

Draft

The Past Hypothesis and the Knowledge Asymmetry

Presented at the Rutgers' *Foundations of Statistical Mechanics* Meeting, May 13-14 2009

Nick Huggett, University of Illinois at Chicago
huggett@uic.edu

May 26, 2009

1 The Knowledge Asymmetry and the Past Hypothesis

1.1 Memories and Macroscopic Traces

Why is our knowledge of the past so much more ‘expansive’ (to pick a suitably vague term) than our knowledge of the future, and what is the best way to capture the difference(s) (i.e., in what sense is knowledge of the past more ‘expansive’)? One could reasonably approach these questions by giving necessary conditions for different kinds of knowledge, and showing how some were satisfied by certain propositions about the past, and not by corresponding propositions about the future. I take it that such is the approach of Chapter 6 of *Time and Chance* (T&C). Here’s another such a proposal, similar to that of, but significantly different from T&C; my purpose in this section is to highlight the differences, by showing how this account fails.

*A proposition can only be known if it is made likely by taking a uniform probability over all states compatible with the current macrocondition of the universe, and with its initial and final macroconditions, given the laws of mechanics.*¹

Then, since the final state is (according to the best of our current knowledge) equilibrium in an ever expanding homogeneous space while the initial state (or some very early state) is one of considerable ‘order’, given by the ‘Past Hypothesis’ (PH)², far more statements about the past will satisfy the necessary condition than statements about the future. Thus it is the PH that both

¹I intend for this to be a proposal about all kinds of knowledge of other times identified in T&C, prediction/retrodiction and record-based, in the senses of.

²NB: the PH is not merely that the universe started with ‘low entropy’ (a claim that some say is ‘not even false’) but that it started in some, to be specified by the full PH, macrostate.

explains and articulates the difference between past and future knowledge. As I said, T&C offers a more sophisticated account than this one, but it is a first approximation. Before we see why the more sophisticated proposal is needed, a word on justification.

What arguments could be offered for such a statement being a necessary condition for knowledge of other times? One might claim that we consciously apply such a method, and indeed must. T&C does not claim that. It would be manifestly crazy to do so, because clearly we quite generally have knowledge of other times without knowing anything much about the nature of the early universe, and so presumably has humanity since before the earliest creation myths. Our ability to reconstruct the past is not evidence of our knowledge of the PH! So it is important to bear in mind that T&C does not attempt to offer a general account of the methods by which we obtain knowledge of other times, just a necessary condition on such knowledge. However, with the greatest respect for the book, I have had a hard time understanding how the argument given in T&C could establish its necessary condition unless it assumes that we do apply a method of this kind! Of course, if I'm right, you would have to be crazy to accept the argument! The point of this paper is to explain my concerns, so in a moment we will look at the argument given; first, back to the main goal of this section, showing that the condition above is not necessary, and how it is elaborated by T&C.

Suppose the 'ice pachinko' of T&C (83) is run with a single ice cube in a sealed room with a human observer watching, but making no record of events. The cube falls through the device and randomly ends up in one of the beakers, the leftmost one say. Eventually, after the ice cube has first melted and then evaporated, there will be no trace of which beaker it fell into left in the macrostate of the universe: the water molecules will be evenly distributed in the air of the room, whatever beaker the ice cube fell into. The macrocondition of the universe will be the same whatever happened, and the initial macrostate of the universe seems unlikely to determine the outcome (certainly it isn't supposed to do that kind of work in T&C), so the true proposition concerning which beaker the cube fell in does not satisfy the necessary condition. That is, as seems correct in this example, there is no macro-record of which is the beaker, and neither can it be retrodicted from the current macrocondition.

(Of course this is not to say that the *microstate* doesn't contain the necessary information: if the evolution is deterministic, then the earlier state can be recovered from the microstate. But the whole game here is to work with macrotraces – and later what can be known of one's own brain state. The fact that aspects of the microstate – DNA at the crime scene, microscopic mineral traces in a beaker – could be discovered – i.e., made manifest at the macro level using magnifying devices of various kinds – is irrelevant to the discussion. If we start including in the necessary condition any microfacts that could potentially be found by observation, then we are on a slippery slope to including the entire microstate – what couldn't be found in principle? – and clearly no interesting asymmetry between past and future will be found if we consider what can be inferred from the microstate using bi-deterministic laws. *It will be important to bear in mind this parenthetical discussion as we proceed: the fact that we could access this or that part of the microstate is not relevant to the project at hand.*)

But of course, that the ice cube fell in the leftmost beaker is perfectly knowable. It is already known by the observer, and can be known by anyone else simply by asking her. So the condition is not necessary, and it cannot be appealed to in an account of the knowledge asymmetry. (And neither does this example seem especially *recherche*. Surely macrorecords are being erased all the time, and surely the observer is relying on no strange procedures for obtain knowledge – just watching!)

1.2 ‘Direct Introspection’

As I said, the account just refuted is not that offered in T&C, for that calls for conditioning on the ‘presently surveyable condition’ of the world, which includes ‘whatever (perhaps macroscopic³ [sic]) features of the present condition of the *brain* of the observer in question may be accessible to her by means of direct introspection.’ (114) The immediate questions are what is accessible by ‘direct inspection’, and whether the pachinko observer’s knowledge satisfies the new condition so understood. I should note that I found the appeal to ‘direct introspection’ quite off-putting for quite a while; the phrase suggests a questionable armchair approach to psychology and philosophy of mind. So I rather ignored the idea. (Moreover, T&C places very little emphasis on the idea, giving little discussion of what role is played by this part of the presently surveyable condition: really just the sentence quoted above.) But I think I have come to understand better what the idea is, and how it might address the missing pachinko knowledge. I hope that David will tell me if I am on the right track – and if not, how he would address the putative counter-example?

‘Direct introspection’ may be intended to be more expansive, but in the case in hand all we need add to the surveyable condition are our memories – or rather, since we will have to worry about their veracity, our putative memories. (Here is a way in which I find ‘direct introspection’ a misleading way to put things: it sounds as if what matters is that we can come to have ‘second order’ knowledge of what our putative memories are, when all that actually matters is that we have the memories. So worries about our ability to know our psychology by introspection are besides the point; what goes into the surveyable state is whatever is implied about the physical state of the world by what we believe, for instance, not by what we can discover about what we believe!) So we include whatever follows for the state of the world (including its microstate) from the fact that the pachinko observer has a (putative) memory of the ice cube falling into the leftmost beaker.

By ‘putative’ we mean here that all that is included is the fact that the observer’s brain is in a state compatible with her truly remembering the beaker; of course being in that state is no guarantee, given the laws of physics, that the memory is veridical. That is, we must bear in mind that the background to this discussion is the reversibility objection: if we accept that future entropy increase is overwhelmingly likely, conditional only on the present surveyable state and the laws, then the symmetry of the laws implies that the conditional probability for past entropy *decrease* is equally likely, contrary to the second law.⁴ So in the present discussion we need to entertain the possibility that putative memories are not veridical, but instead the result of extremely improbably fluctuations in the microstate of the world. It’s not very likely, but with respect to a uniform probability of microstates, much more likely than the earlier, lower entropy states that they purport to represent. (A note on terminology: the unmodified term ‘memory’ means ‘veridical memory’, and I will only use modifiers in cases where I want to eliminate any possible misunderstanding.)

Of course if the presently surveyable state included whatever follows about the world from the fact that we presently *faithfully* remember some event, then the presently surveyable condition will contain facts about the past – that the event really happened. For example, that the pachinko observer presently faithfully remembers that the ice cube fell in the leftmost beaker, entails that it did. It wouldn’t be an oxymoron if the presently surveyable state included specifications of the state at other times, for it is not the ‘surveyable present state’. The problem is rather that the

³Clearly ‘microscopic’ is intended – cf. p.96.

⁴Obviously I am making the standard assumption of an even probability function over microstates: I will everywhere, often without further comment. Also, this objection is sometimes called the ‘course-graining’ objection, to distinguish it from the far more trivial observation that the laws are time-reversal symmetric.

reversibility objection casts doubt on whether the brain state underlying the putative memory was genuinely formed by a process likely to produce veridical memories. The presently surveyable state is supposed to be immune from those worries, so we include only those facts about the current state that follow from the observer apparently remembering the outcome of the experiment.⁵

So the necessary condition for knowledge is just as in the previous section, except that we further condition on whatever follows for the state (micro or macro) of the world from the observer's putative memory of the ice cube falling in the leftmost beaker. It might seem rather counterintuitive that just having an appropriate brain state could make a sufficient difference to make it highly likely that the ice cube fell that particular way. After all, there are lots of ways in which such belief formation might go wrong – optical illusions, tricky magicians, evil scientists, and so on. But we need to remember what game we are playing: we are just worrying about one thing that might undermine beliefs here, the reversibility objection. Then the idea is that conditioning on the PH plus the laws makes it very unlikely that our brain states are random fluctuations, leaving their faithfulness as the only other hypothesis on the table. (I know that some in the audience are sceptical of this claim, but I plan to assume it without further defense here, in order to touch on some other interesting issues.)

In addition, there is also the prospect that many of the processes that could lead to non-veridical memories should also fail the necessary condition. To give a fanciful example, in the Manchurian Candidate a group of GIs is 'brainwashed' to hold false memories about the heroic actions of Raymond Shaw. Applying the laws in reverse to those memories (consistent with the PH) thus underdetermines the past: did he really behave heroically, or was it all planted in their minds? But which occurred makes a difference to the present surveyable condition – beyond those memories – in the memories of the Soviet observers, in records kept by the Koreans in messages sent to conspirators in the US. So there is reason to hope that the condition will rule out many skeptical worries.⁶

In other words, the modification to the condition in T&C, to include 'direct introspection', does, if I have correctly fleshed out the idea, address the counterexample: the proposition that the cube fell in the leftmost beaker does satisfy the necessary condition. If I have the wrong end of the stick, I would like to hear how the counterexample is to be accommodated in the framework of T&C.

However, whether or not this account works or fits the intentions of T&C, the main point of this section has been to show that if we are going to tell a story about our knowledge of other times, we will need to include our beliefs about past experience, not just what we can learn about the current macrocondition of the world. Indeed, if we are to capture all our knowledge of other times we will want to include in the currently surveyable state beliefs beyond the kind of direct memories we are considering now. That is because our knowledge of other times includes such things as inductive generalizations from our experiences (not to mention testimonial knowledge)

⁵At this point, and on, I am extrapolating from T&C. However, the point just made is in line with the skeptical 'catastrophe' (166). Part of the argument is that the experiment evidence for Newton's laws *cannot* be inferred by assuming a uniform probability over states compatible with the currently surveyable condition of the world plus those laws. But if the currently surveyable condition included whatever follows from our veridical memories, then it would, either because we are Wren, Boyle or Newton, or because we did such experiments in school. That is, we can only survey that it 'seems as if' we recall performing such experiments, not that we did.

⁶Notice something counter-intuitive about the approach of giving necessary conditions in this way. (Not that I take that as much of an argument against the approach: intuition is answerable to science, not the other way around.) The goal is not to find a necessary condition for *true* beliefs about other times, but for those beliefs not contradicted by any macroscopic features of the world. The idea, that is, to give an account of what we *think* the world is like (corrected for misinformation about the macrostate), not what it is actually like. The cost of making our true beliefs likely, is that a certain kind of false ones are too.

and inferences drawn from them. Unless one thinks that the PH plus the dynamical laws will suffice to make all our inductive generalizations about any topic likely true (are there any limits to the power of the PH?), then they will need to be included in the surveyable state. Then we can hope that then such knowledge will satisfy the necessary condition, as including putative memories is supposed to bring veridical memories within the condition. (Let me add that I am now having a harder time being confident that this will work out.)

1.3 Records and the PH

Having established the crucial role played in the story by memories (both our need for them and something of their functioning), I want to lay out the argument of T&C for the necessary condition. In the following sections, I want to lay out a different approach, one which (I believe) undermines that of T&C.⁷

Knowledge of other times is divided into two kinds: first there is knowledge that can be had by prediction or retrodiction from knowledge of the current surveyable condition using the laws of mechanics assuming a uniform probability over microstates.⁸ Second, there is the kind of knowledge that can be had by ‘reading records’: comparing the current state of a system with its state at another time (the ‘ready state’), to infer (using the dynamical laws and uniform probability) the occurrence of an interaction in between the two times. (Note that the description is neutral between past and future ‘records’.) Clearly, reading a record in this way requires knowledge of the ready state.⁹ As we shall shortly see, this second kind of knowledge of other times is the source of the knowledge asymmetry.

Now, it is crucial to bear in mind something that I have had the hardest time seeing, despite its being a rather simple point. (However, my disagreement with T&C still lies close to this point.) As I pointed out at the start, these characterizations are not intended as portrayals of the processes by which we come by knowledge of the two kinds but simply as necessary conditions for us to have such knowledge: it’s not predictive/retrodictive knowledge unless it follows from the laws plus present state alone, and it’s not of the second kind unless it *could* be had by the reading of a record – *whether or not it is actually obtained by any of those methods*. Thus the argument is that satisfying this new condition requires satisfying the earlier one (modified to involve the currently surveyable state): hence it is a necessary condition. My point in the rest of the talk is going to be that these new conditions are not necessary, and so provide no reason for thinking the earlier one is either. Of course, we have just seen an argument that as a matter of fact it is satisfied by our memories and not future beliefs: the point of course is that explaining the knowledge asymmetry in this way also requires showing that all knowledge of other times *must* satisfy the condition. That is what I dispute is shown by the following considerations (though I am sympathetic to the idea).

Now we come to the source of the asymmetry in knowledge of the record reading type.

⁷One thing I think I learned from discussions after I presented this paper at Rutgers is that it is too strong to say that the passage quoted in this section was intended as a direct argument for the necessary condition. Instead, it is just intended to indicate a looser idea that the existence of past records rests on the PH. Fair enough: in that case my discussion can be seen as some ground work towards a more detailed account of just how the existence of records might be connected to the PH – or perhaps a demonstration that it cannot.

⁸Actually T&C has a really nice elaboration of this idea, which is very helpful, taking account of the fact that we have a second kind of knowledge of other times. Instead of just the current state, retrodiction infers from whatever is known about the future, while prediction infers from whatever is known about the past.

⁹Question: is everything a record? It will be if we just require that the present state and the state at some other time allow such an inference: something can always be inferred, if only that some change occurred – or not. But maybe the idea of record has an epistemic dimension: the earlier state has to be epistemically accessible.

The sort of inference one makes from a *recording* is not from one time to a second in its future or past (as in prediction/retrodiction), but rather from *two* times to a *third* which lies *in between them*. . . .

And the puzzle is about how it is that we ever manage to *come* by [information about the ready state of a record bearing system]. It can't be by means of retrodiction/prediction It must be because we have a *record* of that other condition! But how is it that the ready condition of this *second* device is established? And so on (obviously) as infinitum. There must be . . . something we can be in a position to *assume* about some other time . . . the mother (as it were) of all ready conditions. And this mother must be *prior in time* to everything of which we can potentially ever *have* a record, which is to say it can be nothing other than the initial macrocondition of the universe as a whole.

And so it turns out that *precisely* the thing that makes it the case that the second law of thermodynamics is (statistically) true . . . is *also* the thing that makes it the case that we can have epistemic access to the past which is not of a predictive/retrodictive sort [i.e., the PH]. (T&C 117-8)

That is, T&C gives a regress argument to conclude that knowledge of the record reading kind requires the existence of suitable boundary conditions.

Now, there is a temptation to take this passage as a description of the *process* by which we read records (at any rate, I have found that temptation irresistible until recently). After all, it asks how we 'manage' to learn the chain of ready states, and apparently what it is we 'assume' when we read records. But as I have repeatedly said, if the argument is read that way, it leads to madness – to the idea that any record reading *requires* us first literally to assume the PH, then read a series of records of subsequent 'daughter' ready states, until we get to the one in which we are interested.

T&C does not want the argument to be taken that way, but rather (I take it) as establishing the necessity of the PH, on the basis of a necessary condition for knowledge of the kind in question. Here's my best stab at how such an argument goes; I hope that I am close enough that my later arguments concerning it are tolerant to any modifications that their author would make! The argument runs as follows: we can't get information of the kind in question from the presently surveyable state by *any* means, unless we could *also* get it, in principle, by a process of record reading, something that would only be possible if we knew the ready state of the relevant system. But we *can't* know that by *any* process unless we could *also* in principle know it by record reading, which would only be possible if we knew an even earlier ready state. And so on back to the PH: getting the information in question from the presently surveyable state by a process of record reading would require knowledge of the PH, and so the existence of such knowledge – of the type that could be obtained by a hypothetical process of record reading – requires the truth of a PH. (I am fully aware that I have largely repeated the quoted argument, but I hope to have sufficiently emphasized the *hypothetical* character of the process of record reading versus whatever actual process is used – compared to the categorical language of T&C.)

In short, it's not that we actually obtain knowledge of the past in the way described, but that (some of) our knowledge of the past is such that we could in principle obtain it that way – that hypothetical process would require the PH, and so the existence of that type of knowledge is only possible because of the PH. (At least that is how I can best see to reconcile the argument of T&C with the claim that it does not commit us to having already to know the PH!) Now, of course we can define a kind of knowledge to be that which satisfies such a condition, but I think we should question whether ours is such; indeed that is the question of the next section.

Before that, note that the concept of a mother is intrinsically temporally asymmetric, so we might ask whether we have begged the question on the knowledge asymmetry in some way. Of course not, what information can be extracted from records depends on the nature of the ready state, or their collective mother. It's not as if don't have a future hypothesis, given the COBE data it is of a universe ever more diffuse and ever closer to equilibrium. And we can use that to read the future record of any non-equilibrium systems evolution to equilibrium! The point about the PH is that it is (one hopes) a far more informative ready state, because it constrains the possible initial microstates so much more. That is, our expansive knowledge comes from the possibility of reading records, itself made possible by the nature of the PH (compared to the FH, say).

A couple of comments about the argument. First, I believe that it can be simply modified to show not just that the reading of any record in the present surveyable state requires the PH, but also that the PH suffices. For reading a record requires knowledge of both the ready state and the later 'record state'. So we can ask, for each of the records, where in time the record state lies, for instance the record state that gives the ready state which is the daughter of the PH – i.e., that for which the PH is the ready state. If it is part of the presently surveyable state, then the PH suffice immediately. If not, then we will need another record (whose ready state is the PH) to tell us the record state, and the record state of that new record better lie in the present state, else we will need another record . . . and so on until we get a record state in the presently surveyable condition. And so on for all the other daughters and granddaughters of the PH.

Now, approaches to the knowledge asymmetry that appeal to the Second Law are familiar, and T&C grounds the Second Law on the PH, so one might suspect that we have here an appeal to entropy in understanding the knowledge asymmetry. (Especially since T&C contains another, similar regress argument – the one involving the ice pachinko – to argue that the PH grounds the Second Law.) But the argument is supposed to be a novel one, breaking from views which assume that records are inevitably entropy increasing. First, such an assumption is not made in the argument. (Though it does seem, as we noted above, that if the records that the PH permits are to tell us anything much, then the PH better be low entropy or something akin.) Second, in rejecting Reichenbach, T&C points out that a system that doesn't change – so one of constant entropy – carries a record of nothing happening!

So things are set up for the final sections now. First, we saw that the kind of account of the knowledge asymmetry developed in T&C must include our (putative) memories in the 'currently surveyable state', if there is any chance of recovering our knowledge of the past. In short, there can readily be things we know of the past that don't leave any traces in the present macrostate. Now we have seen how T&C characterizes knowledge of other times that is not of the prediction/retrodiction sort – not in terms of the actual processes by which we know the past, rather by reference to a (potentially) hypothetical process of record reading. But how do we actually know things about the past? In particular, how should we understand memory in this framework? It seems to me that until we have thought about the actual processes it is not obvious *a priori* that being knowable through hypothetical record reading is a necessary condition. In general, I would prefer to approach the issue, more traditionally, through consideration of what we do when we make inferences about the past.

2 An IGUS Remembers

I want to try a different approach to the question of the knowledge asymmetry, that rather than giving some conditions on what can be known, starts with an obviously simplified model of the

physical basis of our mnemonic knowledge. I have in mind some version of Gell-Mann’s (1994, *Quark and Jaguar*) “information gathering and utilizing system”, or *IGUS* (see also Hartle 2008), though none of the specific details of that account are very important here. In general terms, such a device has sensors capable of responding to its environment and devices capable of affecting its environment: inputs and outputs broadly construed. The system takes inputs from its environment (visual, for example), processes the received information and stores it in memory; that’s gathering. Then it has the capacity to operate algorithmically on the contents of its memory and its current inputs to determine a course of action – what its output devices will do (emit a beep or move away, for instance); that’s utilizing. Clearly such a system can be realised as a digital computer; though the intent is that an IGUS has greater autonomy and consistency than the typical functioning of the personal computer. A little more specifically, I want to consider an IGUS rather like us: especially responding to the same kinds of features of the world that we do, drawing inductive inferences from them¹⁰, and drawing on those inferences for action in the pursuit of specific goals. Below I shall say a little more about the operation of the IGUS, specifically its algorithm.

Insofar as the system has reliable procedures for forming veridical memories, insofar as it draws true inductive inferences, and insofar as its actions realise its goals on the basis of those inferences, I say that the IGUS models an important aspect of our knowledge. Computationally it is relevantly like us, and it could even be physically implemented like us. I trust it is clear enough and uncontroversial enough why I say that. But more importantly, the IGUS has knowledge that is not of the retrodictive/predictive kind, *or of the record reading kind described in T&C*. Let’s focus on memories.

First it is worth seeing that the IGUS need not use explicit record reading in its mnemonic functioning – not because T&C says that it must to get beyond prediction/retrodiction (it does not say that) but to get clearer on how such a device might function. So, we might ask whether an algorithm could be constructed to compute explicitly the significance of the contents of registers by comparing them with their ready state and computing what the cause must have been? Perhaps. After all, in some sense that is what we do when we look at external records: we compare the current state of the dented car with its state this morning and infer a collision. Of course we don’t in general explicitly apply the laws of mechanics – for instance, many of the records we read rely rather on rules about human intentions, and various kinds of inscribing processes – but the story is in the vicinity of the truth about how we read the records we find around us. However, such an algorithm would be a grossly inefficient way of implementing the IGUS’ reasoning capacities.

Insofar as I can envision such a procedure, it would do something like form representations of the current register state (perhaps the register is its own representation in that regard), and *also* of the registers’ ready state, and then compute a representation of the cause of the difference. That third, derived, representation would then serve as a computational input for any decision making processes related to the memory.¹¹ But what would the point of the computation be? A more efficient algorithm would simply assign to the original contents of the registers whatever computational role the derived representation plays in the algorithm discussed. That is, the states of the registers themselves are perfectly capable of playing the computational role directly, without any interpretational process. But without some process such as I described, no record *reading* of the kind described in T&C has occurred, because no role has been played by knowing, or assuming

¹⁰For recent work on algorithmic inductive reasoning see Waltz and Buchanan, *Science* 3 April 2009: Vol. 324. no. 5923, pp. 43 - 44.

¹¹To be honest I am not clear that even this procedure would really count as reading a record in the sense of T&C.

the ready state.¹² In the sense that no kind of inference – no part of the algorithm plays the role of ‘interpreting’ the registers – is involved, the knowledge is ‘direct’.

The point just made is the crucial one: drawing on memories is nothing like T&C’s record reading. (Even in cases in which the IGUS’ goal is to produce an explicit representation such as a diagram or verbal description, the algorithm can take the contents of the registers as inputs, without any representation at all of ready state.)

Maybe it’s worth justifying here the claim that the IGUS ‘knows’ anything; in what sense can we call the contents of the registers ‘knowledge’? It will suffice here to point out that the IGUS can satisfy a wide range of behavioural criteria for knowing: actions depending on the memories, production of representations of the memories, communicating the memories to other IGUSs, for example. I do not mean to rely on a behaviouristic account knowledge – these are just evidence, not constitutive.

And bear in mind that the point of the example is to argue that humans have knowledge of a similar kind, which is distinct from the knowledge that could be obtained by reading records (or by retrodicting). So my purpose is not at all to settle whether the IGUS has knowledge, just that the broad outlines of its functioning match the mechanics of our mnemonic knowledge. To argue that, after all, *our* memories are a matter of record reading would require coming up with some property of knowledge that can only be obtained by record reading. I can’t see what aspect of our knowledge might require it to involve explicit record reading as conceived in T&C – next I will argue that the IGUS’ knowledge does not satisfy the necessary condition for record reading either, and I would further say that I can’t imagine any aspect of our knowledge that would change that result either. So I will generally proceed as if the IGUS has knowledge, and leave the defender of record reading to show why the differences between the IGUS and us are relevant to the discussion.¹³

Now for the punch line to much of the discussion of the last couple of sections: the IGUS’ memories constitute knowledge that does not satisfy the necessary condition for knowledge of the record reading type (nor of the prediction/retrodiction type). We can take this in two stages (in fact as many stages as there are in the regress argument, but two will suffice). First, the condition requires that even though the IGUS does not possess its memories by an explicit process of record reading, the same information could be obtained by a process of explicit record reading. With the example of my pachinko observer – the observer could be our IGUS – it had better be the case that the record in question is the one constituted by the ready and current states of the IGUS’ registers. But why should we think that those (plus the rest of the surveyable condition) constitute sufficient boundary conditions for the laws of dynamics to determine their intentional content, what they are memories of? What the contents of the registers are knowledge of for the IGUS does not depend in that way on the laws of physics, but rather on their computational role according to the algorithm – crudely, what the machine table says about them. That is what connects them to any cognitive process, or any action the IGUS could undertake. (Put another way, the contents of the registers plus their ready state could at most imply what internal process caused the records, not their computational significance, their content, the external events of which they are memories.)

¹²I want to register some general reservations about the assumption that the running of an IGUS’ algorithm *directly* represents our conscious ratiocination. Clearly much of the brain’s processing is not conscious.

¹³I also say that the IGUS need not utilize the contents of its registers to obtain knowledge of the past by retrodiction. I can imagine an algorithm for an IGUS that attempted to employ retrodiction to interpret the contents of its registers, but it doesn’t seem as if it would be very successful. The problem is that the IGUS only has access to its computational state, which will likely not provide sufficient information about its physical state; i.e., the computational state will not determine the microstate. (Of course, it would also be a grossly inefficient algorithm, for the same reasons as the record reading algorithm.)

But – here’s the second step – the regress argument shows that the necessary condition is far more stringent than that: if the IGUS’ memories are knowledge, then it also has to be the case that ready state of the registers be knowable by a record. For remember, the whole point of the regress is to show that ultimately the PH must be assumed to explicitly read any records given just the current surveyable state. The argument must work something like that if it is to establish the PH as a necessary condition for knowledge of other times beyond retrodiction and prediction. But there is no reason at all to think there will be such a record. If we built an IGUS, we might well expect the currently surveyable state to include books or human memories that constitute such records (given a ready state), but what if, say, the IGUS is the last artifact (a space probe say) of a civilization that was destroyed by a supernova? Why should there be anything, anywhere in the whole macrostate of the universe plus whatever follows from direct introspection of all sentient creatures that amounts to a record of the ready state? (And if you are not convinced yet, ask about the record for the ready state of that second record, and so on.)¹⁴

Given what the IGUS can do, and given its similarity with us, I assert that its memories are knowledge of the past, and the preceding points, showing that its knowledge could not be had by explicit record reading, do not undermine that claim at all. Instead they show that knowledge of other times need not be of any of the kinds listed in T&C (though some may be), the condition is not necessary. In that case, the account of the knowledge asymmetry fails, just like the one that appealed only to the macrocondition of the universe, not direct introspection. That is, the asymmetry cannot be pinned on the PH in the way proposed, because the condition for knowledge is not necessary after all.

Of course, that is not to say that the PH isn’t involved at all. In particular, we still have to worry about reversibility – the PH is will guarantee that the state of the IGUS’ registers is not a random fluctuation from equilibrium, but veridical. And that is not simply to fall back to Reichenbach’s account; for it does not (by itself) rule out mnemonic processes that do not increase entropy. But what it does suggest is that the knowledge asymmetry may be more of a piecemeal matter than is generally assumed. There may not be one cover-all story about why the past can be known in a way that the future cannot. Maybe what we need to do is consider models of increasing sophistication – like the IGUS – to say in rather more specific manifestations what prevents this or that kind of device from getting access to the future. Indeed, by invoking the inductive capacities of the IGUS, I

¹⁴Adam Elga made the following suggestion in response to the talk, if I understand him correctly. (In fact he claimed that the following example was mine, but since he had to explain it to me, I think he deserves some credit, if he wishes to have it.) Suppose indeed that the civilization which created an IGUS has been destroyed without (other macro) trace: and suppose it is the IGUS that witnesses which beaker the ice cube falls into (the leftmost one), information it stores in its registers. Suppose also that the IGUS has stored in its registers a memory that of all IGUSs of similar origin only one in ten is reliable when it comes to ice-and-beaker related memories (and that all are generally reliable about other memories): but that our IGUS is one of the rare reliable ones. Then, assuming that the story of section 1.2 about the PH and the presently surveyable condition all pans out, then most likely the ice cube did not fall into the leftmost beaker – according to the currently surveyable state, the memory is most likely not veridical. But the IGUS is (unbeknownst to it) reliable, and so plausibly does know that the ice cube fell into the leftmost beaker. So here we have a counterexample to the claim that all knowledge satisfies the modified necessary condition involving the PH. (And the story just makes things stark: there’s no reason the same point couldn’t be made in a more mundane situation.) Of course there is a response: the IGUS doesn’t know which beaker the ice fell in, *because it doesn’t know that it knows* (in fact, the way things are set up, it has good reason to believe it doesn’t know, but the story could be modified so that the record of unreliability was external). But that response only works if one is prepared to accept an internalist view of epistemology: that knowing X requires knowing that you know X . So the upshot of this example is that this approach to the knowledge asymmetry commits one, surprisingly, to an internalist epistemology. That will be a reason for some to reject the approach; but at very least one should wonder whether physics should be able to settle this issue in epistemology.

think we can draw some interesting further lessons about the knowability of the past, beyond what is immediately known through memory.

References

Albert, D. Z (200) *Time and Chance*: Harvard University Press, Cambridge, MA.

Gell-Mann, M. (1994) *The Quark and the Jaguar*: W. Freeman, San Francisco.

Hartle, J. (2008), "The Physics of Now", *Am. J. Phys.*, 73, 101-109 (2005), gr-qc/0403001.