The Reversibility Meta-Objection

Christopher J. G. Meacham

Abstract

Statistical mechanical probabilities are usually understood in one of two ways: as measures of rational indifference, or as lawfully governed chances. Two standard objections to the former are that such probabilities can't plausibly be derived from indifference considerations, nor can they plausibly provide explanations of thermodynamic phenomena. This brief note adds a third objection: the indifference approach to statistical mechanics cannot provide a satisfying response to reversibility worries.

1 Introduction

Statistical mechanics is generally taken to account for thermodynamic phenomena. It assigns high probabilities to the phenomena we see (ice cubes melting in warm water, milk diffusing in coffee) and low probabilities to phenomena we don't see (broken eggs coalescing into whole ones, objects spontaneously sliding across counters).

But what are these probabilities? The traditional stance is to take them to be measures of rational indifference.[1] On this view statistical mechanical probabilities are the credences that an agent should have regarding a system, if she knows only that the system has the stated macroscopic properties. An alternative stance is to take them to be chances, akin to those that appear in quantum mechanics.[2] On this view, statistical mechanical chances stem from the fundamental laws that govern the world.

There are two standard criticisms of the traditional indifference understanding of statistical mechanical probabilities. First, there is skepticism about the tenability and plausibility of recovering statistical mechanical probabilities from indifference considerations. Second, there are worries about whether these indifference-based probabilities can ground statistical mechanical explanations. This brief note adds a third worry: the indifference approach cannot provide a satisfying response to reversibility worries.[3]

[1] For example, see Jaynes (1983).

[2] For example, see Albert (2001) and Loewer (2001).

[3] Reversibility worries have occasionally been flagged as a problem for proponents of the indifference approach who are unwilling to adopt something like the Past Hypothesis, though not a problem for the indifference approach in general; see Sklar (1993). The worry considered here, on the other hand, persists even if the Past Hypothesis is adopted.

2 The Reversibility Objection

Let's restrict our attention to the simplest case: classical statistical mechanics and point particles. Here's a first take on how statistical mechanics assigns probabilities. First fix the static features of the world: its spatial extension, the number of particles, their masses, and the like. Second, construct the phase space representing all of the possible positions and velocities that the particles could have at a given time given these static constraints. We can call the points of this space *microstates*, and regions of this space *macrostates*. Given that the world is in some region B of this space, the statistical mechanical probability of it being in some region A is equal to the proportion of the B region that A occupies, as measured by the Liouville measure.

This characterization of statistical mechanics gets a lot right. If we look at the cup containing a half-melted ice cube in hot water, it will correctly predict that five minutes from now it's overwhelmingly likely that the ice cube will be completely melted, and the water in the cup will be lukewarm. And it predicts this because, according to the Liouville measure, the vast majority of the microstates compatible with the macroscopic description just given will evolve into a no-ice-cube-and-lukewarm-water macrostate.

Likewise, this characterization of statistical mechanics provides successful explanations for many of the entropy-increasing processes which we encounter: why milk diffuses when poured into coffee, why cups don't spontaneously slide from one part of a table to another, and so on. These explanations will have the same form as the explanation just given. High entropy macrostates are much larger than low entropy macrostates. So an arbitrarily wandering microstate is much more likely to wander into a higher entropy macrostate than a lower entropy one.

That said, this characterization of statistical mechanics also gets a lot wrong. Consider the question of what a cup containing a half-melted ice cube in hot water was like five minutes ago. According to the Liouville measure, the vast majority of the microstates compatible with this macrostate will have evolved from a no-ice-cube-and-lukewarm-water macrostate. So this characterization of statistical mechanics predicts that five minutes ago the cup was filled with lukewarm water. But this is the wrong answer: the correct answer is that five minutes ago there was an unmelted ice cube in a cup of very hot water.

The problem stems from the fact that the characterization we've provided is time symmetric. It predicts that systems had a higher entropy in the past in the same way that it predicts that systems will have a higher entropy in the future. But the world isn't time symmetric. Entropy increases toward the future, not toward the past. So it seems that this characterization of statistical mechanics will make all of the wrong predictions about what the past is like.

One might appeal to our evidence of what the past was like to avoid this conclusion: we *remember* putting the ice cube in warm water, for instance. But if we believe the statistical mechanical probabilities, none of the evidence we have access to—memories, photographs, books, fossils—should be believed. After all, statistical mechanics tells us that it's much more likely that our memories spontaneously formed from a skull full of brain soup than it is that they're reliable indicators of a lower entropy past, and that it's much more likely that our history books spontaneously formed from a higher entropy state than it is that they're accurate descriptions of a low entropy past. So, if statistical mechanics is to be believed, none of our evidence about the past should be trusted.

Call this the *Reversibility Objection*. If statistical mechanics is to be a tenable theory, we need a way to overcome this objection.

Two points before we proceed. First, one might argue that if we come to disbelieve our memories, then we won't trust our evidence for believing in statistical mechanics in the first place. So we can't coherently come to doubt our memories because of the statistical mechanical probabilities. Fair enough. But it will still be rationally unacceptable to believe *both* that statistical mechanics is true *and* that our memories are true. And this is an unhappy state of affairs.

Second, one might wonder whether we can mount a case for trusting our evidence after all. While we may not be justified in trusting this book, considered by itself, or these photographs considered by themselves, or various of our memories considered by themselves, maybe the fact that all of them happen to agree with each other gives us reason to believe they're veridical. After all, it seems immensely unlikely that they'd all agree with each other if they spontaneously formed at random from a higher entropy state.

Unfortunately, this response doesn't work. While it's true that such a coincidence would be immensely unlikely, it's even more unlikely that the past was, in fact, in a low entropy state. We can see this by noting two things: (i) higher entropy states are larger than lower entropy ones, (ii) classical dynamics preserves the volume of states over time. So consider the macrostate composed of all the microstates compatible with what we currently believe to be true, including all of the highly correlated macroscopic evidence we have access to. This macrostate has a higher entropy than the one we believe we came from, so it's much bigger than the one we believe we came from. But the dynamics preserves the volume of these states as they evolve over time. So, since the volume of the state we believe we came from is a tiny fraction of the volume of our current macrostate, the vast majority of the microstates in our current macrostate must have evolved from somewhere else. Thus it's overwhelmingly likely that we didn't come from the lower entropy macrostate we believe we came from. So regardless of what our macroscopic evidence is, or how well correlated it turns out to be, it's much more likely to have been fabricated than to have been produced in the way that we're initially inclined to think it was.

If our credences match the statistical mechanical probabilities described above, then we'll make correct predictions about future thermodynamic behavior, but incorrect predictions about thermodynamic behavior in the past. If we want to believe that the world is like we think it is—that our memories are true, and so on—then we'll need to add something to our initial characterization of statistical mechanics.

The canonical response to the Reversibility Objection is to add a *Past Hypothesis*: a law that requires the initial condition of the world be some particular low entropy macrostate. When we add the Past Hypothesis to the canonical probability distribution, everything falls into place: we get the right predictions about both the past and the future, we can trust our memories, and so on.[4]

Note why it's important for the Past Hypothesis to be a *law*, as opposed to, say, a brute stipulation about what the past is like. The problem, in its most forceful form, is that classical

[4]Some have argued that the stipulation of a low entropy initial condition is not enough to overcome the Reversibility Objection (see Winsberg (2004), Parker (2005)). If this is true, further assumptions about the initial state will be needed. But this is orthogonal to the issues we're concerned with here.

statistical mechanics seems to *rationally compel us* to believe something crazy.[5] If we accept non-dynamical chances and classical statistical mechanics, then we'll be rationally compelled to line up our credences with the statistical mechanical chances, and these require us to believe that our memories are probably wrong. Likewise, if we accept that we should be appropriately indifferent over the possible initial conditions compatible with the laws, then given classical mechanics, we'll be compelled to believe that our memories are false.[6] Adopting the Past Hypothesis gets us out of these worries: with the troublesome initial conditions ruled out, statistical mechanics will no longer assign problematic chances, and indifference over initial conditions compatible with the laws will no longer yield a problematic prescription. But this fix won't work if the Past Hypothesis is *just* a hypothesis—we'll be rationally compelled to believe it's false, and the Reversibility Objection will remain unscathed.

3 The Reversibility Meta-Objection

Adopting the Past Hypothesis addresses the Reversibility Objection as posed. But one might wonder whether we can raise the problem in another way.

We can avoid the Reversibility Objection by adding the Past Hypothesis to the laws; call the resulting set of laws CSM_L.[7] Now consider a different set of laws, in which the standard statistical mechanical laws are conjoined with a different law about the initial conditions, one that entails that the entropy of the initial state of the universe was very high; call these laws CSM_H. Likewise, we can contemplate a third set of laws, in which the standard statistical mechanical laws are conjoined with a law about the initial conditions which entails that the entropy of the initial state of the universe was some moderate value (CSM_M), and so on.

What should our initial credences in these various laws be? To make things simple, assume that we're certain about the features (number of particles, etc.) needed to pick out a particular phase space, and that one of these CSM laws holds. Here's one way we could assign our credences. Assign a credence to each CSM proportional to the Liouville measure of the initial conditions they require. If we assign credences this way, and we're Bayesian, then what should we believe the world is like?

Although we'll believe that *if* CSM_L is true then it's likely that our memories are reliable, our credence in CSM_L will be so low compared to its competitors that this will have little impact on our overall credence. The laws that will dominate our credence are the high entropy initial condition ones like CSM_H, which predict that it's much more likely that we coalesced from a high entropy state than that we evolved from a low entropy state. So we'll believe,

[5] That we're so compelled is straightforward on the indifference approach. If we understand statistical mechanical probabilities to be chances, then we'll be compelled if we adopt something like David Lewis' (1986) Principal Principle.

[6] A further worry here is that if the proponent of indifference only requires indifference assignments that mirror the assignments of statistical mechanics, then he'll be left without an 'ignorance' prescription as to how many particles we should take there to be, or what the spatiotemporal extension of the world is, and so on. And this pick-and-choose attitude toward indifference seems implausible. But for present purposes, I'll put this worry aside.

[7] On the indifference approach CSM_L will consist only of classical mechanics and a law about the initial conditions. On the "chancy" understanding of statistical mechanical probabilities CSM_L will include laws regarding the chances of initial conditions.

again, that our memories are overwhelmingly likely to be false. Call this the *Reversibility Meta-Objection*.

The Reversibility Meta-Objection shouldn't worry proponents of the "chancy" approach to statistical mechanics. After all, there's no reason for them to assign initial credences to the CSMs in the manner suggested. They can, with good conscience, assign a high initial credence to CSM_L, and a low initial credence to the other versions of statistical mechanics. This may not be egalitarian, but so what? We've known that we need to do something like this in order to ground our inductive biases since Goodman's (1954) famous grue cases. If we want to end up believing the world is like we think it is, our initial credences will have to be biased towards the theories we actually believe in certain ways.

But the Reversibility Meta-Objection *is* a problem for proponents of the indifference approach. The proponents of the indifference approach use something like an Indifference Principle to recover statistical mechanical probabilities. Given that an agent knows only that the initial state lies in a given region of the phase space, they require her credences over these possibilities to be proportional to the Liouville measure. But if this is how we ought to be indifferent with respect to the initial conditions, it's hard to see why we shouldn't be indifferent in precisely the same way with respect to which initial conditions are lawfully required. The structure of our ignorance with respect to the lawfully required initial conditions mirrors the structure of our ignorance over the initial conditions. So it seems our indifference with respect to the lawfully required initial conditions should mirror our indifference with respect to the initial conditions as well.

To put this another way, consider a particular phase space S, and consider a partition of this space into various macrostates M_i. Let P_i be the proposition that the initial state is in M_i, and let L_i be the law that requires P_i to hold. We can construct a space S^* isomorphic to S by replacing the microstates in each M_i with the corresponding worlds where L_i holds. The proponents of indifference argue that our initial credences in the various P_is should be proportional to the values assigned them by the canonical measure over S. But if they're to avoid the Reversibility Meta-Objection, they need to simultaneously deny that our initial credences in the various L_is should be proportional to the values assigned them by the canonical measure over S^*. And this is hard to justify, since the standard arguments given for the canonical measure of indifference—symmetry, maximizing (informational) entropy, etc.—apply just as well to the latter as to the former.

Proponents of the indifference approach are often motivated by the desire to avoid the metaphysical commitments that statistical mechanical chances would require. They attempt to get around such commitments by providing plausibility arguments for putting strong constraints on what we're rationally allowed to believe. But plausibility arguments can cut both ways. And by employing plausibility considerations to recover statistical mechanical probabilities, the proponents of the indifference approach sabotage their ability to respond to the reversibility worries.

References

Albert, David. 2001. *Time and Chance*. Harvard University Press.

Goodman, Nelson. 1954. *Fact, Fiction and Forecast*. Harvard University Press.

Jaynes, Edwin. 1983. *Papers on Probability, Statistics and Statistical Physics*. Dordrecht: Reidel.

Lewis, David. 1986. A Subjectivist's Guide to Objective Chance. In *Philosophical Papers, Vol. 2*. Oxford University Press.

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

Parker, Daniel. 2005. "Thermodynamic Irreversibility: Does the Big Bang Explain What It Purports to Explain?" *Philosophy of Science* 72:751–763.

Sklar, Lawrence. 1993. *Physics and Chance*. Cambridge University Press.

Winsberg, Eric. 2004. "Can Conditioning on the "Past Hypothesis" Militate Against the Reversibility Objections?" *Philosophy of Science* 71:489–504.