that to understand some subject matter is to organize one's information about it in the right sort of way. (Of course, actual explanatory practice in mathematics is the place to look, for more serious examples.)

Here is the key upshot. One broad approach sees an explanation of some phenomenon as consisting in information about *objective constraints that apply to it* (that's what's "special"): thus, an explanation of an event might detail its *causes*; an explanation of some unbroken regularity might cite the *laws* that issue in it; an explanation of some robust statistical patterns might appeal to underlying *chances* that generate them. In each of these cases, the conceptual distance between what is to be explained and what does the explaining argues in favor of (without outright implying) an *anti-Humean* account of the physical modalities: the core idea is that since these modalities are, or reflect, explanatory constraints on the (non-modal) phenomena, they cannot be identified with, or even reduce to, such phenomena.[11] But another broad approach see an explanation of some phenomenon as succeeding to the extent that it fits that phenomenon into some suitably broad, organizing framework; following one of the few worked-out philosophical accounts of explanation along these lines (Kitcher 1989), we might call this "pure unificationism" about explanation. What makes such a framework *succeed*, when it comes to conferring understanding, is not that it contains information about "underlying constraints", but that it bears the right sort of logical or more broadly epistemic relationship to the particular phenomena we wish to

[11] Many – most, really – contemporary accounts of scientific explanation fall under this broad approach. See for example Woodward 2005 and Strevens 2009.

understand. And that is an approach that argues in favor of (again, without outright implying) a *Humean* account of the physical modalities.

Zeroing in on the concept of "chance", now, we should appreciate that, on a very natural (if not inevitable) way of developing Humeanism and anti-Humeanism, each proceeds from a *profoundly* different understanding of the *point* of such a concept. For the Humean, it is a concept that plays a role in the construction of framework particularly suited to *organizing statistical information*, to *locating* particular statistical phenomena within some more comprehensive scheme. For the anti-Humean, it is a concept of a metaphysically distinctive kind of *partial causal-explanatory constraint*, one that must be invoked in order to properly understand (among other things) how statistical phenomena *are generated*. This quite stark difference in orientation has, as we'll see in the next three sections, a profound effect on the detailed shape of Humean vs anti-Humean accounts of chance.

§4 Anti-Humean chances

We'll start with some examples of dynamical chances, described from a distinctively anti-Humean perspective – a perspective, remember, that sees chance as *partial causal constraint*. We'll then draw attention to eight features such chances display. These features will turn out to look very different – or look nonexistent – when viewed from the perspective of a Humean approach to chance.

§4.1 Examples

To find real-world examples of physical theories that posit objective chances, we could turn to quantum mechanics – in particular, versions of quantum mechanics that treat the "collapse of the wave function" as an objective physical phenomenon.[12] But these examples require quite a bit of setup, so I am going to make do with a pair of toy examples instead. (These examples will also make it easy to illustrate a few key points.)

For the first example, consider a world consisting of point-particles, that interact when they come within a certain distance d of each other. These particles can be in any of four possible intrinsic states, which we'll call 'dormant', 'excited', 'red', and 'green' (thus: the 'DERG' world). Here is the key interaction: When two *dormant* particles come within d, each becomes excited. From that point on, each of the pair of excited particles has a certain fixed probability per unit time of 'decaying'; when it

[12] For an excellent overview, see Ghirardi 2016.

does so, it becomes either red or green, with a certain chance of each. Two *fundamental constants* thus characterize the objective probabilities in this world: the probability per unit time of decay; and the probability that decay will result in green.[13]

For the second example, consider a world consisting of particles that sometimes enter what we'll call "T-junctions". When a particle enters such a junction, it has a certain chance r of exiting to the right, and a chance (1 - r) of exiting to the left. Again, we should think of r as a fundamental constant that helps characterize the laws at this world.[14]

Finally, we'll take for granted that the fundamental dynamical laws of these two worlds do more than simply specify the chances just described; in each case, they determine, when given a complete physical state for the world at some time t, a probability distribution over the physically possible post-t futures. (That is, as it were, their primary *job*: to say how the world can evolve, and with what likelihood.) So there is quite a bit more structure to these chances than can be read off, just from the descriptions given above. For example, from the fact that (say) any excited particle in our DERG world has a 50% chance of decaying into green, all that follows about the chance that a *pair* of interacting particles will subsequently *both* decay into green is that it lies in the interval [0, 0.5]; so it's up to the dynamical laws to settle, precisely, which value in this interval is the correct one. Below, several important points will emerge concerning the additional chance structure the laws yield, by filling in such gaps.

§4.2 Law

Chances, in our examples, bear a particularly intimate connection to laws. For now, we can bring this out by noting what the connection *isn't*: It's not that there are, along with all the other non-nomic facts, some non-nomic facts about *chances*, and the laws apply to the latter just as they do to the former. Put another way, it's not that laws *govern* chances in the same way that they govern, say, particle motions. It's rather that chances have their *ontological ground* in probabilistic laws: *what it is* for a proposition P to have, at time t, a chance x of coming true, is for the fundamental dynamical laws, applied to the state of the world at time t, to endow that state with a certain objective propensity to bring about a future in which P. For one illustrative upshot, notice that, given a world w governed by the laws in our first toy example, there will be a world v with (i) no laws whatsoever, but that (ii) perfectly duplicates the motions and intrinsic

[13] Or red: it obviously doesn't matter which outcome we choose to focus on.
[14] And yes, "right" and "left" are shorthand for directions that are physically distinguished somehow by the structure of the T-junction.

properties exhibited in w. (Remember that we're working, here, in an anti-Humean setting; so there is no problem in positing such a world v.) What there *won't* be is a world v* with (i) no laws whatsoever, but that (ii) perfectly duplicates the motions, intrinsic properties, *and chances* exhibited in w. In a slogan: *no chance without law.*

§4.3 Time

Because chances, in the first instance, characterize the way in which a complete physical state lawfully generates successive physical states, chances are (on the conception we're now investigating) had relative to times.[15] So the structure provided by the "ur-chance" framework set out in §2 carries a lot of excess representational baggage. Still, there is no difficulty in using that framework to *recover* chances-at-times, as we saw: letting H_t be a proposition that completely characterizes the (non-modal) history of the world up through t, and letting S_t be a proposition that specifies the complete physical state at t, we can identify the t-chances either with $urch(\bullet \mid H_t)$ or with $urch(\bullet \mid S_t)$.

However, the ur-chance framework *imposes* structure on dynamical chances that is not explicitly guaranteed, merely by construing them on the "partial causal constraint" model. Suppose we say this much, and no more: the fundamental dynamical laws, when applied to the complete state of the world at any given time t, yield a probability distribution over a set of possible t-futures. Then, for all we've said, there need be no connection whatsoever between these various probability distributions.[16] But as soon as we impose the ur-chance framework, we get such connections. Suppose, for example, that we identify the t-chances with $urch(\bullet \mid S_t)$. Let S_1 specify the state at time 1, S_2 the state at time 2. Let F be any proposition. Then

$$ch_1(F \mid S_2) = urch(F \mid S_1 \& S_2) = ch_2(F \mid S_1)$$

We should expect more temporal structure than this, though. To capture the usual understanding of probabilistic laws in fundamental physical theories, we should, minimally, add a condition to the effect that *the present screens off the past from the future*. Suppose that F is a proposition entirely about post-t history, P a proposition entirely about pre-t history. Then, stated within our framework, the condition is as follows:

---

15 Generalized to a relativistic setting, chances are had relative to Cauchy-surfaces.

16 Well, there will be this *minimal* one, assuming the dynamical laws are time-translation-invariant: if the state of the world at t_1 is exactly the same as the state of the world at t_2, then the t_1-chances will be the same (modulo the time-translation) as the t_2-chances.

$urch(F \mid S_t \& P) = urch(F \mid S_t)$

Notice that with this condition in place, it will follow (on either way of identifying t-chances) that when $t_2 > t_1$, the t_2-chances of post-t_2 events are equal to the result of conditionalizing the t_1-chances on the intervening history.[17]

Why build in the screening-off condition? Principally because doing so reflects the view that any causal influence of earlier conditions on later conditions must be *fully mediated* by the intervening history; there cannot, in other words, be any causal influence *across a temporal gap*.

§4.4 Explanation

Suppose that the following large-scale statistical pattern obtains in a T-junction world: Over all of history, very close to 61.8% of exits from T-junctions are left-exits. Suppose further that the distribution of right- and left-exits is random. This is the sort of statistical pattern that is – in principle, at least – *explicable*. If, for example, the fundamental dynamical laws assign a chance of $(\sqrt{5} - 1)/2$ that a particle entering a T-junction will exit to the left, and if in addition these laws treat these events as *independent* (see below, §4.6), then the statistical pattern we've identified will be explicable by direct appeal to these laws.

Quite a good philosophical question now arises: What detailed philosophical account of explanation will vindicate the claims of the last paragraph? (A causal account? Not obviously. A Hempel-style covering-law account? Yes, but those have there own serious problems.) I want to set that question aside. Instead, I'd like to draw attention to a *contrast* between the explicability of *this* pattern, and the *inexplicability* of other sorts of statistical patterns. I have two cases in mind.

First case. As before, we're in the T-junction world, and as before, the fundamental laws assign a chance of $(\sqrt{5} - 1)/2$ to left-exits. But something highly (objectively) unlikely happens: just 2.78% of exits are left-exits. Why? What explains this massively improbable outcome? Nothing, I think. We can say this much: Given the laws, this

[17] A quick proof. Let F be about post-t_2 history. Let H_1 completely describe history up through t_1, and H_{12} completely describe history after t_1 and through t_2. Let S_1 describe the state at t_1, S_2 the state at t_2. On the first identification, $ch_2(F) = urch(F \mid H_1 \& H_{12}) = ch_1(F \mid H_{12})$. (So the screening-off condition isn't necessary.) On the second identification (and this time making use of the screening-off condition), $ch_2(F) = urch(F \mid S_2) = urch(F \mid S_2 \& S_1 \& H_{12}) = ch_1(F \mid H_{12})$.

outcome is *possible*. But that's it – and that's not much at all, by way of genuine explanation.

Second case (a bit more involved). We're in a different world this time, one that is mostly boring: it contains billiard balls, floating around in empty space, occasionally undergoing perfectly elastic collisions (but otherwise not interacting at all). But there's one interesting detail. In a certain corner of space, there is a time-travel entrance portal, with the exit portal nearby, facing it. At a certain time, a billiard ball is about to sail right between the entrance and exit portals. What will happen? Well, two things *can* happen, given the laws and the spacetime structure and the exact positioning of the ball. The ball could sail through in a straight line. Or, as it approaches the line connecting the two portals, it could happen that a ball sails out of the exit portal, and collides with our ball so as to send it into the entrance portal. Whence, of course, it is the very same ball, colliding with itself. Notice that there is nothing whatsoever in the fundamental structure of this world that could generate an *objective chance* for either outcome. And so if the ball undergoes self-collision, and we ask "why?", there is *no answer to give* beyond "well, that's one of the two things it *could* do". This sad verdict won't change, if we add more billiard balls. So suppose a gazillion go sailing through this region, and that exactly 30% of them – randomly distributed – undergo self-collision. What could explain this statistical regularity? Nothing. Nothing whatsoever. The contrast with our first example from the T-junction world is stark: for in that case, we can see the statistical behavior of the exits as being *constrained* – not perfectly, but highly – by the probabilistic laws. Not so, with our time-traveling billiard balls.

In sum: on the conception of chance we're now exploring, it is unambiguously possible for the world to feature a statistical regularity – large-scale, with appropriately random structure – that has no explanation.

§4.5 Precision

Pick some proposition P and time t. On the conception of chance we're now exploring, the fundamental dynamical laws specify a *perfectly exact* chance of P, at t. In our DERG world, the two fundamental constants – that give the probability per unit time of decay, and the probability of decay into red or green – are themselves perfectly precise. In our T-junction world, the chance of left-exiting is perfectly precise. In the Ghirard-Rimini-Weber version of quantum mechanics, the probability-per-unit-time of a "collapse" is perfectly precise.

Two points. First, nothing about our formal framework prevents us from representing *imprecise* chances. All we need do is take the chances at a world to be represented not by a *single* ur-chance function, but by a *set* of them.[18] We'll see soon enough a good reason to take advantage of this flexibility (albeit in the *Humean* setting). Second, that a chancy world must be precise in this way seems to me not at all a conceptual truth (I'm not sure how it *could* be), but a substantive ontological assumption, albeit one deeply entrenched in physics. To appreciate *how* deeply, we should notice that this assumption of precision is not really about chances *per se*, but applies across the board. Consider, for example, Newtonian gravitational mechanics: when we first learn this theory, almost all of us take for granted – without giving it a moment's thought – that the gravitational constant G has some *perfectly precise* value. And this, even though there is no technical obstacle whatsoever to supposing it *doesn't*. For example, we might (i) take G to have an imprecise value represented by a *set* of precise values $\{G_i\}$; (ii) interpret the gravitational force law as describing a *constraint* on the gravitational force one particle exerts on another: this force must belong to the set $\{G_i m_1 m_2/r_{12}^2\}$. The result is a kind of radical non-determinism. Once we've combined the gravitational force law, the law of composition of forces, and Newton's second law, we get a theory that 'says' that, given such-and-such an initial state, the particles must move in one the following ways: …. But it says nothing more than this.[19]

I think it's safe to say that such a version of Newtonian gravitational mechanics would strike most if not all practicing physicists as *deeply* unattractive. But that is not, I think, merely because it fails to be *deterministic*. (After all, ordinary Newtonian mechanics is *already* non-deterministic: see for example Xia 1992.) No, the offending feature is its lack of *precision*.

§4.6 Independence

We expect chances, on the conception we're now exploring, to be characterized by certain kinds of basic *independence* principles – and these principles get their intuitive force, I think, from the background conception of chance as partial causal constraint. We've already seen one such principle: since the past cannot act on the future except by way of causally intermediary states, the chances should obey the screening-off

[18] In the Bayesian literature on imprecise credences, this is a familiar move: take a rational agent's opinion to be represented not by a *single* probabilistic credence function, but by a *set* of them. See for example Joyce 2010.

[19] The same point applies to *mass*, in Newtonian mechanics: there is no technical obstacle to setting up a variant of the theory that posits *imprecise* masses for particles. For example, given some (say) perfectly precise net force F applied to some particle with imprecise mass m, represented by a set S of precise values $\{m_i\}$, we understand Newton's second law as saying that whatever acceleration the particle has is constrained to be one of the values F/m_i. So, again, the theory is (fairly radically) non-deterministic.

condition described in §4.3. But there is another such principle, famous from discussions of Bell's Inequalities. I will state the principle in a slightly non-standard way.[20]

Consider two regions of spacetime R_1 and R_2, and some propositions P_1 and P_2, where P_1 is entirely about R_1, P_2 entirely about R_2. Suppose that R_1 and R_2 are so separated that it is impossible for what transpires in one to have any influence on what transpires in the other. (For example, they might be spacelike separated.) For all that, there may well be *probabilistic* connections between them; e.g., it may be that $urch(P_1 \mid P_2) \neq urch(P_1)$. Now consider some time t, that precedes both R_1 and R_2; let proposition S completely describe the state at t.[21] Then our second independence principle says that

$$urch(P_1 \mid P_2 \,\&\, S) = urch(P_1 \mid S)$$

To see the force of this principle, return to our DERG world. At time t_0, dormant particles A and B interact, and thereby become excited. At later time t, A and B are very far apart (and guaranteed to remain so); each is still in the excited state. Let G_A be the proposition that A decays into the green state, G_B the proposition that B decays into the green state. Let S be the state of the world at time t. What is $urch(G_A \,\&\, G_B \mid S)$? As the start of an answer, observe that

$$urch(G_A \,\&\, G_B \mid S) = urch(G_A \mid G_B \,\&\, S) \bullet urch(G_B \mid S)$$

If we now add the assumption that G_A and G_B are entirely about suitably causally separated regions of spacetime, then we may conclude that

$$urch(G_A \,\&\, G_B \mid S) = urch(G_A \mid S) \bullet urch(G_B \mid S)$$

Our laws, remember, posit a certain fixed chance g that any given excited particle will decay into the green state. So, if this independence principle holds, it follows that

[20] Aficionados will recognize that I can state the principle in the way that I do because we are, here, presupposing that the probabilities we are dealing with are chances that arise from fundamental dynamical laws. The standard formulation of Bell's "locality" condition, by contrast, achieves greater generality, by allowing for a non-trivial probability distribution over the relevant initial conditions. The key point is that the failure of our independence condition entails the failure of Bell's locality condition, but not vice versa.

[21] Or, in a relativistic setting, some Cauchy-surface that does duty as a "time".

any given pair of interacting particles have a chance of g^2 of decaying in the pattern green-green.

Of course, it might *not* hold. For example, suppose that the fundamental laws of our DERG world dictate that, while the chance that any given excited particle decays into green is, say, 0.5, the chance that an interacting *pair* decay into the same state is 0 – even though the two excited particles in the pair are *exactly physically alike*. Now, if you were a scientist in this world, you wouldn't believe it. Given the perfect correlation, you would take for granted that there are two physically distinct *kinds* of 'excited' state: one which *guarantees* that its particle will decay into green, the other of which guarantees red. And you would propose that, as a matter of fundamental dynamical law, the interaction between two particles puts one into the first kind of excited state, and the other into the second (with a certain chance of each outcome). Certainly, nothing about the empirical data in this world would prevent you from theorizing in this way. But for all that, you would be *wrong*, as nothing whatsoever distinguishes the states of our two particles, post-interaction.

It appears, rather shockingly, that in our very own actual world we have experimentally verified examples of robust probabilistic correlations between goings on in distant spatiotemporal regions, where these correlations *cannot* be accounted for by any theory that takes them to be screened off (in the way our principle describes) by a prior physical state. (These are the famous violations of Bell's inequalities; see Maudlin 2011 for a thorough and expert treatment of their significance.) Of course, that doesn't mean that our principle *fails*: it may not *apply*, namely, if what happens in one of these distant regions in fact *can* influence what happens in the other. But that's shocking enough. Either way, we appear to have a genuine mystery: causal influences across arbitrarily wide spatiotemporal distances; or a kind of non-causal coordination between chancy outcomes in distant regions, built into the very fundamental laws.

§4.7 Asymmetry

It is part of our ordinary way of thinking about chance that chances are *future-directed*. Sometimes, this prejudice finds expression in the thesis that "what is past is no longer chancy" – i.e., if P is a proposition entirely about pre-t history, then P's chance at time t of being true is 1 if it *is* true, 0 otherwise. But a more modest position is that the fundamental dynamical laws, as applied to the state of the world at time t, only yield determinate chances towards the future of t, falling silent (or perhaps highly indeterminate) when it comes to pre-t history. If we think of fundamental dynamical laws as being in the business of supplying a probability distribution as output, when fed a complete instantaneous state as input, then this more modest position is surely preferable: for on the *less* modest view, we would be forced to suppose that the laws were able to determine, as it were, the complete truth about what happened before t, just from the state of the world at t. And if the fundamental dynamics are probabilistic, that's likely absurd, since it could easily happen that *distinct* past states could lawfully evolve into the *same* present state.

At any rate, it's genuinely unclear to me whether there is any merit to *either* position – to *any* view of chance that builds in some kind of temporal asymmetry. The time-symmetry of our most famous examples of worked-out fundamental theories (Newtonian mechanics, Maxwellian electrodynamics, Schrödinger quantum mechanics) suggests that we shouldn't give up such time-symmetry lightly. And there is no serious *technical* obstacle to seeing the dynamical laws as yielding precise probabilistic constraints towards the past, as well as the future; just observe that the ur-chance framework has no temporal direction built into it.[22] Finally, as I noted in §2, our prejudice in favor of future-directed chance may derive from a deeper prejudice toward seeing the past as metaphysically "fixed", by contrast with the "open" future. If you suspect that this deeper prejudice is misguided, then you may find yourself with no reason to object to past-directed chances.

§4.8 The initial state

Chances are, on the current conception, *transition* chances: in the first instance, they characterize the *change* from one state to another. Accordingly, the only sense that can be attached to an assignment of chances to the *initial* state is this: at some *later* time t, the chance that the initial state is such-and-such is x. In our framework, that is captured by an equation of the form urch(I | S_t) = x, where I is a proposition about the initial

[22] Though it's another matter whether the simplest way to implement this framework will yield up *time-symmetric* probabilistic laws. Here is one way to proceed, if we desire such laws. Suppose we have some laws that determine a set of complete, nomologically possible histories for the world. Suppose we have some well-defined sense in which one complete history H* counts as the *time-reverse* of some other complete history H. And suppose that, so far, we have implemented time-reversibility to this extent: a complete history H is nomologically possible iff its time-reverse H* is as well. Next, if S is a set of complete histories, we can define the 'time-reverse' S* or S to be the set of histories consisting of all and only time-reverses of histories in S. Then a natural way to capture the thought that the *chances* are time-symmetric is via the requirement that, for any set S of nomologically possible complete histories such that urch(S) is defined, urch(S*) is likewise defined, and urch(S) = urch(S*). Note, though, that making this work even for toy examples is hardly straightforward. For instance, the laws of our DERG world fail to be time-symmetric, even in what they say about what is nomologically possible.

state. So – if the only kinds of objective probabilities there *are* are these transition chances – the *unconditional* values urch(I) have no representational significance.[23]

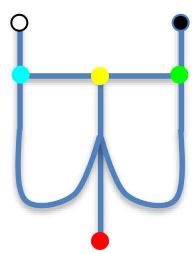
That may be a cost. Albert, for example, has argued persuasively that in order to gain significant predictive power, even from a fully *deterministic* dynamical theory, we must augment that theory (2015b, p. 5):

> The patched-up picture, then, consists of the complete deterministic microdynamical laws and a postulate to the effect that the distribution of probabilities over all of the possible exact initial microconditions of the world is uniform, with respect to the Lebesgue measure, over those possible microconditions of the universe which are compatible with the initial macrocondition specified in the past hypothesis, and zero elsewhere.

If Albert is right, then we need an understanding of what these probabilities *are* that are uniformly distributed over a certain range of initial states. And the conception of chance as partial causal constraint fails to provide any such understanding. (For an argument that this is no cost – and that the work Albert identifies can be done in other ways – see Maudlin 2007.)

§4.9 Imperialism

Consider our T-junction world. Suppose that the fundamental dynamical laws assign a chance of left-exit of 0.9. Suppose these laws treat each exiting event as independent of all others. Then consider the following structure, which we'll call a "Maze":



When a particle enters a Maze (at the red dot), it first encounters the central T-junction (yellow dot); it can either (i) go right to the T-junction at the green dot, then right again, in which case it returns to the central T-junction; (ii) go left to the T-junction at the blue dot, then left again, in which case it returns to the central T-junction; (iii) go left then right, in which case it exits at the white dot; (iv) go right then left, in which case it exits at the black dot. Suppose each of these outcomes takes exactly 1 second

[23] In which case we could represent the chances at a world w equally well by any two functions $urch_1$ and $urch_2$ that agree on all probabilities conditional on the propositions S_t.

to unfold. Then observe that, once the laws governing behavior at T-junctions are settled, *so too* are the chances pertaining to Maze behavior. If a particle enters a maze at time t, then its chance of having exited by t + n seconds is exactly $1 - (0.82)^n$; and its chance of exiting at the white dot is exactly 0.5. This case illustrates a general point: the chances pertaining to fundamental events *settle*, in imperialistic fashion, the chances of any non-fundamental events. (Here, an exit from a T-junction counts as a fundamental event, an exit from a Maze as non-fundamental.)

It's important to recognize that there are ways of understanding "objective probability" that *don't* have this feature. Suppose, for example, that we adopted a certain kind of simple frequentist approach to chance, according to which the chance that a particle left-exits a T-junction is *identified* with the frequency with which particles left-exit T-junctions, and the chance that a particle white-dot-exits a Maze is likewise *identified* with the frequency with which particles white-dot-exit Mazes. Then these two chances will be almost *completely independent of each other.* Almost, because in the limiting case where the frequency of left-exits is 0 or 1, particles that enter mazes won't ever leave. But for any other left-exit frequency f_1, and any desired white-dot-exit frequency f_2, we can always design the history of particle behavior so that both f_1 and f_2 obtain.[24]

The resulting conception of chance is, by contrast with the one we are exploring, strikingly egalitarian, non-hierarchical, and pluralistic. We'll have reason to return to it in the next section, where we explore the revisionist character of a Humean approach to chance.

[24]Proof: First, it is enough to handle the case where f_1 and f_2 are both rational (with $f_1 \in (0, 1)$ and $f_2 \in [0, 1]$). When a particle enters a Maze, it will execute a certain number of "left loops" (left at the central T-junction, then left again) and a certain number of "right loops" (right at the central T-junction, then right again), in some order, before finally exiting the maze by going left at the central T-junction followed by right at the blue-dot junction, or right at the central junction followed by left at the green-dot junction. Let A be the total number of left loops, B the total number of right loops. Let L be the total number of left exits, R the total number of right exits. Then L = 2A + 1; R = 2B + 1. We may assume that f_1 can be expressed as n/(n+m), where n is odd, and n and m share no common factors. (For if n is even, then we could just consider the frequency of *right*-exits instead.) There are now two cases. First case: m is odd as well. Then, since A and B can be freely chosen, we simply stipulate that, every time a particle enters a Maze, A = (n-1)/2 and B = (m-1)/2. So in each such case, L = n and R = m; it follows that the frequency of left-exits is n/(n+m). Second case: m is even. This time, we stipulate that the total number of Maze entries is *even*, and that in exactly half of these cases B = m/2, while in the other half B = (m-2)/2; in *all* cases, A = (n-1)/2. Then for any pair of Maze entries (one of each type), L = 2n and R = m+1+m-1 = 2m; so again, the frequency of left-exits is n/(n+m). Finally, observe that the frequency of white-dot-exits depends *only* on the order of the last two T-junction exits in the Mazes; so we can set it to any rational value we want, provided only that there are enough Maze entries.

§5 Humean revisionism about chance

The substantial literature on the Humean/anti-Humean divide has not, I think, fully appreciated how consequential that choice is, for our understanding of objective probability. I'll attempt, in this section, to draw attention to some of the most interesting revisionary consequences of a Humean approach. (I say "revisionary", because to the extent that we share a pre-theoretical, intuitive understanding of chance, that understanding is best captured by the anti-Humean conception we explored in the last section.) First we'll need to lay out the core features of that approach.

§5.1 The core "best-system" idea

Given that a Humean approach to physical modalities will see them as constituted, somehow, by patterns in the non-modal phenomena, a natural approach to chances, in particular, will see them as constituted by *statistical* patterns in the non-modal phenomena. No surprise, then, that early Humean approaches[25] saw chances as simple frequencies, probabilistic laws as nothing more than statistical (as opposed to universal) generalizations. By now, however, we have available to us much more sophisticated and nuanced Humean accounts – and this, thanks to Lewis's introduction of the "best system" account, and its further development over the last 20 years or so.[26]

To get the key idea into view, we should start with best system accounts of *non*-probabilistic laws. The aim of such an account is to provide a method for selecting, from among all the true statements concerning the non-modal facts, some elite set of statements that collectively deserve to be singled out as "laws". What distinguishes this elite set? Well, not that they most directly reflect the underlying modal constraints or physical necessities that govern our world – for if Humeanism is correct, there are no such constraints or necessities. Instead, it is some sort of *epistemic* virtue that the laws collectively possess; and for Lewis and many authors following him, it has seemed clear that the virtue in question was a kind of optimal combination of simplicity and informativeness.[27] Albert conveys the idea with characteristic verve (Albert 2015b, p. 23):

---

[25] As in Hempel 1965b; note that this was early enough that the label "Humean" had not yet taken hold.
[26] See for example Lewis 1994, Beebee 2000, Loewer 2004.
[27] Note that for "simplicity" to be a useful criterion, we must impose some restriction on the *language* in which these sentences are expressed; else we can say anything we want in a perfectly simple manner, just by gerrymandering a language with suitable predicates. We'll return to this point below REF. See Lewis 1983.

> Here's the idea. You get to have an audience with God. And God promises to tell you whatever you'd like to know. And you ask Him to tell you about the world. And He begins to recite the facts: such-and-such a property (the presence of a particle, say, or some particular value of some particular field) is instantiated at such-and-such a spatial location at such-and-such a time, and such-and-such *another* property is instantiated at such-and-such *another* spatial location at such-and-such another time, and so on. And it begins to look as if all this is likely to drag on for a while. And you explain to God that you're actually a bit pressed for time, that this is not all you have to do today, that you are not going to be in a position to hear out the whole story. And you ask if maybe there's something meaty and pithy and helpful and informative and short that He might be able to tell you about the world which (you understand) would not amount to everything, or nearly everything, but would nonetheless still somehow amount to a lot. Something that will serve you well, or reasonably well, or as well as possible, in making your way about in the world.
>
> And what it is to be a law, and *all* it is to be a law, on this picture of Hume's and Lewis's and Loewer's, is to be an element of the best possible response to precisely this request – to be a member (that is) of that set of true propositions about the world which, alone among all of the sets of true propositions about the world that can be put together, best combines simplicity and informativeness.

Let me just observe that while it is certainly natural for a Humean to privilege simplicity and informativeness in this way, doing so is neither obligatory nor without potentially serious costs. (On the costs, see Hall 2010 and Hicks 2014; on an important alternative – though one that remains firmly committed to the idea that what distinguishes laws is a certain epistemic virtue – see Hicks 2014.) For now, though, we'll stick with the Lewisian conception of 'bestness'.

How shall we extend the best system account so as to accommodate probabilistic laws? Return to our example of the T-junction world. If Humeanism is correct, then all there is to this world, fundamentally, is a totality of (non-modal) facts about particle motions, and in particular facts about the directions particles take out of the T-junctions they enter. (So: no facts about *chances*, that a candidate system might capture.) What must a candidate system do, to achieve an epistemically optimal summary of these facts? It can keep things simple but uninformative: "Whenever a particle enters a T-junction, it exits either to the left or to the right." It can pack in a lot of information – but, since there are no simple-to-state regularities that distinguish left-exits from right-exits, such information will cost a lot in simplicity. It can occupy a middle ground: "61.8% of exits are left-exits." But though simple, that's much less informative than it might seem, as it implies nothing about *correlations* (e.g., what percentage of left-exits are followed by right-exits?).

What to do? This: Let a candidate system consist of two parts. The first part is just a set of true sentences. The second is a measure – one with the mathematical structure of a *probability* measure[28] – over the set of worlds compatible with these sentences. If a candidate system comes out "best", then the worlds compatible with *its* sentences thereby qualify as the "nomological possibilities", and *its* measure counts as encoding the "ur-chances". Fine; but what exactly does "bestness" come to, for systems with this extra ingredient? Lewis thought that to answer this question, we needed to add a third desideratum: That system is best which optimally balance simplicity (of both its set of sentences and its measure), informativeness (of its set of sentences), *and* the extent to which its measure "fits" the actual world – namely, by assigning that world a high value. But it is probably a better idea to stick with the twin criteria of simplicity and informativeness, and treat a measure as "informative", to the extent that it could be successfully used as a basis for prediction. (For a defense of something close to this idea, see Loewer 2004 and Albert 2015b.) I'll simply assume that this can be done – and that ur-chances, so understood, will typically bear some close relation to frequencies. (Thus, we should expect the ur-chances for our T-junction world to assign exit probabilities equal to the exit frequencies.)[29]

Our sketch of a best-system approach to chance in hand, let's walk through the eight features of dynamical chance we canvassed in §4, in order to see how different these features appear, when viewed through a Humean lens.

§5.2 Law

Our Humean account connects chance directly to law – but in a way strikingly different from the way our anti-Humean account connected the two. In the Humean setting, chances do not at all reflect the way the fundamental dynamical laws dictate the world's course of evolution; rather, they arise as part of a two-pronged strategy for conveying a great deal of predictively useful information about the world, in a maximally efficient manner.

To bring out some consequences of this difference, let's start with a point of superficial similarity. On our anti-Humean conception, there is no such thing as a possible world with chances, but with no fundamental dynamical laws. So too on our Humean conception: a world with chances *has* them only because it has a "best system" that *assigns* them; and so it must have laws. But there is a certain looseness to

[28] I.e., it conforms to the Kolmogorov axioms.

[29] Note, though, that these will be *conditional* probabilities: e.g., the probability that a certain particle left-exits a T-junction just after time t, given that it entered the junction at t, is 0.618.

the connection between best systems, nomological possibilities, and chances that has no counterpart, on the anti-Humean conception. On that latter conception, the fundamental laws of a world w delimit a set of nomological possibilities (of which w is, of course, a member); furthermore:

• if world v is one of these nomological possibilities, then the fundamental laws of v are *the same* as those of w;

• if two nomologically possible worlds v and v* feature the same complete state at some time t, then they have exactly the same t-chances.

The Humean conception supports neither of these features.

As to the first, suppose the best system for w contains the set of sentences S. A world v counts as "nomologically possible" just in case all the sentences in S are *true* in v; but it does not follow – and can easily fail to be the case – that the *best system* for v likewise includes the set S. For example, the best system for a T-junction world will count as a nomological possibility a world in which *every* particle exits left; the best system for *that* world will, accordingly, be different (since it will include this very regularity in its set of sentences).

As to the second, suppose again that the best system for w contains the set of sentences S. Suppose v counts as nomologically possible, relative to w. Then even if the best system for v contains the very same set S, its ur-chance function may be *different*: for v may feature frequencies of fundamental events very different from those in w, and so may require a different ur-chance function in order best to capture information about them. For example, suppose w is a T-junction world in which 61.8% of exits are left-exits, and v one in which just 27.8% of exits are left-exits. The *sentences* that appears in the best systems for w and v may well be the same, but the *chances* they assign will be different. Notice, here, that it is not just that world v is *possible*, relative to w; in addition, world v may have, in w, a *non-zero chance* of obtaining. Thus we see the source of the famous "undermining futures" that have been thought (mistakenly: see Hall 2004) to spell serious trouble for Humeanism.[30]

[30] One caveat to all this. It is perfectly possible to artificially construct an account of laws and chances that fits the *letter* of Humeanism, and that builds in the two features we're focusing on here. The trick is, essentially, to "conditionalize out" those nomological possibilities that disagree with w, either with respect to the set of sentences, or with respect to the ur-chance function. Doing so removes all threat of "undermining", without compromising the core Humean supervenience thesis. But it's also highly unmotivated. See Arntzenius and Hall 2003 for details.

§5.3 Time

Unlike the anti-Humean chances explored in the last section, Humean chances have no essential connection to time. Humean chance is, fundamentally, a kind of unconditional probability over a set of nomological possibilities that is distinguished by its epistemic role. As such, these chances may well be had relative to a world (see the last subsection), but they are not had relative to a time. Yes, there are still things we can reasonably call "chances at times", but these are nothing more than one among many species of conditional chance.

§5.4 Explanation

Which sorts of statistical patterns are explicable, and which inexplicable, on a Humean account of chance? To start with, it's worth bearing in mind that the account of explanation that goes along with a worked out Humean conception of the physical modalities is likely to be a "pure unificationist" one, that sees explaining as purely a matter of fitting that-which-is-to-be-explained into a suitable kind of organizing framework. So we shouldn't expect to answer our question in the same way we would, if we were working within an anti-Humean framework. And indeed, the answer looks to be quite a bit different. Just how different I shall leave up in the air. But we can quickly see that the two obvious kinds of inexplicable pattern we saw in §4.4 have a markedly different status, in this Humean setting.

For our first example, we imagined a T-junction world in which the frequency of left-exits diverged dramatically from the underlying chance. For our second, we imagined a robust statistical pattern in a world whose combination of boring laws and funky spacetime structure gave rise to non-probabilistic but indeterministic evolution. But given our current setting, we are in each case imagining the impossible. In the first case, the best system for the given world must, in fact, include a different ur-chance function from the one we were (incorrectly) taking to characterize that world. In the second case, the robust statistical regularity in the pattern of self-collision makes it the case that the best system for the given world will, in fact, include an ur-chance function reflecting the de facto frequencies. So in each case, we actually have available just the sorts of chances we need in order to explain the regularities – though it absolutely *must* be remembered that "explain", here, means something rather different than it meant, in our anti-Humean setting. Again, this verdict concerning what is and isn't explicable shows how profoundly different the Humean approach is from the anti-Humean alternative.

§5.5 Precision

There is nothing about the Humean approach that suggests that chances must have precise values. It might seem otherwise: for won't a system that specifies a unique ur-chance function win out – namely, with respect to informativeness – over a system that introduces imprecision by merely specifying a range of (equally permissible) functions?

No, for two reasons. (Actually, there are three – but we'll have to wait until §5.9 to see the third. REF) First, informativeness of an ur-chance function is partly a matter of how accurate that function is, as a representation of the world in question.[31] Now, it's not immediately clear how we should understand or analyze the relevant notion of "accuracy". (It can't mean, in this setting, "correct in what it says about the objective chances.") But we need such a notion, and can easily recognize an intuitive basis for it. To see this, consider a T-junction world in which 60.8% of exits are left-exits, and in which there are no interesting correlations between exit events. Consider two candidate systems, that differ only in their ur-chance functions. The first system's function assigns a chance of left-exiting of .608; the second assigns a chance of 0.1. Both functions are otherwise quite simple – in particular, because each treats exit events as independent of one another. But clearly the first conveys more information – in some sense yet to be made precise! – about the way our world actually is, thanks to the match between its chances and the actual frequency of left-exits.

Now consider an odd variant on our T-junction world. This time, there is a certain *pattern* to the frequency of left-exits. For the particles in this world all possess values for a pair of distinct intrinsic magnitudes m_1 and m_2. What's more, there are many different values x for m_1 such that

• a very large number of particles with value x for m_1 enter T-junctions;

• the frequency of left-exits among these particles is very closely approximated by the function e^{-kx}, where k is some constant.

There are likewise many different values x for m_2 such that

• a very large number of particles with value x for m_2 enter T-junctions;

• the frequency of left-exits among these particles is very closely approximated by the function e^{-k^*x}, where k* is some constant *different* from k.

[31] So – in case this wasn't obvious – the informativeness of an ur-chance function urch is *not* to be measured by calculating the sum ∑(-urch(w)log[urch(w)]). It's not that there is anything wrong with this way of measuring the information contained in a probability function. But since it is a way which makes no reference to which world is *actual*, it cannot serve the needs of a Humean account of chance.

And this is all perfectly consistent, because those particles with a well-represented value for m_1 do *not* have a well-represented value for m_2, and vice versa.

What sort of ur-chance function should the winning system assign, in a case like this? It's hardly clear – but one defensible view is that it should assign a *pair* of ur-chance functions, one of which sets the chance of left-exit equal to e^{-km1}, the other of which sets it equal to e^{-k*m2}. By so doing, it in effect says that it is *indeterminate* what the chances are – it is determinate that they are given by one function or the other, but not determinate *which*. There are, of course, other options: A candidate system might ignore the evident functional dependency of left-exit frequency entirely; or it might privilege one of the two functions over the other; or it might explicitly say that if the particle's value for m_1 lies in a certain set, then the first function applies, and otherwise the second applies; or it might explicitly say that if the particle's value for m_2 lies in a certain set, then the second function applies, and otherwise the first applies; and so on. Each of these options carries a serious cost, either with respect to simplicity or with respect to informativeness – which is why I say that one defensible position is that the best system will leave the ur-chances indeterminate. For – arguably – it thereby does the best it can do with respect to informativeness, at very little cost in simplicity.

The second reason why we shouldn't automatically assume that a system that specifies a unique ur-chance function wins out, with respect to informativeness, over a system that introduces imprecision is that uniqueness may cost a lot in simplicity, even when we're not dealing with statistical patterns as strange as those recounted in the last paragraph. Consider a T-junction world in which infinitely many particles enter T-junctions, and there is a limiting relative frequency of left-exits equal to r. Alas, the number r is itself *very* difficult to specify, since it is not *computable*. That might be a powerful reason to count any candidate systems whose ur-chance function sets the probability of left-exit equal to r as highly non-simple. ("Might", because the exact standards on "simplicity" are still up in the air.) If so, an alternative system that sacrifices precision by specifying a small *range* of ur-chance functions (a range containing r, obviously) may win out.[32] So here we have a second reason for taking seriously the possibility of imprecise Humean chances.

[32] There is an added complication, which I will just note briefly: it may be that the demand for simplicity in a candidate system should not apply to that system's specification of fundamental constants. Consider, for example, a world whose laws are those of Newtonian gravitational mechanics. Should the Humean account of these laws build in the assumption that the gravitational constant G is (in, we assume, simple-to-specify units) itself a "simple" number? (I.e., at least *computable*?) Maybe so. But maybe *not*: for in *this* case, there is not really a deep problem for the Humean account, as a candidate system could simply include the statement that there is a fixed-over-time value for G, such that F =

This contrast between the perfectly conceivable imprecision of Humean chances and the seemingly non-negotiable precision of anti-Humean chances directly reflects their profoundly different ontological status. Anti-Humean chances are part of the fundamental furniture of the universe – and, for whatever reason, we operate with a strong prejudice that this furniture must have, as it were, perfectly sharp edges. Humean chances, by contrast, are part of an attempt to impose on the world a certain kind of epistemically valuable organizational framework. The standards of success for such frameworks may well place a great deal of value on precision. But not so much that it cannot be traded off against other values.

§5.6 Independence

Other things equal, a candidate system that includes an ur-chance function which incorporates the independence principles we discussed in §4.6 will win out over one that does not. And that is because the former sort of function is simpler to specify than the latter. But not by much. Appreciating this point will allow us to see how easy – and entirely unproblematic – it is for these principles to be violated, in a Humean setting.

Two examples will help make the point. First, consider a T-Junction world in which the frequency of left-exits is 0.608, but in which every left-exit either immediately precedes or immediately follows another left-exit. A winning system for this world will include this very sentence in its set of sentences, and thereby guarantee that its ur-chance function fails the temporal screening-off condition. Second, consider the version of our DERG world discussed above, in which exactly one of every pair of interacting particles decays green, while the other decays red. Here, a winning system will include an ur-chance function that violates the "locality" condition of §4.6. In each case, a significant increase in informativeness comes at a very slight cost to simplicity.

Now, we ordinarily attach a great deal of significance to both of these independence conditions. Just consider, here, the massive literature on the experimental violations of Bell's inequalities and their physical and philosophical significance; for that matter, consider how astonishing it would be for a physicist to offer, as a serious proposal for fundamental dynamical laws, ones which violated the

Gm_1m_2/r_{12}^2. The result is that the value for G is nomologically *contingent* – but that need not be a high price to pay, particularly given that the facts about particle motion in such a world will *settle* what the value of G in that world is.

temporal screening off condition. But if Humeanism is correct, the significance of both conditions is – to put it bluntly – being massively overblown.

It is crucial, here, to remember how utterly spartan the fundamental metaphysical structure of the world is, if Humeanism is correct. There is a space-time, qualitatively characterized by a pattern of instantiation of certain fundamental, non-modal properties and relations. And that's it: there is nothing "underlying" this mosaic, and it is "explicable" only in the sense that we may render it more comprehensible to ourselves by imposing on it some kind of organizing framework. We may discover that the best such framework incorporates chances that violate our independence conditions. So what?

Here, I think, is the most that can be said. Begin with a familiar if controversial claim, which I am simply going to assume without argument[33]: The rationality of our inductive inferences requires that we be rationality entitled to believe (or to assign high credence to) a wide range of contingent propositions about our world, even *antecedent* to having any relevant empirical evidence. We come equipped, in short, with a wealth of rational biases from which empirical inquiry must begin. But though rational and a priori, some of these biases are defeasible – as witness the fact that, historically, there are examples of ones that have been defeated[34]. If Humeanism is correct, then a proper and complete statement of the *content* of these biases can be given in purely non-modal terms: that is, there will be a set of what we might call "categorical constraints on rational credence", and these categorical constraints will collectively settle the nature of rational empirical inquiry.[35] (If anti-Humeanism is correct, then not only will there be additional, non-categorical constraints, but the content of the categorical constraints themselves may differ from their Humean counterparts; we'll return to this issue in §6.) Now, I think it's quite likely that our two independence conditions figure in some of these categorical constraints. That is just to say that it is a priori rational for

[33] Though see my 2004 for more details and defense.

[34] For a trivial example of a defeasible a priori bias, let p be the proposition that a coin with an objective chance of h of landing heads is tossed at some time t, and lands heads; if h is very close to 0, then we are (given any plausible credence-chance principle) rationally required to assign an a priori credence very close to 1 to not-p; and this, notwithstanding the fact that we can *learn* that p is true. For a non-trivial example, consider our strong a priori bias in favor of theories that posit a relation of absolute simultaneity.

[35] There is a rather delicate issue that I am running roughshod over, here. Suppose that the One True Epistemology has it that the choice between Humeanism and anti-Humeanism is itself to be settled at least in part empirically. Then Humeanism may be correct, but for all that there may be additional constraints on rational credence beyond the purely categorical ones. I'm going to ignore this issue, henceforth.

us to expect our world to conform to them. And so it should come as a surprise, of sorts, if we discover either of them to fail. But it must be emphasized that this is a "surprise" *only* in the sense that we discover something antecedently unexpected (which, let's face it, happens to us all the time). Given a Humean setting, there is *no* sense in which that which "surprises" us thereby raises some urgent *explanatory* demand.

One final point. We have, thanks to the experimental violations of Bell's inequalities, some reason to doubt the "locality" condition of §4.6. (Only *some* reason, since (i) it is not identical to Bell's own "locality" condition, and (ii) the failure of the former is just one way of accounting for the experimentally confirmed failure of the latter.) We don't, so far as I know, have any reason to doubt the temporal screening-off condition. What's more, anything that could count as empirical evidence of such a failure – say, robust correlations between earlier and later outcomes of some repeated experiment – could (and likely *would*) always be accommodated by positing enough extra structure to the instantaneous states of the world to allow those states to contain appropriate "records" of the past. Thus, instead of positing chances that violate the condition, we posit states enriched just enough to allow us to preserve it. I think this stratagem is perfectly sensible, *given an anti-Humean perspective*: for it incorporates the anti-Humean commitment to a relation of causal influence that *cannot* hold, unmediated, across a temporal gap. But it is genuinely unclear to me why, given *Humeanism*, the stratagem is appropriate. If an Oracle tells us that Humeanism is correct, then we may still reasonably expect that the simplest, most informative summary of our world's contents will build in the temporal screening-off condition – since, other things equal, it will thereby prove simpler than its rivals, and we may reasonably (though *defeasibly*) expect our world to be simple. But if the empirical evidence appears to show that other things *aren't* equal – that we can gain a great deal in informativeness at little cost in simplicity, by abandoning the condition – then why not do so?

§5.7 Asymmetry

Given the particular understanding we have been operating with as to what makes a candidate system "best", a Humean approach provides even less reason than our anti-Humean one for thinking of chances as in any way temporally asymmetric. The chances that are most directly served up, on the Humean approach we have been investigating, are *ur*-chances; and these have no temporal directionality built into them whatsoever. It is perfectly true that we can *define* from them temporally asymmetric chances (by identifying t-chances with ur-chances, conditional on initial stretches of history). But that is neither here nor there. The underlying conception of what chances are *for* – which is to serve as part of a predictively powerful summary of what the world as a whole is like, non-modally speaking – marks no essential difference between past and future.

But there are perfectly reasonable ways to tweak our understanding of "informativeness", that can introduce a kind of temporal asymmetry. It won't be the kind of asymmetry that we seem to have in mind when we are thinking in our ordinary (and, let's face it, pretty inchoate) causal terms – where it strikes us as part of the essential nature of causal influence that it happens in one temporal direction. It will have a rather different source.

The picture is this. All there is to the world is a vast totality of non-modal facts. The very rationale for our concepts of law and chance is that they allow for some kind of cognitively effective organization of information about that totality. We do *not*, at this point in our collective philosophical inquiry, have a clear, iron-clad grip on just what "cognitively effective organization" amounts to, here. One *proposal* is that it amounts to exactly what is enshrined in the standard "best system" analysis of law and chance: a simple and informationally rich *summary* of what is the case. But if we look again at the extended quote from Albert in §5.1, we'll see a slightly different proposal on display. In particular, right here (emphasis added): "…you ask if maybe there's something meaty and pithy and helpful and informative and short that He might be able to tell you about the world which (you understand) would not amount to everything, or nearly everything, but would nonetheless still somehow amount to a lot. Something that will serve you well, or reasonably well, or as well as possible, *in making your way about in the world*." We can read Albert as drawing attention to a distinctive feature of the human predicament, which is that we are *spatiotemporally embedded creatures*, for whom the difference between past and future is deeply practically significant. (Notice, by contrast, that the standard best-system account more or less explicitly adopts an *atemporal* perspective on the value of knowledge of the laws and chances.) So the kind of summary that will be most useful to *us* may need to be structured in such a way as to build in a temporal asymmetry. What we'll see in the next sub-section is a way of understanding Humean chances according to which this is precisely the case.

§5.8 The initial state

In the last sub-section, we saw two different conceptions of the kind of informativeness that matters, when determining which system is "best". On one conception, what matters is simply that we get as much information as possible about what the world as a whole – past, present, and future – is *like*. On the second conception, what matters – in addition, though not necessarily *instead* – is that we get as much information as possible *of the sort that we can use to make our way around in the world*. Now, we are creatures that interact with the world primarily at a certain *scale*. And we interact in a way that makes it particularly important for us to know how the *future* depends on what we do in the *present*.[36] Let see, now, how these facts make a difference.

Suppose an Oracle presented us with a candidate best system for our *own* world. Suppose, to begin, that what she gives us is a complete set of simple, exact dynamical laws, together with a simple specification of the sorts of instantaneous states our world can occupy. (Thus, she presents us with a canonical fundamental physics.) Suppose, finally, that these dynamical laws are *deterministic* (in a strong sense: given the specification of the complete state of the world at any time, the laws determine a complete *history*, past, present, and future). Then – even if she fails to include in the system she gives us any specification of an ur-chance function – it might seem that we are golden: she has given us a *wealth* of information about our world, in a compact, elegant form. And if *all* that mattered to us was, as it were, the satisfaction of our curiosity about the sort of world we inhabit, that would be correct.[37]

But, as Albert (2015b) expertly and persuasively argues, there is – from a practical perspective – too much that we *couldn't* do, equipped just with such a system. To cite a canonical example: we could not, by means of it, predict that when we place an ice cube in a glass of hot water, it will melt. Here, in brief, is why: The specific macro-

[36] Why is that, exactly? One option: our naive metaphysical prejudices are in fact correct, and the past is (unlike the future) metaphysically *settled*, so that it depends *not at all* on what we do in the present. Another option: the physical structure of our world is such that, given the kinds of creatures we are, the past does not depend on what we do in the present in any way that is *both knowable and exploitable*. For a development of this second option, see Albert 2000.

[37] Though there is a real worry here: if that is what we care about – if that is even an important *part* of what we care about – then she could do even better by us, if she could include in her system a simple specification of a single instantaneous state; for then she would have given us enough information to deduce *every* truth about our world. Alas, it would follow that the distinction between what is nomologically necessary and what is nomologically contingent *collapses* – and, I think, we *really* don't want it to collapse! See my 2010 for a development of this worry, and footnote 39 for a possible response to it.

information we have upon which to base a prediction (namely: that we've put an ice cube of such-and-such dimensions in a glass of water with such-and-such temperature and volume) is compatible with complete micro-states that, given the laws our Oracle has supplied us, entail that the ice cube *grows*. (Or, for that matter, jumps out of the water and turns into a flower, or ….) And there is an additional problem, which arises from the time-symmetric character of these dynamical laws: given this symmetry, how could they possibly underwrite our evident ability to predict very different sorts of things about the past than we do about the future? Encountering a partially melted ice cube in a glass of warm water, for example, we may confidently predict not only that the ice cube *will* melt, but that it *was* less melted. Whence this asymmetry?

In fact, Albert argues, this predicament is both quite general and quite severe. First, our seemingly powerful "best system" is almost *completely useless*, when it comes to predicting macro-behavior on the basis of macro-information: for our macro-information will almost always be compatible with too wide a range of "pathological" complete micro-states, whose collective predictive upshot is far too broad and heterogeneous. Second, its built-in time-symmetry guarantees that it cannot underwrite the very different sorts of macro-predictions we make about past and future. The upshot is that if what we care about is, in part, having the kind of information that will enable such macro-prediction, then the system our Oracle has bequeathed us is *not* best, after all.

In that case, we should return to her with a request: specifically, for a simple-to-describe ur-chance function that will (i) assign to "pathological" states a vanishingly small chance, so that – for purposes of macro-prediction – we may safely ignore them; (ii) capture the very different ways in which macro-past and macro-future are predictable, given information about the macro-present. And this request, Albert argues, she can readily grant. First, she augments the set of sentences she offered us with a sentence asserting that the universe is, on one temporal end, in a particular *low-entropy macrostate*.[38] Second, she supplies an ur-chance function that assigns a suitably "flat" probability distribution to the set of complete microstates *compatible* with the designated low-entropy macro-state. (We will almost certainly need to turn to the

[38] Note that there are two ways to understand this move. First way: the added sentence *specifically identifies* the low-entropy macrostate that our world began in. Second way: the added sentence simply *existentially quantifies* over such initial macrostates. There's probably not much that hangs on the choice. Strictly speaking, the first way makes it a nomological necessity that our world started out in the macrostate it did – which might seem odd. But given the highly metaphysically deflationary character of "nomological necessity", on a Humean approach, I think this is *merely* odd, and not particularly objectionable.

physics encoded in the rest of her 'best system' for guidance as to what "flat" should mean.) Since the dynamical laws are deterministic, that assignment generates, via those laws, an ur-chance function over the nomologically possible worlds. From these materials, we can show – or at any rate very plausibly argue – that the chance is overwhelmingly high that the world will display features that make possible the very kind of powerful and temporally-asymmetric macro-prediction we appear to be capable of.[39] (Along the way, we get an explanation of just why this low-entropy end deserves the label "past".)

There are two points worth emphasizing here. One is that in working out a Humean account of chance, it matters a great deal just what kind of information it is that a "best" system is supposed to optimally balance against simplicity. The second is that there is no conceptual difficulty whatsoever, on a Humean approach, with attaching chances – chances *not* had relative to times, mind you – to the initial state of the universe (or to any other state, for that matter). The contrast with the anti-Humean approach of §4 is striking. That approach gave pride of ontological place to *transition* chances, and thereby rendered unintelligible non-time-relative chances of initial states. The Humean approach treats transition chances as just one among many *conditional* chances – and the underlying unconditional chances can straightforwardly apply to initial states.

I want to close this subsection, however, by observing that this difference provides no reason whatsoever to consider a Humean approach more plausible than an anti-Humean approach (the rhetoric of some Humeans notwithstanding). It may *seem* otherwise: For doesn't a Humean approach allow us both to *predict* and *explain* macro-level regularities (e.g., the regularity that ice cubes placed in hot water melt, etc.)? And doesn't this predictive and explanatory power precisely hinge on its ability to make sense of chances over initial conditions? So, since an anti-Humean account *cannot* make sense of such chances, it must thereby be predictively and explanatorily handicapped.

[39] For example, the chance is overwhelmingly high that entropy will increase, in the direction away from the low-entropy state. Similarly, conditional on there being a partially melted ice cube in a glass of warm water at time t, the chance is very high that it is *less* melted in the temporal direction closer to the low-entropy end of the universe, but *more* melted in the other temporal direction. Note, in addition, that by demanding that the "best" system supply the sort of information that will enable macro-prediction, we may be able to answer the worry raised in footnote 37. For a system whose set of sentences is so informationally rich that it determines the complete history of the world will, for *that* reason, make a certain kind of macro-prediction impossible: namely, prediction about what *would* happen, given different possible courses of action we might undertake.

Well, no. Suppose we learn from our Oracle that anti-Humeanism is correct. And so the *only* laws that there are, are those contained in her original offering to us: the fundamental dynamical laws, plus the principles that delimit the range of nomologically possible instantaneous states. There is no "law" to the effect that the world started out in a low-entropy macrostate, and there are certainly no "chances" that its initial state was one thing or another. All the same, we are perfectly free to adopt, as a high-level empirical hypothesis, the claim that the world started out in a low-entropy macrostate. We are perfectly free to observe that there is an easily mathematically definable, suitably "flat" measure over the possible complete microstates compatible with this macrostate. We are perfectly free to use exactly the arguments Albert does to show that the overwhelming majority of these microstates (as measured by our "flat" measure) will evolve forward in such a way that entropy will increase, ice cubes placed in water will melt, etc. And so we are perfectly free to postulate that the complete initial microstate of our world just *is* one of these "typical" states.[40]

So much for prediction. As to explanation, there are two points to make. The first is just a reminder: if Humeanism is correct, then "explanation" *just* amounts to a certain kind of organization of information. There is simply no sense, then, in which the Humean explanation of our world's macro-regularities proceeds by showing how our world is "constrained" to exhibit them; instead, it proceeds by showing how these regularities fit into a certain kind of organizing scheme. So, as a purely dialectical matter, the anti-Humean should be thoroughly unimpressed by any claim that the Humeanism she rejects is explanatorily superior: the "explanation" that position has to offer just isn't the kind she cares about. Second – and a bit more speculatively – she *may*, after all, be able to offer an explanation-by-constraint of our world's macro-regularities.[41] For, since all but a tiny proportion of the complete initial microstates compatible with our world's initial macrostate guarantee, given the laws, that these regularities will obtain, there is a sense in which that macrostate itself, given the laws, *virtually* guarantees that they will obtain. Granted, though: this notion of less-than-perfect constraint needs closer investigation. But even without it, the first point stands: there is just no relevant sense in which the Humean approach is explanatorily superior to the anti-Humean.

[40] This paragraph essentially recapitulates the argument in Maudlin 2007.
[41] Again, the key idea here is inspired by Maudlin 2007.

§5.9 Imperialism

Imperialism *may* – pending an issue to be addressed shortly – be entirely optional, on the Humean approach. You can build it in easily enough, simply by insisting that a candidate system include just one ur-chance function. But you may not have to build it in – and in some circumstances, there may be powerful reason not to.

As an illustration, consider a T-junction world populated by a gazillion Mazes. Recall from §4.9 that the frequency of left exits from T-junctions is independent of the frequency of white dot exits from Mazes. Now, if imperialism holds, and the one ur-chance function treats exit chances as constant and independent, then it's easy to see that the chance that a particle entering a Maze exits at the white dot is exactly 0.5. So suppose that the *frequency* of white-dot-exits is quite a bit different from this – say, 0.278. Meanwhile, the frequency of left-exits is (say) exactly 0.5. And now let us consider how a candidate system might, in its choice of ur-chance function *or functions*, aim to be informative about one or both of these statistical patterns.

At this point, it is vitally important to pause over an issue that we've mostly kept off stage, up to now. A candidate system specifies both a set of sentences that delimit the nomological possibilities, relative to the given world w, and an ur-chance function over these possibilities. These sentences belong to some *language* – not English or any other natural language, surely, but something much more precisely regimented. And it is by now a commonplace in the literature on Humean accounts of law that there must be some *restrictions* on this language, else the set can oh-so-simply pin down w exactly, by including just the sentence "everything is F", where the (primitive) predicate "F" is understood to express a property had by a thing x in a world v exactly if v = w. In the same way, the candidate system must also draw on some language in order to specify its ur-chance function; and for the same reason, there must be some restrictions on *this* language as well.

Now, it may be that these restrictions are so severe that, in the case at hand, only *one* language will meet them: namely, a language whose primitive vocabulary can express such properties as "being a T-junction" (along with whatever other properties and relations characterize the constituents of this world at the most micro-physical level), but *not* such properties as "being a Maze".[42] If so, then imperialism (as applied to this world, anyway) will likely follow. For either the properties of being a Maze,

---

[42] Lewis 1983 suggested a quite severe restriction: there is really just one appropriate language, namely that language whose primitive vocabulary expresses the most perfectly natural properties and relations instantiated in the given world. For a much more permissive approach, see Callender & Cohen 2009.

exiting at the white dot, etc. will be definable (in the one acceptable language) in a relatively *simple* manner, or not.[43] If the former, then a system may score better on informativeness by specifying an ur-chance function that says (in translation!) that the chance that a particle entering a Maze exits at the white dot is 0.278. (Though note that this will still entail a good deal of added complexity: for example, this ur-chance function cannot treat particles' behavior at T-junctions as probabilistically independent.) If the latter, saying this would cost too much in simplicity, and so the winning system must rest content with an ur-chance function that sets the chance of left-exiting at 0.5, treats these chances as independent, and therefore *also* implies (albeit implicitly) that the chance that a particle entering a Maze exits at the white dot is exactly 0.5.

But suppose the restrictions aren't so severe. Suppose, in fact, that it is perfectly permissible for the content of an ur-chance function to be specified directly in terms of the chances it assigns to Maze behavior – and this, even though (say) expressing "is a Maze" in the language of fundamental physics is no simpler than expressing "is a human". That is, a candidate system may make use of (at least) two distinct languages, one of which has among its primitive vocabulary such predicates as "is a T-junction" and "exits left", and the other of which has among its primitive vocabulary such predicates as "is a Maze" and "exits at the white dot". And it may thereby earn extra points for informativeness, by including one ur-chance function that conveys a great deal of information about the statistics of T-junction exits, and *another* that conveys a great deal of information about the statistics of Maze exits.

Of course it will be tempting to ask, "okay, but which ones are the *real* chances?". That temptation just reflects the deep intuitive pull of imperialism. But for a Humean, the right answer is that, in such a situation, there just is no such thing as "the" chances. Nor is what we have here just another instance of the failure of precision (cf. §5.5). There is nothing indeterminate about the chances, in the world we are imagining. No; there are, perfectly determinately, two chances, for any given event.[44]

Worlds like the one we are now considering – worlds whose statistics support, given Humeanism, anti-imperialist chances – are much more than a mere curiosity. For they reveal an added and fairly serious complication that will confront any attempt

---

[43] Note that we have not said enough about what, precisely, distinguishes Mazes to settle which option obtains. All we've said is that every Maze contains within it a certain network of linked T-junctions and channels; but for all that, there may be much more, micro-physically speaking, to being a Maze.

[44] With the caveat that there will likely be events to which the second ur-chance function assigns no value.

to explain the connection between Humean chances and rational credence. The next section explores this connection.

§6 The status of credence-chance principles

Lewis's groundbreaking "Subjectivist's Guide" (1980) made a persuasive case that a rational agent's credence in a given proposition should bear a certain connection to her credence concerning that proposition's chances of coming true. Unfortunately, Lewis added an unnecessary complication, insisting that a proper statement of this connection needed to make use of a notion of "admissibility": Lewis thought, roughly, that one's credence in a proposition P ought to reflect one's credence concerning P's chances, *provided* that one possessed no evidence that was relevant to P not merely by being relevant to its chances. But if chance merits this kind of epistemic respect because it is, in effect, a certain kind of expert, then the right thing to do is to simply incorporate any "inadmissible" information into the statement of our credence-chance principle. Easily done. Within our ur-chance framework, an exact expression of Lewis's key idea is simple. Let Cr be some rational "initial" credence: a function that represents one possible way that perfectly rational opinion could be, prior to the incorporation of any evidence. Let A and E be any two propositions. Let Pr be some probability function, and let X be the proposition true at a world w iff w's ur-chance function is Pr. Then

(*) $Cr(A \mid E \,\&\, X) = Pr(A \mid E \,\&\, X)$

(See Hall 2004 for more discussion and defense of this principle.)

Shortly, we'll see one reason why, if Humeanism is correct, this principle won't do, after all (a reason you may have already spotted, given the foregoing discussion of imperialism). We'll also see that an adequate replacement may need to reintroduce the notion of admissibility. But before getting to all that, let's pause to unpack the principle a bit.

It may seem far from obvious how to apply this principle in practice. Happily, given certain background assumptions, it will imply something more user-friendly:

(**) $Cr(A \mid E \,\&\, ch_t(A) = x) = x$, where "$ch_t(A) = x$" says that the time t chance of A is x, and E is entirely about history prior to t.[45]

What should we make of principle (*)? Well, it certainly seems that there must be *some* intimate epistemic connection between credence and chance, and (*) unquestionably serves as an elegant and plausible option for what this connection might be. But Lewis claimed much more for it (or rather, for the version he preferred, which incorporated his notion of "admissibility"). To begin, he claimed that this principle "captures all we know about chance" (1980, p. 86) – very likely a mistake, and at any rate not the sort of thing one ought to just baldly assert, without argument.[46] Just as confused, I think, was his insistence that anti-Humean accounts face an insuperable obstacle in their inability to justify (*). Consider for example this passage (which is, unfortunately, rather shameful in its substitution of rhetoric for genuine argument):

> Be my guest – posit all the primitive unHumean whatnots you like. (I only ask that your alleged truths should supervene on being.) But play fair in naming your whatnots. Don't call any alleged feature of reality "chance" unless you've already

---

[45] To begin the derivation, we'll assume that the t-chance of some proposition P is given by $urch(P|S_t)$, where S_t is the proposition correctly and completely describing the state of the world at t. Now consider two partitions on the space of possible worlds: $\{X_i\}$ and $\{S_j\}$, where each S_j completely describes a possible time t state of the world, and X_i is true at a world iff that world's ur-chance function is Pr_i. (Purely for convenience, we'll assume both partitions are countable.) So
$Cr(A \mid E \,\&\, ch_t(A) = x) = \sum_{ij} Cr(A \mid E \,\&\, ch_t(A) = x \,\&\, X_i \,\&\, S_j) \bullet Cr(X_i \,\&\, S_j \mid E \,\&\, ch_t(A) = x)$,
where the sum is taken over all combinations (X_i & S_j) such that
$Cr(E \,\&\, ch_t(A) = x \,\&\, X_i \,\&\, S_j) \neq 0$.
But any such combination must in fact entail that $ch_t(A) = x$, so
$Cr(A \mid E \,\&\, ch_t(A) = x) = \sum_{ij} Cr(A \mid E \,\&\, X_i \,\&\, S_j) \bullet Cr(X_i \,\&\, S_j \mid E \,\&\, ch_t(A) = x)$
$= \sum_{ij} Pr_i(A \mid E \,\&\, X_i \,\&\, S_j) \bullet Cr(X_i \,\&\, S_j \mid E \,\&\, ch_t(A) = x)$.
If we add the background assumptions that ur-chances always treat themselves as certain (i.e., there is no "undermining"; see §5.2), then
$Pr_i(X_i) = 1$,
hence
$Pr_i(A \mid E \,\&\, X_i \,\&\, S_j) = Pr_i(A \mid E \,\&\, S_j)$.
And if we add the temporal screening off condition from §4.3, then, since E is entirely about history prior to t,
$Pr_i(A \mid E \,\&\, S_j) = Pr_i(A \mid S_j) = x$.
Therefore,
$Cr(A \mid E \,\&\, ch_t(A) = x) = \sum_{ij} x * Cr(X_i \,\&\, S_j \mid E \,\&\, ch_t(A) = x) = x$.

[46] As to why it is likely a mistake – at least, by anti-Humean lights – here is one reason: We know, I think, that if anti-Humeanism is correct, then objective chances are perfectly precise; but there is no way to derive this conclusion from a credence-chance principle. For other reasons, see Arntzenius and Hall 2003.

> shown that you have something, knowledge of which could constrain rational credence. I think I see, dimly but well enough, how knowledge of frequencies and symmetries and best systems could constrain rational credence. I don't begin to see, for instance, how knowledge that two universals stand in a certain special relation *N** could constrain rational credence about the future coinstantiation of those universals. (1994, p. 484)

But there is, I think, no good case that anti-Humeanism is in trouble here. And this, for a very simple reason: unless we wish to embrace an utterly insane kind of inductive skepticism, we need to recognize that rational opinion must be constrained by substantial a priori commitments concerning what the world is like, in contingent respects (cf. our discussion in §5.6). If anti-Humeanism is correct, the content of some of these basic epistemic constraints will be given directly in terms of the metaphysical structure that distinguishes anti-Humeanism from Humeanism. And we have, as yet, no good reason for denying that principle (*) *just is* one of these basic constraints. Now advances in epistemology might uncover such reasons – reasons, that is, for thinking that (*) or something like it can only be acceptable as a *derivative* principle of epistemology. But unless and until they do, it is really just pointlessly obstructionist to pretend that the very tenability of anti-Humeanism hinges on whether it can provide some independent justification of the credence-chance connection.[47]

Matters will be different, should Humeanism turn out to be correct. For in that case, the only a priori constraints that rational opinion must respect will concern the purely categorical, non-modal structure of the world. And so we may proceed as follows: describe a hypothetical situation; specify the categorical information about that situation available to some imagined perfectly rational agent; ask what, on the basis of that information and the purely categorical constraints on rational credence, she ought to believe; check to see whether this result conforms to our credence-chance principle. When we do so, I think we discover two reasons for doubting that, in a Humean setting, (*) is the correct such principle.

To bring the first reason into view, consider a T-junction world in which each particle possesses a value for some intrinsic physical parameter, which we'll call "mass".[48] Let us suppose, as usual, that there are a gazillion instances of particles entering T-junctions. What's more, the exit statistics display, to a very close approximation, a very simple dependency on mass; as a consequence, the best system

[47] Nor should we hastily assume – as Lewis seemed to – that it cannot provide such a justification. See the discussion of what I called "primitivist hypothetical frequentism" in my 2004, for an attempted proof of concept.

[48] What follows recapitulates an argument I gave in my 2004.

assigns a chance of exiting left equal to e^{-km} (for some constant k). Suppose our rational agent knows all this; that is, she knows that the exit statistics are such as to make an ur-chance function yielding this assignment "best". Next, she knows that a particle with mass 1 is about to enter a T-junction; so the best system assigns it a chance (say) of 0.1 of exiting left. And finally, she knows that this situation is a bit of a statistical outlier, in that in no other case does a particle with mass 1 ever enter a T-junction. What ought her credence be that this particle will exit left?

One answer is given by a principle of indifference: She knows that there are just two possible outcomes, so she should assign credence 0.5 to each. Another answer is given by (*): she should assign credence 0.1. And of course there might be other answers. The question before us is why, given Humeanism, the answer provided by (*) should trump all others, and in particular the answer given by an indifference principle.

To appreciate the question's difficulty, we might imagine our agent's information coming in a certain order. First she learns that whenever a particle enters a T-junction, it exits either left or right. Then she learns that in all intrinsic respects, this particular instance of a particle entering a T-junction is the only one of its kind. So far, it is at least rationally permissible – and quite possibly rationally obligatory – for her to assign a credence of 0.5 to a left exit. Next, she acquires a very general kind of information about other T-junction interactions: she learns that for very many values m of mass, very many particles with mass m enter T-junctions, and very close to e^{-km} of them exit left. What remains unclear[49] is why, exactly, this new information should have any bearing on her credence concerning the instance she is considering – indeed, a bearing specific enough that she should change her credence in a left exit from 0.5 to 0.1. I will leave this issue unresolved, save to note that it can't be handled this way: The extra statistical evidence available to our agent provides her with excellent reason for judging that T-junction interactions are *governed by a probabilistic law*; that law assigns a chance of 0.1 to a left exit on this occasion; so she should conform her credence in a left exit to this chance. That won't do, in a Humean setting, since no "governing" is going on. There is no conceptual distance, in this setting, between a probabilistic law and the high level summary of statistical behavior we have already stipulated our agent to possess. So what we need to understand is how this high level summary can have the very specific evidential force required by (*) *without* mediation through an inference concerning "underlying" chances. That, I submit, remains a mystery.

[49] *Genuinely* unclear; this isn't that annoying, cheap rhetorical use of the term.

There is a second, closely connected mystery, one which may be much harder to solve – at least, if we insist on trying to make Humeanism compatible with principle (*). The credit for spotting this mystery goes to Ittay Nissan-Rozen, who in a very important recent paper (REF) argues that a credence-chance principle suitable for Humeanism must make use of a notion of admissibility, and – more surprisingly still – there will be cases where purely historical information is not admissible. While I will set up the difficulty somewhat differently, I want to gratefully acknowledge Nissan-Rozen's quite powerful insight (and encourage you, the reader, to consult his paper for a careful, detailed development of it and exploration of its consequences).

Imagine a rational agent, convinced of the truth of Humeanism, living in a T-junction world that is also replete with Mazes. One hypothesis open to her to consider is that the best system for this world, *written in the language of fundamental physics*, asserts that each particle has a chance of exactly 0.5 of exiting left from any T-junction it enters, and that these events are all independent of one another. More generally, she might entertain the hypothesis that the fundamental laws of her world are L, where these laws entail the foregoing claim about the chance behavior of particles in T-junctions.

Consider some time t. Let the proposition B state that, at t, a certain particle will enter a certain Maze. Let the proposition A state that this particle will eventually exit via that Maze's white dot. Let E be our rational agent's total evidence, at time t. Let Cr be a reasonable initial credence function such that her time-t credence is just Cr(- |E). And now let us ask: what is Cr(A | E & B & L)?

That should be an easy question to answer. L and B together entail that the *chance* that the particle exits at the white dot is ½.[50] The example involves no threat of 'undermining'; so – even for a Humean – this conditional credence should likewise be ½.

But matters are not so simple. Suppose that our agent's evidence includes the results of a massive amount of investigation of the behavior of particles in Mazes. That evidence has revealed an extraordinarily stable statistical pattern: about 27.8% of particles exit from the white dot, with no discernible correlations between different 'trials'. Nothing about the conjunction (E&B) distinguishes *this* trial as in any way different from the ones about which our agent has so much evidence. So, since our

[50] And this, just from the fact that, according to L, the behavior of any particle at any T-junction is probabilistically independent of the behavior of any other particles at any other T-junctions.

agent is rational – in particular, perfectly able to *learn from experience* – it must be that Cr(A | E & B) ≈ 0.278. Why should Cr(A | E & B & L) be any different?

The anti-Humean has an obvious, straightforward answer. Given L, the past behavior of particles in Mazes can only be a massive fluke: something *allowed* by the laws, but judged by them to be extraordinarily objectively improbable. Fluke that it is, this behavior ought not – *given* the truth of L – to constrain rational credence about the behavior of the particle described by B. In short, adding L to what is being conditionalized on *undermines* the evidential significance that E would otherwise have for A.

But our imagined rational agent – convinced, remember, of the truth of Humeanism – cannot avail herself of this answer. One way to see why not draws on our earlier discussion of imperialism. Suppose we grant that a candidate system may draw on distinct vocabularies, and thus may specify more than one ur-chance function. In the present case, one of these functions is picked out in the language of T-junctions; *that* function assigns a chance of 0.5 to the proposition A. But for all our agent knows, the best system for her world *also* specifies – this time, in the language of Mazes – a function that assigns a chance of 0.278 to A. Supposing it does, then while it is true that *one* of the chance-functions that captures statistical patterns in her world treats her evidence as "flukish", there is *another* such function that doesn't. So without additional premises, we can't just explain away the difference between Cr(A | E & B) and Cr(A | E & B & L) by observing that, conditional on L, her evidence E becomes irrelevant-because-flukish.

In fact, the underlying issue remains, even if we build imperialism into our Humean account of chance. Suppose, for example, that we decide on reflection that it's an analytic truth about "objective chance" that there is at most *one* ur-chance function per world; we might then insist that a best system may specify only one ur-chance function, and must do so in a language suitable to capture the fundamental micro-physics of the given world. In the present case, then, this version of Humeanism will only recognize an objective chance of 0.5 for A. But even if this is the version of Humeanism our imagined rational agent endorses, that makes absolutely no difference to the underlying epistemological issue. Consider that in order to be able to learn from experience at all, her credence function must have various biases built into it – roughly, she must treat it as antecedently highly likely that she lives in a suitably "uniform" world. What we confront, in the present case, is a sharp question as to what "uniform" means – and whether it can mean what it *needs* to, in order to secure the result that Cr(A | E & B & L) = 0.5.

To bring this question into focus, let's begin with a very plausible assumption about the nature of some of the rational biases built into our agent's credence function. Let L^- be the *non-probabilistic* part of L; that is, L^- says that it is a fundamental law that there are particles moving in such-and-such ways, where the "such-and-such" includes the information that every particle entering a T-junction exits either to the left or to the right, but L^- specifies no ur-chance function. Then we may assume that our agent considers it overwhelmingly likely, *given* L^-, that any two sufficiently large, reasonably naturally shaped regions of spacetime feature very similar statistics in the behavior of particles in T-junctions.[51] Now, what L *adds* to L^- is, in effect, the information that the simplest probabilistic summary of the global exit statistics assigns a uniform chance of 0.5 to each left-exit, and treats these events as independent. So our agent should, plausibly, consider it overwhelmingly likely, *given* L, that in any sufficiently large, reasonably naturally shaped region of spacetime, the fraction of left-exits will be close to 0.5, and, more to the point, that there will be no correlations between exits *that are simple to specify in the language of T-junctions* (i.e., in the language suitable for directly describing the world at the most fundamental microphysical level).

But there will – of *course* – be correlations galore, that are *complicated* to specify in this language. And we've already stipulated that there is no simple way to describe Mazes in this language. So it is perfectly open to add the further hypothesis M that there are just those incredibly-complicated-to-specify (in the "fundamental" language) correlations needed to make it the case that (i) the fraction of white-dot-exits from Mazes (over all of space and time) is 0.278; (ii) there are no simple-to-specify correlations in Maze behavior (simple to specify, that is, in the language of Mazes). Perfectly open, that is, in the following epistemic sense: our agent's rational bias toward believing that nature is uniform at the level of T-junction behavior is not in the least threatened, if she comes to believe (or give a very high credence to) – specifically, on the basis of her considerable store of evidence E – the hypothesis M. So what sort of irrationality could she be guilty of, if $Cr(A \mid E \& B \& L) = 0.278$?

That's not a rhetorical question. Indeed, the first point to emphasize is that the rational biases just described *can't* be the whole story about what grounds the rationality of our agent's credences; just observe that she can conform to these biases while, for example, considering it certain that the first T-junction exit that happens every Monday is a left-exit. So we already knew that there must be further constraints on her credences, in virtue of which they count as rational. Perhaps it is one or more of these

[51]The patches need to be large enough to include sufficiently many instances of particles entering T-junctions. And they should be reasonably naturally shaped, since our agent knows for certain that, unless all or almost all particles exit in just one direction, there will be many scattered, oddly shaped patches that each feature many exits but that differ dramatically in their statistics.

further constraints that guarantees that, on pain of irrationality, Cr(A | E & B & L) must equal 0.5.

But there is a real puzzle about what these further constraints could be – what, in detail, a Humean epistemology could look like that will deliver the result that Cr(A | E & B & L) = 0.5. I'll consider and reject two options, and then leave it open where to look for more successful alternatives. The first option is to combine imperialism with a bald insistence that principle (*) just is one of the non-derivative, non-defeasible categorical constraints on rational credence. (You *need* to combine (*) with imperialism, else the principle becomes either inconsistent or too ambiguous to deliver the wanted result.) But we've already seen that imperialism is in some tension with Humeanism. And its combination with principle (*) raises an additional challenge: for we'll need to find some principled basis for settling which candidate ur-chance function is the "real" one, while simultaneously selecting the ur-chance function best suited to constrain credence in the way (*) requires. Finally, to simply invoke (*) as one of the fundamental categorical constraints on rational credence built into a Humean epistemology is to ignore the *explanatory* demand raised by the scenario we've been considering. It's clearly rational for our agent to have Cr(A | E & B) ≈ 0.278. (Stronger: it's *irrational* for to have a substantially *different* credence.) Why is it, exactly, that conditionalizing on L requires her to change this credence to 0.5? A good answer should make it clear why a certain claim about statistical patterns visible at the level of T-junction behavior should render *irrelevant* her wealth of information about statistical patters visible at the level of Maze behavior. Baldly insisting on (*) as a non-derivative constraint on her credence simply punts on that demand.

A second, more interesting option invokes a kind of symmetry constraint on her credence (one similar to exchangeability). Here is a natural proposal about the form of such a constraint: Suppose worlds w1 and w2 both have L^- as the non-probabilistic part of their best systems. And suppose they agree on their numbers of left-exits and right-exits from T-junctions. Then our agent's initial credence must be such that Cr(w1) = Cr(w2). Given such a principle, we can plausibly argue not just that Cr(A | E & B & L) = 0.5 but, more strongly, that Cr(A | E & B & L^-) = 0.5.[52] Unfortunately, this turns out to be too much of a good thing. For the derivation in fact shows that for *any* L* that extends L^- by specifying some ur-chance function, correlations between

[52] Divide the worlds in which (E & B & L^-) holds into those in which A is true, and those in which A is false. Then we need a modest assumption: there is a one-one mapping between these two sets of worlds such that, when w1 is a world in the first set, its image under the mapping w2 features exactly the same numbers of left- and right-exits. (The idea is that we get from w1 to w2 by swapping the last two T-junction exits inside the given Maze, and holding all other T-junction exits – whether inside or outside this Maze – fixed.) Then given our symmetry constraint on the agent's credence, it will follow that Cr(A & E & B & L^-) = Cr(not-A & E & B & L^-), hence Cr(A | E & B & L^-) = 0.5.

T-junction behavior that L* "sees" will make no difference to the agent's credence. Consider, for example, a world in which particles oscillate statistically in their T-junction behavior: it's true about 90% of the time that after a particle left-exits, its very next exit is a right-exit; likewise, it's true about 90% of the time that after a particle right-exits, its very next exit is a left-exit. The best system for such a world will pick up on this pattern, and build it into the specification of the ur-chance function, by assigning to each particle a chance of 0.9 of "flipping" from one exit to the next. Let L* be this best system. Let X be the proposition that a certain particle has just executed a left-exit. Let *left-right* be the proposition that its next two exits, in order, will be left, then right; left *right-left* be the proposition that its next two exits, in order, will be right, then left. Principle (*) yields the result that Cr(*left-right* | X & L*) = (0.1)•(0.9), while Cr(*right-left* | X & L*) = (0.9)•(0.9). But our symmetry principle yields the result that these two conditional credences must be equal. So far from providing a basis for deriving (*), our symmetry principle in fact conflicts with it.

At this point, it may be that the best hope for a plausible Humean epistemology of chance lies not in trying to vindicate principle (*), but in modifying it – specifically so as to allow that Cr(A | E & B & L) = Cr(A | E & B) ≈ 0.278. That is Nissan-Rozen's important suggestion: we should return to Lewis's original idea that an agent's beliefs about chances ought to guide her credences, *provided* she possesses (or is conditionalizing upon) no inadmissible information; and we should recognize that in the present case, the agent's knowledge of past Maze behavior counts as inadmissible information.

Now, principle (*) was already written in a way that allowed it to accommodate *ordinary* sorts of inadmissible information. Suppose, for example, that A is the proposition that a certain particle that entered a T-junction at time t exited left from it. Suppose it's a bit later, and our agent's total evidence E includes the proposition that 58% of this particle's exits from time t on have been left-exits (and includes nothing else relevant). Suppose she knows that the chance of each left-exit is exactly 0.5. Even so, her credence in A should surely be 0.58, and not 0.5. So it can seem, naively, that her beliefs about chances don't constrain her credence in this case, thanks to her possession of inadmissible information.

Not so. Look again at the statement of (*):

Let A and E be any two propositions. Let Pr be some probability function, and let X be the proposition true at a world w iff w's ur-chance function is Pr. Then

(*) Cr(A | E & X) = Pr(A | E & X)

Applied to this case, we simply get the result that her credence in A should be set by (what she knows to be) the *conditional* chance of A, given E. And this conditional chance will equal 0.58. The key idea is quite simple: the relevance of the agent's evidence is, at it were, simply reflected in the structure of the conditional objective chances.

What Nissan-Rozen points out is that matters are fundamentally different, in the case we have been considering. Crucially, if Pr is the ur-chance function picked out at the level of T-junction behavior – that is, the function that assigns to each left-exit a constant and independent chance of 0.5 – then, since A concerns an outcome that is wholly distinct from the stretch of history that the evidence E is about, Pr(A | E & X) = Pr(A | X) = Pr(A) = 0.5. The *agent* sees (quite properly) that her evidence bears on A; but the *chances* do not.

That means that (*) itself must be qualified by an admissibility requirement (bracketing the option of outright rejecting it): it describes a constraint that our agent's credences must respect, *provided* that the proposition E contains no information that is inadmissible with respect to the given proposition A and chances Pr. The intuitive idea, as we've just seen, is that E is inadmissible just in case Pr fails to recognize its evidential bearing on A. Obviously, the crucial piece of unfinished business here is to replace this intuitive idea with something more precise, and more illuminating.

§7 Open questions

I'm going to close with some open question that strike me as especially worth further investigation.

1. We've seen above that a certain Humean reductionist approach to chances is, along a number of different dimensions, highly revisionary. But recall that the approach we were considering treated the "bestness" of a system as a function, in part, of how *informative* it is. That's not the only way to go. Following Michael Hicks's (2014) lead, for example, we might privilege systems that contain truths about the world that are (i) within our epistemic reach, in the sense that the kinds of evidence we can acquire can, at least if we're careful enough, direct us towards these truths; and (ii) powerful enough that we can draw on them as a basis for making accurate predictions (or retrodictions, etc.) about aspects of our world that would otherwise be inaccessible to us. Granted, being informative is one way for a system to be epistemically useful; but so is being Hicksian – and the latter sort of utility might provide a better rationale for our concepts of law and chance than the former. Suppose so; or suppose that

some *other* development of the Humean approach, not wedded to Lewisian informativeness, is correct. How revisionary will *that* account of chances end up being?

2. What's the best way to understand the "informativeness" of an ur-chance function?

3. Why is the assignment of a *probability* measure the most effective way to convey information about statistical patterns?

4. Given the difficulties canvassed in §6, what is the right form for a credence-chance principle, for a Humean?

5. Can there be chances completely divorced from laws?

6. What other forms of anti-Humean chances can there be – in particular, forms which abandon the laws-as-describing-generation paradigm?

§8 References

Albert, David 2000. *Time and Chance*, Cambridge MA, Harvard University Press.

Albert, David 2015a. *After Physics*, Cambridge MA: Harvard University Press.

Albert, David 2015b. "Physics and Chance", in Albert 2015a: 1 - 30.

Armstrong, D.M. 1983. *What Is a Law of Nature?*, New York: Cambridge University Press.

Arntzenius, Frank and Ned Hall 2003. "On What We Know About Chance", *British Journal for the Philosophy of Science* 54: 171-179.

Beebee, Helen 2000. "The Non-Governing Conception of Laws of Nature", *Philosophy and Phenomenological Research* 61: 571-593.

Callender, Craig and Jonathan Cohen 2009. "A Better Best System Account of Lawhood", *Philosophical Studies* 145: 1-34.

Carnap, Rudolf 1945. "The Two Concepts of Probability", *Philosophy and Phenomenological Research* 5: 513-32.

Ellis, Brian 2002. *The Philosophy of Nature: A Guide to the New Essentialism*, Oxford: Routledge.

Ghirardi, Giancarlo 2016. "Collapse Theories", *The Stanford Encyclopedia of Philosophy (Spring 2016 Edition)*, Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/spr2016/entries/qm-collapse/>.

Hale, Bob and A. Hoffmann (eds.) 2010: *Modality: Metaphysics, Logic, and Epistemology*, Oxford: Oxford Univ. Press.

Hall, Ned 2004. "Two Mistakes about Credence and Chance", *Australasian Journal of Philosophy* 82: 93-111.

Hall, Ned. 2010. "Humean reductionism about laws of nature." Unpublished manuscript.

Hempel, C. 1965a. *Aspects of Scientific Explanation and Other Essays in the Philosophy of Science*, New York: Free Press.

Hempel, C. 1965b. "Aspects of Scientific Explanation", in Hempel 1965a, 331–496.

Hicks, Michael 2014. "The Epistemic Role Account of Lawhood", unpublished manuscript.

Joyce, James 2010. "A Defence of Imprecise Credences in Inference and Decision Making", *Philosophical Perspectives* 24: 281-323.

Kitcher, Philip, 1989. "Explanatory Unification and the Causal Structure of the World", in P. Kitcher and W. Salmon (eds.), *Scientific Explanation*, Minneapolis: University of Minnesota Press: 410-505.

Lewis, David 1980. "A Subjectivist's Guide to Objective Chance", in Lewis 1986a, pp. 83-113.

Lewis, David 1983. "New Work for a Theory of Universals", *Australasian Journal of Philosophy* 61: 343-377.

Lewis, David 1986a. *Philosophical Papers, Volume II*, Oxford: Oxford University Press.

Lewis, David 1994. "Humean Supervenience Debugged", *Mind* 103: 473-90.

Loewer, Barry 2004. "David Lewis' Humean Theory of Objective Chance", *Philosophy of Science* 71: 1115-25.

Maudlin, Tim 2007. "What could be objective about probabilities?", *Studies in History and Philosophy of Modern Physics* 38: 275 - 291.

Maudlin, Tim 2011. *Quantum Non-Locality and Relativity, third edition*, Oxford: Blackwell.

Nissan-Rozen, Ittay 2017. "On the Inadmissibility of Some Historical Information", *Philosophy and Phenomenological Research.*

Rosen, Gideon 2010. "Metaphysical Dependence: Grounding and Reduction", in B. Hale & A. Hoffmann (eds.), *Modality: Metaphysics, Logic, and Epistemology*, Oxford: Oxford Univ. Press 109-36.

Strevens, Michael 2009. *Depth: An Account of Scientific Explanation*, Cambridge: Harvard Univ. Press.

Woodward, J. 2005. *Making Things Happen: A Theory of Causal Explanation*, Oxford: Oxford Univ. Press.

Xia, Zhihong 1992. "The Existence of Noncollision Singularities in Newtonian Systems." *Annals of Mathematics* 135: 411-468.