Emergent Chance

Christian List and Marcus Pivato*

February-April 2013, revised June 2014

Abstract: We offer a new argument for the claim that there can be non-degenerate objective chance in a deterministic world. Using a formal model of the relationship between different levels of description of a system, we show how objective chance at a higher level can coexist with its absence at a lower level. Unlike previous arguments for the level-specificity of chance, our argument shows, in a precise sense, that higher-level chance does not collapse into epistemic probability, despite higher-level properties supervening on lower-level ones. We show that the distinction between objective chance and epistemic probability can be drawn, and operationalized, at every level of description. There is, therefore, not a single distinction between objective and epistemic probability, but a family of such distinctions.

1. Introduction

There has been much debate on whether there can be objective chance in a deterministic world. The "orthodox view" is that non-degenerate objective chance ("true randomness")

* This paper was presented at the 10th Annual Formal Epistemology Workshop, Rutgers University, 5/2013, at the Workshop on Deterministic Chance, University of Groningen, 1/2014, and at a seminar at Lund University, 4/2014. We thank the participants at these events as well as Rosa Cao, Hannes Leitgeb, and Kai Spiekermann for helpful comments and suggestions. We are especially grateful to an anonymous referee and the editors for exceptionally detailed and helpful comments. Our work was supported by a Leverhulme Major Research Fellowship and an NSERC grant (#262620-2008).

is incompatible with determinism, and that any use of probability in a deterministic world is purely epistemic, reflecting nothing but an observer's lack of complete information. This view was held by Popper (1982) and Lewis (1986) and has recently been defended by Schaffer (2007). Other authors defend "compatibilist views", according to which there *can* be non-degenerate objective chance in a deterministic world (e.g., Hoefer 2007, Ismael 2009, Sober 2010, Glynn 2010). They employ a variety of argumentative strategies, ranging from an appeal to statistical mechanics (e.g., von Plato 1982, Loewer 2001, Frigg and Hoefer 2010) to a semantic approach linking chance to ability (Eagle 2010).

One strategy is to argue that the objective chance of an event depends on the level of description (e.g., Loewer 2001, Glynn 2010, Strevens 2011). According to this strategy, saying that, macroscopically described, a coin toss has an objective chance of ½ of landing heads is consistent with saying that, microscopically described, the initial state of the coin determines the outcome. Furthermore, as Glynn (2010) argues, such level-specific chances can play the role we expect "objective chance" to play. However, no existing version of this strategy has been sufficiently immunized against the objection that so-called "higher-level chances" are best understood, not as true objective chances, but as expressing the observer's uncertain degrees of belief about the events in question, given his (or her) informational limitations.[1]

We develop an account of objective chance as an emergent phenomenon that answers this objection. Our account is based on a formal model of the relationship

[1] Another view, defended by Lyon (2011), is that higher-level probabilities, such as those we find in classical statistical mechanics or evolutionary theory, are neither objective chances nor credences, but *counterfactual probabilities*, whose primary role is to convey certain counterfactual information in explanations.

between different levels of description of a system (drawing on List 2014 and Butterfield 2012) and shows how indeterminism and chance at a higher level can coexist with determinism and the absence of chance at a lower level.[2] We identify a precise sense in which higher-level chance does not collapse into epistemic probability and show that the distinction between the two can be drawn and operationalized at *every* level of description. It is therefore misleading to draw a single overall distinction between objective chance and epistemic probability. There is an entire *family* of such distinctions: one for each level.

The key insight underlying our account is that different levels of description of a system correspond to different specifications of the system's state space and its set of possible histories, at different levels of "coarse-graining", which induce different "algebras of events" on which probabilities are defined. Far from overcomplicating matters, this insight allows us to develop a parsimonious criterion of what separates objective chance from epistemic probability. What we are suggesting is no doubt implicit in earlier work on the topic (e.g., von Plato 1982), but the literature does not yet contain a

[2] Butterfield (2012) and List (2014) discuss emergent indeterminism in different contexts, not the context of chance. As we were revising this paper, a new defence of level-specific chance by Frigg and Hoefer (2013) came to our attention, which is somewhat similar in spirit to ours, though without an explicit model of branching histories or level-specific algebras of events. Frigg and Hoefer defend a "chance-rule pluralism" according to which "[p]robability rules can be formulated in terms pertaining to different levels of discourse such as macro physics, chemistry, genetics, mechanical engineering and meteorology, and probability rules formulated in such terms have equal right to be considered for inclusion in a Best System package of rules, alongside micro-level rules" (ibid., 8).

satisfactory account of why the objective-epistemic distinction can be drawn at every level and how different levels are insulated from one another so as to permit objective chance as a higher-level phenomenon, despite "chancy" higher-level world histories supervening on "non-chancy" lower-level ones.

2. The basic setup

We model a system whose state evolves over time.[3] *Time* is represented by a set T of points that are linearly ordered. The *state* of the system at each time is given by an element of some set S of possible states, which we call the *state space*. A *history* of the system is a temporal path through the state space, formally a function h from T into S, where, for each time t in T, $h(t)$ is the state of the system at t.

In this model, histories play the role of possible worlds. We write Ω to denote the set of all histories deemed *possible*. This could be either the set of all logically possible functions from T into S or, more plausibly, a proper subset of that universal set, so as to capture the fact that the laws of the system permit some histories while ruling out others. Possibility (in Ω) can then be understood as nomological possibility.[4]

It is helpful to view the states in S as the different possible physical states that the system could be in, and the histories in Ω as the different possible physical histories.

[3] We use and extend a formalism that has previously been used in a different context – that of agency and free will – by List (2014). Structurally similar branching-history models include Butterfield (2012) and, in agential contexts, Belnap, Perloff, and Xu (2001).

[4] The laws of the system may go beyond specifying modal facts (facts about what is and is not nomologically possible); the set Ω only encodes those modal facts. A family of objective chance functions, when it exists, may encode additional, probabilistic facts.

Later, in Section 5, we introduce more "coarse-grained" sets $\mathbb{S}$ and $\mathbb{\Omega}$ of "higher-level" states and "higher-level" histories, but S and Ω as introduced here should be understood as containing only states and histories at a single, "lower" level.

To define determinism and indeterminism, some further terminology is needed. For any history h and any time t, we write h_t to denote the *truncated history* up to time t (defined as the restriction of the function h to all points in time up to t in the relevant linear order). A history h is *deterministic* (in Ω) if, for every time t, its truncation h_t has only one possible continuation in Ω, where a *possible continuation* of h_t is a history h' such that $h'_t = h_t$. A history h is *indeterministic* (in Ω) if, for some time t, its truncation h_t has more than one possible continuation in Ω. Thus indeterministic histories allow *branching*, while deterministic histories do not. Note that a history's property of being deterministic or indeterministic is defined relative to the set Ω of possible histories and thereby relative to the underlying laws (which induce Ω). These laws can be said to be *deterministic* if Ω contains only deterministic histories (i.e., there is never any branching), and *indeterministic* otherwise (i.e., there is sometimes branching).

Probability functions, irrespective of their interpretation, are always defined on algebras of events. An *event* is a collection of histories, i.e., a subset of Ω. An *algebra* is a collection of events that is closed under union, intersection, and complementation. One example of an algebra is the set of all possible events (i.e., the power set of Ω). However, when Ω is infinite, it is technically useful to work with smaller algebras. Typically, the structure of Ω dictates a canonical choice of algebra, which we label $\mathcal{A}(\Omega)$.[5] A *probability function* is a function Pr from $\mathcal{A}(\Omega)$ into the interval from 0 to 1 with standard properties;

---

[5] For example, if Ω has a topology, $\mathcal{A}(\Omega)$ is usually the Borel sigma algebra.

$\Pr(E)$ denotes the probability of event E. The function is *non-degenerate* if some events have probability greater than 0 and less than 1.

There can be different probability functions on the same algebra, indexed to different "locations" or "vantage points". It is widely agreed, for example, that any objective chance function, when it exists, is indexed to a particular history and time (Lewis 1986, Schaffer 2007). To indicate this, we use the notation $\Pr_{h,t}$. Chance assignments thus take the form "event E has objective chance p in history h at time t" (in short, $\Pr_{h,t}(E) = p$).[6] Epistemic probability (or credence) functions are indexed to agents and their informational states (and optionally histories and times). Assignments of epistemic probability (or credence) thus take the form "agent A with information I (in history h at time t) has degree of belief p in event E".[7] For the moment, we do not need any explicit notation for epistemic probability functions.

---

[6] An alternative approach, also consistent with our analysis, is to take conditional chance as basic (e.g., Hájek 2003a,b). On it, one need not explicitly *define* a chance function $\Pr_{h,t}$ for each history h and time t, but can *derive* it from a family of conditional probability functions $\Pr(\bullet|\bullet)$ by conditionalizing on h_t (i.e., on the event $\{h' \in \Omega : h' \text{ is a continuation of } h_t\}$); for any event E, $\Pr_{h,t}(E) = \Pr(E \mid h_t)$. If we take $\Pr(\bullet|\bullet)$ as basic, the functions $\Pr_{h,t}$ inherit Bayesian restrictions (e.g., if t is after t', $\Pr_{h,t}(E)=\Pr_{h,t'}(E|h_t)$), a slight (but perhaps plausible) loss of generality. For analyses of chance via conditionalization on histories, see Loewer (2001, esp. 618), Hoefer (2007, esp. 562-565), and Glynn (2010, esp. 78-79).

[7] If one takes agent A's information I to "screen off" the history h and time t where A is located, one can drop the indices h and t. This would slightly reduce generality by ruling out the possibility that, in different histories or at different times, the same information I may lead to different credences, e.g., due to different priors or psychological states of A.

3. Objective chance

When does a history-and-time-indexed probability function $\Pr_{h,t}$ qualify as an objective chance function? Consider a family of such functions, $\langle \Pr_{h,t} \rangle$, with h ranging over the histories in Ω and t ranging over the times in T. Schaffer (2007) proposes six desiderata that this family must satisfy to play the "objective chance" role.[8] These express the idea that chance must relate in the right way to various other pertinent concepts, such as credence, possibility, the future, intrinsicness, lawfulness, and causation. For present purposes, we accept Schaffer's claim that whichever family of probability functions "best" satisfies these desiderata represents objective chance. We call such a family an *(objective) chance structure* on Ω. A lot more could be said about the desiderata than space constraints allow us to say here, but since they are the starting point of Schaffer's critique of deterministic chance, they also serve as a useful starting point for us. Adapted to our framework,[9] the desiderata are as follows:

---

[8] Schaffer speaks of a single function with three arguments: a proposition (event), a world (history), and a time. Technically, only the projection of this function for a fixed history and a fixed time is a probability function; so it is more correct, though equivalent to Schaffer's usage, to speak of a *family* of history-and-time-indexed probability functions.

[9] Schaffer speaks of each world (history, in our terms) having laws. In our model, laws enter in two ways: (i) they impose modal constraints on what histories are nomologically possible, determining Ω; (ii) they determine the chance structure. As in Schaffer's analysis, some properties can be attributed to histories only relative to the laws. For instance, whether a history is branching depends on the laws encoded in Ω. Similarly, the chance of an event E in history h at time t depends on the laws encoded in the chance structure.

> **The chance-credence desideratum:** If an agent, in history h at time t, were to receive the information that the objective chance of some event $E \subseteq \Omega$ is p, he or she would assign a degree of belief of p to E, no matter what other admissible information he or she has (where this information is formally represented by a subset $I \subseteq \Omega$, containing precisely the histories consistent with the information).[10]

This is a version of Lewis's "Principal Principle", which is commonly accepted as a key constraint on the role of chance. The second desideratum is equally natural: only events that are possible can have non-zero chance.

> **The chance-possibility desideratum:** A necessary condition for an event $E \subseteq \Omega$ to have non-zero objective chance in history h at time t is that E is *possible* in h at t, meaning that E contains a continuation of h_t.

This implies further that only contingent events can have non-degenerate chance in history h at time t, where an event E is *contingent in history h at time t* if E and its

---

[10] We adopt a permissive definition of "admissible information", deeming any information about the past admissible: $I \subseteq \Omega$ is *admissible* in history h at time t if I contains at least all possible continuations of h_t. A more restrictive definition would only strengthen our (positive) conclusions, by weakening the chance-credence desideratum. Since the laws of the system are encoded in Ω (and the chance structure $\langle \Pr_{h,t} \rangle$), the admissible information $I \subseteq \Omega$ can also convey information about the laws. An agent in a deterministic history can thus, in principle, fully predict the future *if* he or she learns which truncated history he or she is in.

complement are each possible in h at t. (Note that this is a notion of contingency *relative to a history and time*.) The third, and related, desideratum says that only future events can have non-degenerate chance.

> **The chance-future desideratum:** A necessary condition for an event $E \subseteq \Omega$ to have non-degenerate objective chance in history h at time t is that E is "properly in the future" in h at t.

Spelling out what it means for an event E to be "properly in the future" in history h at time t is a non-trivial task, but for a variety of criteria the chance-future desideratum is a consequence of the chance-possibility desideratum. This is easy to see for a simple but perhaps unsatisfactory criterion, according to which an event is in the future in history h at time t just in case it is contingent in h at t. However, the implication also holds for more sophisticated criteria, provided the set of contingent events in history h at time t is a *subset* of the set of future events in h at t.[11] The chance-future desideratum could be

[11] One criterion of the future is the following. Let S^T be the set of *all* functions from T into S, i.e., the set of all histories that are *logically* possible, given the state space S and the set of time points T; S^T is a superset of Ω. Any event $E \subseteq \Omega$ is (non-uniquely) representable as $E = E^* \cap \Omega$ for some $E^* \subseteq S^T$. Heuristically, E^* is a (merely) *logically* possible event, while E is the set of all *nomologically* possible histories in E^*. For any time t, let $B(t)$ be the set of all times up to and including t, and $A(t)$ the set of all times after t. Let $S^{B(t)}$ be the set of all functions from $B(t)$ into S, and $S^{A(t)}$ the set of all functions from $A(t)$ into S. Relative to S and T, $S^{B(t)}$ is the set of all *logically* possible truncated histories up to and including time t, and $S^{A(t)}$ the set of logically possible future histories after time t. Then

problematic in a relativistic space-time, where the distinction between "past" and "future" depends on the reference frame of the observer, but we set this complication aside.

The fourth desideratum captures the idea that the objective chance of any event is determined by relevant properties of the event itself, not by extrinsic or relational properties.

> **The chance-intrinsicness desideratum:** For any histories *h*, *h'*, any events E, $E' \subseteq \Omega$, and any times *t*, *t'*, if the triple (E, h, t) is an intrinsic duplicate of the triple (E', h', t'), the objective chance of E in h at t is the same as that of E' in h' at t'.

The precise definition of an "intrinsic duplicate" is difficult and raises a number of philosophical issues beyond the scope of this paper.[12] Informally, if all intrinsic

---

$S^T = S^{B(t)} \times S^{A(t)}$. An event E is *settled in the past* of t if $E = (P \times S^{A(t)}) \cap \Omega$ for some $P \subseteq S^{B(t)}$; E is *properly in the future of t* if it is not settled in the past of t. (If all histories in Ω are *deterministic*, any event $E \subseteq \Omega$ is settled in the past in this sense.)

[12] Schaffer restricts the desideratum to event-world-time triples in which the world is the same. In his model (where worlds are endowed with laws), the same laws then apply to those triples. Since in our model the laws are encoded in $\boldsymbol{\Omega}$ and the chance structure, the same "global" laws apply to any triples. Our formulation of the chance-intrinsicness desideratum can be made more or less demanding, depending on our criterion of when (E',h',t') and (E,h,t) are intrinsic duplicates. One operationalization involves (i) specifying a set Π of permutations (one-to-one, onto functions π) on the state space S such that each π in Π induces a permutation π^* on S^T satisfying $\pi^*(\Omega)=\Omega$, and (ii) interpreting the

properties of (E, h, t) are exactly replicated in (E', h', t'), for instance in two separate runs of the same experiment, then the objective chance facts should be the same.

The fifth desideratum requires that objective chances must be determined by the laws of the system, as opposed to, for instance, the attitudes of the observer.

> **The chance-lawfulness desideratum:** There is a set of laws at the level of Ω that determines the chance structure on Ω.

For example, there are physical laws that imply that a photon has a chance of ½ of passing through each of the two symmetrical slits in the classic double-slit experiment.

The final desideratum, as stated by Schaffer (2007, 126), requires that "if a given chance is to explain the transition from cause to effect, that chance must concern some event targeted within the time interval from when the cause occurs, to when the effect occurs". However, formalizing this is difficult, and several authors have criticized the desideratum as either unclear or implausible (e.g., Glynn 2010, Frigg and Hoefer 2013).

---

elements of Π as transformations of the system's state that preserve all causally relevant features. Different specifications of Π encode different specifications of what those features are. Examples of π in Π might be shifting all particles in the universe five metres in a specific direction or assigning a unique integer number to every electron and exchanging the even- and odd-numbered electrons. Given a specification of Π, we can define (E',h',t') to be an *intrinsic duplicate* of (E,h,t) if there is some permutation π in Π such that $h'(t')=\pi(h(t))$ and $E'=\pi^*(E)$. The chance-intrinsicness desideratum then becomes the requirement that the permutations in Π be symmetries of the system, where π is a *symmetry* if, for any h in Ω, t in T, and $E \subseteq \Omega$, $\Pr_{h',t}(\pi^*(E)) = \Pr_{h,t}(E)$ with $h' = \pi^*(h)$.

As Glynn notes, moreover, Schaffer's formulation is not "the obvious candidate for *the* platitude connecting causation and chance", and a better candidate is the following: "causes (tend to) raise the chance of their effects" (Glynn 2010, 76). We therefore replace Schaffer's desideratum with a variant of Glynn's alternative:

> **Chance-causation desideratum:** If, in history h at time t, some event C is positively causally relevant to another event E, then (except in a case of redundant causation) the chance of E, conditional on C, is greater than the unconditional chance of E.[13]

Note that this desideratum only says that C's raising the chance of E is a *necessary* condition for C's positive causal relevance to E (setting aside redundant causation). It does not say that C's raising the chance of E is a *sufficient* condition for positive causal relevance. On most accounts, causation goes beyond a simple probabilistic relationship.[14]

---

[13] One might further require that if C is negatively causally relevant to E, then (except in a case of redundancy) the chance of E, given C, is lower than its unconditional chance.

[14] We are greatly indebted to a referee for helping us improve our discussion of the six desiderata. For completeness, here is our adaptation of Schaffer's "causal transition" desideratum: if some event C is probabilistically causally relevant to another event E in history h at time t, then C must happen after time t and before E. (On the time of an event, see footnote 11.) This is a consequence of the chance-future desideratum if we adopt the following simple definition: *C is probabilistically causally relevant to E* in history h at time t (before E's occurrence) if C occurs before E and $Pr_{h,t}(E|C) \neq Pr_{h,t}(E)$. The "causal transition" desideratum then asserts, in essence, the *Markov property* of a stochastic

If we accept the six desiderata, we obtain the following conclusion, as noted by Schaffer (2007).

> **Observation 1:** There can be no non-degenerate objective chance in a deterministic history.

To see this, let h be a deterministic history, and consider, for example, the chance-credence desideratum. If an agent were to receive the information that some event E has non-degenerate objective chance p in history h at time t, he or she would have to assign a degree of belief of p to E, no matter what other admissible information he or she has. However, the full information about the truncated history up to time t is certainly admissible.[15] Formally, this is the subset I of Ω consisting of all possible continuations of

---

process: all "background causes" before time t are encoded in the truncated history h_t and no longer count as probabilistically causally relevant *once time t in history h is reached*. Four caveats are due. First, the claim that the "causal transition" desideratum is implied by the chance-future desideratum holds if we disallow $\Pr_{h,t}(C)=0$ (to avoid complications raised by conditioning on possible but zero-probability events). Second, we require C to occur before E to rule out backwards causal relevance (though this may be contested; see, e.g., Price 2008). Third, the probabilistic causal relation depends on a history and time because certain background conditions may be necessary for C to become probabilistically causally relevant to E; e.g., lighting a match cannot "cause" a fire without flammable material nearby. Fourth, *probabilistic* causal relevance must be distinguished from causal relevance *simpliciter*, since the latter may go beyond the former. See also Pearl (2000).

[15] Proponents of higher-level objective chance typically deny this, deeming only higher-level information admissible (e.g., Loewer 2001, Hoefer 2007, and Glynn 2010). In Section

h_t. But h is deterministic, so I is the singleton set containing only h itself. Thus, conditional on I, the agent will assign credence 0 or 1 to E, depending on whether E contains h or not. This contradicts the chance-credence desideratum, which mandated a credence p strictly between 0 and 1.

Similarly, consider the chance-possibility desideratum. We have already noted that it implies that only events that are contingent in history h at time t can have non-degenerate chance in h at t. But if h is deterministic, no event is contingent in h at t. This is because the truncation h_t at any time t has only one continuation, namely h itself, and thus any event E is possible in h at time t if and only if E contains h itself, in which case the complement of E is impossible. So, in a deterministic history, the chance-possibility desideratum rules out non-degenerate objective chance.

Finally, consider the chance-future desideratum. In a deterministic history, all events are "settled in the past" (as technically explicated in footnote 11), and thus no event counts as being "properly" in the future. This would also follow if we defined an event's being "properly in the future" simply as requiring that it be contingent in the relevant history at the relevant time; as we have seen, in a deterministic history, no event has this property. Either way, the necessary condition for non-degenerate objective chance, according to the chance-future desideratum, cannot be met under determinism.

6, we offer a subtly different proposal on how higher-level descriptions can be "insulated" from lower-level descriptions. Instead of defining probabilities on a single algebra of events and restricting the criterion of admissible information, we move from the original, fine-grained algebra to a more coarse-grained, higher-level algebra of events. *Relative to each level*, we can then preserve a permissive criterion of admissible information.

(This is not to deny that, in a more richly described ontology of events, some events could count as being "in the future" in some other sense, even under determinism.)

The question of whether, in a deterministic history, the other three desiderata – chance-intrinsicness, chance-lawfulness, and chance-causation – can be met by a non-degenerate chance structure is less straightforward. But in any case, it is clear that our package of six desiderata cannot be satisfied in its entirety. By contrast, in indeterministic histories, the desiderata pose no such restriction.

> **Observation 2:** There can be non-degenerate objective chance in an indeterministic history.

To see this, it suffices to consider an example. Take a toy universe containing only one particle, whose state is fully described by its location. Space and time in that universe are both discrete, and space is one-dimensional. Thus, spatial positions can be represented by integers, and points in time by positive integers. In other words, we can write $S=\{...,-3,-2,-1,0,1,2,3,...\}$ and $T=\{1,2,3,....\}$. For the purposes of our example, we suppose that the set Ω of nomologically possible histories consists of all histories where the particle begins at some spatial position s in S at time $t=1$, and then moves exactly one spatial position (either left or right) in each time period. These histories are non-deterministic, because any truncated history of length t can be extended in Ω to a truncated history of length $t+1$ in two ways. For example, the truncated history (0,1,2) can be extended both to (0,1,2,3) and to (0,1,2,1). We complete the example by supposing that the laws of this toy universe, over and above inducing the set Ω, determine the following chance structure.

For any time t and any spatial position s, let $E_{[s \text{ at } t]}$ be the set of all histories h in Ω such that $h(t)=s$. Now let h be a specific history in Ω, and suppose $h(t)=s$. We suppose that the laws then specify $\Pr_{h,t}(E_{[s+1 \text{ at } t+1]}) = ½$ and $\Pr_{h,t}(E_{[s-1 \text{ at } t+1]}) = ½$. In other words, the particle has an equal chance of moving right or left in each period. The chance function $\Pr_{h,t}$ is given by multiplying these "one-step" chances in the obvious way. To be precise, for any positive number n, there are 2^n possible extensions of any truncated history h_t of length t to a truncated history of length $t+n$, and the chance function $\Pr_{h,t}$ will assign probability $1/2^n$ to each of these possible extensions. There is no barrier for the resulting family of chance functions to satisfy all of the six desiderata listed above. This should, of course, be uncontroversial.

4. Epistemic probability

We have seen that non-degenerate objective chance can exist only in indeterministic histories in Ω. This does not mean, however, that non-degenerate probability assignments are never appropriate in deterministic histories: they may reflect our uncertain degrees of belief, given incomplete information. In fact, whether or not a history is deterministic, an agent's *epistemic* probability (or credence) function will typically be non-degenerate, unless the agent has complete information *and* there are no chance events in the world.

For example, after having watched the first half of an old recorded football match, we may assign probability $^2/_3$ to our favoured team's winning – a non-degenerate probability assignment – despite understanding that the outcome of the match is long settled: we just do not know what it is. This probability assignment simply expresses our uncertain degrees of belief; we are even dealing with a past event. On the other hand, we may also hold non-degenerate epistemic probabilities in cases involving real chance. Consider our disposition to bet on the outcome of tomorrow's football match. Ordinarily, we think that, over and above the players' skills, there is some real randomness involved here. Generally, therefore, epistemic probabilities (or credences) reflect a mix of incomplete information and objective-chance hypotheses.

Clearly, epistemic probabilities need not, and typically will not, satisfy the six desiderata on objective chance. Instead, they must satisfy the following desideratum, which ensures compatibility between an agent's epistemic probabilities (or credences) and his information:

> **Epistemic probability-possibility desideratum:** A necessary condition for an event $E \subseteq \Omega$ to be assigned non-zero epistemic probability by agent A with information I (at time t in history h) is that E is *epistemically possible* from A's perspective, meaning that $E \cap I$ is non-empty.

Furthermore, an agent's epistemic probabilities may be constrained by the chance-credence desideratum if they are to respect the agent's knowledge of chance facts. One may or may not wish to impose other desiderata on epistemic probabilities, such as Bayesian ones, but we need not take a stand on this here. The key point is that even in the limiting case of a completely deterministic history, non-degenerate epistemic probability is still possible – and typically entirely rational, given an agent's incomplete information. As long as the information set I is non-singleton, there is no conflict between non-degeneracy and the epistemic probability-possibility desideratum. The earlier example of the pre-recorded football match illustrates this: none of the conceivable outcomes of the match is ruled out by our information after watching the first half.

The example also motivates a simple criterion for distinguishing "pure" epistemic probability from probability assignments that are driven, at least in part, by objective-chance hypotheses. If we had complete information about the history of football, we would already know the outcome of the pre-recorded match before we even watched it: there would be no room for non-degenerate epistemic probabilities. However, *even* with complete information about football history, we would *still* not know the outcome of tomorrow's match. We would continue to entertain non-degenerate epistemic probabilities here, which arise from objective-chance considerations.[16] (For those who prefer a microphysical example: we certainly do not know which of the two slits in the classic double-slit experiment a photon will pass through. But even with complete information about the past, we would still assign non-degenerate probabilities to the two possibilities, since – at least on standard interpretations – we are dealing with objective chance.)

In general, a sufficient condition for a non-degenerate probability assignment made in history h at time t to some event E to be *purely epistemic* is that it becomes degenerate, once we conditionalize on complete information about the truncated history up to time t.[17] Formally, this yields the following test:

---

[16] There is no guarantee, of course, that such full-information epistemic probabilities will match the "true" objective chances. The fact *that* we assign non-degenerate epistemic probabilities to the different outcomes here is driven by objective chance; the question of *which* probability values we assign reflects our beliefs about the laws of the system.

[17] Thus an assignment of non-degenerate probability counts as "purely" epistemic if it stems from incomplete information alone, rather than any chance considerations. In an earlier analysis of epistemic probability, Skyrms (2000, 26) defines the "*epistemic*

Test for pure epistemic probability: Let Pr be a probability function held by an agent A in history h at time t, and suppose $\Pr(E)$ is non-degenerate for some event E. A *sufficient condition* for Pr to be *purely epistemic* with respect to E is that $\Pr(E|h_t) = 0$ or 1. Here, $\Pr(E|h_t)$ is shorthand for $\Pr(E|I)$, where I is the information that the truncated history is h_t; formally, $I = \{h' \in \Omega : h'$ is a continuation of $h_t\}$.

5. Emergent indeterminism

So far, we have described the system of interest at only one level, interpretable as the *lower level*. The state space S can be understood as the set of all possible microphysical states, and Ω as the set of all possible microphysical histories. Often, however, we wish to employ *descriptions at some higher level*, for example by describing the state of water as liquid or frozen, rather than as a complex configuration of individual molecules, or by describing a tossed coin as landing heads or tails, rather than as following a particular finely specified physical trajectory. We now focus on some such higher level of description, for example the one used in some special science. (Of course, there can be many different levels of description; more on this later.)

---

probability of a statement" as the "*inductive* probability of that argument which has the statement ... as its conclusion and whose premises consist of ... our stock of knowledge". Our test for "pure" epistemic probability is consistent with this, since, on Skyrms's definition, a non-degenerate epistemic probability that is *solely* due to incomplete information also becomes degenerate if we conditionalize on complete information.

We assume that (i) higher-level states and histories *supervene* on lower-level states and histories (meaning that there cannot be any variation at the higher level without variation at the lower level), and (ii) higher-level states are typically *multiply realizable* by lower-level states. There are many different configurations of water molecules that each instantiate the same state of liquid water. Similarly, there are many different physical trajectories of a coin that all correspond to landing heads.

The relationship between the lower and higher levels can be formally captured by the idea of *coarse-graining*: each higher-level state corresponds to an equivalence class of lower-level states, consisting of all its possible lower-level realizations. The higher level of description thus induces a partition of the lower-level state space S into some set of equivalence classes. We call such a partition a *coarse-graining* of S, and write $\mathbb{S}$ to denote the set of all equivalence classes under it. Each $\mathbb{s}$ in $\mathbb{S}$ represents one higher-level state, so that we can interpret the set $\mathbb{S}$ as the higher-level state space (note the outlined letters for higher-level objects).[18] Let σ denote the function that maps each lower-level state s in S to the corresponding higher-level state $\mathbb{s}$ in $\mathbb{S}$ (formally, the equivalence class to which s belongs). The function σ can also be interpreted as a supervenience relation that maps subvenient lower-level states to their supervenient higher-level counterparts.

---

[18] The representation of higher-level states as equivalence classes of lower-level states (adapted from List 2014) is the simplest mathematical way of capturing the assumption that higher-level states supervene on lower-level states and are multiply realizable by them. We need not take a stand on whether higher-level states *are* equivalence classes of lower-level states or whether they are merely *mathematically represented by* such equivalence classes. Our analysis is compatible with either interpretation.

Next consider histories. Just as a lower-level history is a temporal path through the lower-level state space, formally a function h from T into S, so a higher-level history is a temporal path through the higher-level state space, formally a function $\mathbb{h}$ from T into $\mathbb{S}$. For each time t in T, $\mathbb{h}(t)$ is the higher-level state of the system at t. Note that the supervenience relation σ applies not only to states, but also to histories. For each lower-level history h in Ω, the corresponding higher-level history $\mathbb{h}$ is the function from T into $\mathbb{S}$ such that, for each t in T, $\mathbb{h}(t) = \sigma(h(t))$. The set of all possible higher-level histories is the projection of the set Ω of all possible lower-level histories under the function σ, formally $\mathbb{\Omega} = \sigma(\Omega)$. Further, just as the set Ω of lower-level histories can be associated with an algebra $\mathcal{A}(\Omega)$ of lower-level events, so the set $\mathbb{\Omega}$ of higher-level histories can be associated with an algebra $\mathcal{A}(\mathbb{\Omega})$ of higher-level events.[19]

With these definitions in place, all the concepts, definitions, and observations from the previous sections carry over, *without any formal changes*, to the system described at the higher level, where $\mathbb{S}$ is now the state space, $\mathbb{\Omega}$ the set of possible histories, and $\mathcal{A}(\mathbb{\Omega})$ the algebra of events. All the relevant symbols in those sections must simply be replaced by their outlined counterparts.

For example, we can define determinism and indeterminism in higher-level histories in exact analogy to determinism and indeterminism in lower-level histories: a higher-

[19] Formally, we require the function σ to be *measurable* with respect to $\mathcal{A}(\Omega)$, meaning that the inverse image $\sigma^{-1}(\mathbb{h})$ (i.e., $\{h \text{ in } \Omega : \sigma(h)=\mathbb{h}\}$) is an element of $\mathcal{A}(\Omega)$ for every $\mathbb{h}$ in $\mathbb{\Omega}$. We can then define $\mathcal{A}(\mathbb{\Omega}) = \{\sigma(E) : E \text{ is any event in } \mathcal{A}(\Omega)\}$.

level history $\mathbb{h}$ is *indeterministic* (in $\mathbb{\Omega}$) if, for some time *t*, its truncation $\mathbb{h}_t$ has more than one possible continuation in $\mathbb{\Omega}$, and *deterministic* (in $\mathbb{\Omega}$) otherwise. A *possible continuation* of $\mathbb{h}_t$ is defined, as before, as a history $\mathbb{h}'$ in $\mathbb{\Omega}$ such that $\mathbb{h}'_t = \mathbb{h}_t$. Similarly, a higher-level event $\mathbb{E} \subseteq \mathbb{\Omega}$ is *possible* in history $\mathbb{h}$ at time *t* if $\mathbb{E}$ contains a continuation of $\mathbb{h}_t$.

A key point to note is that determinism at the lower level (in Ω) is fully compatible with indeterminism at the higher level (in $\mathbb{\Omega}$).

> **Observation 3:** For suitable σ (and sufficiently large Ω), there can be indeterministic histories in $\mathbb{\Omega}$ even when all histories in Ω are deterministic (List 2014; for a structurally similar result, see Butterfield 2012).

Since this observation is – in mathematical terms – a possibility result, it can be proved by giving an explicit example of a set Ω of deterministic lower-level histories and a coarse-graining function σ for which the resulting set $\mathbb{\Omega}$ of higher-level histories, where $\mathbb{\Omega}$=σ(Ω), contains some indeterministic histories. Figure 1 provides such an example. Part (a) shows a simple system at the lower level of description (Ω). Time is plotted on the horizontal axis (*T*={1,2,3,...}), and the state of the system on the vertical one. Here the state space *S* is the set of all real numbers. The figure displays five deterministic histories, from time *t* = 1 to time *t* = 25. Part (b) shows the same system at a higher level of description ($\mathbb{\Omega}$), obtained by coarse-graining the state space *S*. Specifically, $\mathbb{S}$ is the set of all integers. The function σ maps each real number *s* in *S* to the closest integer $\mathbb{s}$ in $\mathbb{S}$ (with the usual rounding convention). In this coarse-grained description, there are now five indeterministic histories, supervenient on the lower-level deterministic ones. In particular, they all coincide up to time *t* = 9 before diverging from one another.

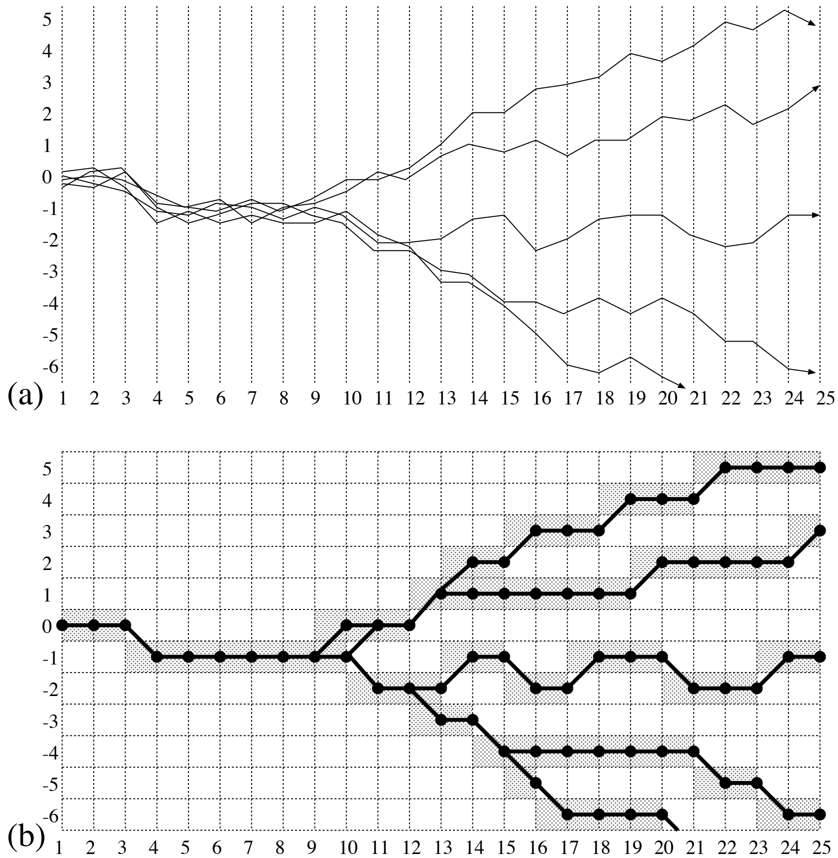

Figure 1: Emergent indeterminism

6. Objective chance at a higher level

Observation 3 shows that while a system may be deterministic at a lower level of description, indeterminism can emerge at a higher level: while the set Ω may contain only deterministic histories, a suitable coarse-graining may yield a set $\mathbb{\Omega}$ of indeterministic histories. But then Observation 2, applied to the level of $\mathbb{\Omega}$ rather than Ω, shows that $\mathbb{\Omega}$ may admit a non-degenerate objective chance structure. Thus, non-degenerate objective chance is possible at a higher level of description, *even if the system is totally deterministic at a lower level of description.*

Corollary of Observations 2 and 3: There can be non-degenerate objective chance in a higher-level history (in $\mathbb{\Omega}$), even when all lower-level histories (in Ω) are deterministic. (A necessary condition for this is that the higher-level history is indeterministic, which is compatible with lower-level determinism.)

At first sight, this conclusion may seem puzzling. Have we not established that when the histories in Ω are deterministic, only degenerate objective chance structures can meet the six desiderata? However, the key insight is this: when evaluating chance and (in)determinism at a higher level of description, *only higher-level language is available.* The relevant family of history-and-time-indexed probability functions now consists of functions defined on the algebra $\mathcal{A}(\mathbb{\Omega})$ of higher-level events rather than the lower-level algebra $\mathcal{A}(\Omega)$, and the index $\mathbb{h}$ now ranges over $\mathbb{\Omega}$ rather than Ω (while t continues to range over T). To make this explicit, we write $\langle \mathbb{P}\mathrm{r}_{\mathbb{h},t} \rangle$ to denote the family of higher-level probability functions (on $\mathcal{A}(\mathbb{\Omega})$, with $\mathbb{h}$ ranging over $\mathbb{\Omega}$), as distinct from the family $\langle \mathrm{Pr}_{h,t} \rangle$ of lower-level probability functions (on $\mathcal{A}(\Omega)$, with h ranging over Ω). Our entire analysis from the previous sections, including the desiderata, must then be re-applied at the level of $\mathbb{\Omega}$ rather than Ω.[20]

[20] Note that this is subtly different from the approach to higher-level chance in, e.g., Loewer (2001), Hoefer (2007), and Glynn (2010), which consists in taking a probability function on the original algebra $\mathcal{A}(\Omega)$ and conditionalizing on higher-level information. This still yields probabilities defined on $\mathcal{A}(\Omega)$, though conditional on higher-level information. By contrast, the algebra $\mathcal{A}(\mathbb{\Omega})$ on which we define higher-level probability functions (indexed to histories in $\mathbb{\Omega}$) is isomorphic to a *sub-algebra* of $\mathcal{A}(\Omega)$.

Past arguments for the *incompatibility* of higher-level objective chance and lower-level determinism tended to make a conceptual error: they supposed that, when evaluating the chance of some higher-level event $\mathbb{E} \subseteq \mathbb{\Omega}$, we could employ a lower-level probability function $\mathrm{Pr}_{h,t}$, indexed to a lower-level history h, or conditionalize on a lower-level event, as in expressions of the form "$\mathrm{Pr}_{h,t}(\mathbb{E}) = 0$" or "$\mathrm{Pr}_{h,t}(\mathbb{E}|E) = 0$". But it should now be clear that this is misguided. Such expressions involve a category mistake: they mix two different levels of description.[21]

The obstacle here is conceptual, not epistemic. There are, of course, a number of epistemic questions about whether, and why, we should employ higher-level descriptions (lower-level information may or may not be accessible to us, higher-level descriptions may or may not be "reducible" to lower-level ones, and so on). We turn to these issues in Section 8. However, the conceptual point is that *when we are operating at the higher level of description, lower-level language is unavailable*. The chances of higher-level events are given by functions of the form $\mathbb{Pr}_{\mathbb{h},t}$, indexed to a higher-level history $\mathbb{h}$ and defined on the higher-level algebra $\mathcal{A}(\mathbb{\Omega})$. So, our claim, at this stage of the argument, is conditional: *if* we are operating at the higher level, we must stick to it.

As further evidence of the pitfalls of mixing levels, note that expressions like "$\mathbb{Pr}_{\mathbb{h},t}(E) = 0$" or "$\mathbb{Pr}_{\mathbb{h},t}(\mathbb{E}|E) = 0$" are not even mathematically well-defined when the

[21] Even Glynn's (2010) defence of "indeterministic chance", whose claim about the level-specificity of chance in response to Schaffer (2007) we agree with, preserves the quantification over lower-level histories (worlds) and introduces different levels only via level-specific laws, not via explicit coarse-graining of states and histories (worlds) and the move to a higher-level algebra.

probability function $\mathbb{P}\mathrm{r}_{\hbar,t}$ is indexed to a higher-level history $\hbar$ and defined on the higher-level algebra of events $\mathcal{A}(\mathbb{\Omega})$, while the event E to which a probability is assigned or on which the probability is conditionalized is described at the lower level. Technically, lower-level events are not in the domain of the probability function $\mathbb{P}\mathrm{r}_{\hbar,t}$.

In Sections 2 and 3, we laid out a theory of objective chance in the setting of indeterministic histories in Ω. But this theory applies equally well to $\mathbb{\Omega}$: simply replace every symbol with its outlined counterpart. As we have seen, when the higher-level analysis (in $\mathbb{\Omega}$) is correctly insulated from lower-level descriptions (in Ω), higher-level indeterminism can coexist with lower-level determinism. So, the possibility of higher-level non-degenerate objective chance follows immediately from the "outlined letter" version of the framework in Sections 2 and 3.

We now give a simple example of emergent chance (familiar from the dynamical-systems literature).[22] Consider a system whose state space S is the interval of all real numbers between 0 and 1. Time is given by the set of positive integers, $T=\{1,2,3,...\}$. The system changes its state from one time period to the next via a *transition rule*, which is formally a function f from S into itself. If s is the state at time t, then $f(s)$ is the state at time $t+1$. Thus, starting at any state s in S, we obtain the following history:

$$h(1) = s,\ h(2) = f(s),\ h(3) = f(\,f(s)\,),\ h(4) = f(\,f(\,f(s)\,)\,), \text{ and so on.}$$

The set Ω is the set of all histories that can be obtained in this way. The system is clearly deterministic.

---

[22] For other, similar examples, see Winnie (1998, 310-314), Suppes (1999), and especially Werndl (2009).

More specifically, suppose that f is defined as follows (as illustrated in Figure 2):

$$f(s) = \begin{cases} 2s \text{ if } 0 \leq s \leq ½; \\ 2\text{-}2s \text{ if } ½ < s \leq 1. \end{cases}$$

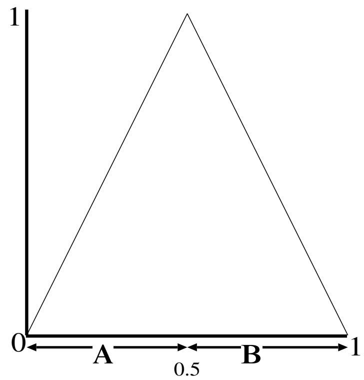

Figure 2: The transition rule *f*

Now we introduce a coarse-graining of this system. Let A and B be symbols representing higher-level states, and let $\mathbb{S} = \{A,B\}$. Define the function σ from S to $\mathbb{S}$ by setting

$$\sigma(s) = \begin{cases} A \text{ if } 0 \leq s \leq ½; \\ B \text{ if } ½ < s \leq 1. \end{cases}$$

By implication, σ maps each lower-level history h to a higher-level history $\mathbb{h}$ that takes the form of a sequence of As and Bs. For example, if we begin in the lower-level state s=1/7, we obtain the lower-level history h=(1/7, 2/7, 4/7, 6/7, 2/7, 4/7, ...). After coarse-graining via σ, this becomes the higher-level history (A,A,B,B,A,B,...). Let $\mathbb{\Omega}$ be the set of all higher-level histories that can be obtained in this way. Then $\mathbb{\Omega}$ contains *every* possible function from T into $\{A,B\}$. Thus, there is not only indeterminism in $\mathbb{\Omega}$, but "maximal" indeterminism: *every* truncated history up to time t can be extended to two truncated histories up to time t+1, *four* truncated histories up to time t+2, and so on.

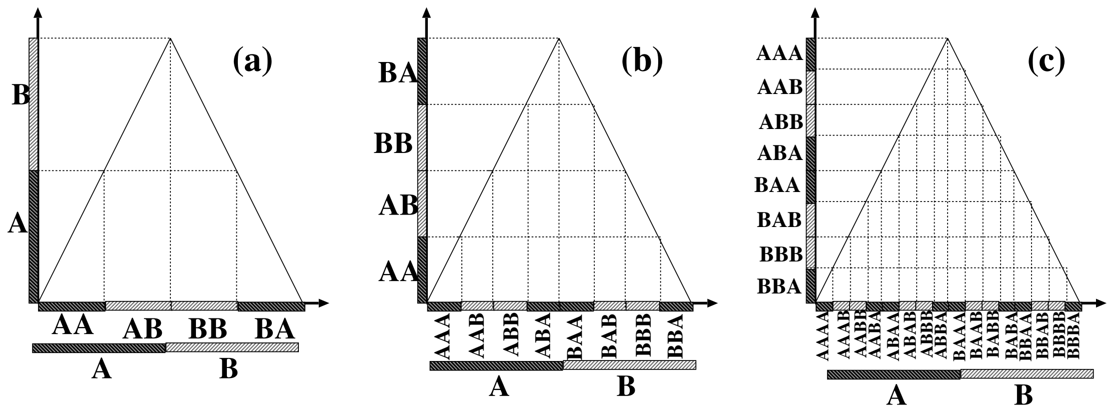

Figure 3: Higher-level histories generated by the function *f* and the partition {A,B}

To see how a non-degenerate chance structure on Ω arises in a very natural way, consider Figure 3. Suppose a higher-level history $\mathbb{h}$ begins with A at time $t = 1$. Then $\mathbb{h}$ must be the coarsened counterpart of a lower-level history h beginning at some state s between 0 and ½. There are two possibilities: either $0 \leq s \leq$ ¼, or ¼ $< s \leq$ ½. In the first case, $f(s)$ (and thus $h(2)$) must be between 0 and ½, and so $\mathbb{h}(2) = A$. In the second case, $f(s)$ (and thus $h(2)$) must be between ½ and 1, and so $\mathbb{h}(2) = B$. Similarly, if $\mathbb{h}$ begins with B at time $t = 1$, its lower-level realizer must begin with some state between ½ and 1. Here, either ½ $< s <$ ¾, or ¾ $\leq s \leq 1$. In the first case, $\mathbb{h}(2) = B$; in the second, $\mathbb{h}(2) = A$.

So, depending on where exactly in the interval S the lower-level state falls at time $t = 1$, we obtain higher-level histories beginning with (A,A), (A,B), (B,B), or (B,A). These correspond exactly to four sub-intervals of S, each of length ¼, as shown in Figure 3(a).

What happens at time $t = 3$? We must now consider *eight* sub-intervals of S, each of length ⅛, as illustrated in Figure 3(b). Which of these we start in determines the higher-level history up to time $t = 3$. For example, if ¼ $< s \leq$ ⅜, then ½ $< f(s) \leq$ ¾, and so ½ $< f(f(s)) \leq 1$. Thus $\sigma(s)=A$, while $\sigma(f(s))=B$ and $\sigma(f(f(s)))=B$. It follows that $\mathbb{h}(s)$ begins with (A,B,B).

To determine the higher-level history up to time $t = 4$, we must consider sixteen sub-intervals of S, each of length $^1/_{16}$. These correspond to the sixteen possible truncated histories of length 4, as illustrated in Figure 3(c).

By iterating this argument, we see that, for any time t, the interval S can be subdivided into 2^t subintervals, each of length $1/2^t$, which correspond to the 2^t possible truncated histories of length t in Ω. This symmetry suggests a chance structure for the higher-level system, where each of these 2^t truncated histories has an equal chance of occurring. A higher-level history can then be seen as a sequence of random choices between A and B, both having probability ½, and with different choices independent of one another, as in a sequence of fair coin tosses. In other words, the higher-level chance structure is that of a classic *Bernoulli process*. There is clearly no barrier for this chance structure to satisfy the six desiderata on objective chance.[23]

One might object that the emergence of non-degenerate objective chance in this example is an artifact of the excessively coarse partition of S into only two sub-intervals, from 0 to ½ and from ½ to 1, which we labelled A and B. But non-degenerate objective chance also emerges from finer partitions. Suppose, for example, we partition S into four

[23] The set Ω of higher-level histories as formally defined here admits many other non-degenerate chance structures. The one we describe has some claim to being the most "natural" one, as it is invariant under an exchange of the symbols A and B; it can thus be motivated by some reasonable assumptions about the symmetries of the system. But it is not necessary for our purposes to show that this chance structure is in any way unique or canonical. What matters is that it is consistent with the six desiderata (at the level of Ω).

sub-intervals of length 1/4 each, labelled $\{AA,AB,BB,BA\}$, as in Figure 3(a). Then an argument similar to the one just given shows that, for any higher-level history $\mathbb{h}$ (now a function from T into $\mathbb{S} = \{AA,AB,BB,BA\}$), if $\mathbb{h}(t) = AA$ (for example), we must have $\mathbb{Pr}_{h,t}[\mathbb{h}(t+1)=AA] = ½$ and $\mathbb{Pr}_{h,t}[\mathbb{h}(t+1)=AB] = ½$. Similar points apply if $\mathbb{h}(t)$ is AB, BB, or BA. If, instead, we partition S into eight sub-intervals of length 1/8 each, labelled $\{AAA,AAB,ABB,ABA,BAA,BAB,BBB,BBA\}$, as in Figure 3(b), then non-degenerate chance emerges again: for any higher-level history $\mathbb{h}$ (now over an even finer $\mathbb{S}$), if $\mathbb{h}(t) = BBB$ (for example), we have $\mathbb{Pr}_{h,t}[\mathbb{h}(t+1)=BBA] = ½$ and $\mathbb{Pr}_{h,t}[\mathbb{h}(t+1)=BBB] = ½$.

Indeed, a non-degenerate chance structure emerges for *any* finite partition of the interval *S*.[24] The reason is that lower-level histories of the system are extremely sensitive to small perturbations. To see this, suppose that *s* and *s'* are two points in the interval *S*, which generate lower-level histories *h* and *h'*, corresponding to higher-level histories $\mathbb{h}$ and $\mathbb{h}$' via some coarse-graining function σ. Suppose *s* and *s'* are very close together. The distance between *f*(*s*) and *f*(*s'*) will then typically be *twice* the distance between *s* and *s'*.[25] And the distance between *f*(*f*(*s*)) and *f*(*f*(*s'*)) will typically be *twice* that between *f*(*s*) and *f*(*s'*) (hence *four* times the distance between *s* and *s'*),[26] and so on. In this way, the lower-level histories *h* and *h'* will rapidly diverge from each other. This, in turn, will lead the corresponding higher-level histories $\mathbb{h}$ and $\mathbb{h}$' to come apart eventually. Even if two higher-level histories $\mathbb{h}$ and $\mathbb{h}$' agree for their first two million entries, there is no reason for $\mathbb{h}$(2,000,001) to be the same as $\mathbb{h}$'(2,000,001).

[24] See Werndl (2009, Section 4.2) for similar observations.

[25] This is true as long as *s* and *s'* are both in sub-interval *A*, or both in sub-interval *B*.

[26] This is true as long as *f*(*s*) and *f*(*s'*) are both in sub-interval *A*, or both in sub-interval *B*.

7. The objective-epistemic distinction at every level

In Section 4, we drew the distinction between objective chance and epistemic probability at the lower level of description (i.e., in Ω). However, the *same* distinction can be drawn at the higher level (in $\mathbb{\Omega}$) and, indeed, at *every* level of description. Objective chance at any level is represented by whichever family of history-and-time-indexed probability functions "best" satisfies the six desiderata *at that level*. An epistemic probability function is only required to satisfy the epistemic probability-possibility desideratum at the relevant level and – if it is also constrained by level-specific chance information – the chance-credence desideratum. While an objective chance function can be non-degenerate *only if* there is indeterminism at the level at which the function is defined, an epistemic probability function can be non-degenerate even if there is determinism at that level.

Importantly, our earlier operational test for *pure* epistemic probability applies at each level. A non-degenerate probability assignment at a given level meets the sufficient condition for being *purely epistemic* – not driven by any chance hypotheses – if it becomes degenerate once we conditionalize on complete information about the truncated history *at that level*.

> **Test for pure epistemic probability, where $\mathbb{\Omega}$ is the level-specific set of histories:** Let $\mathbb{Pr}$ be a probability function held by an agent *A* in history $\mathbb{h}\in\mathbb{\Omega}$ at time *t*, with $\mathbb{Pr}$ defined on $\mathcal{A}(\mathbb{\Omega})$, and suppose $\mathbb{Pr}(\mathbb{E})$ is non-degenerate for some event $\mathbb{E} \subseteq \mathbb{\Omega}$. A *sufficient condition* for $\mathbb{Pr}$ to be *purely epistemic* with respect to $\mathbb{E}$ is that $\mathbb{Pr}(\mathbb{E}|\mathbb{h}_t) = 0$ or 1. As before, $\mathbb{Pr}(\mathbb{E}|\mathbb{h}_t)$ stands for $\mathbb{Pr}(\mathbb{E}|\mathbb{I})$, where $\mathbb{I}$ is the information that the truncated history is $\mathbb{h}_t$; formally $\mathbb{I} = \{\mathbb{h}' \in \mathbb{\Omega} : \mathbb{h}' \text{ is a continuation of } \mathbb{h}_t\}$.

Consider the example of two dice being thrown onto a gaming table. Perhaps the system of tumbling dice admits a microphysical description Ω that is completely deterministic. However, as explained in Section 6, the system may also admit a higher-level description $\mathbb{\Omega}$ in which the objective chance that the gambler is about to throw snake-eyes (a pair of ones) is $^{1}/_{36}$. Now suppose the gambler has already thrown the dice, but the result is hidden from your view by a barrier. The gambler can see the dice, but you cannot. There is no longer any non-degenerate objective chance here; either the dice came up snake-eyes, or they did not. The objective chance of this event is now either zero or one. But from *your* perspective, with limited information, the epistemic probability of the event (your credence) remains $^{1}/_{36}$. Once the barrier is removed, however, you will assign probability 0 or 1. In this story, there is both objective chance (about how the dice will land in the future) and epistemic probability (about how the dice have already landed).

For another example, consider the sort of uncertainty confronted by meteorologists. Perhaps the Earth's atmosphere admits a microphysical description Ω that is completely deterministic. However, this system also admits a higher-level description $\mathbb{\Omega}$, in which the objective chances of future weather events are non-degenerate. Meteorologists gather data from a large array of weather sensors (thermometers, hygrometers, barometers, etc.) and analyze it with computers to predict tomorrow's weather. These predictions are uncertain, and this uncertainty arises in part from the existence of non-degenerate objective chance about tomorrow's weather *at the level of* $\mathbb{\Omega}$. Indeed, meteorologists model the weather as a stochastic system. However, meteorologists also confront another kind of uncertainty. Their network of sensors is sparse. The current meteorological conditions at some location X between two sensors are unknown. But the meteorologists can assign a probability distribution to the current conditions at *X*. This is a purely *epistemic* probability; if there *had* been a sensor at *X*, the epistemic probability for the conditions at *X* would be degenerate, because the meteorologists would know the actual conditions at *X*.

Finally, consider an example from the social world. The police in a big city wish to forecast crime rates in various neighbourhoods, in order to organize effective patrols. Whether or not there is some physical or neuropsychological level (Ω) at which each individual crime is pre-determined, at the ordinary human or social level ($\mathbb{\Omega}$) the police will have to treat patterns of crime as involving non-degenerate objective chance. The chance of various crimes happening will differ from neighbourhood to neighbourhood: there is a higher chance of petty theft and pickpocketing at the railway station than on a quiet residential street. The probabilities in question would not become degenerate even if the police had complete information about the human and social history up to now. Contrast this with a murder investigation in which the police assign probability $^1/_3$ to the hypothesis that Jones did it. This probability is purely epistemic. Conditional on complete historical information (at the level of $\mathbb{\Omega}$), it would collapse into 0 or 1, since the relevant history would settle the matter.

Our claim that there is a well-defined distinction between objective chance and epistemic probability at every level of description, and that non-degenerate objective chance can be a level-specific phenomenon, does not depend on our reasons for employing descriptions at different levels. In particular, the question of *why* we should describe a system at a particular level – say a higher level – is distinct from the question of whether, *at that level*, the system admits non-degenerate objective chance. Recall that, *once we are describing a system at a given level*, the level-specific chance facts are represented by whichever family of history-and-time-indexed probability functions best satisfies the six desiderata *at that level*. What makes them objective chance functions, *relative to that level*, is simply the satisfaction of the six desiderata. Their status as level-specific objective chance functions does not depend on any claims about the "objectivity" of the level itself, and indeed we make no such claims. Keep in mind that the identified functions are not interpreted as representing "objective chance *simpliciter*", independently of the level of description. They represent objective chance *relative to that level*.

On the present picture, one could consistently hold that (i) higher-level descriptions are needed because of our informational (or, more broadly, cognitive) limitations, and yet that (ii) *relative to the higher level*, there can be non-degenerate *objective* chance as distinct from *epistemic* probability. It is a consequence of what we have argued that, even if our reasons for employing higher-level descriptions were *entirely* epistemic, this would not undermine the distinction between objective chance and epistemic probability *relative to that level*; that distinction is drawn solely on the basis of level-specific criteria.[27]

[27] Our point that there can be objective chance relative to a given level of description, independently of our reasons for employing descriptions at that level, is reminiscent of a point that Cohen and Callender (2009) make about (in our terminology) level-relative laws. They argue that relativizing laws to levels "does not entail subjectivity in the sense of making scientists (or others) invulnerable to errors about the laws" (ibid., 30). The reason is that there are observer-independent facts about which generalizations satisfy the appropriate level-relative criteria for lawhood (ibid.). Similarly, we suggest that the question of whether a given family of history-and-time-indexed probability functions satisfies the desiderata for objective chance *at a particular level* does not depend on the observer or on why he or she is interested in that level.

8. The reasons for employing higher-level descriptions

Although the *well-definedness* of the distinction between higher-level chance and higher-level epistemic probability does not depend on our reasons for employing higher-level descriptions, its *interest-value* does. If higher-level descriptions could easily be dispensed with, then the claim that there can be non-degenerate objective chance, relative to the higher level, would be, at most, an idle theoretical curiosity. However, if higher-level descriptions are indispensable in practice, then higher-level chance is a phenomenon of theoretical as well as practical interest. In what follows, we briefly review some of the reasons for employing higher-level descriptions and suggest that, in the case of many systems, such descriptions may actually be indispensable. Of course, a full defence of this claim – familiar from many discussions of the special sciences – is beyond the scope of this paper. But it is also not required for the theory of higher-level chance we have developed. The point of this section is merely to summarize, in broad outline, why higher-level descriptions are needed in many contexts. We have already presented the case for the possibility of level-specific chance itself.

The simplest reason for using higher-level descriptions is that we often lack, and cannot acquire, complete information about the lower-level state of a system. We have seen that, even if this were the only reason for using higher-level descriptions, it would still not undermine the distinction between objective chance and epistemic probability, *relative to the higher level*. However, incomplete information about a system's lower-level state is not the only reason for using higher-level descriptions. Arguably, whether or not we could acquire complete information about the lower-level state, higher-level descriptions are often necessary for agents such as ourselves to avoid the unmanageable complexity of the lower level.

First, the lower-level dynamics of many systems are *chaotic*. Very small errors in our measurement of the current state of a system can lead to very large errors in our predictions of the system's future behaviour. (A simple example is the system shown in Figures 2 and 3 in Section 6.) Since some tiny amount of measurement error is inevitable in practice, prediction may not generally be feasible at the lower level. By contrast, under a suitable coarse-graining, the chaotically diverging trajectories at the lower level can perhaps be amalgamated into a single, predictable trajectory at some higher level – or at least, into a higher-level stochastic process with a manageable amount of randomness; weather forecasting is an example.

Second, even if we could make *perfect* measurements, *or* even if the lower-level dynamics were *not* chaotic, lower-level predictions might still be uneconomical or unparsimonious due to computational complexity, and often not what we are ultimately interested in. For example, imagine that we pour a few drops of blue dye into one part of a water tank, undisturbed by any movement. How will the dye diffuse? If viewed through a microscope, each of the trillions of jostling, jiggling blue dye particles would exhibit Brownian motion, and wander along some convoluted, labyrinthine path through the tank, which, in turn, is the result of a deterministic kinetic-molecular process. All this is extremely hard to model.[28] At a macroscopic level, however, the system admits a very

[28] Note further that the apparent "randomness" of each dye particle's Brownian motion (though not our focus here) is arguably yet another instance of emergent chance: it is due to the way each dye particle is constantly buffeted by millions of much smaller and faster water molecules. At the microscope's level of resolution, a dye particle's motion is best described by an elegant probabilistic model called the *Wiener process*. But it supervenes on the deterministic dynamics of the underlying, more fine-grained molecular collisions.

simple and informative description: if we write down a function describing the three-dimensional density distribution of the dye in the water at time zero, then this function evolves predictably under a partial differential equation called the *heat equation*, which is often amenable to a relatively easy computational solution at all future times.[29] Similarly, to give a more informal example, the dynamics of trillions of molecules of water and other organic compounds ricocheting around a tea cup are hard to model at a microphysical level, while what we are ultimately interested in is how long it takes for the tea to brew or how strong it is, for which we have simple rules of thumb. The examples illustrate that it is often more economical, parsimonious, and informative to use a coarse-grained model, perhaps a statistical one, at a higher level of description.[30]

Third, in many systems, there are robust regularities among higher-level properties. We have just mentioned two such regularities, in the heat equation governing diffusion processes and in the familiar rules of thumb telling us how long it takes for a tea to brew. Further, proponents of "higher-level causation" defend the view that the causes of higher-level effects are sometimes other higher-level properties, and not always the token realizers at the lower level (see, e.g., the contributions in Ellis, Noble, and O'Connor 2012). This view is particularly plausible when (as common in the special sciences)

---

[29] For instance, if the concentration at time zero is a multivariate standard normal distribution with variance σ^2, then the concentration at time t will be a multivariate standard normal distribution with variance $Ct + \sigma^2$, where C is a constant describing the diffusion speed, determined by the water temperature, the mass of the dye particles, etc.

[30] The case for using higher-level statistical models is frequently made in the literature on deterministic chance (e.g., Loewer 2001, Frigg and Hoefer 2010, Albert 2012).

causation is understood as difference-making (List and Menzies 2009, Raatikainen 2010). The difference-making cause of an agent's action, for example, may be, not her brain state, microphysically described, but her mental state, described at a higher level. Similarly, the difference-making cause of an increase in inflation may be a set of actions of the central bank and the government, together with other macro-economic properties, all described at a higher level, rather than their token realizers at the individual level, let alone at the microphysical one (e.g., Sawyer 2003, List and Spiekermann 2013). If these claims are correct, then higher-level descriptions sometimes yield better causal explanations than lower-level ones. Indeed, the special sciences have taught us that, for many phenomena, the most explanatorily illuminating level of description is not a microphysical one, but a chemical, biological, psychological, or social one.[31]

Furthermore, although we are often interested in higher-level information – both in ordinary life and in the special sciences – recovering this information from a complete lower-level description of a system is sometimes not merely uneconomical, but not possible at all, given our cognitive limitations. The "coarse-graining" map σ may be a well-defined mathematical object, but there is no reason to assume that it admits a simple (or even just finite) description in any formal language available to us. Via σ, each higher-level history $\mathbb{h}$ corresponds to an equivalence class H of lower-level histories.

[31] For a related discussion, see Callender and Cohen (2010), who argue that "although all events supervene on a fundamental level, there is no one unique locus of projectibility; rather there are a large number of loci corresponding to the different areas (ecology, economics, solid-state chemistry, etc.) in which there are simple and strong generalizations to be made" (ibid., 427).

Unfortunately, the simplest description of H may be just an enumeration of its elements. If H contains infinitely many elements, it may not even be describable by *any* finite sentence. This is not an outlandish possibility; there is a sense in which, if Ω is itself infinite, "almost all" subsets of Ω admit no finite description.[32] And even if H *is* finitely describable, the shortest description of H could be astronomically large: it may contain as many symbols as there are atoms in the Milky Way galaxy.

Finally, and more speculatively, even if the description of H is finite and of manageable length, it is still possible that it is *formally undecidable*[33] or *computationally*

---

[32] The class of all subsets of Ω that admit a finite description is *countable.* But the class of *all* subsets of Ω is *uncountable*. So the former is a very small subclass of the latter.

[33] This means there is no algorithm which, given a precise specification of the low-level history h, will always correctly determine whether h is in H, after a finite computational process. Many apparently "simple" questions of the form "Is object a of type B?" are formally undecidable. For example: Given an initial point in an orbit of a dynamical system, will this orbit enter a certain region of the state space? Given an arbitrary finite collection of polygonal tiles, can we tile the Euclidean plane with non-overlapping copies of them? Given a geometric object obtained by gluing together polygons, polyhedrons, and their higher-dimensional counterparts (a *simplicial complex*), does this object have a hole in it like a donut? (Is it *simply connected*?) The claim is not that *no* problems of these types can be solved. For many special cases, there are solutions. The claim is that there is no general algorithm for *all* problems of these types. For an introduction to formal undecidability and other examples, see Moore and Mertens (2011, Ch. 7).

intractable[34] whether any particular lower-level history belongs to *H* or not. If this is the case, then some questions about a system's higher-level history cannot be answered on the basis of a lower-level description of the system alone, even a complete description.[35] Instead, we may need to use higher-level descriptions as "primitives".

In sum, there are both practical and theoretical reasons why higher-level descriptions are often indispensable. As should be evident, the present arguments are in line with familiar arguments against the reducibility of higher-level properties to lower-level ones (among other things, due to multiple realizability, as argued, e.g., in Fodor 1974 and Putnam 1975), and for non-reductive physicalism in the special sciences (e.g.,

---

[34] Informally, this means that, although there *is* an algorithm to determine whether *h* is in *H*, it may take trillions of years for it to produce an answer. Thus it is unknowable, *in practice*, whether *h* is in *H*. The best-known (but not the only) formal notion of computational intractability is *NP-completeness*. Many apparently "simple" problems are NP-complete. For example: Given an arbitrary expression in propositional logic, can we assign truth-values to the propositional letters to make it true? Given a finite set of points {**p**,**q**,**r**,**s**,**t**,...} in Euclidean space, is **p** a weighted average of **q**, **r**, **s**, **t**, ... ? Given a list of numbers, can we split it into two sub-lists that add up to the same value? Given an arbitrary network of vertices and links, is there a path that goes through every vertex exactly once? NP-complete problems are only the *bottom* of an infinite hierarchy of increasingly intractable (but decidable) problems. See Moore and Mertens (2011, Ch. 5-6).

[35] By contrast, given the solution to one formally undecidable problem, we can solve any other formally undecidable problem at the same level of the relevant hierarchy; given the solution to one NP-complete problem, we can solve any other NP-hard problem relatively easily, via a "polynomial-time reduction" (Moore and Mertens 2011, Sect. 7.2.4 and 5.1).

Jackson and Pettit 1990, Pereboom 2002, List and Menzies 2009). Non-reductive physicalism is the view that higher-level properties (i) supervene on lower-level properties, (ii) are non-identical to them (and admit no type reduction), and (iii) play a causal role in the world. Our case for the need to employ higher-level descriptions can be regarded as an instance of the case for non-reductive physicalism more generally.

9. An objection

We have seen that different levels of description of a system correspond to different algebras of events on which probabilities are defined, and that the distinction between objective chance and epistemic probability can be drawn at every level. A necessary condition for the occurrence of non-degenerate chance at a given level is indeterminism *at that level*. Since higher-level indeterminism is consistent with lower-level determinism, the latter does not rule out non-degenerate chance at a higher level.

A critic might still object that, when there is lower-level determinism, such higher-level chance is not "true" objective chance. "True" objective chance, the critic might say, requires indeterminism at the lowest or most fundamental level of description. This is what Schaffer (2007) seems to suggest. So formulated, the objection makes the assumption that there is a lowest or most fundamental level, which can be challenged.[36] But there is a second version of the objection, which does not make this assumption. It asserts that if there is *some* level at which the system is deterministic, then any non-

[36] For example, in an earlier article, Schaffer (2003) argues that there is "no evidence" in support of that assumption (ibid., 498) and that there could plausibly be "an infinite descending hierarchy of levels" without any bottom (ibid., 499). In subsequent work, however, Schaffer implies that there is a fundamental level (see especially 2010, 37).

degenerate probability at a higher level can only be epistemic. Or relatedly: if there is non-degenerate objective chance at a higher level, then all lower levels must also admit non-degenerate objective chance.[37] What can be said in response to this objection?

First, the objection, in both versions, overlooks the level-specific nature of the distinction between objective chance and epistemic probability. On the picture we have defended, the hallmark of an objective chance structure at a particular level is the satisfaction of the six desiderata *at that level*, not any facts about other levels. As our formal results have shown, non-degenerate objective chance, understood in this level-specific way, requires indeterminism at the same level, not at others.

Second, in the absence of indeterminism "all the way down", the critic endorsing the objection is compelled to lump all higher-level probability together into an "epistemic" (as opposed to "objective") category. More precisely, the critic is committed to the view that if there is determinism at either the fundamental level (for the first version of the objection) or some level below the level of interest (for the second), then any higher-level probability can only be epistemic. Objective chance at the higher level would have to be accompanied by indeterminism at the lower level. For example, if a system of tumbling dice has deterministic microphysical foundations, the critic loses the ability to distinguish between the *chance* of the dice coming up snake-eyes in the next round (a future event that is "random" from a higher-level perspective) and our *credence* in the hypothesis that, in the last round, the outcome was snake-eyes (a past event that is already settled, but which we have not yet observed). The second of these probabilities would become degenerate if we received complete information about the truncated

[37] We are grateful to the editors for helping us sharpen this objection.

higher-level history up to now. The first probability would not. Indeed, it would satisfy the desiderata for playing the "chance role" at the higher level. Our distinction between higher-level chance and higher-level epistemic probability captures this point naturally. The critic is unable to make the distinction.

10. The mutual embeddability of deterministic and indeterministic systems

Two final results should give further pause to those who think that we can draw a single distinction between objective chance and epistemic probability *simpliciter*, or between indeterminism and determinism *simpliciter*, instead of drawing a family of level-specific such distinctions.

> **Observation 4:** *Any* deterministic system can be expressed as emerging from a more fine-grained indeterministic system.
>
> **Observation 5:** *Any* indeterministic system can be expressed as emerging from a more fine-grained deterministic system.

Observation 4 shows that, no matter how fine-grained the level is at which a given system displays determinism, we can never rule out the possibility of indeterminism at an even lower level. Similarly, Observation 5 shows that, no matter how fine-grained the level is at which a given system displays indeterminism (and admits non-degenerate chance), we can never rule out the possibility of determinism at an even lower level.[38]

To prove Observation 4, let T be any set of times, $\mathbb{S}$ any state space, and $\mathbb{\Omega}$ any set of deterministic histories (i.e., functions from T into $\mathbb{S}$). Now let $S = \mathbb{S} \times \{0,1\}$. In other

[38] See Winnie (1998, 305-306) and Werndl (2009, Section 3.2) for results similar to Observation 5. See also Suppes (1993, 250-252 and 254) for an earlier discussion.

words, S is the set of all ordered pairs of the form ($\mathbb{s}$, b), where $\mathbb{s}$ is an element of $\mathbb{S}$, and b is either 0 or 1. Any S-valued history (i.e., any function h from T into S) is thus a combination of two functions: a function $\mathbb{h}$ from T into $\mathbb{S}$, and a function β from T into {0,1}. Let Ω be the set of all histories ($\mathbb{h}$, β), where $\mathbb{h}$ is any element of $\mathbb{\Omega}$, and β is any possible function from T into {0,1}. It is clear that Ω is a set of *indeterministic* histories; any length-t truncated history ($\mathbb{h}_t$, β_t) in Ω has two possible extensions to a truncated history of length t+1: one where $\beta(t+1)=0$, and one where $\beta(t+1)=1$. Now we define the function σ from S to $\mathbb{S}$ by setting $\sigma(\mathbb{s}, b) = \mathbb{s}$ for any $\mathbb{s}$ in $\mathbb{S}$ and b in {0,1}. Then σ is a coarse-graining map that converts the (indeterministic) histories of Ω into the (deterministic) histories of $\mathbb{\Omega}$.

To prove Observation 5, let T be any set of times, $\mathbb{S}$ any state space, and $\mathbb{\Omega}$ any set of indeterministic histories (i.e., functions from T into $\mathbb{S}$). Now define $S = \mathbb{\Omega} \times T$. In other words, S is the set of all ordered pairs of the form ($\mathbb{h}$, t), where $\mathbb{h}$ is any element of $\mathbb{\Omega}$, and t is some point in time. For any history $\mathbb{h}$ in $\mathbb{\Omega}$, we define a function h from T into S by setting $h(t) = (\mathbb{h}, t)$, for all t in T. Clearly, this is a completely deterministic history. (Heuristically, the lower-level world consists of a "book" and a "clock". The "book" is a complete record of the entire history of the higher-level world, both past and future. It is represented by $\mathbb{h}$, and it never changes. The "clock" is represented by the t-coordinate, which simply records the current time, and thus evolves in an entirely predictable way.) Let Ω be the set of all lower-level histories obtained in this way; then Ω is a deterministic system. Finally, we define a function σ from S to $\mathbb{S}$ by setting $\sigma(\mathbb{h}, t) = \mathbb{h}(t)$, for any $\mathbb{h}$ in $\mathbb{\Omega}$ and t in T. Then σ is a coarse-graining map that converts the (deterministic) histories of Ω into the (indeterministic) histories of $\mathbb{\Omega}$.

Of course, these are purely mathematical constructions, which only provide a proof of possibility. We do not claim that the lower level of any physical system would have the structure described in the previous two paragraphs. In reality, the lower level would presumably be some system of interacting particles and fields, of the kind described in modern microphysical theories. But these examples illustrate that there is no necessary entailment from indeterminism at a higher level to indeterminism at a lower level, or vice versa. Similarly, there is no reason to assume that when a system admits a non-degenerate chance structure at a higher level, it must also admit a non-degenerate chance structure at a lower level, or vice versa.

Furthermore, the present constructions can be iterated indefinitely; it is perfectly possible to have a deterministic higher-level system that is a coarse-graining of an indeterministic lower-level system, which is in turn a coarse-graining of an even *lower*-level deterministic system, and so on. There could be an *infinite* hierarchy of such systems, with no "rock bottom" level; it could be "turtles all the way down".[39]

---

[39] As mentioned above, Schaffer (2003) argues that an infinite hierarchy of levels without any bottom is a coherent possibility, though he does not examine the question of determinism in such a setting. It is worth noting that the metaphysical picture in Schaffer's 2003 article supports the reality of higher-level properties. He criticizes the "ontological attitude according to which the entities [for our purposes: properties] of the fundamental level are *primarily* real, while any remaining contingent entities [properties] are at best derivative, if real at all" (ibid., 498). He suggests that if we drop the assumption of a fundamental level while retaining "a hierarchical picture of nature as stratified into *levels*" (a picture he takes to be "reflected in the structure and discoveries

Observation 6: An infinite hierarchy of levels, forever alternating between deterministic and indeterministic sets of level-specific histories, is a coherent (and unfalsifiable) possibility.[40]

In this scenario, it would make no sense to ask whether the system was deterministic or indeterministic "*simpliciter*". There would be no level-independent answer to this question.

11. Conclusion

Objective chance, along with indeterminism, should be understood as a level-specific phenomenon, which stands in no conflict with determinism at other levels, both lower and higher.[41] Objective chance is only incompatible with determinism *at the same level*. There is not a single distinction between objective chance and epistemic probability, but a separate distinction for each level.

---

of the sciences"), we arrive at "a far more palatable metaphysic in which … all entities [for us: properties] are equally real" (ibid.).

[40] Mathematically, this can be obtained through an *inverse limit* construction.

[41] Throughout this paper, our use of the conventional terminology of "levels" of description should not be taken to imply that there must generally exist a *linearly ordered* hierarchy of levels. Our approach of representing different levels of description in terms of different algebras of events, related to one another via coarse-graining, permits the existence of an entire *lattice* of "levels", *partially ordered* by the coarse-graining relation.

References

Albert, David. 2012. "Physics and Chance." In *Probability in Physics*, ed. Yemima Ben-Menahem and Meir Hemmo, 17-40. Heidelberg: Springer-Verlag.

Belnap, Nuel, Michael Perloff, and Ming Xu. 2001. *Facing the Future: Agents and Choices in our Indeterminist World*. Oxford: Oxford University Press.

Butterfield, Jeremy. 2012. "Laws, Causation and Dynamics at Different Levels." *Interface Focus* 2 (1): 101-114.

Callender, Craig, and Jonathan Cohen. 2010. "Special Sciences, Conspiracy and the Better Best System Account of Lawhood." *Erkenntnis* 73 (3): 427-447.

Cohen, Jonathan, and Craig Callender. 2009. "A better best system account of lawhood." *Philosophical Studies* 145 (1): 1-34.

Eagle, Antony. 2011. "Deterministic chance." *Noûs* 45 (2): 269-299.

Ellis, George F. R., Denis Noble, and Timothy O'Connor, eds. 2012. *Top-down causation*. Theme issue, *Interface Focus* 2 (1).

Fodor, Jerry. 1974. "Special sciences (or: The disunity of science as a working hypothesis)." *Synthese* 28 (2): 97-115.

Frigg, Roman, and Carl Hoefer. 2010. "Determinism and Chance from a Humean Perspective." In *The Present Situation in the Philosophy of Science*, ed. Stadler, F., 351-371. Heidelberg: Springer-Verlag.

Frigg, Roman, and Carl Hoefer. 2013. "The Best Humean System for Statistical Mechanics." *Erkenntnis* (in press).

Glynn, Luke. 2010. "Deterministic chance." *British Journal for the Philosophy of Science* 61 (1): 51-80.

Hájek, Alan. 2003a. "What conditional probability could not be." *Synthese* 137 (3): 273-323.

Hájek, Alan. 2003b. "Conditional Probability Is the Very Guide of Life." In *Probability is the Very Guide of Life: The Philosophical Uses of Chance*, ed. Henry Kyburg Jr. and Mariam Thalos, 183-203. Chicago: Open Court.

Hoefer, Carl. 2007. "The third way on objective probability: A sceptic's guide to objective chance." *Mind* 116 (463): 549-596.

Ismael, J. 2009. "Probability in deterministic physics." *Journal of Philosophy* 106 (2): 89-108.

Jackson, Frank, and Philip Pettit. 1990. "Program Explanation: A General Perspective." *Analysis* 50 (2): 107-117.

List, Christian. 2014. "Free will, determinism, and the possibility of doing otherwise." *Noûs* 48 (1): 156-178.

List, Christian, and Peter Menzies. 2009. "Non-Reductive Physicalism and the Limits of the Exclusion Principle." *Journal of Philosophy* CVI (9): 475-502.

List, Christian, and Kai Spiekermann. 2013. "Methodological Individualism and Holism in Political Science: A Reconciliation." *American Political Science Review* 107 (4): 629-643.

Lewis, David. 1986. "A Subjectivist's Guide to Objective Chance." In *Philosophical Papers*, Vol. II, 83-132. Oxford: Oxford University Press.

Loewer, Barry. 2001. "Determinism and chance." *Studies in History and Philosophy of Modern Physics* 32B (4): 609-620.

Lyon, Aidan. 2011. "Deterministic probability: neither chance nor credence." *Synthese* 182 (3): 413-432.

Moore, Cristopher, and Stephan Mertens. 2011. *The Nature of Computation*. Oxford: Oxford University Press.

Pearl, Judea. 2000. *Causality: Models, Reasoning, and Inference*. Cambridge: Cambridge University Press.

Pereboom, Derk. 2002. "Robust Nonreductive Materialism." *Journal of Philosophy* 99 (10): 499-531.

Price, Huw. 2008. "Toy models for retrocausality." *Studies in History and Philosophy of Modern Physics* 39 (4): 752-761.

Popper, Karl. 1982. *Quantum Theory and the Schism in Physics*. London: Hutchinson.

Putnam, Hilary. 1975. "Philosophy and our mental life." In *Mind, Language and Reality*, 291-303. Cambridge: Cambridge University Press.

Raatikainen, Panu. 2010. "Causation, Exclusion, and the Special Sciences." *Erkenntis* 73 (3): 349-363.

Sawyer, Keith. 2003. "Nonreductive Individualism Part II – Social Causation." *Philosophy of the Social Sciences* 33 (2): 203-224.

Schaffer, Jonathan. 2003. "Is there a fundamental level?" *Noûs* 37 (3): 498-517.

Schaffer, Jonathan. 2007. "Deterministic chance?" *British Journal for the Philosophy of Science* 58 (2): 113-140.

Schaffer, Jonathan. 2010. "Monism: The Priority of the Whole." *Philosophical Review* 119 (1): 31-76.

Skyrms, Brian. 2000. *Choice and Chance*. Fourth edition. Belmont, CA: Wadsworth.

Sober, Elliott. 2010. "Evolutionary Theory and the Reality of Macro-Probabilities." In *The Place of Probability in Science*, ed. Ellery Eells and J. H. Fetzer, 133-161. Heidelberg: Springer-Verlag.

Strevens, Michael. 2011. “Probability out of determinism.” In *Probabilities in Physics*, ed. Claus Beisbart and Stephan Hartmann, 339-364. Oxford: Oxford University Press.

Suppes, Patrick. 1993. “The Transcendental Character of Determinism.” *Midwest Studies In Philosophy* 18 (1): 242-257.

Suppes, Patrick. 1999. “The Noninvariance of Deterministic Causal Models.” *Synthese* 121 (1/2): 181-198.

Von Plato, Jan. 1982. “Probability and Determinism.” *Philosophy of Science* 49 (1): 51-66.

Werndl, Charlotte. 2009. “Are deterministic descriptions and indeterministic descriptions observationally equivalent?” *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics* 40 (3): 232-242.

Winnie, John A. 1998. “Deterministic Chaos and the Nature of Chance.” In *The Cosmos of Science*, ed. John Earman and John D. Norton, 299-324. Pittsburgh, PA: University of Pittsburgh Press.