AUSTRALASIAN PHILOSOPHICAL REVIEW
2024, VOL. 8, NO. 3, 216–233
https://doi.org/10.1080/24740500.2022.2155200

Routledge
Taylor & Francis Group

LEAD ARTICLE

The Open Universe: Totality, Self-reference and Time

Jenann Ismael

Columbia University

ABSTRACT

Before the twentieth century, the Universe was usually imagined as a large spatially extended thing unfolding in time. The past was fixed and the future was open; unfolding was conceived as an asymmetric process of coming into being. Relativity introduced a new vision in which space and time are presented together as a single four-dimensional manifold of events. That, together with the fact that the fundamental laws of our classical theories are symmetric in time, made understanding why the past and future present themselves so differently in our experience one of the central challenges of physics. The last two centuries have seen a great deal of progress in understanding various of the so-called arrows of time: the thermodynamic arrow, the cosmological arrow, the arrow of knowledge or information. There remains an outstanding piece of this puzzle that has seen little progress, one that Roger Penrose described in his beautiful paper Singularities and Time-Asymmetry in 1979 as: 'The arrow most difficult to comprehend … namely the feeling of relentless forward temporal progression, according to which potentialities seem to be transformed into actualities.'

I will propose that the insight needed to resolve the problem involves taking into account that we are part of the universe and that any attempt to model it as a totality involves self-reference. I will argue specifically that self-reference, against the background of a thermodynamic gradient, creates an instability in an embedded agent's ability to know the future or even treat it as a potential object of knowledge. That instability captures the sense in which the future remains for her perpetually open and the passage of time resolves openness into the fixity of fact.

KEY WORDS Self-reference; agency; temporal progression; thermodynamics; temporal asymmetry

Self-reference

Any attempt to describe the Universe as a totality inevitably involves self-reference. This isn't something that one often confronts in physics. In most day-to-day physics, one is modelling other systems: for example, cells, gases, organisms. One maintains a provisional separation of subject and object, or of investigator and system being investigated. Cosmology is the physics of totality so one might have expected self-reference should become explicit in that setting, but it is customary in cosmology to maintain the imaginative fiction that we are sitting outside of the Universe looking down.

This article was originally published with errors, which have now been corrected in the online version. Please see Correction (https://doi.org/10.1080/24740500.2024.2526935)

Ultimately, however, we are part of the Universe. We exist.[1] However we regiment the Universe, whatever regime we work in, if we aim for a theory of the Universe as a whole, self-reference is unavoidable. It's not going to matter in some contexts; it matters in others.

The people that have unavoidably encountered it are people that are (for example) trying to program an artificial general intelligence. They want to program a system with a bunch of general knowledge and the ability to model the world, and they are coming up against the fact that there are going to be problems of self-reference. The problems arise because some of what happens is stuff that the computer does. Let me give you a little simple example. Suppose you want to program a computer: it's just a simple little device that is going to be a grand overarching database or repository of information about everything. So we set about programming it with as much factual information as we can about the world: the laws of physics and all of the scientific knowledge we've amassed, facts we know about the monkeys of Costa Rica and the vast reaches of space.

The goal is to be able to put any question of physical fact to it and the answer will appear in the output channel. It is not, however, hard to find a question that it can't answer truthfully. Ask: '**Is the answer to this question that's about to be displayed in the output channel 'no'?**'. This might seem like a logical glitch. These kinds of problems are familiar to philosophers and certainly to computer scientists. The problem is that what the computer does in giving the answer interferes negatively with what the answer says. Computer scientists know that self-reference leads to paradox and we don't want to build the potential for paradox into an artificial general intelligence. So the first reaction is to try to suppress it or keep these questions from arising. This is more subtle to accomplish than you might think, since there's nothing wrong with the question that is being asked. It is a question about a perfectly mundane matter of physical fact. Any other computer could answer truthfully. The computer itself could answer truthfully at any other time. Another reaction is to acknowledge the difficulty and try to quarantine it. Again, this is more difficult than one might think. If the domain is connected, the word 'no' appearing on screen at some moment is going to affect the truth values of other propositions. Maybe a camera in the room records what appears on the screen; a person in the room looks at the screen and reports what they see to others; electricity flowing through circuits and static in the air will all be affected. Do we withhold truth values from propositions about all of these things as well? There is not in general a way of quarantining the effects of withholding truth value from one proposition on others in a connected network. It all depends on what other propositions are included and how they are connected. If propositions describe events in the world, the effects will propagate indefinitely.

However one deals with it, self-reference is usually thought of as a pathology. One wants to get as close to minimizing any disturbances that the computer's own activity has on the truth values of sentences in the field that it is representing. The reason that it arises is clear. By saying things, the computer is *doing* things and that keeps it from being able to stabilize the facts it is representing independently of how it represents them.

[1] If 'the Universe' is used explicitly in the all-inclusive sense to mean 'the totality of all that there is', then the fact that the Universe includes us follows from the fact that we exist. If 'the Universe' is used to mean 'all of material existence' then whether the Universe includes us (i.e., whether self-reference arises in any description of the Universe) will seem to depend on whether we are ourselves physical things. Philosophers have sometimes used 'the Universe' ambiguously, sometimes meaning 'all of existence' and sometimes 'all of material existence', treating physics as providing us with (at least provisionally) an explicit representation of the Universe. If one is a dualist, a slightly more roundabout way of showing that self-reference will arise is needed. All that is needed for what follows is that you are an embodied intelligence, that your actions are part of what happens, and that your thoughts make a difference to what happens. See the section: "Taking Self-Reference into Account".

Paradoxes of self-reference involve negative interference. The computer can't stabilize the fact that the answer is meant to describe (the word that appears on the screen) independently of giving the answer and the situation is set up so that no matter which answer the computer gives, what it does in giving the answer conflicts with what it says. Negative interference can arise in other ways: suppose, for example, I offer you a $10000 bet for $1 to predict the output of a simple deterministic switch that will go up or down depending on the input. The glitch is that you have to reveal your prediction to the device, and it is programmed to do the opposite of what is predicted.[2] Positive interference can arise too. A positive interference effect is a self-fulling prophecy. A man's insecurity about keeping his commitment-shy lover drives her away. A favorable prediction by an influential analyst drives stock prices higher.

Analysis

Any system that is acting in the domain it is representing is going to encounter interference. If the domain includes its own activity then some of what happens is stuff that it does; and the more connected the domain, the more difficult it will be to quarantine the effects. Interference impedes pure knowledge acquisition. If you are trying to understand the intrinsic structure of some domain, you want to keep your own activity outside the frame. We ignore it for good reason when we are doing physics, but it arises naturally and inevitably for a system that is part of the domain that it is representing. That is a feature, not a bug. And I'm going to suggest it helps to understand something that is otherwise very difficult to understand. It is an important piece of a very complicated puzzle in understanding our relationship to time.

Time

There are lots of routes in, but I will start with some remarks that Penrose made back in a beautiful paper from 1979: 'Singularities and Time-Asymmetry'. I will show how this simple logical point, set against the right physical backdrop, can help us understand a feature of our experience that he put particularly clearly. Penrose is talking about the prospects for explaining the various arrows of time in physical terms and he writes:

> "The arrow most difficult to comprehend is, ironically, that which is most immediate to our experiences, namely the feeling of relentless forward temporal progression, according to which potentialities seem to be transformed into actualities."[3]

I'll call this process of transformation of potentiality into actuality 'Becoming'. The term is used in other ways as well, but I'll use it here to refer strictly to this process of transformation of potentiality into actuality. Everyday common sense tends to universalize Becoming, that is, to think that the Universe as a whole is a single synchronous process of Becoming. As Penrose says, relativity theory rules that out. The

[2] This is the set up in the paradox of predictibility. See "An Essential Unpredictability in Human Behavior", by Michael Scriven, from the book *Scientific Psychology: Principles and Approaches*, edited by Benjamin B. Wolman and Ernest Nagel,1965. Also Ismael, (2019) "Determinism, Counterpredictive Devices, and the Impossibility of Laplacean Intelligences" in *The Monist*, Special Issue edited by Gordon Belot.

[3] Penrose, R. (1979) "Singularities and Time-Asymmetry". In: Hawking, S.W. and Israel, W., Eds., *General Relativity: An Einstein Centenary Survey*, Cambridge University Press, Cambridge, 581-638.

geometry of a relativistic space-time does not permit a global notion of now and cannot be properly understood as a single global process of Becoming. It does not however rule out the idea that temporal progression for an individual creature that is actively producing its future is properly understood as a transition from potentiality to actuality. It just means that there is going to be some asynchrony between the world-lines of creatures at a distance from one another.

I'm going to suggest that when we see how interference plays out against the background of a thermodynamic gradient, we get an interpretation of this most difficult of arrows.

Statistical Mechanics

The foundations of statistical mechanics are a heavily contested area. The last couple of decades has seen an enormous amount of attention. I'm going to assume the neo-Boltzmannian account of those foundations proposed in Albert (2000), though one can substitute one's own favored account.[4] It shouldn't make a difference to what I say. Albert's account has three first principles:

The classical dynamical laws.

The Statistical Postulate (SP):	assigns a probability to a system's being in a given microstate, given its macrostate. It assumes the standard Lebesgue measure over phase space and asserts that if A is the macro-state of a system, then the probability that the system's *micro*-condition is located in any particular subregion of the volume associated with A is proportional to the volume of that subregion.[5]
The Past Hypothesis (PH):	assumes that the Universe was in a state of very low entropy in the distant past.

[4] There has been a mass of literature on the foundations of statistical mechanics since Albert's book: some technical, some philosophical. There is widespread agreement that the classical laws together with some form of a probability distribution and some sort of hypothesis about the past will underwrite thermodynamic generalizations, but dispute about the form the probability distribution and the Past Hypothesis should take. See, for example, Wallace, D., "The Logic of The Past Hypothesis", http://philsci-archive.pitt.edu/8894/, Earman, J. (2006). The 'Past Hypothesis': Not even false. *Studies in the History and Philosophy of Modern Physics* 37, 399–430. Myrvold, W. C. *Beyond Chance and Credence*, Oxford University Press, 2021. Winsberg, E., "Laws and Statistical Mechanics", *Philosophy of Science* 71 (5):707-718 (2004). Frigg, R. (2008). "Typicality and the approach to equilibrium in Boltzmannian statistical mechanics". http://philsci-archive.pitt.edu, and the wide-ranging commentary devoted specifically to Albert's program in *The Probability Map of the Universe: Essays on David Albert's Time and Chance,* Edited by Barry Loewer, Brad Weslake, Eric Winsberg, Harvard University Press, forthcoming. Another source of dispute concerns the fact that Albert presents the account at the global level, but there are (good) reasons for thinking it might be best told at the level of local adiabatically isolated subsystems of the world: those are the systems to which thermodynamics is applied and the probabilities have a clear statistical interpretation when applied to local subsystems. See Reichenbach's *The Direction of Time* (Dover, New York, 1956), Rovelli, "Back to Reichenbach", http://philsci-archive.pitt.edu/20148/, Fernandes, http://philsci-archive.pitt.edu/20955/. Leeds, Stephen. 2003. Foundations of Statistical Mechanics—Two Approaches. *Philosophy of Science* 70(1): 126–144. Albert is building on a tradition that goes back to Boltzmann, and much of which was present in There is also an alternative Gibbsian tradition. E.T., Jaynes, Gibbs vs Boltzmann Entropies, *American Journal of Physics* 33, 391 (1965); https://doi.org/10.1119/1.1971557, Gibbs and Boltzmann Entropy in Classical and Quantum Mechanics Sheldon Goldstein, Joel L. Lebowitz, Roderich Tumulka, and Nino Zangh, June 2, 2019, https://arxiv.org/pdf/1903.11870.pdf

[5] Phase space is a continuous space containing six dimensions for every particle in the system, one for every dimension of position and momentum. The Boltzmannian entropy of a system is proportional to the log of the measure of microstates in phase space compatible with its macrostate, using the standard Lebesgue measure.

The classical dynamical laws are the familiar time-symmetric laws of Newtonian physics. They are presumed in every classical account of thermodynamics. The Statistical Postulate is a strict addition to the laws. It doesn't follow from them although there are many attempts to motivate it dynamically; it is the distribution that matches observed statistics. The Past Hypothesis is a contingent hypothesis about the early history of the world. It says that the Universe was in the distant past in a state of low enough entropy to make thermodynamic generalizations applicable for the roughly 15 billion years we think these generalizations have held.

The three postulates work together as follows. The Newtonian laws delimit the space of physically possible worlds. The Past Hypothesis eliminates worlds that don't start in a low entropy state. If you take the initial microstates of those worlds and apply the Newtonian dynamical laws, you will see that some of the remaining worlds will be on entropy-increasing trajectories, and some will be on entropy-decreasing trajectories. The statistical postulate provides a probability distribution over remaining worlds that overwhelmingly favors those on an entropy-increasing trajectory. The result is that the most probable history of the Universe is one in which entropy rises.

While there are disputes about elements of Albert's account, what I will say should float largely free of disputed details. What's really crucial for our purposes (and what any adequate account of the foundations of statistical mechanics must agree on) is that the Past Hypothesis and Statistical Postulate jointly make the macroscopic environment rich with information about its macropast. That sets the stage for the emergence of creatures that use information to guide behavior. The conditions that are needed to support the emergence of complexity and life are not well understood. We know that they are satisfied in the biosphere and are special enough that we have not discovered other planets where they seem to have emerged. Given those conditions, given competition for resources and mechanisms for heritable selection, and given also the mix of regularity and randomness that we have in our environment, one can see how evolution would reward increasingly sophisticated ways of using information. The tangible signs of predator and prey will be available to any creature moving through a forest or field; such a creature doesn't need to actually represent time. But a creature that stores large bodies of information organized temporally, that can extract predictive regularities and use them to avert, promote, facilitate or defray events, a creature that can plan and strategize with an eye to specific ends at some temporal distance, will have an advantage.[6] That's the kind of thing you are.

What we are interested in here is what the lived experience of a creature like this is; what the world would look like from its point of view, over time. Until now, I've been using the notion of agent in a small-a sense, to refer to anything that acts in the small-a sense (i.e., anything whose behaviors are part of what happens). Any material object from a rock rolling down a hill to a snowflake melting in the sun is an agent in this sense, because some of what happens is stuff that it does. From here on, I'm going to capitalize the A's using 'Agent' and 'Act' in the sophisticated and developed form that it takes in you and me. So I'm going to leapfrog over agents in the medium a senses—steps along the scale of increasing cognitive sophistication, from dogs

[6] Microscopic knowledge will not typically help to reliably increase our macroscopic predictability. Out in the wild, microscopic variables are too sensitive and fragile to too many factors to be stabilized effectively as a source of prediction.

moving around their neighborhood with their nose to the ground, ants leaving pheromone trails to mark paths to the location of food, and squirrels burying nuts for the winter—right to systems like you and me, systems that explicitly represent time and choose behaviors with an eye to the traces they will leave.

All of the systems along this scale are making use of the record-bearing properties of their environment, and it is the thermodynamic gradient that creates the conditions for their existence as agents in these intermediate senses. We don't, however, know a lot about the ways other systems manage temporal information.[7] We know a lot more about how we manage temporal information, at least at the level of conscious processing, so it is easiest to start with this fully articulated framework and then see shadows of this structure in creatures that are doing less explicit representation.

We can bracket the most difficult questions about human mentality. We model the Agent as a physical system getting information through perceptual channels and controlling the voluntary movements of its body. The natural thought that a lot of people have had to try to ground phenomenological time in the arrow of entropy is that we can see that entropy creates an asymmetry in the amount of information that we have respectively about the past and future. On Albert's treatment we represent the information that is available to an agent with access to the macroscopic state of its environment.[8] We assume an idealized Agent in the setting of thermodynamic gradient. She has access to the current surveyable macroscopic state of her environment. Assuming the classical laws and low entropy past, the information about the future that is available to her is derived by taking the currently surveyable macrostate, applying the Statistical Postulate to derive a probability distribution, and evolving that distribution forward using the dynamical laws. The distribution in question is the one that is uniform with respect to the standard measure on phase space over all of the possible microscopic *histories* of the world which are compatible with that macrostate. That distribution will tell the agent generically to expect thermodynamic behavior. She can use it together with other bits of specific knowledge to make predictions about how particular systems will behave under controlled conditions. That means that she will know to expect cream to disperse in coffee, ice to melt in warm water, and people to age. It means that she will know *not* to expect footprints to spontaneously form on an undisturbed beach, ice cubes to spontaneously form in warm water, and so on. But she won't be able to predict, for example, what path a skier will take across a particular patch of snow in the next ten minutes or whether some animal will find its final resting place in a particular patch of earth in the next decade.

To get information available to her about the past, we evolve that same distribution backwards, conditionalizing this time on a low entropy state in the distant past. And it's easy to tell from tracks in the snow which path the skier took and that an animal was buried in the place where you find bones a decade later. In general, conditionalizing on the low entropy past (or whatever plays the role of the low entropy past in one's

[7] This is an active, but nascent area of investigation. See Hoerl, Christoph & Teresa McCormack (2018) "Animal Minds In Time: The question of episodic memory" in Kristin Andrews & Jacob Beck (eds.) *The Routledge Handbook of Philosophy of Animal Minds* Routledge pp. 56-64, Templer, Victoria L. & Robert R. Hampton (2013) "Episodic Memory in Nonhuman Animals" *Current Biology* 23 R801-R806, Tramacere, A., Allen, C. Temporal binding: digging into animal minds through time perception. *Synthese* 200, 1 (2022). https://doi.org/10.1007/s11229-022-03456-w. There are obvious difficulties of studying animal consciousness, because our knowledge is indirect and there is a bewildering variety from elephants and crows to gophers and fish.

[8] See *Time and Chance*, Chapter 6 and *After Physics*, "The Difference Between the Past and Future".

preferred version of the Past Hypothesis) is going to make the current macrostates of semi-ordered adiabatically isolated systems in the environment carry traces of earlier events, specific information about their pasts of the kind we don't generally have about the future. Footprints, fossils, melting ice cubes and decaying bones will be remnants or traces of an even more ordered state in the past, whose details can be projected backwards from the current one.[9]

A view that wanted to ground the sense that the future is open on the epistemic arrow would say the transformation of possibility into actuality is only at the level of epistemic possibility. On such a view, the sense of openness comes from the fact that we *know less about* the future. Penrose considers and (rightly I believe) rejects this. He writes: 'the direction of psychological time … is not just a question of the past being (apparently) more certainly knowable than the future … It is not the ease in inferring the past that is relevant here, but the feeling that the past is unchangeable. … it is not the difficulty that we might have in guessing … the future that concerns us, but the feeling that we can affect [it] … ' (p. 594-596)[10]

The kind of unfixity in question is something more than the merely epistemic sense and you can see this by thinking of the difference between time and space. When you cast your eyes across a landscape, you think of it as a fixed object that comes into view in stages. Most of us don't think of the future that way. We think of future as *coming into being* as it is experienced. If one is a cosmologist or a physicist who works on the large-scale structure of spacetime, one may have a disengaged view of time, but for people actively engaged in the process of living, this is *the* most basic pre-theoretic difference between time and space. It is fundamental to the experience of an embedded Agent.

Should we say so much the worse for these squishy intuitions? I don't think so. It is clear that there is something that we don't understand very well. If we are just doing physics, for many purposes, it won't matter. If we are looking to physics to help us understand ourselves and our place in the Universe, it is worth trying to understand this better. That sense of a universe coming into being as it is experienced (rather than just coming into *view*) might be telling us something important.

Taking self-reference into account

I am going to suggest that the problem with the epistemic reading is that it doesn't take into account that we are a *part* of the Universe and our knowledge-gathering activity is itself connected in the world that it represents. Instead of suppressing interference and treating ourselves as observers outside the Universe, let's see what happens if we take it into account without presupposing anything more than the ingredients above: viz. the classical laws, low entropy past and Statistical Postulate.

Go back to the computer answering factual questions about the domain in which it is situated. The computer can't truthfully answer the question 'is the word that is going

[9] Any adequate account of the foundations of statistical mechanics is going to have to deliver this result; it is going to have to support the generic expectation of thermodynamic behavior to the future and underwrite inferences from records to the past. Note that these aren't meant to describe self-conscious inferential practices of agent, but rather to make explicit the physics that underwrites their typically unreflective reliance on records.

[10] After saying that the epistemic account is unsatisfactory, Penrose leaves off this train of thought and moves to a different way of trying to understand the intuitive nature of the difference between past and future, in terms of thinking of the past as providing reasons (or, as he says 'causes') of future events. "My attitude has been that a low entropy at one time may be regarded as providing a 'reason' for precise correlations.

to appear on the monitor "no"?'[11] or the predictor that can't predict the behavior of a simple algorithm. The problem in those cases wasn't ignorance or a lack of information. It was that there is negative interference between what it is doing and what it is saying. The potential for interference arises for any system that is acting in the domain that it is representing. If you are trying to build a system that is just going to be a passive repository of knowledge, this is a logical glitch that you want to suppress. But if you want to embody the intelligence so that it can *use* its knowledge to *guide* its own behavior – the whole point of intelligence from a natural perspective - you realize that it is not a logical glitch. A system like that *knows* that its own knowledge has interaction effects in the domain that it is representing, and instead of suppressing or nullifying those effects, it anticipates and exploits them.[12]

This is a good place to pause to be explicit about a couple of things. It is hard to recognize self-reference looking at a physical description of the world for a lot of reasons that anybody who has spent time thinking about mind-body identity statements knows. You don't look at the activation of some pattern of neurons in an image of your brain and think '*that* event is *this* thought'. But your mental life is implicitly included in an objective description of what is happening in the world, in the form of a functional description of processes mediating the perception to action pathways in your body. You refer to mental events naturally with other people when you describe or predict what they are thinking and doing, and you do it for yourself when you vocalize your thoughts; 'here's what I'm thinking'. The reason in part that I told the evolutionary story was to get you into the frame of mind of viewing what is going on in your brain as just an ordinary part of physics, described at the functional level that calls attention what it was selected to do. The underlying processes in your brain were arranged to support intelligent processes in the way that the hardware in a computer is designed to support information processing. This much is widely agreed nowadays. The mind-body problem in its modern form is focused on the status of the specifically phenomenal properties. The reason that the phenomenal properties are problematic—that is, the reason that it is argued they aren't captured by the functional description of states that mediate perception and action, the reason that phenomenal consciousness is supposed to be something extra, something over and above what is captured by the functional description—is (the arguments go) you can fix the behavior and fix all of the physical/functional duplicates and not fix whether a system is phenomenally conscious (or what its conscious life is like).[13] For our purposes here, one can take whatever view

[11] The indexical here is for ease of exposition: we can replace it with an explicit specification of location and time. Notice, however, that the indexical is crucial to *recognizing* that the answer can't be given truthfully.

[12] A brief recap might be useful: self-reference arises naturally and inevitably for a system that is part of the domain that it is representing. An agent using information to guide behavior knows that some of what happens in the world it is representing is stuff that it does. It knows that its own knowledge-gathering activity is connected in the domain by way of its effect on behavior. And it knows that its behavior produces records. And it exploits the interaction effects of its beliefs with an eye to the records they will leave. If one can't stabilize the answer to a question independently of *giving* the answer – i.e., if whether the answer that is *true* is going to depend on the answer that one *gives* - that creates interference. Interference can be immediate or attenuated. It can be negative or positive. Both examples above were examples of negative interference. But it is easy to change the examples to get positive interference. Just as there's no way to be right in the question displayed in the original example, there's no way to be wrong if we ask the computer 'is the word that is about to appear on the monitor yes?'. Or better 'what word is about to appear on the monitor?'. It doesn't matter how you answer – yes or no – what you say will be true. The attenuated effects are going to depend on how the answer delivered is connected in the wider domain.

[13] Chalmers, D., *The Conscious Mind: In Search of a Fundamental Theory*. New York: Oxford University Press, 1996.

one likes about the status of phenomenal consciousness. When we are talking about thoughts/perceptions etc., we are talking about the kinds of representational states that mediate perception and action in information-gathering and utilizing systems. These happen to be conscious in us, but we don't have to have any commitment to what their being conscious amounts to.

Also, when we are representing the world in day-to-day life, we are not normally representing our own occurrent internal processes. We do it sometimes, but usually our focus is outside; we're looking at what other people are doing and our own bodies. That's also why I put things in terms of recognizing that our own internal processes are *connected in the domain they are representing* by way of their effects on the motions of bodies. Even if one is just representing macroscopic events in the publicly available landscape, if one is thinking about what to do, one is recognizing that one's own thoughts are connected, and one is exploiting the interaction effects. So implicitly or explicitly, you are locating your own occurrent thoughts and decisions in the dynamics of the domain you are representing. If we were to make the logic of your representation fully explicit, we would find self-reference.

Still, you might think: 'Look, I don't get it. You can describe everything just in terms of physics; one event and then another. No trouble, no paradoxes, no self-reference. The same is true with the computer. We can describe the computer at the level of hardware. Electricity flowing through transistors, pixilations on screens.' Of course, that is correct. There is no threat of contradiction or constraint at the level of physics. The problem is entirely at the representational level. If one is dealing with a system that is in the business of representing the world—whatever you think that amounts to, and whether the system in question is a computer or in a human mind—and the system falls under its own scope, there is going to be self-representation at the semantic level: i.e., in the *content* of its representation. Most of the examples of self-reference in logic are static; people are thinking about logical relations among eternal propositions. But here I'm interested in physics, and in representation as an *embodied activity*. We are assuming the Universe is a field of events (though this should transfer to other regimes) and self-reference is arising because some of what the system is representing is also what the system is doing. That will create interference between the two levels—the representational and the physical. And that will allow us to rig things (as we did with the self-refuting question that we asked it) so that the interference was negative. This isn't going to affect what the system can do at the *physical* level. It does place constraints on what the system can truthfully represent and it does lead to an essential incompleteness in the worldview of an embedded Agent.[14]

So, let me repeat this now in application to you. You are an embodied intelligence and you are in the business of representing the world. You are acting in the domain you are representing and are negotiating two levels: the physical level and the semantic level. There's no avoiding that there's going to be interference between them. Philosophers often picture representation in a detached way: as an abstract relation between eternal propositions and the world. But it is not; representation is an embodied activity. It affects what happens. And everything that happens registers at the representational level. We are seeing how the interference between those two levels plays out in the mind of the creature representing. And let me emphasize that this isn't just

[14] There will be an essential incompleteness on pain of contradiction. That is the familiar Gödelian conclusion; any logical system powerful enough to permit self-representation is either inconsistent or complete.

psychology—it's not just about understanding what time is for us—intelligent activity is part of physics; it moves earth and builds cities. Intelligent activity hacks natural processes and becomes a dominant dynamical force, one that has altered the very ecology of our planet.

The practical arrow

Picking up where we left off: let's go back to the epistemic arrow of time and see whether we can illuminate the practical arrow. By the practical arrow, I mean the sense that we can affect the future but not the past, that is, that what you do in the here and now makes a difference to the future but not the past. It turns out that the practical arrow is just the flip side of the epistemic arrow. To see this, we take the same ingredients: the current macrostate, the canonical microprobability distribution, and the low entropy past. We consider an agent looking into a future about which she knows only the kinds of general things that can be derived from those ingredients and we ask: what should she expect the future to be like, conditional on some macroevent of the kind that we ordinarily take to be under voluntary control. So we ask what should she expect if she, for example, walks across a sandy beach, drops an ice cube into a glass of water, or scrapes her knee? To answer we take the probability distribution derived from applying the probability postulate to (what she knows of) the current surveyable macrostate of the world and evolve it forward. We find that she can form some relatively reliable expectations. These things will leave an imprint: scars on the world/records of their occurrence. If she walks across a sandy beach, she will leave footprints. If she drops an ice cube into a glass of warm water, she can expect it to be half-melted and cooling the water in the next few minutes. What happens if she digs a ditch, or builds a house, that is, does some work to create an ordered state in the environment? Forged in more durable materials the traces will take longer to decay, but they aren't different in kind.[15] An intelligent Agent will exploit this strategically, choosing her actions with an eye to their expected results. That's what intelligent agency is: acting with foresight and letting the expected results guide your actions.

How do things look to an Agent like this over time as she makes choices and acts? Exactly as you'd expect: the future is a field of open possibility transformed into the thin hard line of fact by decisions. Notice that in deriving the practical arrow, I made no appeal to counterfactuals, only hypotheticals and that I drew inferences that depend only on the thermodynamic gradient. I assumed no causal arrow. The only thing that I've added to the ingredients above is the observation that embodied knowledge is not detached but engaged. It has interference effects and exploits them. It exploits them in the most direct way by anticipating the immediate results of its actions and in an indirect way in the wider landscape, using its actions to forestall, avoid and deter things it doesn't like and promote, encourage, arrange and facilitate ones that it does. The interference effects are centered on the Agent's body, so they

[15] This point is indifferent to disputes in the foundations of statistical mechanics. Whatever one's preferred account of the foundations of statistical mechanics, that account will have to provide an understanding of the physics of traces. The practical asymmetry derives from the physics of traces; effects are expected records of occurrences. See Rovelli, C. "Memory and Entropy", https://arxiv.org/abs/2003.06687 for a similar account of the relationship between records and agency. The paper also provides a very general account of the nature of traces from a statistical mechanical point of view.

are localized in space. They run in the future direction for the same reason that records carry information about the past. The interference effects of one's actions are precisely records of their occurrence.

I want to look now at what one's attitude to the future is going to be. It is stronger than simply that one's own decisions are going to provide one with information about the future that one couldn't derive otherwise. It is that *the potential for negative interference is going to keep one from being able to stabilize beliefs about the future in advance*. Any prediction one makes about the future one can act to forestall, deter, or derail, and that makes it impossible to stabilize the future as an *object to be known*. It does this in the same way and for the same reason that the computer can't stabilize the word that will appear on its screen as an object to be known. The crucial thing here is that any prediction made in advance can be used as fuel to do something else. The same phenomenon arises in predictions of the stock market and political elections: public predictions guide the behavior of the system itself. And there are people in the market that are trying to outsmart it and that is going to create negative feedback. Any prediction can itself be used as fuel for overturning it. That's the whole point of knowledge. Knowledge is not passive. It is embodied and efficacious; you are always anticipating the future, acting to deter, deflect and counteract what you don't like, and to promote, encourage and facilitate what you do.

Of course, how strong the interference is depends on how one's knowledge is connected in the world. My knowledge is directly connected to my own voluntary behavior. Somebody tries to predict what I will do and I can overturn it. Control over the wider environment is uncertain and derivative. I'm not going to be able to deter the next eclipse from happening but who knows what effect a decision of mine will make to the future of the world. The ripples of our actions propagate indefinitely, outward and into the future; the farther away they get, the more difficult to predict. You know for yourself when you are making your own decisions, up until the last second, you can overturn any earlier expectations. The next thought can overwrite the first thought, and the one after that. You go back and forth all night and decide to take the conservative route on some big decision. But then your partner says 'Oh, I knew you'd do that', and in a burst of anger you do the opposite. It is the potential for negative feedback that animates the sense that it is not over until the facts are behind you.

This is the source of the phenomenology so vividly described by William James:

> The great point is that the possibilities are really *here*. Whether it be we who solve them, or [God] working through us, at those soul-trying moments when fate's scales seem to quiver, … is of small account, so long as we admit that the issue is decided nowhere else than *here* and *now*. *That* is what gives the palpitating reality to our moral life.

Putting all of this in terms of interference effects sounds a little perverse. The way you would put it from an embedded perspective is that my decisions make a difference to the future and so the future is not fixed (already). What I'm showing here is how to recover that way of speaking—see why it arises and what it signals—in terms that presuppose only statistical mechanics and interference effects.[16]

[16] There is a metaphysical picture of the world that frames much of common sense. We imagine that we live in a universe that is unfolding in time, that causal relations run from past to future and that is why we regard the past as fixed and the future as open. This metaphysical picture has not found a hospitable home in physics. The actual metaphysics of post-relativistic classical physics portrays the Universe as a four-dimensional

If I'm sitting outside a system and simply trying to predict its behavior, it's a different matter. It might be hard to predict what the system is going to do because it is complex or because there is missing information. But what is going on here is a different kind of thing; it's interference, it's the inability to stabilize the system independently of our predictions about what it is going to do. That's the point that was made by the self-defeating answers of the computer and the self-defeating attempt to predict. In those cases, the problem wasn't ignorance. It wasn't missing information. It was interference.[17] The potential for negative interference prevents you from stabilizing the future as a potential object of knowledge.

It sounds odd to our everyday ways of speaking, but from a statistical mechanical point of view what is happening here is that your actions will leave future records, and you are choosing them with an eye to the records they will leave. In principle, anything in the choice-dependent region downstream, that is, anything in your future light-cone can bear the marks of present action. In practice, the interference effects are strongest (and easiest to control) in the immediate aftermath of action and you build on them. You are always thinking about what to do, acting on your decisions, living in the wake of your decisions, plotting your next move. When you are thinking about your future, you are actively engaged in trying to shape it: trying to avoid, forestall, avert … Of course, it's true that nothing you are going to do is going to make a difference to whether the sun comes up tomorrow or whether a pregnant panda in the mountains of southwest China has a boy or girl. But your day-to-day life is centered on the regions of the world in which the interference effects of your knowledge are strongest.

All of this is true not because you are a magical thing that comes from outside of nature and interrupts the orderly flow of events, but because you are *part* of nature; your own thoughts and decisions are part of that law-governed flow. They interrupt it from the inside. Your ability to act on your knowledge is going to interfere with your ability to know things in advance.

All agents are relying on the thermodynamic gradient to some extent: using information contained in records to guide behavior whose rationale depends on its effects. That's simple to see if you are just thinking of animals responding to the immediate signs of predators and prey. Sophisticated cognitive Agents harness that, self-consciously choosing actions with an eye to the records they will create.

So, interference effects arise because our information-gathering activity is connected in the world. It is embodied and guides action in the very domain we are getting information about. If agency is acting in the domain you are representing, *intelligent* agency is realizing that you are acting in the domain you are representing and *exploiting interaction* effects to create a desirable future. The effects are explicitly perspectival. The strength of interference between your choices and other events will depend on how your choices are connected in the world. They taper off quickly in space and they are asymmetric in time.

From a physical perspective, minds are just part of nature. You and I making choices and decisions about what to do is just part of the fabric of what happens. From our own perspective, we are making decisions about how things go and precisely

manifold of events with no global conception of time. There is no unfolding and the laws don't incorporate any intrinsic direction of determination. That means that we need a new way of understanding it, and that new way of understanding it is bound to seem alien and unintuitive.

[17] This is for the same reason that the computer can't truthfully answer questions about what will appear on its monitor without giving an answer. No amount of information is going to overcome the problem.

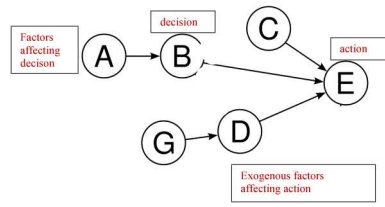

FIGURE 1: Causal Diagram with Decision

because our current thoughts and experiences are an essential and integral part of the fabric of nature, they are going to interfere strongly with what lies ahead.

A brief look at the logic of the situation here from your own perspective; you are trying to predict what you will do. Your decision is probabilistically connected to your own action in a way that is not screened off by other variables, so any prediction about your action has to *pass through* your decision. Here's what the situation looks like in causal terms. B is your decision, E is your action. C and D are other factors that influence your action in a way that bypasses decision (for example, mood, hunger, unconscious biases, strength of will …). A collects all of the facts about your prior history whose influence on action is by way of decision. Whatever probabilistic contribution your decision makes to how you will act, you will need to determine by predicting your decision. The strength of the dependence can vary and will tell you how strong the interference is. Sometimes what I end up doing doesn't depend strongly on what I decide (people can decide to stop gambling, or to leave an errant lover and then never do so). Sometimes the action is almost wholly dependent on decision. If you try to predict your decision you'll run into interference of the most direct kind; any information you might try to gather about what you will decide is irrelevant because it is screened off by your decision. You can ignore any evidence of that kind, because you can't go wrong epistemically here. 'I will decide to p at time t' is equivalent (at time t) to 'I decide to p', and whatever you put in for p will be self-fulfilling.[18]

The upshot is that you can't predict what you will do without predicting your decision and the decision will screen off any information you might use in predicting it. There is nothing to do but simply decide. The only way out, as they say, is through.

[18] On the logic of this, see my, *How Physics Makes Us Free* (New York, NY: Oxford University Press, 2016) and "Decision and the open future," in Adrian Bardon, ed., *The Future of the Philosophy of Time* (New York, NY: Routledge, 2012), 149–168. Routledge. James Joyce, "Levi on causal decision theory and the possibility of predicting one's own actions," *Philosophical Studies* 110:1 (2002), 69–102. James Joyce, "Are Newcomb problems really decisions?," *Synthese*, 156:3 (2007), 537– 562. Huw Price, (2012)., "Causation, chance, and the rational significance of supernatural evidence," *Philosophical Review* 121:4 (2012), 483–538. David Velleman, "Epistemic freedom," *Pacific Philosophical Quarterly* 70 (1989), 73–97. Wlodek Rabinowicz, "Does practical deliberation crowd out self-prediction?," *Erkenntnis*, 57:1 (2002):91–122. Brian Skyrms, *The Dynamics of Rational Deliberation* (Cambridge, MA: Harvard University Press, 1990). Isaac Levi, *The Covenant of Reason: rationality and the commitments of thought* (Cambridge, UK: Cambridge University Press, 1997). Yang Liu, Yang and Huw Price, "Ramsey and Joyce on Deliberation and Prediction," *Synthese* https://doi.org/10.1007/s11229-018-01926-8 (2018). Yang Liu and Huw Price, "Heart of DARCness," *Australasian Journal of Philosophy* DOI: 10.1080/00048402.2018.1427119 (2019). Alison Fernandes (2016) Varieties of Epistemic Freedom, *Australasian Journal of Philosophy*, 94:4, 736-751, DOI: 10.1080/00048402.2015.1116015. Anscombe famously asked what the difference is between prediction and intention, e.g., between 'I'm going to be sick' and 'I'm going to go for a walk'? In epistemic terms, the signature of an intention is that it is a self-affirming act that interferes positively with what it predicts: Elizabeth Anscombe, *Intention*, (Cambridge, MA: Harvard University Press, 1957).

Here you can see the signature of teleology. If decisions are made strategically, that is, by looking ahead and seeing the effects of potential actions, in the context of decision, the decision depends on the expected outcome.

Time Again

Let's go back to examining the contours of this and see why it matters to getting the psychological arrow of time right. There are two ingredients: self-reference + the thermodynamic gradient. The point about self-reference is entirely time-symmetric. The asymmetry comes from the thermodynamic gradient. Whether you are explicitly representing your own knowledge-gathering and decision-making procedures or just representing the publicly observable movements of bodies, your knowledge is going to be connected in the world by way of its effects on the movements of your body. And the ripples produced by those movements (the records they create) are going to run into the future in a way that keeps you from stabilizing your beliefs in advance.

For an Agent like this looking into her future, there isn't a merely epistemic kind of unfixity, but a directly practical kind of unfixity. I can myself act to forestall, avert, derail any antecedent expectations about what will happen. The future can be (for me) no more stable than my decisions about what to do.

Fitting the pieces together

Why does this matter? It doesn't matter for the practice of physics. Most often in physics we are modelling other systems. In cosmology, we are modelling the Universe as a whole, but nothing you do in the here and now is going to realistically make a bit of difference to features of the Universe cosmologists study.

It does give us a better understanding of the psychological arrow of time, so it helps us slot in an important piece in a complicated structure that relates physical time to human time (and by doing it in a way that preserves the current understanding of the physics of time, solidifies the foundation). It also helps us understand ourselves and what time is *for us*. Our own personal view of the world is totally organized around that instability. The scale is much smaller, and explicitly centered on the choice-dependent region in which the interference effects are very strong. And it is not an accident of course, that that is where our attention is focused. The whole point of your intelligence from a naturalistic point of view—that complex loop of neurally supported computational activity that interrupts the perception-to-action pathways in your body—is to use knowledge to interfere.

Philosophical assessment

I've put all of this in the most neutral way possible and refrained from drawing philosophical conclusions except to present it as capturing the logic of the relationship between the internal view of the Universe and the view from the outside.[19] From the outside the Universe is a seamless 4-dimensional manifold; the manifold as a whole just *is*. Notions of past and future, fixity and openness

[19] Of course, there is no view from outside; I mean simply the view of the Universe that is invariant under transformations between internal perspectives.

all have to be understood from an embedded perspective. From the perspective of a decision-making agent inside the Universe, the Universe is rightly seen as coming into being as it is experienced and the future can't be any more stable than her own decisions about what to do next.

Let me caution against reading this in a too *thinly* epistemic way. There's temptation to think that the transformation of potentiality into actuality is purely epistemic.[20] At least in a deterministic setting, it might seem that it is only in our minds that there is a plurality of possibilities. It might seem, that is to say, that we are working with macro-structures and so forward prediction is limited, but from a physical point of view, in fact, once the initial conditions are in place there is only one possible future. One part of this is correct. From a detached perspective there is only what is. What will be will be. Relativity unequivocally rules out the idea that there is any global directed process of the Universe itself coming into being, so it rules out one way of trying to be a realist about the transformation of potentiality into actuality. But if we restrict attention to relativistically well-defined internal notions of past and future, and we consider the relationship between absolute past and absolute future at any point p along the world-line of a system in any of the kinds of local deterministic theories that we have examples of, even if we fix the complete contents of the back light cone at p, there are indefinitely many nomologically possible futures. That is because you need to take the total state of the Universe into account, and that will include events outside of its past light cone.[21] The reduction of potentialities, moreover, happens in real time. The set of events nomologically sufficient to determine an event e is not contained in the back light cone of any moment [absolutely] prior to e itself). That means that even if we gave an agent full microknowledge of the events in her past light cone, that wouldn't reduce the number of possibilities for her future. So, it is wrong that the idea of having many future possibilities is simply an artifact of the macroperspective.

The reason that our senses are sensitive to macroscopic structure, rather than microscopic, is that cognition works on records and traces. That is where the opportunities for using information to guide behavior are located. It's the record-bearing properties of the macroscopic environment (the fact that it bears the imprint of its macroscopic past and that the future will bear the imprint of the macroscopic present) that agency exploits. So whether it is described in microscopic or macroscopic terms, as we trace a path up your world-line, we see that your choices and actions are part of the process that transforms potentiality into actuality. Where you're looking into the future on the precipice of a difficult decision, the future hangs in the balance while you deliberate, and is brought to a close in the moment of decision.

Second, perhaps this doesn't need to be said, but the kind of interference we are talking about when we say that the ripples of our own actions interfere with our ability to know the future, isn't the kind of interference that arises when one object blocks another from view. Nor is it like what happens when you turn on the light to

[20] To be clear on what the 'transformation of potentiality into actuality' can mean. There is one world. The transformation of some particular event from potentiality to actuality has to be something that happens over time, so we have to be talking about how particular events look from different temporal perspectives. So if we are looking at some particular event – say who wins the NCAA tournament in 2025– *pro*spectively, at the beginning of the tournament, there are many possibilities; *retro*spectively, at the end, there is only one.

[21] See Ismael, "Rethinking Time and Determinism: what happens to determinism when you take relativity seriously, *Time and Science*, edited by Remy Lestienne and Paul Harris, World Scientific Publishing, forthcoming.

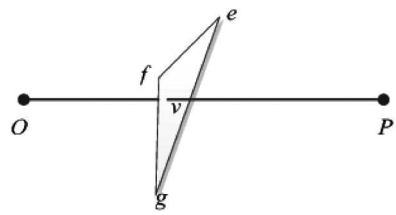

FIGURE 2: Object Hidden behind an Obstacle

check if your children are sleeping, or you try to measure something that is altered in the course of measurement. In those cases, there is a fact of the matter that is well-defined independently of your attempt to ascertain it and your attempts are simply unsuccessful. In this case there is not. It is more like you are trying to reach for a floating object in water and the motion of your arm pushes it out of reach.

To get a proper understanding of this we can't treat ourselves as we mostly do in physics, as sitting outside the Universe looking down. We have to absorb our own lives—including the thoughts and deliberations—into the fabric of the Universe as an essential and integral part of what happens. We can't infer much about our futures from our past (or maybe much of what we care about in day-to-day life) without passing through our own decisions, and that means that all of the unfixity, unsettledness, and contingency of our decisions, bleeds out into the world. Getting this piece right shifts other things that matter—maybe not so much to the actual practice of physics, but to places where the shoe pinches philosophically.

There's another way to think of this in a too-thinly epistemic way. One thinks of this as an explanation of why the Universe appears to an embedded Agent as coming into being as she experiences it, but that there is some alternative view—the way things *really are*—that reveals that this is *mere* appearance, that it is a mistake or error. I challenge that alternative view. Insofar as you exist—that is, insofar as you are part of the Universe and your activity is part of what happens – there is no well-defined conception of the Universe as a whole that is (for you) there *anyway*, *already*, or *independent* of your actions. The claim that there is no mind-independent world has been a credo just for idealists, but it is true for materialists as well. There is no mind-independent world insofar as the world *includes* minds and the effects of mental activity. That is the crucial step down from 'mere appearance' to actuality. Knowledge is embodied and engaged and makes a difference to what is the case. There can be no pure and complete conception of the way that world is independently of ourselves and our representational activity.

Conclusion

All of this is a reminder that we are part of the Universe and an important correction to the tendency to think about the Universe from a transcendent point of view. There's a strong tendency to remove ourselves from the Universe when we model it in physics and then to think that the differences between the way things are and the way they seem somehow reveals the way they seem to be illusory. But when it comes to understanding time as we know it, we need to abandon the fiction that we are outside the world. This is what it is to *be*, to be a part of the Universe, not a spectator, but one of the players on the field.

References

Albert, David (2015) 'The Difference Between the Past and Future', in *After Physics*: 31–70. Harvard University Press.

Albert, David (2000) 'The Asymmetries of Time and Intervention', in *Time and Chance*: 113–130. Harvard University Press.

Anscombe, Elizabeth (1957) *Intention*. Harvard University Press.

Chalmers, David (1996) *The Conscious Mind: In Search of a Fundamental Theory*. Oxford University Press.

Earman, John (2006) 'The "Past Hypothesis": Not even false', *Studies in the History and Philosophy of Modern Physics* **37**: 399–430. doi: 10.1016/j.shpsb.2006.03.002

Fernandes, Alison (2016) 'Varieties of Epistemic Freedom', *Australasian Journal of Philosophy* **94**: 736–751. doi:10.1080/00048402.2015.1116015.

Fernandes, Alison (2022) 'How to explain the direction of time', preprint. http://philsci-archive.pitt.edu/20955/.

Frigg, Roman (2009) 'Typicality and the approach to equilibrium in Boltzmannian statistical mechanics', *Philosophy of Science* **76**: 997–1008. http://philsci-archive.pitt.edu

Goldstein, Sheldon, Joel L Lebowitz, Roderich Tumulka, and Nino Zanghi (2020) 'Gibbs and Boltzmann Entropy in Classical and Quantum Mechanics' in Valia Allori, ed., *Statistical Mechanics and Scientific Explanation*: 519–581. World Scientific. https://arxiv.org/pdf/1903.11870.pdf

Hoerl, Christoph and Teresa McCormack (2018) 'Animal Minds in Time: The question of episodic memory' in Kristin Andrews and Jacob Beck, eds., *The Routledge Hand- book of Philosophy of Animal Minds*: 56–64. Routledge.

Ismael, Jenann (2012) 'Decision and the open future', in Adrian Bardon, ed., *The Future of the Philosophy of Time*: 149–168. Routledge.

Ismael Jenann (2016) *How Physics Makes Us Free*. Oxford University Press.

Ismael, Jenann (2019) 'Determinism, Counterpredictive Devices, and the Impossibility of Laplacean Intelligences', *Monist* **102**: 478–498. doi: 10.1093/monist/onz021

Ismael, Jenann (2023) 'Rethinking Time and Determinism: what happens to determinism when you take relativity seriously', in Remy Lestienne and Paul Harris, eds., *Time and Science*: 147–172. World Scientific Publishing.

Jaynes, E T (1965) 'Gibbs vs Boltzmann Entropies', *American Journal of Physics* **33**: 391–398. doi:10.1119/1.1971557

Joyce, James (2002) 'Levi on causal decision theory and the possibility of predicting one's own actions', Philo*sophical Studies* **110**: 69–102. doi: 10.1023/A:1019839429878

Joyce, James (2007) 'Are Newcomb problems really decisions?' *Synthese* **156**: 537–562. doi: 10.1007/s11229-006-9137-6

Leeds, Stephen (2003) 'Foundations of Statistical Mechanics—Two Approaches', *Philosophy of Science* **70**: 126–144. doi: 10.1086/367873

Levi, Isaac (1997) *The Covenant of Reason: rationality and the commitments of thought*. Cambridge University Press.

Liu, Yang and Huw Price (2018) 'Ramsey and Joyce on Deliberation and Prediction', *Synthese* **197**: 4365–4386. doi: 10.1007/s11229-018-01926-8.

Liu, Yang and Huw Price (2019) 'Heart of DARCness', *Australasian Journal of Philosophy* **97**: 136–150. doi: 10.1080/00048402.2018.1427119

Loewer, Barry, Brad Weslake and Eric Winsberg, eds., (2023) *The Probability Map of the Universe: Essays on David Albert's Time and Chance*. Harvard University Press.

Myrvold, Wayne C (2021) *Beyond Chance and Credence*. Oxford University Press.

Penrose, Roger (1979) 'Singularities and Time-Asymmetry', in Stephen Hawking and Werner Israel, eds., *General Relativity: An Einstein Centenary Survey*: 581–638. Cambridge University Press.

Price, Huw (2012) 'Causation, chance, and the rational significance of supernatural evidence', *Philosophical Review* **121**: 483–538. doi: 10.1215/00318108-1630912

Rabinowicz, Wlodek (2002) 'Does practical deliberation crowd out self-prediction?', Erkenntnis 57: 91–122. doi: 10.1023/a:1020106622032

Reichenbach, Hans (1956) *The Direction of Time*. Dover.

Rovelli, Carlo 'Back to Reichenbach', http://philsci-archive.pitt.edu/20148/

Rovelli, Carlo (2003) 'Memory and Entropy', https://arxiv.org/abs/2003.06687
Scriven, Michael (1965) 'An Essential Unpredictability in Human Behavior', in Benjamin Wolman and Ernest Nagel, eds., *Scientific Psychology: Principles and Approaches*: 411–425. Basic Books.
Skyrms, Brian (1990) *The Dynamics of Rational Deliberation*. Harvard University Press.
Templer, Victoria L and Robert R Hampton (2013) 'Episodic Memory in Nonhuman Animals', *Current Biology* **23**: R801–R806. doi: 10.1016/j.cub.2013.07.016
Tramacere, Antonella and Colin Allen (2022) 'Temporal binding: digging into animal minds through time perception', *Synthese* **200**: 1–24. doi:10.1007/s11229-022-03456-w.
Velleman, David (1989) 'Epistemic freedom', *Pacific Philosophical Quarterly* **70**: 73–97. doi: 10.1111/j.1468-0114.1989.tb00370.x
Wallace, David (2023) 'The Logic of The Past Hypothesis', in Barry Loewer, Brad Weslake and Eric B Winsberg, eds., *The Probability Map of the Universe: Essays on David Albert's Time and Chance*: 76–109. Harvard University Press.
Winsberg, Eric (2004) 'Laws and Statistical Mechanics', *Philosophy of Science* **71**: 707–718. doi:10.1086/425234