Present Records of the Past Hypothesis

Abstract

A striking feature of our world is that we only seem to have records of the past. To explain this 'record asymmetry', Albert and Loewer claim that the Past Hypothesis induces a narrow probability density over the world's possible past macrohistories, but not its future macrohistories. Because we're indirectly acquainted with this low-entropy initial macrostate, our observations of records allow us to exploit the associated narrow density to infer the past. I will argue that Albert and Loewer cannot make sense of why this probabilistic structure exists without falling back on the very records they wish to explain. To avoid this circularity, I offer an alternative account: the 'fork asymmetry' explains the record asymmetry, and this in turn explains the narrow density - not vice versa.

Acknowledgements: Thanks to David Papineau and Barry Loewer for helping trigger these thoughts in Bochum, and thanks to two anonymous referees for helping bring them to fruition.

1 Introduction

Our world exhibits a 'record asymmetry' in the sense that it contains many records of the past - that temporal 'side' of us we feel like we're *flowing away* from - but not of the future.[1] However, its fundamental dynamical laws are time-symmetric. In these circumstances, how might the record asymmetry be explained? If time-asymmetry doesn't enter the picture via the dynamics, then it must do so via some sort of boundary condition. The standard candidate is the 'Past Hypothesis', the posit that the universe began in some particular very-low-entropy macrostate. But it's one thing to acknowledge this bare logic, and quite another to actually join the dots between the universe's initial state and the record asymmetry.

To this end, Albert and Loewer argue that because the initial macrostate occupies a specific tiny corner of phase space, it imposes a narrow probability density over the universe's possible past macrohistories, but not over its future macrohistories. By utilising records in the present, we exploit this narrow density to infer the past, for this utilisation is steered by an evolutionarily hardwired, indirect acquaintance with the initial macrostate. Hence, the Past Hypothesis explains the narrow density over past macrohistories, and this in turn explains the record asymmetry.

Although this picture has faced numerous objections, many misinterpret it as a literal account of human thought when dealing with records, and perhaps many more are aimed at the validity of the Past Hypothesis in the first place.[2] In this paper I attempt to analyse Albert and Loewer's account

[1] In a paper about the record asymmetry, this a more useful definition of "the past" than "the temporal side of us that's recorded" (I thank an anonymous referee here). Having made this definition however, my analysis will *specifically* target records, and not our subjective sense of temporal flow/directionality. See Hemmo and Shenker (2022c) for some surprising connections between the two.

[2] See Earman (2006) for a seminal criticism in the second camp.

in a way that's both charitable towards its content and sympathetic towards their background framework. Nevertheless, whilst I agree that there exists a narrow probability density over past macrohistories, I do not think they can make sense of it without helping themselves to the very records they wish to explain.

I will argue that this circularity is avoided if we explain the record asymmetry through a different route: via the 'fork asymmetry', a probabilistic structure connecting localised events. There are good reasons to think this phenomenon goes a long way to underwriting the effectiveness of records as windows on the past. This will warrant some brief discussion, but my main focus will be on why it's a better way of linking the record asymmetry to the universe's initial state. Neither Albert nor Loewer imagines their theory of records to be a complete one, and Albert (2016, 58n4) in particular suspects that the fork asymmetry plays some role. But as we shall see, my way of looking at this structure leads to a very different understanding of the record asymmetry.

This paper proceeds as follows. In section 2 I introduce the statistical mechanical framework used by Albert, Loewer, and myself, and use it to sketch a roadmap of our agreements and disagreements to come. In section 3 I present their rendition of the 'ready state model' of records, and show how it links up to the universe's initial state - a state we've allegedly evolved to become obliquely aware of through its statistical impact on past events. In section 4 I show how they characterise this impact as a narrow probability density over past macrohistories, and argue that they cannot make sense of this structure without falling back on present records - for them, a circular move. In section 5 I explain the record asymmetry by appealing to a different structure, namely the fork asymmetry, showing how it avoids this circularity and brings other advantages. In section 6 I sketch the fork asymmetry's link to the initial state, and explain the narrow density by making unfettered use of present records. I conclude in section 7.

2 The Updated Mentaculus

My way of explaining the record asymmetry is in much the same spirit as Albert and Loewer's. We employ the same statistical mechanical framework, we acknowledge that the initial state lies at the root, and we agree that the record asymmetry is best understood in terms of a time-asymmetric probabilistic structure. Our disagreement, however, is about what this explanatory structure *is*. In this section I will set the scene by outlining our shared framework, clarifying our common ground, and signposting where it will end.

The Past Hypothesis, the posit that the universe began in a particular, very-low-entropy macrostate (the 'Past State'), was originally devised to recover the time-asymmetric character of the Second Law. But armed with the full statistical mechanical apparatus, it spills far beyond thermodynamic territory. Fixing a uniform distribution (the 'Statistical Postulate') over the Past State induces a uniform distribution over all possible sequences of microstates ('microhistories') the universe could evolve through.[3] Since each microstate inhabits a single macrostate, some of which contain many more microstates than others, this a fortiori induces a *non*-uniform distribution over all possible sequences of *macros*tates ('macrohistories') that the universe could evolve through.[4] And because each of these macrostates embeds

[3]As is standard, I mean 'uniformity' with respect to the Lebesgue measure. The questions of what justifies a) this choice of volume measure, and b) its adoption as a probability measure, are well-known issues that I won't enter into. See Hemmo and Shenker (2012, Ch. 8; 2015) for discussion, and also an interesting circularity objection against using these assumptions to explain the record asymmetry; see Stradis (2021) for a response.

[4]Hemmo and Shenker (2012, Ch. 5; 2016) observe that from a purely mechanical perspective, all possible partitions of phase space into macrostates are on an equal footing. But as they also observe (2022b), standard portrayals of the Past State (low-entropy, hot, dense, etc.) single out *one* partition as privileged: that associated with thermodynamic

all subsystems in the universe, this framework attaches a probability to every possible state that every possible object could ever evolve into at the initial time. Albert and Loewer call this framework the 'Mentaculus', aptly subtitled 'a probability map of the universe'.

If the record asymmetry can be explained within the Mentaculus, then the initial state (Past State + Statistical Postulate) must provide the foundation, for this is the time-asymmetric ingredient. However, it isn't immediately obvious how our everyday use of records fits into this rather abstract framework. To start figuring this out, we need to decide on what we actually *mean* when talking about 'records'. This is notoriously hard to pin down in a precise way, but for my purposes two simple points will suffice.

Firstly, it seems reasonable to think that records are macroscopic states of subsystems. If this is right, then the ones we encounter today are part and parcel of the universe's current macrostate (the 'Current State'). Secondly, since records are generally reliable, they reveal the objective probabilities of events at other times. These are the statistical mechanical probabilities found by conditionalising the Mentaculus on the Current State; I call this modified framework the 'Updated Mentaculus', subtitled 'a probability map of the universe *from today's standpoint*'.[5] Although they don't refer to it as such, Albert and Loewer clearly have the 'updated' version in mind when discussing

macrovariables. Absent a justification for this preference, this is a weak spot in Albert and Loewer's theory, as well as my own. Hemmo and Shenker's 'Flat Physicalism' (2022a, 2022b) is a rival to the Mentaculus that avoids this commitment, but further investigation here would take us too far afield.

[5]Throughout this paper I am assuming that statistical mechanical probabilities are 'objective' in the sense that that they appear in the laws of nature. Although determinism and objective chance have historically been seen as incompatible, Loewer (2001) offers a reconciliation. See Frigg (2016) for an overview of recent work on 'deterministic chance', and see Myrvold (2016) for a wider discussion on how to interpret statistical mechanical probabilities.

the Mentaculus, so this will form the basis of the foregoing discussion.[6]

This minimal portrayal of records puts us in a much better position to start understanding the record asymmetry. By observing records, we are somehow tapping into the objective statistical mechanical probabilities encoded in the Updated Mentaculus, and more records we observe, the more our credences converge towards those values. But it just so happens that records are only ever of the *past*, which implies that the Updated Mentaculus attaches much higher probabilities to past events than future events. Hence, one way of understanding the record asymmetry is in terms of a time-asymmetric probabilistic structure within this framework.

So, what exactly is this probabilistic structure? Different answers to this question will amount to different explanations for the record asymmetry. Naturally, these accounts will agree that the initial state lies at the root. But they may offer very different ways of understanding how and why the record asymmetry emerges from it, and it is in this sense that I offer a 'different explanation' to Albert and Loewer's. This is the sense of 'explanation' that I have in mind throughout this paper - as *accounting for*, and not as something stronger like 'causing' or 'grounding'. Albert and Loewer's explanation runs via what I call the 'density asymmetry', the fact that the Updated Mentaculus is dense over few past macrohistories but diffuse over many future macrohistories (see Figure 1).[7] By contrast, my explanation runs via the

[6] This is clear when Albert (2000, 96) writes: "The Newtonian statistical-mechanical contraption for making inferences about the world consists, in its entirety, of three laws and one contingent empirical fact." He continues: "The empirical fact is the one about what the macrocondition of the world currently happens to be", i.e. about the Current State - or at least the portion of it that we observe. See also Albert (2016, 17).

[7] Admittedly, some future macrohistories - or at least lower-dimensional portions of them - are extremely likely, such as the sun's rising tomorrow. But such cases are relatively uncommon, and in any case lack the immense detail characterising probable *past* macrohistories. See Stradis (2021) for more on the reliable, detailed, far-reaching, and easily accessible character of what's disclosed by records.

fork asymmetry, a structure their account omits. Whilst I agree that the density asymmetry exists, I do not think it explains record asymmetry. On the contrary, I think we can only make sense of the density asymmetry by appealing to present records - a move that's off-limits for Albert and Loewer on pain of circularity.

Hopefully, this panorama of the similarities and differences between Albert and Loewer's account and my own will help orientate the reader for the sections to come. With this in place, let's take a closer look at their account, starting with how they think about records themselves.

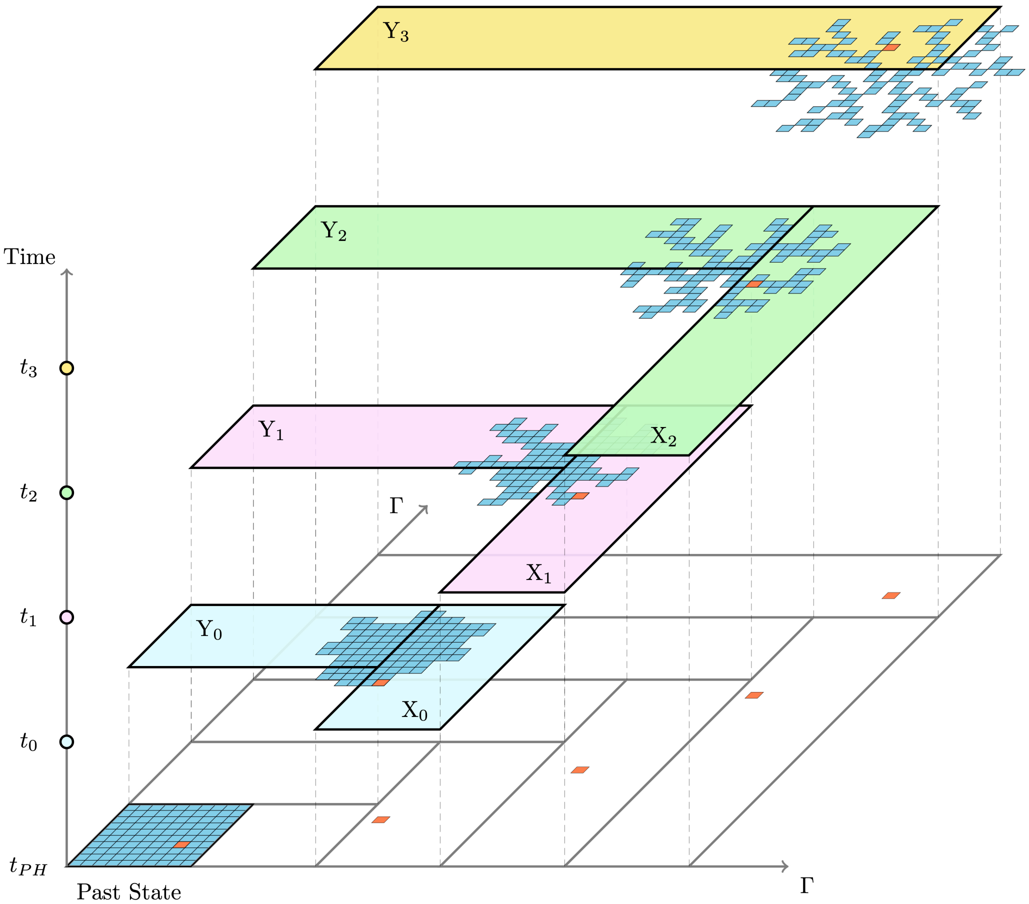

Figure 1: A pixelated diagram of the Mentaculus. The xy-plane represents phase space, the z-axis represents time, small squares represent microstates, and big blocks represent macrostates - projected upwards in pastel. Discarding all microstates that *don't* end up in the Current State (Y_3) yields the Updated Mentaculus. The remaining microstates would overwhelmingly follow the $Y_0 \to Y_1 \to Y_2$ macrohistory; those like the red one which follow $X_0 \to X_1 \to X_2$ are highly atypical. This, combined with the fact neither Y_i nor X_i is especially favoured after t_3, is the density asymmetry.

3 Records and Ready States

Albert (2000, Ch. 6) connects the reliability of records to the Past Hypothesis using a much-discussed regress argument, to which Loewer (2007, 2011, 2012, 2020, 2023) is sympathetic. However, each step of the regress utilises a certain model of records that calls for careful examination. I begin this section by laying out this 'ready state model' in general terms. I then show how Albert levers it into a regress towards ever-earlier times, and explain why the Past State is meant to terminate it.

The ready state model, echoing an earlier theory proposed by Wolpert (1992), is best understood through an example. Let's suppose we observe some record in the present (at t_3) - say, a footprint on a beach. How do we use it to infer a past event (at t_2) like a stroller? At the very least, we need to assume it began in a well-calibrated state at some *more remote* time (at t_1), in this case some smooth sand. This is vital because if the sand was already imprinted at t_1, then the record at t_3 would be spurious, for a stroller needn't have visited at t_2 to form the footprint. Hence, Albert (2000, 117) writes: "The sort of inference one makes from a recording is not from one time to a second in its future or past... but rather from *two* times to a *third* which lies *in between them*." Calibrated state like the smooth sand at t_1 are called 'ready states'.

Because this is meant to be a realistic example, the ready state of course lies in our past. But as it stands, this model doesn't preclude us from inferring future events (at t_4) with the aid of future ready states (at t_5). So, why can't we seem to do this? Whatever else might be said about ready states, this much is clear: we require some means of epistemic access to them. According to Albert, this is possible because as a general rule, ready states are *themselves* recorded in the present. In our example, the smooth sand qualifies as a ready state because there exists some *auxiliary* record, perhaps a polaroid photo, logging its prior existence. By characterising records in this way, Albert respects their sheer diversity, for the open-endedness of what sorts of things can be recorded (and hence qualify as ready states) translates to the open-endedness of what sorts of things can qualify as records. This confers a substantial edge over various influential 20$^{\text{th}}$ century accounts that defined records in overly narrow entropic terms.[8]

Nevertheless, the ready state model faces an obvious regress. In order for a given record to be reliable, it needs to have started out in the right ready state. For example, in order for a footprint to genuinely denote a stroller, it needs to have started out as a smooth portion of sand. This ready state only qualifies as such if it has a present record of its own, such as a polaroid photo. But what supports the reliability of this auxiliary record? This calls for yet another ready state, presumably some blank paper that wasn't fed into the camera pre-printed. However, this ready state only qualifies as such if *it too* has a present record, and what makes *that* reliable? A regress is in plain sight, for in order to support the original record, we need to involve ever more ready states along with ever more present records.

Ultimately, Albert thinks the regress is redeemed by a key feature: each ready state must be *more temporally remote* than the previous to be implicated. This progression away from the present is built right into the ready state model, for this tells us that any inferred state - including a ready state - must be sandwiched between a temporally farther ready state and a present record. This pattern is obvious in something like our polaroid example, for we would certainly expect the blank paper to exist before the sand was smooth. But the point is that this isn't just an artefact of that example. Rather, it

[8]See Frisch's (2023) criticism of Reichenbach (1956), which also applies to the related theories of Grünbaum (1963, 282) and Smart (1967, 131). Even Wolpert fell into this trap by characterising ready states as local equilibrium states; this is true of many highly informative records like photos and fingerprints, but certainly not *all* records. I will revisit this point in footnote 20.

reflects a deeper principle about our epistemic reach when reading records, which is that our inferences cannot be cast to more remote times than their ready states.

The upshot is that although the regress implicates ever more ready states and ever more records, the ready states occur at iteratively earlier times. This logic eventually lands on a record whose ready state is the earliest state in the universe, the Past State, which Albert calls the 'mother of all ready states'. If this is going to live up to its name and end the regress, then the ready state model requires us to have some means of epistemic access to it. But that very model also rules out our having *records* of the initial state, for this would require even earlier ready states - and these by definition cannot exist.

In that case, how *do* we access the Past State? Albert (2014, 162) makes the following suggestion:

> "...some crude, foggy, partly unconscious, radically incomplete, but nonetheless perfectly serviceable acquaintance with the consequences of the past-hypothesis and the statistical postulate and the microscopic equations of motion will very plausibly have been hard-wired into the cognitive apparatus of any well-adapted biological species by means of a combination of natural selection and everyday experience and explicit study and god knows what else."

What 'consequences' of the Past State and its trimmings does Albert refer to? It seems to me that the answer is its *statistical* consequences, which have expressed themselves through the past frequencies of various sorts of events. Over billions of years, these frequencies have hammered organisms into having well-adapted wiring whereby they make survival-conducive responses to environmental cues. But these responses by their very nature reflect an ability to *read records correctly*: bees register scents as records of flowers; venus flytraps register their own microscopic hair-movements as records of insects; and we humans, through our innate fears, register snakes and spiders as records of venomous attacks. Our natural familiarity with the causes of records acts as a kind of proxy for our dim familiarity with the universe's initial state.

Albert's (2023) example should help make this a bit more concrete. A 'regular guy', call him 'Bob', walks into a room to find a $1/2$-melted block of ice in a puddle in an undisturbed warm room. Bob knows nothing about statistical mechanics, but he's curious about the room's contents 10 minutes ago. To figure this out, he might draw on two lessons from past experience. Firstly, blocks of ice are extremely unlikely to become less molten over time. And secondly, the prior probabilities of encountering a $1/4$-melted block and a $3/4$-melted block in a warm room are roughly equal. Bob unconsciously plugs these facts into a simple Bayesian calculation which tells him that the block was more likely to be $1/4$-melted than $3/4$-melted 10 minutes ago. But here is the key point Albert seems to be driving at: the probability values that enter Bob's calculation are based on his exposure to past frequencies (whether personal or ancestral), and since these track the relevant events' statistical mechanical probabilities, his ability to use the ice block as a record reveals a kind of indirect sensitivity to the Past State.

This reading of Albert's account provides an easy answer to a widely-held concern. Many have pointed out that we were perfectly capable of using records long before we knew about the Past Hypothesis - indeed, other organisms manage it to this day. This would be a problem if Albert's account was posed as a literal reconstruction of our conscious reasoning when utilising records, for he would then be committed to the untenable idea that this requires an awareness of the Past State as physics actually describes it. But as he has often affirmed, this is not its purpose. Rather than trying to capture *how* we reason using records, Albert's goal is to capture *why* this reasoning generally succeeds, and since this only requires the oblique sort of acquaintance with the Past State described above, this objection falls away.[9]

But I think we can also rule out another, much more viable interpretation than this one. Fernandes (2022) takes Albert to be saying that the Past State grounds the record asymmetry because it's a particular *known* state, and not because of its physical characteristics. As she rightly points out, this is flawed for a number of reasons, not least because it can't explain why a similarly known 'heat death' doesn't permit records of the future. However, this reading doesn't seem to gel with Albert's story above - a story that Fernandes sees as detached from his account of the record asymmetry.[10] According to Albert (2016, 16-17), the sort of acquaintance with the Past State which underlies our record-savviness was instilled "...as far back as when we were fish, as far back (indeed) as when we were *slime*, by natural selection - and lies buried at the very heart of the deep instinctive primordial unarticulated feel of the world" on penalty of extinction. Only relatively recently has this acquaintance been "...amended and expanded, over time, through explicit scientific practice" (Albert 2016, 39) into the conception that physics affords us today - low-entropy, hot, dense, and all the rest. To my mind, the only non-mysterious way we could be instinctively acquainted with the Past State is through its statistical legacy; and to this end its physical characteristics certainly *are* relevant, as these have real consequences for the Updated Mentaculus' probabilities.

In summary, here is how I think everything fits together in Albert's picture. Records of the past are reliable because they began in the right ready states,

[9]Frisch (2007) leaned towards the more literal interpretation, though in fairness, it wasn't totally clear that this *wasn't* Albert's view back in the days of *Time and Chance*.

[10]Fernandes (2022, 389) writes: "While Albert offers an independent account of how we come to know the Past Hypothesis [via natural selection], that account is not part of his explanation of the record asymmetry and does not appeal to the entropic features of the Past Hypothesis."

and this is probable because the universe began in the Past State - something we indirectly access insofar as we read records correctly. Contra Fernandes, the idea isn't that we have reliable records because we *already* know about the Past State. Rather, reliable records exist because of the Past State, and the fact that we've glommed onto these records expresses a sense in which we *indirectly* know about it.

4 The Bottleneck Argument

The above account hinges critically on the idea that the Past State successfully ends Albert's regress. In this section I will argue that Albert and Loewer cannot justify this claim within their picture, undermining their explanation for the record asymmetry. After discussing the need for such a justification, I show how they envisage the density asymmetry providing it. I then show that they cannot make sense of this structure in a non-circular fashion, and finish by relating my objection to one from Frisch.

In order for the Past State to terminate Albert's regress, it needs to shape probabilities in a way that tallies with the content of records, which of course tell us much about the past and little about the future. In saying this, the point isn't merely that the Past State must *fix* the probabilities of all past events, since it fixes the probabilities of all *future* events in precisely the same way. For all that this tells us, past events could be less probable than future events, which would hardly support the idea that the Past State underpins the record asymmetry as we know it. Rather, the point is that the Past State must ascribe *higher probabilities* to past events than future events, for only then would it have the appropriate time-asymmetric relevance.[11]

[11] Blanchard (2023, §2) makes a similar point: together with the Current State, the Past State might determine *what* we know about the past (via records), but it doesn't automatically follow that this amounts to *more* than what we know about the future.

This is a substantial claim, and one that needs to be justified rather than stipulated. What might provide this justification? One possibility is an argument to the effect that Past State induces a density asymmetry within the Updated Mentaculus. If the macrohistory distribution is dense towards our past but diffuse towards our future, it a fortiori ascribes higher probabilities to past events than future events. So, if this structure could be chalked up to the fact that the universe began in the Past State, this would seem to vindicate its time-asymmetric statistical impact, and by extension its role as the 'mother of all ready states'.

Albert and Loewer clearly endorse the idea of a density asymmetry with the Past State as its source, though they express it in slightly different ways. Loewer (2007) is more explicit, claiming that the Past State induces a tree-like structure whereby macrohistories with sizeable probabilities undergo much more branching towards our future than our past, despite determinism at the microhistory level. This is essentially the density asymmetry with highly improbable macrohistories filtered out. Meanwhile, Albert (2000, 82-85) uses the example of ice cubes falling down a Galton board to argue that a low-entropy initial constraint on a subsystem renders a single past macrohistory (but no single future macrohistory) overwhelmingly probable. As I understand it, his message is that a scaled-down version of the Past State induces a scaled-down version of the density asymmetry. Similar ideas are echoed throughout Albert and Loewer's writings and discussions, and many seem to share my interpretation.[12]

How exactly does the Past State induce a density asymmetry? Loewer (2007) answers this with what I call the 'bottleneck argument', an idea which seems implicit in Albert (2000).[13] Because the Past State occupies a tiny region of phase space whereas the Current State occupies a fairly large region,

[12]See for instance Parker (2005), Frisch (2010), North (2011, 328-329), and Lavis (2012).
[13]Again, Frisch (2023) shares my view on this second point.

the world's possible microhistories become more *bunched up* in phase space as we follow their trajectories from the present towards the past, for they need to converge in that temporal direction. By contrast, because there is no low-entropy 'Future State' that microhistories need to end up in, they are free to diverge wildly as we follow their trajectories from the present towards the future. Since past microhistories are relatively constrained in this way, they have a relatively strong tendency to weave their way through the same *macro*histories over and over. And since the Statistical Postulate ascribes microhistories equal probabilities, the upshot is that the Updated Mentaculus will be sharply peaked over relatively few past macrohistories - the ones which are traversed over and over - but thinly spread over relatively many future macrohistories.

At this point, one might wonder why we still need the story about ready states once we have the bottleneck argument. If the end-goal is to explain the record asymmetry, doesn't that story become obsolete once we've established the Past State's time-asymmetric statistical impact? I think the correct response is that the record asymmetry is in the first instance a puzzle about *records* - localised, macroscopic objects - and not the probabilistic structure of the universe. Without a story about how things like fingerprints, photos, and fossils provide windows on the probabilities encoded in the Updated Mentaculus, this account of the record asymmetry would be seriously lacking. The bottleneck argument's potential is therefore not to explain the record asymmetry all by itself, but to shore up a story that *does* link up with individual items - the story about ready states.

Unfortunately, however, I do not think the bottleneck argument achieves this goal, for the very considerations used to rationalise the density asymmetry count against it in equal measure. It seems plausible that the Past State induces a bunching-up of microhistories towards the Past State. It also seems plausible that other things being equal, this would serve to make their corresponding macrohistories more homogeneous. But in this case other things are *not* equal in a crucial respect. Just as the Past State induces a relatively tight clustering of microhistories due to its small volume, the kinds of macrostates they'll enter into towards our past will *themselves* have relatively small volumes, for the universe's entropy was lower during this period. This means they will constitute smaller targets in phase space, making them less likely to be visited over and over by different microhistories. This consideration implies that the macrohistories corresponding to these microhistories will become more *hetero*geneous, contrary to what the bottleneck argument is meant to show.[14]

We can look at this from another angle by running an analogous argument *forwards* in time. Because there isn't a low-entropy Future State, it seems plausible that microhistories sprawl more erratically towards our future than towards our past. It also seems plausible that other things being equal, this would serve to make their corresponding macrohistories more heterogeneous. But once again, other things are not equal: since the universe's entropy is higher during this period, these microhistories will tend to find themselves in larger macrostates, and since these constitute larger targets in phase space, they are more likely to be visited over and over by different microhistories. This implies that the macrohistories corresponding to these microhistories will become more *homo*geneous, again undermining the bottleneck argument.

In summary, the association between 'low/high entropy' and 'being a small/large target in phase space' cuts both ways. If it's reasonable to think the Past State induces a crowding of microhistories towards our past, then

[14] As an anonymous referee has pointed out, this diversification will have to stop at some point, and indeed swap to a coagulation - after all, there is only *one* Past State. But this would presumably happen in an era far earlier than the times we usually infer using records (yesterday, last year, etc.). Since Albert and Loewer seek to explain the effectiveness of precisely these items, this consideration cannot be used to save the bottleneck argument for their intended purposes.

it's also reasonable to think they'll undergo a diversification of macrohistories in this temporal direction. Likewise, if it's reasonable to think microhistories diverge towards our future, then it's also reasonable to think they'll undergo a clumping of macrohistories in that temporal direction. To be clear, I do not deny that the Updated Mentaculus exhibits a density asymmetry. What I *do* deny, however, is the bottleneck argument can explain this fact. Since this structure provides the rationale for thinking the Past State ends Albert's regress, this way of explaining the record asymmetry is problematic.

To avoid this difficulty, one might be tempted to explain the density asymmetry in an altogether different way - not via the bottleneck argument, but via the many records baked into the Current State. At one point, Loewer (2007, 303) seems to do just this: "...there is typically much less branching towards the past. The reason is that the macro states that arise in our world typically contain many macroscopic signatures... of past events but fewer macroscopic signatures of future states/events." However, this move is off-limits for him and Albert. In their picture, the whole point of arguing for a density asymmetry is to justify the claim that the Past State has the right time-asymmetric statistical impact to end Albert's regress, and the whole point of doing *that* is to explain the record asymmetry. If the only way of rationalising the density asymmetry is by falling back on the existence of records, then this line of explanation is rendered circular. They are stuck with having to explain the density asymmetry *without* enlisting records, which (to its merit) is what the bottleneck argument tries to achieve on entropic grounds.

Let me finish by relating my argument to an existing one in the literature. Frisch (2005a, 2010, 2023) has criticised the bottleneck argument by drawing an analogy with a more tractable system than the whole universe: a partially-expanded box of gas. If we're interested in its future, we can predict a uniformly expanded state with near-certainty. But if we're interested in its past, simply knowing that it began in some low-entropy macrostate, which is meant to be analogous to the Past State, will leave us rather clueless about its history: whether it was confined to the left-hand edge, the top-right corner, or whatever else. Hence, if this toy model is anything to go by, the Past Hypothesis should lead us to expect the universe to have a *reversed* density asymmetry along with a *reversed* record asymmetry.

As many have recognised, this objection misfires, for Albert and Loewer characterise the Past State as a *particular* very-low-entropy macrostate. Admittedly, they aren't very clear about what its distinctive features are,[15] but this doesn't necessarily open the door to Frisch's objection. Albert and Loewer could easily claim that conditionalising on *any* particular very-low-entropy macrostate, so long as we select just one of them, will induce some sort of a density asymmetry. The matter of *which* past macrohistories enjoy a high probability density will obviously depend on this choice. But since the bottleneck argument glosses over this level of detail anyway, that vagueness isn't necessarily fatal.

My argument is essentially that the bottleneck argument fails even when granted this explanatory leeway. Even if Albert and Loewer can get away with being a bit vague about the Past State's distinguishing features over and above 'very low entropy', the bottleneck argument fails to establish a density asymmetry at all. Since this argument is meant to justify the Past State's time-asymmetric statistical impact, and hence its ability to terminate Albert's regress, this account of the record asymmetry is undermined.

[15] Loewer (2007, 300) says the Past State must satisfy "certain further symmetry conditions", whilst Albert (2016, 5) adds that it was "simple, compact, cosmologically sensible". This is a somewhat richer description than just 'very low entropy', but it is consistent with countless other macrostates besides the Past State.

So, where does this leave us? It would be rash to conclude that Albert and Loewer's account cannot be made to work. Perhaps the bottleneck argument can be bolstered with further arguments to do with phase space, entropy, and the dynamics. Or perhaps there is some completely different way of justifying the Past State's statistical role besides a derivation of the density asymmetry. If either strategy were to succeed, then Albert and Loewer's account of the record asymmetry would be back on track. But as it's not clear to me how this might be done, I think we are warranted in trying out a different line of explanation within the Updated Mentaculus framework.

5 The Fork Asymmetry

In this section I outline an alternative account of the record asymmetry based on the fork asymmetry, not the density asymmetry. Reichenbach (1956, §19) observed long ago that the fork asymmetry, or what he called the 'Common Cause Principle', closely resembles the statistical relationship between records and past events. This line of thought has undergone significant development in the years since,[16] but in this section I will simply highlight its main advantages. I will begin by identifying some shortcomings of the ready state model, and then show how the fork asymmetry approach avoids these - and has additional virtues. From there, I revisit the role of ready states within this alternative picture.

The notion that records must start out in appropriate ready states certainly looks like a precondition for their reliability. As an actual portrayal of records however, this model has three major limitations. Firstly, it cannot provide a satisfactory account of why some things are obvious, archetypal records, and other things aren't - for instance, why a photo of a tree is, but a random chunk of cloud isn't. Since this model characterises records as items that

[16] The seminal account is Horwich (1987); see Stradis (2021) for a contemporary picture.

start out in ready states, the only conceivable reason is that some things (e.g. photos) *do* start out in ready states, whereas other things (e.g. random chunks of cloud) don't. But it's hard to see why this would provide the right differentiation, for the prior states of non-records seem no less capable of being recorded - and hence qualifying as ready states - than the prior states of bona fide records. Earlier, I pointed out that this thin portrayal of ready states endows Albert and Loewer's account with one of its main strengths, which is that it respects the diversity of records. But we can now see that this goes too far, for it can't seem to distinguish records at all.

The second issue, much more worrisome, is that the ready state model neglects the actual content of records, which generally concerns *other systems* out there in the world. When I observe a footprint, it seems fair to say that I will make the wrong inference if the sand was imprinted all along. But what we really need is a story about how I make the *right* inference - to use Earman's (1974) example, how I infer a human stroller and not an alien spaceship with foot-shaped landing gear.[17] My ability to get this right clearly depends on my background knowledge of what sorts of events tend to produce foot-shaped imprints on sand. But the ready state model doesn't offer a mechanism of how we lever this information into specific encounters with records, and therefore can't account for the actual subject-matter of our inferences.

This leads on to a third issue, which is that the ready state model suggests records always outdate their content. According to this picture, the inferences we make using records must be sandwiched between even earlier ready states and present records. But in actual fact, it is only the record-forming *interaction* that must be temporally confined in this way, and not the inferred event itself. A photo of an ancient supernova illustrates the dif-

[17] Earman directed this criticism at Grünbaum (1963, 282), but I think it applies here too.

ference: the paper must be prepared before the light enters the camera (the interaction), but certainly not before the explosion itself (the inferred event). This distinction is masked when the inferred event is only recent, for this leaves impossibly little time for the ready state to prepared. If for example I take a polaroid photo of a tree, it's hard to imagine a realistic scenario where during the light's brief journey from tree to camera, I manage to prepare the paper, stuff it in, and press the shutter button. But when the inferred event is very remote, as in our supernova example, the distinction between 'interaction' and 'inferred event' becomes stark, for the time-interval between them becomes wide.[18]

In what follows I argue that the fork asymmetry can explain the record asymmetry in a way that avoids all three of these difficulties. After describing this structure in general terms, I will sketch the standard view of how it underwrites the record asymmetry. Next, I show how we can get considerable extra mileage out of it, and then revisit the three issues above.

The fork asymmetry can be described as follows. Whenever we find a correlated pair of events A and B (e.g. 'footprint' and 'beach towel'), there always seems to be an earlier event C (e.g. 'stroller'), but not a later event, that has the following two properties. Firstly, C raises the probabilities of A and B. This is clear in our example, for a stroller raises the chance of a footprint forming and also of a beach towel being laid down. Secondly, both C and $\neg C$ 'screen off' A and B, which is to say A and B are statistically independent given C and also given $\neg C$. Our example again illustrates this, for although footprints and beach towels are correlated on beaches at large, we'd expect them to be uncorrelated if there was stroller, and also if there wasn't. The fork asymmetry is so-called because there are many 'forward

[18]If this criticism is valid, then the backwards-directionality of Albert's regress no longer seems inevitable, for our inferences needn't be sandwiched between *earlier* ready states and present records.

forks' with C as the earlier fork-point and A, B as the later fork-tips, but no 'backward forks' with the time-reversed structure.

The standard way of having this explain the record asymmetry is by interpreting the fork-tips as records, and the fork-point as the recorded event. Since C raises the chances of A and B, likewise A and B each raise the chance of C, so we have a picture in which multiple items each raise the probability of a single event in their past. When we acknowledge that there can be many more fork-tips than just two, their collective probability-raising can be substantial - just like with real records. Admittedly, C may be accompanied by a background event such that the two together make a future fork-tip highly probable. If for instance someone strolls on a beach *and* the sand is damp, then these will make a footprint highly probable. But because these precursors are not correlated, they are not symptomatic of later fork-tips in the same way that fork-tips are symptomatic of earlier fork-points, and therefore do not constitute records of the future.

This interpretation of the fork asymmetry captures the *collective* effectiveness of records, whereby multiple items triangulate a single past event. But one of the most striking aspects of records is their *individual* effectiveness, so unless this can also be captured, our account will be unsatisfactory.[19] Fortunately, this can be achieved by looking at forward forks in a slightly different way. I have argued elsewhere that if we interpret fork-tips not as discrete records, but as sub-components of a *single* record, they will have the same statistical relationship to the fork-point as the one described above (Stradis, 2021). Following our earlier example, the various indented sand grains that comprise a footprint are correlated with each other, so the presence of one usually involves the presence of many; and because they are all correlated

[19] Horwich (1987, 82) raises this issue to discount Lewis' asymmetry of overdetermination as a theory of records, but as Gołosz (2017, 26-27) observes, it applies equally to the fork asymmetry - at least when viewed conventionally.

with an earlier stroller, their totality renders that event highly probable.[20]

As I said at the start, this isn't the place for an in-depth analysis of this approach. But my cursory sketch is enough to show how it overcomes all three shortcomings of the ready state model. Firstly, it offers a straightforward answer to why some things are records whereas others aren't: some things enter into the requisite correlations whereas others don't. Photos of trees, along with other items like fallen fruit, mycorrhizal fungi, and so on, occupy the role of fork-tips relative to a fork-point, in this case the prior existence of an actual tree. However, the same cannot be said of our random chunk of cloud with respect to any past event.[21] Like the ready state model, this approach respects the diversity of records, for it doesn't ban this or that entity from qualifying as a fork-tip on the grounds that it has the 'wrong' physical properties. Unlike the ready state model however, it is still somewhat discriminating, for such entities must have the right *statistical* properties.

Secondly, this approach far surpasses the ready state model in capturing the content of records. Depending on whether we interpret fork-tips as discrete records or sub-components of a single item, we derive an account which explains their collective or individual informativeness, respectively. On either interpretation however, fork-tips indicate past events by being correlated with them. These correlations are objective probabilistic features of the world, and we are clued into them through our past exposure to the

[20]Back in footnote 8, I hinted that Wolpert was onto something in focussing on highly informative records like photos and fingerprints, whose ready states are local equilibrium states. Unlike Albert's (2000, 117) example of a rolling billiard ball recording a collision, these things *literally resemble* the things they document: a photo of (say) a zebra looks like an actual zebra, and a fingerprint looks like an actual finger's skin pattern, in a way that a rolling billiard ball doesn't 'look like' an actual collision. But we can now understand this difference: local equilibrium states (e.g. blank paper, smooth surfaces) provide fertile ground on which densely-packed fork-tips can form, constituting a highly detailed record.

[21]A cloud can of course be a record of (say) oceanic evaporation, but we're focussing *specifically* on a random, arbitrary chunk.

relevant sorts of events: strollers, beach towels, footprints, or whatever else. The fork asymmetry therefore captures the content of records by bringing this background knowledge to bear on particular encounters with records - something the ready state model cannot do.

Thirdly, the fork asymmetry allows that records can be produced much more recently than the events they document. This is because for a given cluster of fork-tips in the present, the corresponding fork-point can be arbitrarily far in the past. Of course, these fork-tips need to have begun in the right ready states in order to be reliable. But this calibration event needn't happen prior to the event that we actually infer. Hence, there is nothing to say we can't have young records of old events, in accord with common sense.

All this invites the following question: what role, if any, do ready states play in the above picture? As I've said, I admit that records must begin in the right ready states in order to be reliable. However, we no longer need to state this as an independent posit. Once we have the fork asymmetry in place, the fact that records *really do* tend to begin in the right ready states is ensured by the fact that they *really do* correlate with whatever past events they seem to signify. If for example I observe a footprint, the fact that footprints genuinely correlate with past strollers will make it highly probable that this item started out flat. We cannot rule out the possibility that it was imprinted all along, but this is improbable because such bogus footprints are rare, and such things are rare because of the aforementioned correlation. Since the fork asymmetry gives us ready states for free, we needn't put them into the picture by hand.

As an added bonus, this allows us to have an even weaker epistemic relationship to ready states than the one demanded by Albert and Loewer. We often make good use of records without having any idea of the relevant ready state. When my friend shows me a photo of a giraffe on her phone, I take it at face value, but I frankly haven't a clue how the electronics had to be configured in order for the camera to be receptive. Although Albert and Loewer would agree that I am consciously ignorant of the ready state, they must insist that this information does in fact reach me in some very roundabout way through my indirect acquaintance with the Past State. But my account asks for even less: all it requires is that I've had enough experience to twig a correlation between photos of giraffes and actual giraffes - or perhaps more generally, between photos and corresponding worldly objects. In my picture, information about the phone's ready state needn't reach me at all - even indirectly.

By now, I hope to have made a convincing case that the fork asymmetry does a better job of explaining the record asymmetry than the ready state model. My next task is to show how it's rooted in the initial state, and tie up some loose ends in relation to the Updated Mentaculus.

6 The Initial State

There are a number of theories on the table to the effect that the fork asymmetry is underpinned by the character of the initial distribution: the fact that it was random, chaotic, non-conspiratorial, correlation-free, or something along these lines. Once again, rather than rehashing this in great detail, my goal is to show how the general idea supports my overall proposal.[22] I will start by sketching the basic picture, and showing how it allows us to explain the record asymmetry in a non-circular way. I then clarify how it explains the density asymmetry, and make some closing remarks about our epistemic access to the initial state.

[22]The foregoing loosely follows Papineau (1985) and Stradis (2021), but nothing important hinges on this. Related accounts include Arntzenius (1992), Spirtes *et al.* (2000), Pearl (2009), Frisch (2014, 224-8), and Blanchard (2023). See Arntzenius (2010) for an overview of how this approach links up with the wider literature on causal modelling.

To appreciate the link between the fork asymmetry and the initial distribution, let's start by picturing a forward fork with C as the fork-point and A, B as the fork-tips. We know from experience that C guarantees neither A nor B. However, it is a well-known feature of deterministic systems that any given event has a necessitating set of physical circumstances, a 'determinant', at every other moment in time. This means that for a given instance of C, there exist background conditions X_A such that $(C \wedge X_A)$ necessitates a particular instance of A, and also some X_B such that $(C \wedge X_B)$ necessitates a particular instance of B:

$$\Pr(\, A \mid C \wedge X_A \,) = 1 \tag{1}$$

$$\Pr(\, B \mid C \wedge X_B \,) = 1 \tag{2}$$

We can think of background conditions X_i as incomplete determinants that are 'completed' when combined with C. Since putative records can be spurious, there also exist determinants Y_A and Y_B that exhaust all the spurious ways of producing A and B, respectively:

$$\Pr(\, A \mid Y_A \,) = 1 \tag{3}$$

$$\Pr(\, B \mid Y_B \,) = 1 \tag{4}$$

Since $(C \wedge X_A)$ and Y_A exhaust all possible ways of producing A, and since $(C \wedge X_B)$ and Y_B exhaust all possible ways of producing B, (1)-(4) imply:

$$A \Leftrightarrow (\, C \wedge X_A \,) \vee Y_A \tag{5}$$

$$B \Leftrightarrow (\, C \wedge X_B \,) \vee Y_B \tag{6}$$

where the double arrow implies symmetric entailment. C appears in the necessary and sufficient conditions for both A and B, so given C we can expect a correlation between A and B so long as $(X_A \vee Y_A)$ and $(X_B \vee Y_B)$

are statistically independent:

$$(X_A \vee Y_A) \perp\!\!\!\perp (X_B \vee Y_B) \tag{7}$$

We can therefore explain why A, B, and C constitute a forward fork, and not a backward fork, by appealing to the independence of their associated background conditions X_i and determinants Y_i. If we interpret X_A, Y_A, X_B, and Y_B as inhabiting in the early universe, then we arrive at this result: the fork asymmetry is a consequence of the fact that certain correlations did *not* obtain in the universe's initial state.

What sorts of correlations would these be? In general, an event's past determinants are time-slices of its past light-cone. We can therefore expect the relata of backward-fork-producing correlations to be highly convoluted, microscopically specified circumstances. This in turn implies that the correlations *themselves* would be highly conspiratorial, for they require these initial circumstances to perform miraculous choreography: they must evolve and interact in such a way that they eventually produce macroscopic correlations - say, between footprints and beach towels - without any sort of common macroscopic germ. It seems reasonable to think that initial microstates which are rife with these latent correlations would have an extremely small Lebesgue measure, and be sparsely scattered across the Past State in phase space. Therefore, they will be given a low probability by any initial distribution that *doesn't* oscillate across phase space in a carefully orchestrated manner to pathologically favour them. Since the Statistical Postulate is a uniform distribution, this seems to give us what we want.

Of course, if the aim of the game is to explain the record asymmetry, then we still need the Past State to block Loschmidt's objection. Without it, putative records in the form of fork-tips will in all likelihood be fake, having just fluctuated from higher entropy. But the Statistical Postulate closes the explanatory gap by underpinning the fork asymmetry, which is my immediate model of how records work. This points to a division of explanatory labour within the initial state: whereas the Past State helps paint an entropic picture of the universe that's consistent with reliable records, the Statistical Postulate explains why such things positively exist.[23]

This overview sums up my way of seeing the link between the initial state and the record asymmetry, and captures my salient reason for favouring this approach over Albert and Loewer's. In their account, the probabilistic structure used to explain the record asymmetry (i.e. the density asymmetry) cannot be derived from the fact that the Past State occupies a much smaller volume in phase space than the Current State. It looks like the only way of explaining this structure is with reference to records embedded in the Current State, which they cannot enlist on pain of circularity. In my account, the probabilistic structure used to explain the record asymmetry (i.e. the fork asymmetry) can be derived from the initial state directly. Hence, we can explain the record asymmetry without begging the question.

So, what's the status of the density asymmetry in my picture? When we tally up the sum-total of what records in the Current State tell us, it looks very much like this structure genuinely exists: the Updated Mentaculus is dense over past macrohistories in which dinosaurs once roamed the Earth, the Romans invaded Britain, and so on, but thinly spread over many future macroscopic scenarios. But since we've already explained the record asymmetry through a different channel, namely the fork asymmetry, we can help ourselves to whatever present records we want in order to explain the density asymmetry.

[23] See Stradis (2021) for details.

At this point, one might suspect something fishy at play like backward causation: by explaining the density asymmetry using present records, it might look like I'm saying they retroactively *make it the case* that history happened the way it did. However, this would assume a much stronger notion of explanation than the one I committed to back in section 2. The metaphysical root of all time-asymmetric phenomena is the initial state. But some ways of understanding this might be better than others, and there is nothing to say this can't involve using present features of the Updated Mentaculus like records to understand past features like the sharply-peaked distribution over past macrohistories. In this sense, my explanatory arrow here can be seen as 'pointing backwards'. But when we recall that its orientation is due to the record asymmetry, and that this phenomenon is ultimately rooted in the initial state, my account isn't as metaphysically radical as it might first appear.

I wish to finish with some closing remarks about our epistemic access to the initial state. I agree with Albert and Loewer that our natural capacity to read records correctly - a capacity which is displayed even in our most primitive survival-conducive responses to external stimuli - betrays a sort of ancient familiarity with the past frequencies of events. But whereas they see this as a kind of indirect acquaintance with the Past State, I see us having a more conventional, direct acquaintance via records *of the state itself*. This for me is a possibility because whilst I agree that reliable records must start out in the right ready states, I do not require them to exist prior to the recorded event. And come to think of it, this seems like a perfectly natural way of thinking about much of our cosmological evidence. Are all the world's copies of the Hubble Deep Field not records of the early matter inhomogeneities that spawned stars and galaxies? And do these photos, like the individual pixels that make them up, not point to these nascent stirrings in the way that the fork asymmetry describes? If the initial state is epistemologically unusual, it's not because we *can't* have records of it. Rather, it's because we *can* have records of it, and yet it provides the metaphysical foundation of why those very records are reliable.

7 Conclusion

According to Albert and Loewer, the record asymmetry is explained by the fact that records began in the right ready states, a circumstance that's made highly probable by the Past State. By making successful use of present records, we're revealing our long evolutionary exposure to past frequencies, and since these frequencies reflect probabilities, they claim that we're indirectly acquainted with the Past State through its consequences. To sustain this picture, Albert and Loewer need a rationale for thinking that the Past State has the requisite time-asymmetric statistical impact. But the bottleneck argument fails to provide this rationale, and since falling back on present records lands them with a circularity, their account is problematic.

By explaining the record asymmetry via the fork asymmetry, this problem is avoided, for this structure can be derived directly from the initial state. Once this is established, we can make free use of records to explain the density asymmetry. In the big picture of the Updated Mentaculus, this has been my main reason for preferring that approach over Albert and Loewer's. But even when we zoom in and examine its merits as an immediate account of records, it has three further advantages: firstly, it explains why some things are records and others aren't; secondly, it does justice to their collective and individual content; and thirdly, it permits young records of old events. Ready states must indeed be in place, but we needn't stipulate this, as it's already assured by the correlations that constitute the fork asymmetry. Hence, ready states do not play a significant explanatory role in my picture.

In very crude terms, Albert and Loewer's order of explanation runs 'initial state, density asymmetry, record asymmetry', whereas mine runs 'initial state, fork asymmetry, record asymmetry, density asymmetry'. We agree on the fact that the initial state is the metaphysical root of the record asymmetry. But whereas they see it as the 'mother of all ready states', I see it as the 'mother of all forward forks'.

References

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

———. 2014. "The Sharpness of the Distinction between the Past and the Future". In *Chance and Temporal Asymmetry,* ed. Alastair Wilson, 159–174. Oxford: Oxford University Press.

———. 2016. *After Physics.* Harvard University Press.

———. 2023. "Conclusion". In *The Probability Map of the Universe: Essays on David Albert's Time and Chance,* ed. Barry Loewer, Brad Weslake, and Eric Winsberg, 351–376. Harvard: Harvard University Press.

Arntzenius, Frank. 1992. "The Common Cause Principle". *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association* 2:227–237.

———. 2010. "Reichenbach's Common Cause Principle". *Stanford Encyclopedia of Philosophy.*

Blanchard, Thomas. 2023. "Causation and the Time-Asymmetry of Knowledge". *Australasian Journal of Philosophy* (forthcoming).

Earman, John. 1974. "An Attempt to Add a Little Direction to "The Problem of the Direction of Time"". *Philosophy of Science* 41 (1): 15–47.

———. 2006. "The "Past Hypothesis": Not even false". *Studies in History and Philosophy of Modern Physics* 37 (3): 399–430.

Fernandes, Alison. 2022. "How to explain the direction of time". *Synthese* 41 (1): 15–47.

Frigg, Roman. 2016. "Determinism and Chance". In *The Oxford Handbook of Probability and Philosophy,* ed. Alan Hájek and Christopher C. Hitchcock, 460–474. Oxford: Oxford University Press.

Frisch, Mathias. 2005. ‘‘Counterfactuals and the past hypothesis”. *Philosophy of Science* 72:739–750.

———. 2010. ‘‘Does a Low-Entropy Constraint Prevent Us from Influencing the Past?”. In *Time, Chance, and Reduction: Philosophical Aspects of Statistical Mechanics,* ed. Andreas Hüttemann and Gerhard Ernst, 13–33. Cambridge: Cambridge University Press.

———. 2014. *Causal Reasoning in Physics.* Cambridge: Cambridge University Press.

———. 2023. ‘‘Causes, Randomness, and the Past Hypothesis”. In *The Probability Map of the Universe: Essays on David Albert's Time and Chance,* ed. Barry Loewer, Brad Weslake, and Eric Winsberg, 294–311. Harvard: Harvard University Press.

Frisch, Matthias. 2007. ‘‘Causation, Counterfactuals, and Entropy”. In *Causation, Physics, and the Constitution of Reality: Russell's Republic Revisited,* ed. Huw Price and Richard Corry, 351–395. Oxford: Oxford University Press.

Gołosz, Jerzy. 2017. ‘‘Weak interactions: asymmetry of time or asymmetry in time?”. *Journal for General Philosophy of Science* 48:19–33.

Grünbaum, Adolf. 1963. *Philosophical Problems of Space and Time.* New York: Alfred A. Knopf.

Hemmo, Meir, and Orly R. Shenker. 2012. *The Road to Maxwell's Demon: Conceptual Foundations of Statistical Mechanics.* New York: Cambridge University Press.

———. 2015. ‘‘Probability and Typicality in Deterministic Physics”. *Erkenntnis* 80:575–586.

———. 2016. *Maxwell's Demon.* DOI: 10.1093/oxfordhb/9780199935314.013.63: Oxford Handbooks Online.

Hemmo, Meir, and Orly R. Shenker. 2022a. "Flat Physicalism". *Theoria* 88 (4): 743–764.

———. 2022b. "Is the Mentaculus the Best System of Our World?". In *Rethinking the Concept of Law of Nature: Natural Order in the Light of Contemporary Science,* ed. Yemima Ben-Menahem, 89–128. Jerusalem Studies in Philosophy and History of Science. Cham, Switzerland: Springer.

———. 2022c. "The Second Law of Thermodynamics and the Psychological Arrow of Time". *British Journal for the Philosophy of Science* 71 (1): 85–107.

Horwich, Paul. 1987. *"Asymmetries in Time: Problems in the Philosophy of Science".* Cambridge, MA: MIT Press.

Lavis, David A. 2012. "Book Review: Time, Chance, and Reduction: Philosophical Aspects of Statistical Mechanics. Gerhard Ernst and Andreas Hüttemann, Eds., Cambridge University Press." *Studies in History and Philosophy of Modern Physics* 43 (1): 69–71.

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

———. 2007. "Counterfactuals and the Second Law". In *Causation, Physics, and the Constitution of Reality: Russell's Republic Revisited,* ed. Huw Price and Richard Corry, 293–326. New York: Oxford University Press.

———. 2011. "The emergence of time's arrows and special science laws from physics". *Interface Focus* 2 (1): 13–19.

———. 2012. "Two Accounts of Laws and Time". *Philosophical Studies* 160:115–137.

———. 2020. "The Mentaculus Vision". In *Statistical Mechanics and Scientific Explanation: Determinism, Indeterminism, and Laws of Nature,* ed. Valia Allori, 3–29. World Scientific.

Loewer, Barry. 2023. "The Mentaculus: A Probability Map of the Universe". In *The Probability Map of the Universe: Essays on David Albert's Time and Chance,* ed. Barry Loewer, Brad Weslake, and Eric Winsberg, 13–54. Harvard: Harvard University Press.

Myrvold, Wayne C. 2016. "Probabilities in Statistical Mechanics". In *The Oxford Handbook of Probability and Philosophy,* ed. Alan Hájek and Christopher C. Hitchcock, 573–600. Oxford: Oxford University Press.

North, Jill. 2011. "Time in Thermodynamics". In *The Oxford Handbook of Philosophy of Time,* ed. Craig Callender, 312–350. Oxford: Oxford University Press.

Papineau, David. 1985. "Causal Asymmetry". *British Journal for the Philosophy of Science* 36:273–289.

Parker, Daniel. 2005. "Thermodynamic Irreversibility: Does the Big Bang Explain What It Purports to Explain?". *Philosophy of Science* 72 (5): 751–76.

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

Reichenbach, Hans. 1956. *The Direction of Time.* Ed. Maria Reichenbach. New York: Dover Publications.

Smart, John J. C. 1967. "Time". In *Encyclopedia of Philosophy,* ed. Paul Edwards, 126–132. Vol. 8. New York: Macmillan.

Spirtes, Peter, Clark N. Glymour, and Richard Scheines. 2000. *Causation, Prediction, and Search 2e.* Cambridge, MA: MIT press.

Stradis, Athamos. 2021. "Memory, the fork asymmetry, and the initial state". *Synthese* 199 (3-4): 9523–9547.

Wolpert, David H. 1992. "Memory Systems, Computation, and the Second Law of Thermodynamics". *International Journal of Theoretical Physics* 31:743–785.