Forthcoming in a volume ed. By Valia Allori

"If someone points out to you that your pet theory of the universe is in disagreement with Maxwell's equations—then so much the worse for Maxwell's equations. If it is found to be contradicted by observation—well these experimentalists can bungle things sometimes. But if your theory is found to be against the Second Law of Thermodynamics I can give you no hope; there is nothing for it but to collapse to deepest humiliation." Arthur Eddington

The Mentaculus Vision[1]

Barry Loewer
Rutgers University

Thermodynamic laws apply to the melting of ice, the dispersal of smoke, the mixing of milk in coffee, the burning of coal and a myriad of other familiar processes. All of these manifest an arrow of time pointing from past to future. We never see ice spontaneously forming out of warm water, milky coffee separate and so on. The second law of thermodynamics, which covers these processes, says that the entropy of an energetically isolated system never decreases and typically increases until the system reaches a state of maximum entropy, its equilibrium state. If all one knew about the physics of the world was that all systems are composed of material particles (atoms) whose motions conform to classical mechanics then the temporal asymmetry of the second law would or should be deeply puzzling since the dynamical laws of classical mechanics that govern the motions of particles are deterministic and temporally symmetric. Their temporal symmetry consists in the fact that if a trajectory of a configuration of particles is compatible with the classical mechanical laws then the temporally reversed trajectory of the configuration is also compatible with the laws[2]. So, since the trajectories of the particles of ice and water that comprise the melting of ice are described by the laws so are the trajectories that comprise the spontaneous formation of ice out of water. The puzzle is how temporally asymmetric laws emerge from fundamental temporally symmetric laws? The development of statistical mechanics (SM) by Maxwell and Boltzmann partly answered this question. They claimed that

[1] Thanks to Eddy Chen, Isaac Wilhelm, Denise Dykstra, Dustin Lazarovici and especially David Albert for comments and discussion of these issues.
[2] The same symmetry is also exhibited by the laws contained in quantum mechanics, general relativity, and plausibly the dynamic laws of any future physics.

a thermodynamic system, e.g. a gas in a box, is composed of an enormous number particles moving randomly under the constraints imposed by the macroscopic properties of the system. To capture the random motion they posited probability distributions over the positions and velocities of the particles compatible with a system's macroscopic state and they and their followers were able to show for some systems that if the system is not in equilibrium it is enormously likely that it will evolve in conformity with the classical mechanical laws to states of higher entropy until equilibrium is reached. In this way they hoped to demonstrate that a probabilistic version of the second law is compatible with the fundamental time symmetric laws.

Boltzmann's attempt to account for the second law encountered a famous obstacle. If the universe is in a certain macroscopic state now then it follows from Boltzmann's probabilistic posit and the temporal symmetry of the laws that while it is enormously likely that the entropy of the universe will increase toward the future it also enormously likely that its entropy was greater in the past. The same holds for approximately isolated subsystems like the ice cube. While the statistical mechanical probability assumption implied that it is likely that the ice cube will be more melted in the future it also implied that it was more melted in the past. But we know that this isn't so. How to remove this bad consequence from Boltzmann's account while preserving its explanation of the second law is called "the reversibility problem."

Recently, David Albert in *Time and Chance* takes Boltzmann's approach to SM and building on ideas from Eddington, Feynman and many others describes a framework that solves the reversibility problem.

Albert's framework contains three ingredients:

1. The fundamental dynamical laws that describe the evolution of the fundamental microstates of the universe (and the fundamental microstates of its isolated sub- systems). We will assume that the dynamical laws deterministic and temporally symmetric.

2. The Past Hypothesis (PH): A specification of a boundary condition characterization of the macro state M(0) of the universe at a time shortly after the Big Bang. In agreement with contemporary cosmology M(0) is a state whose entropy is very small.

3. The Statistical Postulate (SP): There is a uniform probability distribution specified by the standard Lebesgue measure over the physically possible microstates that realize M(0).

Albert and I have dubbed this framework "the Mentaculus."[3] The Mentaculus is a probability map of the universe in that it determines a probability density over the set of physically possible trajectories of microstates emanating from M(0) and thereby probabilities over all physically specifiable macro histories. Once the dynamical laws are specified the Mentaculus contains an answer to every question of the form "What is the objective probability of B given A?" for all physically specifiable propositions A and B. For example, it specifies the probability that an ice cube will completely melt in the next five minutes given a description of the macro state of its environment and the probability that Trump will be impeached given the current macro state of the universe. Though, to be sure, we cannot extract these conditional probabilities from the Mentaculus.[4] Further, Albert has argued that the Mentaculus not only grounds the

[3] Discovering these conditional probabilities faces two obstacles. One is specifying the sets of micro states corresponding to the macro conditions mentioned in the conditional probability and the second is doing the

thermodynamic arrow of time but also explains other arrows of time. While many of the details of arguments supporting these claims are yet to be worked out the Mentaculus is sufficiently bold and promising to warrant taking it seriously.[5]

This paper has two aims. One is to make a case that the Mentaculus should be taken seriously by reviewing reasons to believe that it explains the second law of thermodynamics, various arrows of time, and why it is arguably a complete scientific theory of the universe. Aim number two to argue that the best way to understand the metaphysics of Past Hypothesis and probabilities in the Mentaculus is a Humean account along the lines of the best systems account of laws and chances developed by David Lewis.

---

[4] "Mentaculus" comes from the Coen Brothers' film "A Serious Man" where it is used by a mentally disturbed characters as a name of what he calls "a probability map of the world." Hence, with the permission of Ethan Coen and in the spirit of self-mockery we appropriated the name.

Before explaining how the Mentaculus accounts for the second law and other temporal arrows I want to address two worries. The first is that my discussion assumes the ontology and dynamics of classical physics. But classical mechanics has been superseded by quantum field theory and general relativity and these are likely to be replaced by a quantum theory of gravity. The second worry is that M(0) is characterized along the lines of current cosmology i.e. that shortly after the moment of the big bang the universe consisted of a rapidly expanding space-time that contained a soup of elementary particles and fields that was very dense, very hot with almost uniform density and temperature, and the entropy associated with M(0) is very low.[6] One problem is that it is vague exactly what counts as the macro state of the early universe. More worrisome is that it is not obvious how entropy of the early universe should be characterized especially if the ontology of early universe is specified by a yet to be formulated theory of quantum gravity.

The reply to the first worry is that any future plausible candidate for fundamental ontology and laws will posit states that are composed of particles and/or fields that possess an enormous number of degrees of freedom and laws that are capable of grounding the motions of material bodies. As

---

calculation.

[5] Rigorous proofs that systems that satisfy the statistical mechanical probability distribution evolve in conformity to probabilistic modifications of thermodynamic laws (e.g. the entropy of an isolated macroscopic system is very likely to increase) have been produced only for systems that satisfy very special constraints for example ergodicity. (Sklar 1993). Arguments for such claims make them plausible.

[6] The claim that the entropy of the early universe was very low may seem surprising since an ordinary gas at a high temperature uniformly spread out in a container is a system whose entropy is high. But while the effects of gravity are miniscule for the gas in the container in the very dense early universe the contribution of gravity to entropy is significant. It is thought that taking gravity into account the state of the early universe has very small entropy. The reason for is that in the presence of gravity a very dense uniform distribution of matter/ energy is very special (i.e. low entropy) and will likely evolve to higher entropy states as matter/energy clump to form stars etc. See Penrose (2005) ch27 and Carroll (2010) for discussion of this point. For complications and some dissent see Callender (2009), Wallace (2010), and Earman (2006).

Eddington implies the second law will survive no matter the developments of future physics. As for the second worry, although it is not known how to characterize entropy for the very early universe all that the Mentaculus requires that at some point in the distant past the entropy of the universe is very low. Whatever the ontology of future fundamental theories may be it is plausible that it will sustain a characterization of entropy analogous to the usual Boltzmannian account.

The claim that the Mentaculus is a probability map of the world is outrageously ambitious. Why should the Mentaculus be taken seriously? The line of reasoning supporting it originates in Boltzmann's account of how the temporally directed second law of thermodynamics emerges from the temporally symmetric fundamental dynamics. The story is familiar but a quick retelling will help to explain the Mentaculus.

Clausius introduced the concept of entropy to characterize the fact that in thermodynamic processes energy involved in the performance of work tended to be transformed into energy that is no longer available for work. For example, the operation of an engine produces heat that cannot be used to do work. Thermodynamic entropy is, roughly, a measure of the energy in a system that is unavailable for work. The original formulation of the second law of thermodynamics says that the entropy of an energetically isolated system never decreases and typically increases until the system reaches equilibrium.

Boltzmann's explanation of the second law in terms of classical mechanics begins with his characterization of entropy as a property of the macro state of a system. The macro state M of a system specifies its total energy, temperature, pressure, density and so on. For each macro state M there is a set of micro states (positions and momenta of the particles that compose the system) that realize M. Boltzmann identified the entropy of M with the logarithm of the volume on the standard Lebesgue measure of the set of microstates that realize M multiplied by a constant (Boltzmann's constant). We can think of the inverse of the hyper-volume of a system's macro state M as specifying how much information M contains about the micro state that realizes it. So greater entropy corresponds to less information. The second law says that an isolated system's macro state evolves from specifying more to specifying less information about its microstate. Boltzmann's next step was to interpret the Lebesgue measure as specifying a probability measure over the space of the microstates that realize M. He then argued that a probabilistic version of the second law followed.[7] His argument, in a nutshell, is that since overwhelmingly most (on the measure) of the micro states that realize the macro state M of a system not at equilibrium are sitting on trajectories that evolve according to the dynamical laws to realize macro states of greater entropy it follows that it is overwhelmingly likely that entropy will increase as the system evolves. The qualification "overwhelmingly likely" is required since there will be some micro states that realize M that don't evolve the system to states of higher entropy. But these states are rare (on the measure) and scattered among the set of states that realize M. That is, in every very small (but of a certain finite size) convex region of phase space almost all states evolve to higher entropy. The neighborhood of each "bad" micro state that doesn't evolve to higher entropy contains mostly "good" micro states that do. A slight disturbance of the particles comprising a" bad" micro state will change it into a good one. Boltzmann and those following in his footsteps have been able to show that it is plausible that if it is a thermodynamic law that system S at time t in macro state M will evolve to be in macro state M* at t* then Boltzmann's account will recover a probabilistic version of this regularity. Boltzmann's probability hypothesis is

[7] The question arises of what "probability" means in this context especially since the fundamental dynamical laws are assumed to be deterministic. I discuss this issue later in the paper.

vindicated by its success in accounting for the second law and other thermodynamic laws.

As mentioned earlier Boltzmann's account ran smack into the reversibility problem. While the account entailed that it is enormously likely that the entropy of the universe (or an isolated subsystem) will increase toward the future, because of the temporal symmetry of the dynamical laws it also entailed that it is enormously likely that entropy increases toward the past.[8] One reaction to the reversibility problem is to construe the Boltzmann probability distribution solely as an instrument for making *predictions* for the future behavior of a thermodynamic system but refrain from using it for retrodictions. This approach is usually combined with the views that statistical mechanical reasoning should only be applied piecemeal to subsystems of the universe and the view that probability should be understood epistemically as a measure of an experimenter's knowledge (or lack of knowledge) of the system's micro state. The package of these views amounts to an instrumentalist understanding of statistical mechanics. It avoids the problem but it leaves us completely in the dark as to why Boltzmann's probability posit works for predictions and in what sense it provides *explanations* of physical phenomena.[9] And unlike the Mentaculus it does not apply to the universe as a whole. In contrast, the Mentaculus is a realist account on which statistical mechanics applies to the entire universe and, as I will later argue, it is compatible with and objective non-epistemic account of statistical mechanical probabilities.

The Mentaculus solves the reversibility problem by conditionalizing the Lebesgue probability distribution on the low entropy initial condition M(0); i.e. the PH[10]. This results in a probability distribution that gives the same *predictions* (inferences from M(t) to times further away from M(0)) as Boltzmann's prescription when applied to the universe as a whole and to its energetically isolated subsystems while avoiding the disastrous retrodictions we found without the PH. The PH and the Statistical Postulate work together to eliminate or make very unlikely trajectories of configurations of the particles in the universe that although they are compatible with the dynamical laws violate thermodynamical laws. The Past Hypothesis eliminates all trajectories that fail to start in the very low entropy macro state M(0) and the Statistical Postulate implies that as the universe evolves it is very likely that the entropy of the macro state that the configuration realizes will increase.

The probabilistic version of the second law not only says that the entropy of the entire universe likely increases as long as the universe is not at equilibrium but also that this holds for typical

---

[8] It is not only absurd but as Albert points our leads to cognitive instability. If the Boltzmann probability posit is applied to the macro condition of the universe at the present time since it implies that it is likely that this macro condition arose out of higher entropy states and in particular this means that the "records" in books, etc. more likely arose as fluctuations out of chaos than as accurate records of previous events. This undermines the claim that there is evidence reported in those books that supports the truth of the dynamical laws and so results in an unstable epistemological situation.

[9] Boltzmann's prescription leads to instrumentalism since it couldn't be literally true as it prescribes incompatible probabilities at different times since the uniform distribution over the macro state at t will differ from the uniform distribution over the macro state at other times.

[10] See Sklar (1993) for a discussion of some other proposals for responding to the reversibility paradox.

energetically isolated or almost isolated sub-systems e.g. the entropy of an ice cube in a glass of warm water will likely increase toward the future but decrease toward the past. It may seem astonishing that conditionalizing on the macrostate of the universe 13.8 billion years in the past results in the second law holding for subsystems at much later times; that, for example, the explanation why ice cubes melt but never unmelt ultimately involves cosmology. Here is a "seat of the pants" argument that makes it plausible that the Mentaculus has this consequence. Suppose that S is a subsystem of the universe that at time t "branches off" from the rest of the universe so as to become approximately energetically isolated. As an example, consider an experimenter that places an ice cube in a tub of warm water. As mentioned earlier, "bad" micro states (i.e. those on a trajectory that is not entropy increasing) are scattered throughout the phase space and surrounded by "good" micro states. Because of this preparing the micro state of the ice cube in warm water system in a bad state would require incredible precision beyond the ability of the experimenter. It is plausible then that the micro state of the ice cube in warm water system at the time it branches off is not bad and as long as it is (approximately) energetically isolated it is very likely that its entropy will increase.

Due to the sparseness of bad micro states in the phase space it is also plausible that almost all (on the Lebesgue measure) the microstates of subsystems that are not prepared by an experimenter but branch off and become approximately energetically isolated as the universe evolves are not bad. Of course this doesn't mean that the entropy of *every* subsystem of the universe is likely to increase. Some subsystems are not isolated but are interacting with other parts of the universe so as make their entropy decrease likely (e.g. a glass of water in a freezer) while the entropy of the larger system increases (the environment of the freezer). Also there may be systems that are specially prepared so that even when they become isolated their entropy will very likely decrease.[11] In these cases the second law doesn't hold. But that is as it should be. The job is to get the second law (and other thermodynamic laws) from the Mentaculus *in so far as* the second law is correct and, arguably, the Mentaculus does exactly that.

Albert pointed out that entropy decreasing micro states are not the only "bad" micro states. There are also micro states compatible with macro states of macro objects that while they need not be entropy decreasing are on trajectories that manifest bizarre behavior. Consider, for example, a comet in orbit around the sun. Almost all the micro states of the particles that compose the comet and are compatible with its macro state M are ones that maintain its integrity as it travels on its orbit. But there will be some microstates compatible with M whose particles' positions and momenta are so arranged that the comet disintegrates or in Albert's example takes the shape of a statue of the royal family. Of course, nothing like this ever occurs. The reason is that the set of such arrangements has measure 0 (or indiscernible from 0) on the Lebesgue measure and so according to the Mentaculus has probability 0 (or indiscernible from 0). The application of classical mechanics to macroscopic objects like comets or baseballs implicitly assumes that the microstate of the enormous number of particles that constitute them is not one of these bad states. Unbeknowst to Newton and Haley the Mentaculus justifies this assumption.

The Mentaculus provides an account not only of the laws of thermodynamics but also an important ingredient in an account of other special science laws. Consider a true regularity of the form a system which is F at t evolves to be G at t'. Since the Mentaculus entails probabilities over all physically possible macro histories and since F and G are, to the extent they are precise, correspond to sets of micro histories it entails the value of P(G(t')/F(t)). For the regularity to be a law the value must be near 1. If it were much less than 1 we should conclude that the regularity was accidental or that it only held under certain conditions which when added to the antecedent result in a conditional probability with a

[11] See *Time and Chance* Ch. 5 for a discussion of how a system may be prepared so that its entropy reliably decreases.

value near 1. The reason for believing that this is a necessary condition for a special science law is a strong as our reasons for believing that the probabilities entailed by the Mentaculus are the objective probabilities. The primary reason for believing that is the Mentaculus success in accounting for thermodynamics. [12]

Albert also makes a case that the Mentaculus can account for the "arrows of time" in addition to the thermodynamic arrow. I want to briefly review part of that case and add a few wrinkles to it.

Time's arrows are pervasive features of our world. The epistemic arrow is that we can know much more about the past than about the future since there are records (including memories) of past events but nothing like that for future events. The influence arrow is that we can have some influence or control over future events but no control over the past. Both these temporal asymmetries are closely related to the fact that causes precede their effects which itself is closely connected to the temporal asymmetry of counterfactuals. The question is where do these arrows come from? In view of the temporal symmetry of the dynamical laws the issue is as puzzling for them as for the thermodynamic arrow.

There are two views that are held about the status of the temporal arrows. Primitivism is the view that time itself (unlike space) is equipped with a primitive direction.[13] Somehow the "flow of time" is behind the temporal arrows. Reductionism is the view that time (like space) has no fundamental intrinsic direction in itself but that the temporal asymmetries arise from the fundamental laws. I won't say anything more about primitivism except that it faces the problem of connecting the primitive arrow to the familiar arrows.[14] Reduction of the temporal asymmetries to the laws encounters the problem posed by the temporal symmetry of the fundamental dynamical laws. The Mentaculus aims to remedy this by adding to the dynamical laws two non-dynamical laws -the PH and the SP- which are not temporally symmetric. The reductionist program is to make it plausible that with these additional laws the Mentaculus can explain the other arrows. One might worry that by including the PH the Mentaculus is assuming a past/future distinction. But this is not correct. The Past Hypothesis is a boundary condition of the universe obtaining at a time close to the time of the Big Bang. If the Mentaculus program to explain time's arrows is successful then that time be shown to in the past and the boundary condition will earn its name.

Suppose that at time t(1) you come upon a closed box in which there is some blue smoke partly spread in the box with most of the smoke near the bottom right hand corner of the box. The usual statistical mechanical probability distribution (SP) is the best you can do to predict the future of the smoke. The SP will predict that the smoke is likely to become more defuse over time and eventually more or less uniformly fill the entire box but it won't enable you to predict exactly what shapes it will take as it spreads. But the SP would be a terrible way of inferring to the state of the smoke at a time t(0) earlier than t(1). The SP implies that at an earlier time the smoke was spread uniformly throughout the box and followed a course that led to the state in which you found it at t(1). i.e. that it would evolve in the time reverse of the way you predict it will evolve from its state at t(1). But if you were informed that the smoke was released into the box from a small hole at the right hand corner at t(0) then conditionalizing on this would lead you to infer that the smoke spread out from the hole to its present distribution. Unlike predictions of the future evolution of the smoke we would be justified in inferring a great deal of detail about how the smoke evolved from the hole to its condition at t(1) since the way it could evolve would be restricted by its having to get from the hole to its state at t(1) in a certain amount of time. The idea is that the PH plays the role for the universe that the fact that the smoke was

---

[12] The reason I have for saying that P(G/F) is near 1 is a necessary but not a sufficient condition for Fs evolve to Gs to be a law is that further conditions, for example being a component of an optimal systematization of the truths expressible in the vocabulary of a science is plausibly an additional requirement for special science law hood. See Callender and Cohen (2009)

[13] Tim Maudlin (2007) is the most effective proponent of this view.

[14] See Loewer (2012)

introduced from the hole at a certain time plays for inferences from t(1) to times between t(1) and t(0).

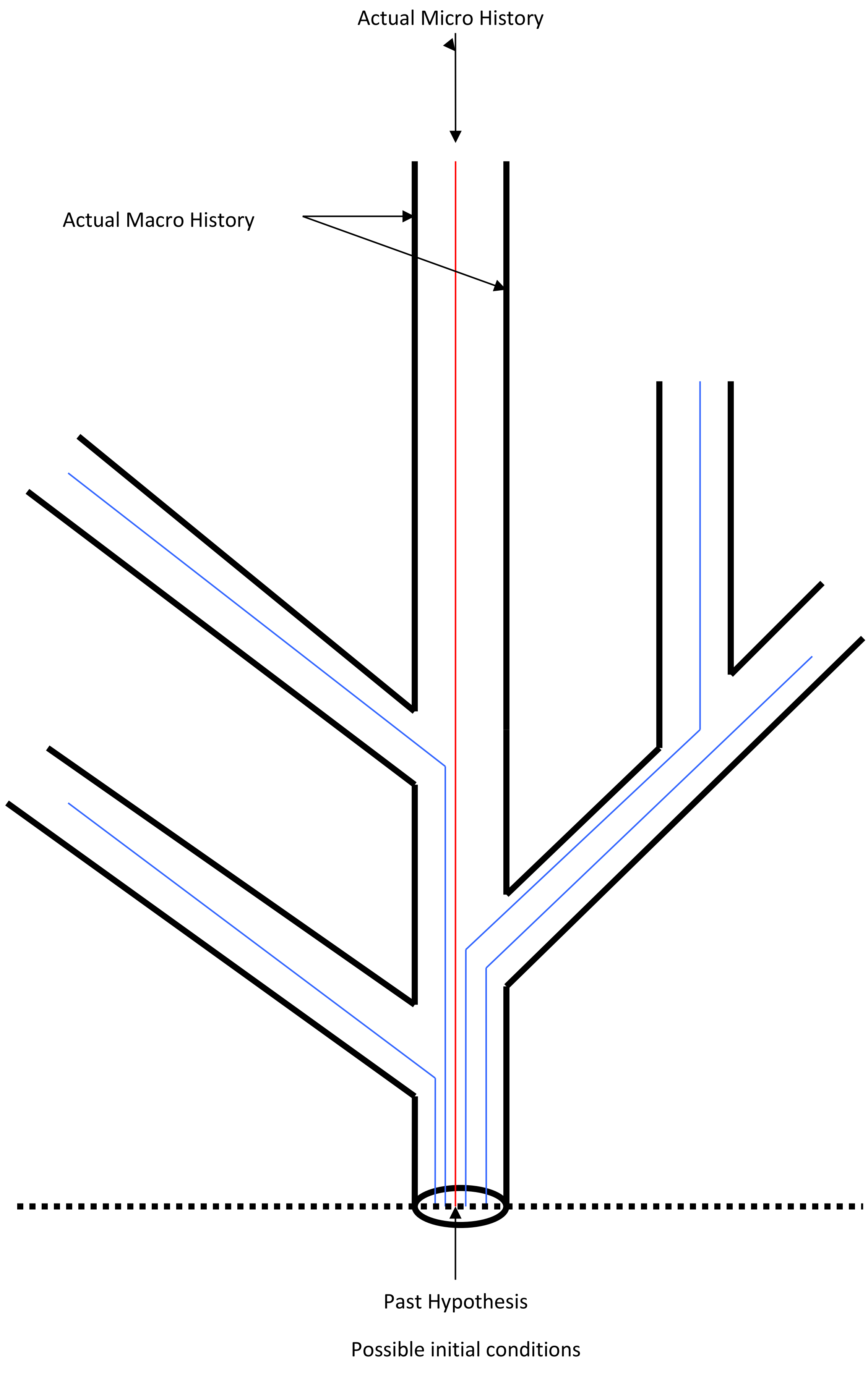
We can think of the condition of the smoke at t(1) as a record of its condition at earlier times and the state of the smoke being released from the hole as a "ready condition" for the record. Albert observes that this reasoning is common to all inferences from what we take to be a record to the state it records. The inference from a record at t to a condition at an earlier time t′ involves the assumption of a ready condition at a yet earlier time t′′. But how can we know at t that the ready condition obtained at t′′? We would know it if we had a record at t(1) that the ready condition obtained but that would require knowing the state at a time before t′′. This begins a regress that Albert says can be stopped only by an assumption concerning the state at the earliest time. The PH is such an assumption. It constrains the way the universe has evolved from its very low entropy condition at t(0) so as to allow records to be produced of conditions at times between the time of the record and the time of the PH. It is important to note that this proposal doesn't require that to be justified in inferring from a record (e.g. tracks in the snow) to what it is a record of (a bear having recently passed by) that one knows the that the ready condition (that the snow was fresh prior to the bear walking on it) obtained or that one knows the truth of the PH. It is sufficient that the ready condition did obtain.[15]

By assuming that the PH is true the Mentaculus provides an explanation of why there can be records of events during the interval between the time of the record and the beginning of the universe. But because there is nothing like the PH assumed about the state of the universe during the time we call "the future" there are no "records" of events occurring during that interval. For predicting the future the best we can do is to is to infer on the basis of the SP and that will leave a great deal of uncertainty. The role of the PH in underlying the account of why we can have records of events during the interval from the big bang until now justifies calling that interval "the past" and the fact that the SP leaves the time from now epistemically much more open justifies calling that time "the future."

The other arrow of time I want to discuss is the counterfactual arrow since it is closely connected to the arrows of influence and causation.

The Mentaculus imposes a probabilistic structure on the possible macro histories of the universe compatible with the PH. While the possible micro histories evolve deterministically macro histories evolve probabilistically. Micro histories that are macroscopically identical from the time of the PH until t can diverge. Macro histories can also converge (eventually if the universe has an equilibrium state all will converge to it) but during this epoch of the universe since it is near the big bang divergence overwhelmingly predominates. So from the macroscopic perspective the evolution of the universe (and typical isolated subsystems) appears to be indeterministic. The resulting branching structure is pictured below.

[15] . In fact Cosmologists infer from observations of the state of the universe now that at a time shortly after the big bang the entropy of the universe was very small in agreement with the PH. The cosmological inference are themselves justified by the PH. If it had turned out that cosmologists inferred from the present state that the state of the universe was incompatible with the PH then Albert's explanation of the existence of records would be sunk.



Possible initial conditions

Counterfactuals are temporally asymmetric. Counterfactual suppositions typically lead to large departures from the future of the actual world but not from the past. For example, if Nixon had pushed the button in 1974 the subsequent course of history since then would have been very different but the prior history would have been pretty much the same. David Lewis proposed an account of counterfactuals that he thought captured this temporal asymmetry.[16] But as Adam Elga pointed out he was wrong.[17] The Mentaculus suggests a better account that captures this asymmetry. Unlike Lewis' account it doesn't involve "miracles" but it is restricted to counterfactuals whose antecedents are compatible with the Mentaculus.[18] As a first stab consider this:

> COUNT) A(t)>B is true at iff conditional probability P(B/Disjunc(A(t)) is close to 1. Where t is the time of A and Disjunc(A(t)) is the disjunction of all the worlds at which A(t) is true. P(B/Disjunc(A(t)) is the conditional probability according to the Mentaculus of B given the set of all the worlds most similar to the actual world at which A(t) and the Mentaculus are true.

Suppose that the actual world at the time referred to in the antecedent, call it t, Nixon is distraught by the prospect of impeachment and is in the Lincoln bedroom down the hall from the location of the button. The most similar worlds to the actual world are those whose macroscopic states are most similar to the actual macroscopic state at t except that Nixon is by the button at t pushing it. According to the Mentaculus the probability that these evolve so that there is a nuclear war is close to 1[19] (assuming that the presidential launch system is working properly) and so the macroscopic future from t is very different from the actual future. But the probability that the macro history of these worlds is just like the history from a time shortly before t will also be high. There is some back tracking since in the counterfactual worlds Nixon will have to have gotten from the Lincoln bedroom to the location of the button. This backtracking seems right since we think that if Nixon had pushed the button at t he would have had to have left the Lincoln bedroom a few minutes before t. In general, the more dissimilar the most similar counterfactual states in which the antecedent is true is from the actual state the more back tracking is required. But the back tracking is limited since there records of the actual macroscopic history that remain in the counterfactual worlds. So the macroscopic histories of the counterfactual world will be much like the actual history as is possible. If this account is on the right track we see how the PH plays a crucial role in accounting for the temporal asymmetry of counterfactuals. It is due to the fact that the PH underwrites the existence of records of the past that the counterfactual past is similar to the actual past. But since there is no "future hypothesis" constraining the future the counterfactual future may be very different from the actual future.

If the relation between counterfactuals and causation is anything along the lines of Lewis' counterfactual account then it reasonable to expect that the Mentaculus also accounts for the temporal asymmetry of causation but it is beyond the scope of this paper to fully develop this suggestion. However, we can see how the Mentaculus can account for the causal influence of decisions.

Here is how COUNT accounts for the fact that we can have influence and control over the future but not the past. Suppose at t Nixon has his finger on the button deciding whether or not to push it. At t COUNT will plausibly judge each of these counterfactuals true

1) If Nixon were to decide to push the button there would be a nuclear war after soon after t

2) If Nixon were to decide not to push the button there would not be a nuclear war soon after t

So at t Nixon can influence whether or not there is a nuclear war soon after t. Furthermore, since Nixon

[16] Lewis "Times Arrow and Counterfactual Dependence" in Lewis (1986).
[17] See Elga (2001) and Loewer (2007)
[18] This account is approximately coextensive with an account that Jonathan Bennett proposes (2003).
[19] Not 1 but close to 1 since as was discussed earlier there is a small statistical mechanical probability of a fluctuation even if the launch mechanism is working properly.

knows that these subjunctives are true he has control over whether there will be a nuclear war. But he has no similar control over any past events. The reason is that decisions correspond to brain events tha involve relatively small number of degrees of freedom. For this reason it is plausible that either decision can be added to the macroscopic state at t while keeping the remainder of the macroscopic state with all its records of the past almost exactly the same. This means that decisions made at t have no almost no effect on the macro history prior to t.[20] Of course the PH plays a crucial role here as well since it is involved in the existence of records.

It has often been suggested that the thermodynamic arrow is at the bottom of the other arrows of time. That is not quite right. Rather it is the Mentaculus that accounts for both the entropic arrow and the other arrows and thus makes for a unified account of temporal direction. While more needs to be done to develop and defend this claim I think it is sufficiently promising to provide good reasons to take the Mentaculus seriously.

Taking the Mentaculus seriously means addressing the questions of interpreting the probabilities posited by the SP and how PH can be understood to be a law. The question concerning probability is vexing since it is widely believed that if determinism holds then probabilities cannot be objective features of the physical world but must be construed epistemically or subjectively. Here is a famous expression of this is view by Karl Popper

> "Today I can see why so many determinists, and even ex-determinists who believe in the deterministic character of classical physics, seriously believe in a subjectivist interpretation of probability: it is in a way the only reasonable possibility which they can accept: for objective physical probabilities are incompatible with determinism; and if classical physics is deterministic, it must be incompatible with an objective interpretation of classical statistical mechanics"[21]

An epistemic understanding of statistical mechanical probabilities raises the question of what is it that makes the credences specified by the SP the right ones. Sometimes appeal to an *a priori* principle like the principle of indifference is deployed to answer this question but aside from all the problems with such principles it is easy to imagine a world in which these credences are not the right ones. That they are right seems to be a contingent feature of our world. Also the Mentaculus probabilities must be understood as objective if they are to be part of the explanation of the laws of thermodynamics and the other temporal arrows. An ice cube doesn't melt because we believe that its microstate is a good one.

Popper thought of the objective physical probability to evolve in into another state as the degree of the propensity for it to evolve into that state and he thought of propensities as ontologically fundamental. It is not surprising then that he concluded that statistical mechanical probabilities must be subjective if determinism is true since a state has a unique evolution. But there is another way of understanding objective probabilities so that they are non-fundamental physical features of the world that can figure in explanations but are compatible with determinism. This other way is a development of an account originally due to David Lewis. Lewis' account of probabilities is part of his best system account (BSA) of laws. The BSA says that fundamental laws are expressed by certain generalizations or

---

[20] There are some complications. Alison Fernandes (2010) and Matthias Frisch (2010) pointed out that if an actual decision is itself a record of prior events the most similar worlds that contain the counterfactual decision will differ from the actual world in what records they contain. Adam Elga pointed out that there is a complication caused by decisions to influence a past event when there is no record of that event in the actual macro state. Say that Atlantis once existed but there is no record in the actual macro state. Then if D is a counterfactual decision it will be true that if D had been made then Atlantis would not have existed. This is influence but not control since the decider doesn't know what his decision might influence.

[21] Karl Popper (1959)

equations that are implied by the best true systematization of the "the Humean mosaic' (HM) i.e. the totality of fundamental events throughout all of space-time.[22] The best systematization is the one that best balances certain virtues that physicists prize in a theory. Lewis mentions informativeness and simplicity as the virtues that should be balanced but there may be others.[23] Objective probability enters the account as a device to abet systematization. The idea is this. Consider a long sequence of outcomes of coin flips hthhthhhtththhtththhhthtthht..... There may be no simple sentence that is very informative about the sequence. But if a probability function is added to the system then "the outcomes are independent with a probability .5" is both simple and informative. Lewis proposed that a probabilistic law should be evaluated in terms of "fit" by which he means the probability assigned by the laws of candidate systems to the HM.[24] The system that best balances simplicity, informativeness is the law and objective probability giving system.

On the BSA propositions about probabilities and laws are objective features of the HM since they supervene on the HM. They are also contingent and not knowable *a priori* since whether they are true or not depends on features of the actual HM. In particular, actual frequencies are especially relevant to probabilities although systematization allows for probabilities and associated frequencies to diverge. It has been suggested that just because BSA laws and probabilities supervene on the HM they can't play the role that they are usually assigned in explanations. The objection is that since on the BSA laws and probabilities depend on the HM calling on them to explain parts of the HM would make the explanation defectively circular in the way that circular causation would be. I think this is to misunderstand how BSA laws and probabilities explain. They explain not by causing parts of the HM as Popper seems to think propensities do but by their role in unifying the mosaic. More needs to be said about this issue but I leave that for elsewhere.[25]

Although their views about the metaphysics of probability are miles apart Lewis agreed with Popper that non- trivial probabilities are incompatible with determinism. But as we will see his reasons differed. Lewis says

> "To the question of how chance can be reconciled with determinism....my answer is it can't be done....There is no chance without chance. If our world is deterministic there is no chance in save chances of zero and one. Likewise if our world somehow contains deterministic enclaves, there are no chances in those enclaves."[26]

The reason Lewis held that objective probability and determinism are incompatible is connected with his claim that credences (and the actions they rationalize) should aim at or be guided by objective probabilities. He initially formulated this as the Principal Principle (PP) which says

(PP) $C(A/P(A)=x\&E)=x$[27]

Where C is a credence function, P(A) is the objective probability of A and E is information that is "admissible". Lewis doesn't define "admissible information" but he includes laws and information temporally prior to P(A) as admissible. It immediately follows that if determinism is true the laws and the past imply either A or –A and so P(A) equals 1 or 0. However, as we will see it turns out that the BSA can

---

[22] Or at least throughout the observable universe and for 100 billion years after the time of the big bang.
[23] Lewis does not say much about how to measure simplicity and his suggestion for how to measure informativeness is defective. Other virtues may be that the system avoids fine tuning, that the best systematizations of typical regions of space time are the same as the best systematization of the entire universe, and that the system is informative about certain macroscopic phenomena e.g. thermodynamic phenomena.
[24] Fit may not be the best way to evaluate a system with probabilities. It is not clear that it is the best way to capture the informativeness of a system that involves probabilities and it runs into trouble if the HM is infinite.
[25] See Loewer (2012), Lange (2018) Emery (2018), Hicks and Van Elswyk (2015)
[26] Lewis (1982)
[27] (PP) expresses how your belief that the objective probability of A is x should guide your credence. A related principle says that your credence in A should aim to match the objective probability of A.

be naturally extended and a rule that does the work of the PP devised so that non trivial objective probabilities are compatible with determinism.

To extend Lewis' BSA of laws and probabilities so that non-trivial probabilities are compatible with determinism all that needs to be done is to allow probabilities to be assigned to initial conditions (or equivalently to physically possible histories). This fits in quite well with Lewis idea that informativeness is one of the criteria for evaluating systems. By adding a probability distribution we might be able to greatly increase the informativeness of a system with deterministic dynamical laws with little cost in simplicity. This is exactly the case for the Mentaculus. To apply this to the Mentaculus we need to add the criteria that make for the best systematization that the system is informative with respect to facts involving macroscopic properties. Fleshing this idea out requires an account of informativeness that is different from Lewis'. Lewis measure information of a system in terms of the set of worlds it excludes. While this is a sense of informativeness it is not the one that is relevant to evaluating a candidate for the best systematization. A better idea is to measure informativeness of a system in terms of the number and significance of the truths that can be deduced from it and include among these truths ones in the vocabulary of thermodynamics.

The BSA (as modified above) fits the Mentaculus like a glove. My proposal is that the Mentaculus (with the correct dynamical laws) is the best system for our world. It follows that the statistical mechanical probabilities it implies are to be understood in terms of the BSA.[28] The BSA also endorses the claim that the PH is a law since adding it to the dynamical laws and the SP greatly increases the informativeness of the system. It plays a law like role in accounting for the fact that thermodynamic regularities hold since the big bang and that they are laws. It is often claimed that since the PH is incredibly unlikely given the SP it "cries out for an explanation" (e.g. Carroll (2010)). But if it is a law as it is a Mentaculus then it is not unlikely and the feeling that it requires explanation is somewhat alleviated. But even if the PH is a law that doesn't mean that it is a fundamental law. It may have an explanation in terms of fundamental dynamical laws as in the model proposed by Carroll and Chen.[29]

I will conclude by discussing a few objections to the proposal that the Mentaculus is the best system (or a framework for the best system) for our world. There are many objections that I (to say nothing of others) can think of so I will limit myself to few seem to me the most pressing. Most of these come from an interesting paper by Frigg and Hoefer[30] that also construes statistical mechanical probabilities in terms of the BSA but in a very different way that I pursued in this paper

Objection 1. The Mentaculus is a wild speculation. Why believe that it is true?

Reply. My aim wasn't to demonstrate that it is true but provide reasons for taking it seriously. Nevertheless that the case that has been made that it can account for the laws of thermodynamics, for the applicability of the dynamical laws to macroscopic systems, for the arrows of time, and cosmological evidence for the PH all argue in favor of its being correct. That the BSA is the best metaphysical account of laws and probabilities has been much discussed in recent philosophy of science. I don't want to defend it more in reply except to remark that the fact that it can make sense of the probabilities that occur in deterministic theories argues in its favor.

Objection 2. Frigg and Hoefer say that in thermodynamics "the systems under investigation are 'laboratory systems' like gases in containers, liquids in tanks, and solids on tables. This is what SM is mostly applied to in the hands of working physicists, and so there is nothing objectionable about this restriction. However, following Albert (2000), Loewer considerably extends the domain of application of the theory, and treats the entire universe as one large SM system."

---

[28] Eddington famously observed that although physicists' views about the correct fundamental ontology and dynamical laws may change the second law will be preserved. I would make a similar claim about the Mentaculus.
[29] Carroll and Chen (2004)
[30] Frigg and Hoefer (2013)

Reply. This is correct. But there is nothing wrong with applying statistical mechanics to the whole universe. This follows Boltzmann and many others.

Objection 3. But even if the Mentaculus applies to the entire universe this doesn't show that it applies to the "laboratory systems" to which thermodynamics is usually applied. Frigg and Hoefer claim that if the Mentaculus applied to "numerous small subsystems of the world, does happen to make predictions with good fit to the patterns in HM, that is a happy accident, but not something that obviously follows from the global probability rule postulated in Loewer's approach."

Reply. Earlier in this paper I sketched an argument following Albert that makes it plausible that the Mentaculus does imply that thermodynamics applies to 'laboratory systems' that can be considered to be approximately energetically isolated. The fact that it unifies all these applications counts in its favor.

Objection 4. There are probability distributions over the initial conditions compatible with the PH that make for a better fit then the uniform distribution. Frigg and Hoefer say "one can show that the fit of a system can be improved by choosing a peaked rather than a flat (Lebesgue) distribution. A peaked distribution, nearly dirac-delta style, over the world's actual initial condition is, qua postulate, just as simple as a flat distribution, but it will assign significantly higher probabilities to the actual world's macro-state transitions, and hence give the system to which it belongs a much higher fit. So the best system is not one with the flat distribution but one which has a distribution that is peaked over the actual initial condition."

Reply. A distribution that peaks at the actual initial condition is not as simple as the flat (i.e. uniform) distribution since specifying it in an informative way (i.e. not by "the actual initial condition") requires specifying the positions and momenta of 10 to the 80 power particles. Smoothing the distribution out a bit may make it simpler to describe but then it will still be more complicated than the uniform distribution and if sufficiently smooth will agree with the uniform distribution in its consequences for thermodynamics.

Objection 5.The SP component of the Mentaculus is not the only probability distribution over the trajectories compatible with the PH that recovers the second law and the accounts of times arrows. By claiming that the uniform distribution is the correct one the Mentaculus is committed to probabilities for all physically specifiable propositions e.g. the probability that given the current macro state of the universe the president of the U.S. will be impeached within the next year. But this is an enormous over commitment. There is no way to derive this conditional probability from the Mentaculus and unlike the case of repeatable events like melting of ice cubes or the outcomes of coin flips no way to have any idea on the basis of the Mentaculus as to what it is.

Reply. It is correct that many other probability distributions will agree with the uniform one on thermodynamics. But the uniform distribution is the simplest and by the lights of the BSA that qualifies it as the lawful one. As a consequence the Mentaculus will assign a probability to every physically specifiable proposition. If one doesn't like this one can retreat to the view that the SP should be replaced by a family of probability distributions each one of these agreeing on thermodynamics. This revision results in an account that is closer to typicality accounts of thermodynamics.[31] But, on the other hand, one might like the idea that all physically specifiable propositions have objective probabilities. We can then think of subjective credences as aiming to match objective probabilities. For example, my credence that the president will be impeached aims to match the Mentaculus probability that he will be impeached conditionally on the information currently available to me.

Objection 6. As mentioned earlier the Mentaculus requires a revision of Lewis' PP. How will that work?

Reply. The PP says C(A/P(A)=x&E)=x where E is "admissible." Lewis says that laws and truths in the past

[31] See Maudlin's paper in this volume for a discussion of typicality. However there remain important differences between the revised Mentaculus and typicality accounts.

of the time of P(A) are admissible. As pointed out earlier if E is the conjunction of the micro state at some time prior to P(A) and the laws then this implies that P(A) can only take the values 1 and 0. To revise the PP all that is needed is to drop the notion of admissibility entirely and replace PP with MP) C(A/P(A)=x)=x and MPP) C(A/B&P(A/B)=x)=x where P is the Mentaculus distribution. These say that given only that the Mentaculus probability of A (or conditional probability of A given B) has value x one's credence should match the Mentaculus probability (conditional probability). If the conditional probability of an ice cube completely melting in 10 minutes given that the ice cube's macro state is x then your credence that the ice cube will be completely melted in 10 minutes should be x. There is no conflict among MPP, non trivial objective probabilities and determinism. MPP permits inference from conditions at a time t to conditions at prior times as well as inference to future times. So, for example, it licenses an inference from the current macro condition of the smoke in the example in the paper to probabilities of its earlier macro state and its later macro states.

I have endeavored to show in this paper that the hypothesis that the Mentaculus is the Humean best systematization of the universe if true provides a unified account of thermodynamics, the other arrows of time, special science laws, and the natures of the laws and probabilities that occur throughout the sciences. Much more needs to be done to fill in gaps in these accounts before we conclude that the hypothesis is true. But I hope enough has been said to take it seriously.

Albert, D. (2015). *After physics*. Cambridge, MA: Harvard University Press.

Albert, D. (2000) *Time and Chance* Cambridge MA: Harvard University Press

Bennett, Jonathan, *a Philosophical Guide to Conditionals*, Oxford, 2003

Callender, C. (2004) There is not Puzzle about the Low-entropy Past. In Hitchcock ed. *Contemporary Debates in Philosophy of Science*. Blackwell. pp. 240-255 (2004)

Callender, C. and Cohen, J.(2009) "A better Best System Account of Lawhood" *Philosophical Studies* Volume 145, Issue 1, pp 1–34

Carroll, S. (2010) From Eternity to Here. Dutton

Carroll,S. & Chen, J Spontaneous (2004 Inflation and the Origin of the Arrow of Time. arXiv:hep-th/0410270

Elga, A (2001). Statistical Mechanics and the Asymmetry of Counterfactual Dependence. Philosophy of Science (suppl. vol. 68, PSA 2000): S313-S324, 2001.

Emery, N. (2018) Laws and their instances. *Philosophical Studies*.

Earman, John. (2006). “The ‘past hypothesis’: Not even false.”, *Studies in the History and Philosophy of Modern Physics* 37

Fernandes,A.(2010) Time, Flies, and Why We Can’t Control the Past

Frigg, R. & Hoefer, C (2103)The Best Humean System for Statistical Mechanics. Erkenntnis 80(S3) pp.551-574.

Frisch, M (2010) Does a low-Entropy constraint prevent us from influencing the Past?, in *Time, Chance, and Reduction*, eds. Andreas Huttemann and Gerhard Ernst, Cambridge University Press (2010), 13-33.

Hicks, M. (2017). Dynamic humeanism. *British Journal for the Philosophy of Science*.

Hicks, M., & van Elswyk, P. (2015). Humean laws and circular explanation. *Philosophical Studies*, *172*, 433–43.

Lange, M. (2018). Transitivity, self-explanation, and the explanatory circularity argument against humean accounts of natural law. *Synthese*, *195*, 1337–53.

Lewis, D. (1986). *Philosophical papers* (Vol. II). New York: Oxford University Press.

Loewer (2007) “Counterfactuals and the Second Law” In Huw Price & Richard Corry (eds.), *Causation, Physics, and the Constitution of Reality: Russell's Republic Revisited*. Oxford University Press (2007)

Loewer, B. (2012). Two accounts of laws and time. *Philosophical Studies*, *160*, 115–37

Maudlin, T. (2007). *The metaphysics within physics*. Oxford: Oxford University Press.

Penrose, R. (2005) The Road to Reality. Knopf Doublday.

Popper, Karl (1959). “The Propensity Interpretation of Probability.” The British Journal for the Philosophy of Science 10, no. 37: 25-42

Sklar,L. (1993) *Physics and Chance: Philosophical Issues in the*

Foundations of Statistical Mechanics (Cambridge University Press, 1993)