Fried Eggs, Thermodynamics, and the Special Sciences

(*forthcoming* in ***British Journal for the Philosophy of Science***)

Jeffrey Dunn

Abstract: David Lewis ([1986b]) gives an attractive and familiar account of counterfactual dependence in the standard context. This account has recently been subject to a counterexample from Adam Elga ([2000]). In this paper, I formulate a Lewisian response to Elga's counterexample. The strategy is to add an extra criterion to Lewis's similarity metric, which determines the comparative similarity of worlds. This extra criterion instructs us to take special science laws into consideration as well as fundamental laws. I argue that the Second Law of Thermodynamics should be seen as a special science law, and give a brief account of what Lewisian special science laws should look like. If successful, this proposal blocks Elga's counterexample.

1 Introduction

In *Counterfactuals*, David Lewis ([1973]) tells us that, in general, it is a fallacy to strengthen the antecedent of counterfactuals. That is, inferring $((p \wedge q) > r)$ from $(p > r)$ is invalid. But when we evaluate counterfactuals in everyday life, we often must do something very much like strengthening the antecedent. If I hadn't set my alarm last night, then what would have happened? Without some way to strengthen the antecedent—to indicate that we don't stray *too* far from actuality—it's far too indeterminate to say.

Lewis's well-known account of how to fill in the missing information in counterfactuals relies on the comparative similarity between worlds. The counterfactual (p > q) is non-vacuously true just in case in the set of all the worlds, W, where p is true there is at least one world, $w \in$ W where q is true, and there is no world $w' \in$ W that is more or equally similar to the actual world than w and where q is false. Thus, the state of the most similar world at which the antecedent is true does the job of filling in the details left unsaid in the antecedent (Lewis [1986b], p. 41).

Similarity is highly context dependent. In certain contexts one feature of a world matters for similarity, in another context, a different sort of feature of a world matters for similarity. The context dependence of similarity is a strong point of Lewis's account: it mirrors the context dependence of counterfactuals. However, in 'Counterfactual Dependence and Time's Arrow' Lewis ([1986b]) goes further with his account. Although our evaluation of counterfactuals can vary with context, Lewis notes that there is a standard context in which counterfactuals are typically evaluated. Within this context there is an asymmetry of counterfactual dependence: roughly, if the past had been different, the future would be different, but if the future were different, that wouldn't change the past. This asymmetry involves two claims; I will be concerned solely with the first in this paper.

Lewis then proposes his familiar metric that determines the similarity of worlds in the standard context. This is to deliver correct truth-conditions for counterfactuals and thus give us the asymmetry of counterfactual dependence:

(1) It is of the first importance to avoid big, widespread, diverse violations of law.

(2) It is of the second importance to maximize the spatio-temporal region

throughout which perfect match of particular fact prevails.

(3) It is of the third importance to avoid small, localized, simple violations of law.

(4) It is of little or no importance to secure approximate similarity of particular fact.

(Lewis [1986b], pp. 47-8)[1]

Lewis's most general insight concerning counterfactuals—that their truth-conditions depend on a comparative similarity relation—is very plausible, and perhaps correct. However, his more specific claims about similarity in the standard context and counterfactual dependence have come under more serious attack. In this paper I will be primarily concerned with one important counterexample to this similarity metric given by Adam Elga ([2000]) and offer what I think is a Lewisian solution. Lewis's similarity metric appeals only to fundamental laws. The solution I will advocate is to tweak the metric so that matching of non-fundamental, special science laws is taken account of when judging the similarity of worlds. I will argue that this solution blocks Elga's counterexample, and allows us to have the asymmetry of counterfactual dependence that Lewis desired.

It is worth noting that this project is of wider interest than the analysis of counterfactuals. For a critical part of my proposed solution is to appeal to special science laws. Lewis, however, never offers an account of such laws, so part of my proposal will be to sketch a Lewisian account of such laws.[2] Though I will only be focused on showing how such an account can be used to answer Elga's counterexample, and thus get the asymmetry of counterfactual dependence, there are other profitable uses to which an account of special science laws could be put.

---

[1] I will refer to these four constraints as the *similarity metric*. It is worth noting that securing a way to evaluate counterfactuals in the standard context is important to other areas of Lewis's philosophy, most notably his analysis of causation.

[2] In "New Work For A Theory of Universals" ([1999], pp. 42-43), Lewis makes a distinction between the fundamental laws, which are the axioms of the best system, and the derived laws that follow fairly straightforwardly from the fundamental laws. This, however, does not get us the special science laws: laws that may have exceptions and that are not obviously entailed by the fundamental laws.

For one, Lewis's account of laws gives us the fundamental laws: the axioms of the simplest, strongest systematization of the phenomena described in a language that references only perfectly natural properties. But there are almost certainly lawlike relations that are not entailed by the fundamental laws. So in looking only at fundamental laws and their entailments, we are apt to miss out on interesting structural facts about the world: that certain collections of fundamental entities move about in generally lawlike ways, ways which are not entailed by fundamental law. By introducing the notion of a special science laws, we are in a better position to track these interesting structural facts about the world.

Another reason for wanting a Lewisian account of special science laws is because they will be helpful in evaluating a certain special class of counterfactuals.[3] Consider the kind of counterfactual that a biologist might utter:

If it were that A, but the biological laws were the same, then C would have been the case.

What we want to do is to evaluate a certain claim in all of the A-worlds where the biological laws are the same. It is often irrelevant to this kind of counterfactual whether or not we have match of fundamental law. It is a substantive and unsupported hypothesis that any world with matching biological laws is one with matching fundamental laws. What is important is match of biological law. But then we need an account of such laws. An account of special science laws will be of particular service when A is, or implies, the denial of a fundamental law. For instance, Marc Lange ([2000a]) considers the 'area law' of island biogeography, which holds that the equilibrium number of species on an island is proportional to the area of the island ($S = cA^z$).

[3] Later in this paper I will propose a new similarity metric for the evaluation of counterfactuals in the standard context. It is important to note that I am not claiming that that proposed similarity metric (for the standard context) is capable of handling the kinds of counterfactuals discussed in this paragraph. It is plausible that counterfactuals like this, which explicitly mention laws, may take us out of the standard context. My claim is that having an account of special science laws will be useful in whatever account gets these kinds of counterfactuals right.

> For instance, suppose there had been birds with antigravity organs that assist them a little in becoming airborne. […] had there existed such creatures, the island-biogeographical laws would still have held. After all, the factors affecting species dispersal would have been no different. For instance, smaller islands would still have presented smaller targets to off-course birds or driftwood-borne seeds and so would have picked up fewer stray creatures as migrants. ([2000], pp. 236-7)

We'd like to be able to deal with situations like this, where the antecedent of a counterfactual implies that the fundamental laws are violated. Though I will not address such counterlegals in this paper, an account of special science laws would be of clear relevance to such a project. What we'd like to have is some way of meaningfully comparing worlds, all of which exhibit massive violations of fundamental laws. An attractive way of dealing with such scenarios is to appeal to special science laws. But again, to appeal to such laws, we need an account of such laws.

So, though my focus will be on Elga's counterexample and the asymmetry of counterfactual dependence, I think there is independent motivation to pursue an account of Lewisian special science laws. I will sketch such an account, and then attempt to show its usefulness in responding to the problem Elga raises for Lewis's account of the truth-conditions of counterfactuals in the standard context. Before we get to Elga's counterexample, however, there is some groundwork to be done.

2 The Asymmetry of Miracles and Overdetermination

To see Lewis's original similarity metric in action consider the counterfactual:

(Nuclear) *If Nixon had pressed the button, there would have been nuclear holocaust.*

For (Nuclear) to be true on Lewis's account, the similarity metric must ensure that there is some button-pressing-nuclear-holocaust world more similar to the actual world than any button-pressing-no-holocaust world. According to Lewis, there is such a world: it perfectly matches the actual world up until just before Nixon's (counterpart's)[4] decision to press the button. Assuming the actual world to be deterministic,[5] and since Nixon is to press the button in this possible world, the laws of the actual world cannot hold in this possible world. There must, then, be a miracle in this possible world to bring about Nixon's pushing of the button. Miracles are relativized to world-pairs: some possible world exhibits a miracle relative to the actual world just in case there is an event in that possible world that violates the laws of the actual world. The possible world under consideration has such a miracle relative to the actual world. After this miracle-induced button-pressing, however, the deterministic laws of the actual world hold, ensuring a nuclear holocaust in that possible world. Call this *nuclear-holocaust world*. Lewis's similarity metric is designed to ensure that nuclear-holocaust world is more similar to the actual world than any button-pressing-no-holocaust world. This yields asymmetric counterfactual dependence in the standard context.

However, consider a possible world that seems to do just as well on the similarity metric, and yet at which the consequent of (Nuclear) is false. The past of this world is nothing like the past of the actual world in matter of fact, though the deterministic laws of the actual world are obeyed during this time. In this world, Nixon presses the button, uncompelled by a miracle. However, immediately after he presses the button, there is a

[4] Given Lewis's well-known views about possible worlds, Nixon does not literally reappear in other possible worlds; instead, he has counterparts in these worlds. For the rest of the paper I will drop talk of counterparts, but it is to be assumed that when I talk of otherworldly Nixons or otherworldly eggs (etc.), that I mean to be speaking of counterparts of these things.

[5] Throughout Lewis's discussion, and subsequent criticism of Lewis's view, it is assumed that the actual world is deterministic according to the Newtonian dynamical laws. Given this stipulation, we aren't considering the *actual* world since the Newtonian laws are false here. Instead we are considering a deterministic Newtonian correlate of the actual world. I ignore the fact that Newtonian dynamics is not *strictly* deterministic (see Earman [1986]).

miracle that brings this world into perfect match with the actual world. Call this world *different-past world*. Thereafter, different-past world and the actual world evolve (in perfect match) obeying the deterministic laws of the actual world. We might wonder why different-past world isn't at least as similar to the actual world as nuclear-holocaust world. Both feature one miracle, and a *long* stretch of perfect match. Lewis considers worlds such as different-past world:

> But are there any such worlds to consider? What could they be like: how could one small, localized, simple miracle possibly do all that needs doing? How could it deal with the fatal signal, the fingerprints, the memories, the tape, the light waves and all the rest? I put it to you that it can't be done! Divergence from a world such as w_0 [the actual world] is easier than perfect convergence to it. Either takes a miracle, since w_0 is deterministic, but convergence takes much more of a miracle.(Lewis [1986b], p. 49)[6]

Lewis calls this the *asymmetry of miracles* and it is this that accounts for the temporal asymmetry of counterfactual dependence in the standard context.

As Jonathan Bennett ([1984]) notes, there is a problem with this reasoning. Consider the actual world where Nixon fails to press the button. Since the laws are Newtonian, they are time-reversal invariant (where a set of laws are time-reversal invariant iff if a sequence of events from t_0 to t_1 is allowed by the laws, then the reversed sequence of events t_1 to t_0 is allowed by the laws). Given this, take the state of nuclear-holocaust world just after the small miracle causing Nixon to push the button and extrapolate this state forwards and backwards in accordance with the laws. Call this world w_2. It perfectly matches nuclear-holocaust world from the button-pushing and in the future, its dynamical laws are identical to those of the

[6] In this passage Lewis is considering *re*-convergence worlds, but the point is the same.

actual world and are never violated. w_2 is a Bennett world. Bennett notes that whereas the miracle in nuclear-holocaust world is a *divergence* miracle from the actual world's perspective, it is a *convergence* miracle from w_2's perspective. Thus, it is not the case that every world obeys the asymmetry of miracles. Some worlds, like w_2, are such that a small miracle can produce convergence to them (see Fig. 1).

From @'s perspective the miracle at w_1 is a miracle of *divergence*, whereas from w_2's perspective the miracle at w_1 is a miracle of *convergence*. Nevertheless, @ and w_2 have the same laws. Thus, there is small-miracle convergence to a world with our laws. So it doesn't seem that there is an asymmetry of miracles: there are worlds with the same laws as our world, and yet where there is small-miracle convergence.

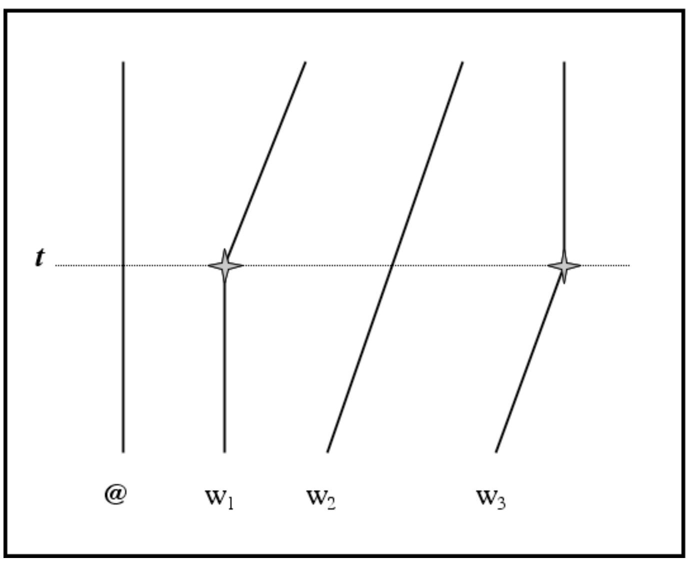

Figure 1: Lines represent worlds, with later times towards the top of the diagram. Four-pointed stars represent miracles. Anywhere that a world's line is straight, the fundamental laws of @ are being obeyed. The actual world is @; w_1 is like nuclear-holocaust world; w_2 is a Bennett world; w_3 is like different-past world.

Lewis answers:

> Same laws are not enough. If there are *de facto* asymmetries of time, not written into the laws, they could be just what it takes to make the difference between a world to which

> the asymmetry of miracles applies and a world to which it does not; that is, between a world like [@] (or ours) to which convergence is difficult and a Bennett world to which convergence is easy. (Lewis [1986b], p. 57)

Lewis claims that w_2 lacks *de facto* asymmetries of time that @ possesses. This accounts for the possibility that a small miracle converges w_1 to w_2, and rules out worlds like w_3 that exhibit small miracle convergences to @. We might be able to get easy convergence to a world with our laws, but the world to which we are converging won't exhibit the *de facto* asymmetries true of our world and therefore won't be as similar to @ as worlds like w_1 that diverge from @.

We might wonder, however, what *de facto* asymmetry guarantees that there will be no small-miracle convergence to the actual world. Lewis tells us that it is the asymmetry of overdetermination. Every fact, Lewis says, 'has at least one *determinate*: a minimal set of conditions jointly sufficient, given the laws of nature, for the fact in question.' (Lewis [1986b], p. 49) If a certain fact has more than one determinate at a time, then it is overdetermined. If the determinate comes before the fact in question, the fact is predetermined, if it comes after, the fact is postdetermined. Our world, Lewis claims, is one where facts are over-postdetermined. That is, there are many determinates of a certain fact, *after* it has happened, each sufficient to determine that fact. However, there are much fewer determinates *before* it has happened. Since we are assuming deterministic Newtonian laws, the complete set of conditions from any instant of time is sufficient for any fact. Thus, in deterministic worlds, every fact has at least one pre-determinate and one post-determinate: the complete determinate. But since our world is one with an asymmetry in favor of post-determinates, this means there must be more of these sufficient sets *after* the fact than there are *before* the fact.

This is supposed to underwrite the asymmetry of miracles as follows. If a world, @, exhibits the asymmetry of miracles, then a small miracle can produce divergence from @, but not convergence to it. Assume that @ exhibits over post-determination, and imagine that we are trying to get a world, w_3, that converges to @ via a small miracle at *t*. This means that at *t*, w_3 and @ match perfectly, even though they do not match before *t*. Since @ has over post-determination, there are lots of determinates after *t* of all the things that happened in @ before *t*. Since w_3 matches @ after *t*, w_3 must have all these same determinates after *t*. But, of course, some of the things that the determinates speak of did not actually happen in w_3: there are many "fake" determinates in w_3 that seem to entail "facts" that didn't actually happen. Every such fake determinate requires a miracle, and so a convergence world will require many miracles (or one big one), rather than one small miracle. So, Lewis claims, if a world exhibited the *de facto* asymmetry of overdetermination, then there is a reason for it to exhibit the asymmetry of miracles. It is the asymmetry of overdetermination that makes our world unlike a Bennett world.

Whatever one thinks of how the asymmetry of overdetermination is meant to ground the asymmetry of miracles, the point is moot: it is false that our world exhibits an asymmetry of overdetermination. To see why, consider an example that Lewis gives of such overdetermination: a spherical wave expanding from a point source to infinity. This happens in our world, says Lewis, but the opposite (a spherical wave contracting from infinity *to* a point source) never does. He writes:

> A process of either sort exhibits extreme overdetermination in one direction. Countless tiny samples of the wave each determine what happens at the space-time point where the wave is emitted or absorbed. (Lewis [1986b], p. 50)

If this were true, then any condition describing a small sample of the wave after it is emitted is a determinate for the emission of the wave from a point source. Since there are many such small samples of the wave after the point of emission, but not before, we would have extreme overdetermination.

Lewis is certainly right that at our world waves do expand to infinity and never contract from infinity, but this isn't an example of overdetermination. Recall, a determinate is *sufficient*, together with the laws of nature, for the fact in question. A set of propositions about a small portion of the wave, however, is not sufficient for its emission from a point. To get sufficiency we must add the further information about what is happening outside this region. Perhaps, for example, the small part of the wave is not a part of a spherical *wave* at all, but merely a part of space that is identical to what this part of the wave would be like, were there a wave. What is sufficient for inferring the point source emission is not some small part of the wave but the complete set of conditions at any time after the fact in question. Such a set is the biggest possible determinate for the fact in question. If the biggest possible determinate is also the minimal determinate, then there is only one determinate for this fact. Thus, there is no overdetermination in either direction, neither past nor future.[7]

Lewis's account is now faced with a problem: The asymmetry of overdetermination would ensure that miracles of convergence to our world are larger than miracles of divergence. This, plus the similarity metric, would give us the asymmetry of counterfactual dependence. But there is no asymmetry of overdetermination at our world. So we are left with no assurance that the asymmetry of miracles is true of our world, and thus no assurance that Lewis's account will yield the desired asymmetry of counterfactual dependence.

[7] After writing this paper, it was brought to my attention that Sanford ([1989], p. 187) makes a similar point.

It is important to emphasize that the claim that the asymmetry of overdetermination is false should not be confused with the claim that there are no *de facto* asymmetries true of the actual world. There do seem to be *de facto* asymmetries at the actual world. One of these is a kind of *epistemic* asymmetry of overdetermination. It *is* true that there are many more conditions after an event, rather than before an event, which we are able to *notice* and that make it reasonable for us to infer that some event occurred. But this asymmetry isn't written into the fundamental laws. Since the similarity metric is only sensitive to fundamental laws, an *epistemic* asymmetry isn't capable of underwriting the asymmetry of miracles.[8]

Given all this, we have no assurance that the asymmetry of miracles is true of the actual world. If it is not, then there is nothing in Lewis's similarity metric that will ensure we get the asymmetry of counterfactual dependence. Lewis's account is primed for a counterexample. Adam Elga delivers one.

3 Elga's Counterexample

Here is Adam Elga's ([2000]) counterexample. In the actual world (@), Gretta cracks an egg at 8:00 am letting it drop onto the hot frying pan in her kitchen. Now, consider:

> (Egg) *If Gretta hadn't cracked the egg, then at 8:05 there wouldn't have been a cooked egg on the pan.* (Elga [2000], p. S314.)

If (Egg) is to come out true, then the closest no-egg-cracking world must be a world where there isn't a cooked egg on the pan. This world perfectly matches @ until just before 8:00, at which point there is a small miracle keeping Gretta from cracking the egg. After this, the laws of @ are obeyed. Call this world w_1.

---

[8] We might ask ourselves *why* there is an epistemic asymmetry overdetermination. It seems that any such *epistemic* asymmetry of overdetermination must be grounded in some *physical* asymmetry. At the end of this paper, I propose one such physical asymmetry.

If the Lewisian strategy is to succeed, we must rule out a competing world, w_3. Before 8:05, w_3 does not match @. In this world the laws dictate that no egg is cracked. However, just before 8:05 there is a small miracle that converges w_3 into perfect match with @ at all later times. If there is such a world, (Egg) is not true.

Lewis claims that there is no w_3 where a *small* miracle leads to convergence with @. Elga, however, gives us a recipe for constructing w_3. We start with the state of @, S_1, at 8:05 just after the egg is cooked, and extrapolate things forward according to the laws. Thus, w_3 matches @ from 8:05 into the future. To construct the before-8:05 section of w_3, we take the velocity-reverse of state S_1 (call it Z_1) and extrapolate backwards according to the laws—but not before making a small miracle change.

Since the process of the egg cooking (from 8:00 to 8:05 in @) is an entropy-increasing process, the reversed process (from 8:05 to 8:00) is an entropy-decreasing process. Now, consider the phase space of the kitchen system, and the subregion of this phase space corresponding to the situation at 8:05: a cooked egg sitting in a frying pan. We know from classical statistical mechanics that the volume of states in this subregion with entropy-*decreasing* futures (states like Z_1) is unbelievably small compared to the volume of the rest of the subregion. Elga concludes that a small miracle moves us from the extremely rare state, Z_1, to a more typical state with an entropy-increasing path. Thus, a small miracle at 8:05 results in normal entropic behavior in the backwards time direction. If we run time forward, we see something strange. At 8:00 in w_3 there is a cooled down and cooked egg in the pan. As we move towards 8:05 the cooked egg spontaneously heats up, as does the pan until it reaches a just-cooked state. Suddenly, there is a small-miracle change bringing w_3 to state S_1. After this the dynamical laws take over ensuring that w_3 perfectly matches the actual world. The egg was never cracked, and yet it is cooked.

4 Possible Solutions

We are now in position to consider solutions that can be made on behalf of Lewis. I will consider solutions that reflect the strategy that Lewis himself articulated in response to Elga's counterexample:

> […] the worlds that converge onto worlds like ours are worlds with counter-entropic funny-business. I think the remedy—which doesn't undercut what I'm trying to do—is to say that such funny business, though not miraculous, makes for dissimilarity in the same way miracles do. (quoted in Bennett [2003], p. 296)

There are several ways to take this comment by Lewis. One option is to construe the counter-entropic funny-business not as violations of *law*, but rather as violations of *kinds* of particular facts. To this end one might invoke criterion 4 of the similarity metric. A different option is to try to construe some law of thermodynamics or statistical mechanics as a fundamental law and thus have the counter-entropic funny-business literally be a miraculous violation of such a law. To this end one might attempt to show that something like David Albert's ([2000]) Past Hypothesis could be a Lewisian fundamental law. A different strategy is to try to construe some law of thermodynamics or statistical mechanics as a *non-fundamental* or special science law, and thus have the counter-entropic funny-business be a non-fundamental miracle. It is this third option that I will pursue.

5 Structure of the Special Science Solution

Call this third option, the *special science solution*. The guiding idea behind this solution is that thermodynamically reversed worlds violate laws of the actual world. However, these violated laws are not *fundamental* laws. Instead, Elga's w_3 will be seen to violate non-fundamental or special science laws. In what follows I will show how one can see thermodynamics as a special science, and thus countenance the Second Law of Thermodynamics as a special science law. If violation of such special science laws is built into the similarity metric, then there is a way out of Elga's argument.[9]

First, however, we need to say something about thermodynamics and particularly the Second Law. Thermodynamics is a science chiefly concerned with macroscopically observable properties of systems. Thermodynamics formalizes the relationship between properties such as temperature, pressure, and volume for macroscopic systems. An introductory text on the topic puts it as follows:

> [Thermodynamics] deals broadly with the conservation and interconversion of various forms of energy, and the relationships between energy and the changes in properties of matter. The concepts of thermodynamics are based on empirical observations of the macroscopic properties of matter in physical, chemical, and biological changes, and the resulting observations are expressed in relatively simple mathematical functions. (Gokcen & Reddy [1996], p. 1)[10]

---

[9] Of course, by adding consideration of all special sciences to the similarity metric (and not just thermodynamics) we'll probably have more than we need to solve Elga's problem. Adding thermodynamics on its own may be enough to get the time asymmetry we're after. So why add in *all* the special science laws? The main reason for adding in all the special science laws is that in the standard context the similarity that special science laws add to a world matter for the evaluation of counterfactuals. For instance, we seem to think that the following is true:

If the apple farmers' crop yield had outrun demand, the price of apples would have gone down.

Now, if we go to worlds where the antecedent is made true by a small fundamental miracle and where there are no other violations of fundamental law or thermodynamic laws, we might only be left with worlds where the economic laws hold. If the fundamental laws together with the thermodynamic laws *entail* that the special science laws stay fixed then this is certainly the case, and so putting them in the similarity metric is superfluous (though it doesn't do any harm). But this might not be the case. And if the thermodynamic laws and the fundamental laws aren't enough to entail the special science laws, then it seems to be a good thing to have them in our similarity metric. Having the special science laws (and not *just* the thermodynamic laws) in our similarity metric does not add anything with respect to Elga's problem. Rather, the idea is that if we add them in to the similarity metric—and it might be nice to do so for reasons just given—then Elga's problem can be taken care of, too. Thanks to an anonymous referee for emphasizing this point.

[10] Another text puts it: 'The task of thermodynamics is to define appropriate physical quantities (the *state quantities*), which characterize macroscopic properties of matter, the so-called *macrostate*, in a way which is as

The Second Law of Thermodynamics is one of the axioms of this scientific theory. Rudolf Clausius formulates it as follows: 'Heat can never, of itself, flow from a lower to a higher temperature.' (Burshtein [1996], p. 277) The Second Law receives a more modern formulation using the concept of entropy.[11] Using this, the Second Law tells us that in a spontaneous evolution of a thermally closed system, the entropy never decreases.[12]

The before-miracle section of w_3 violates the Second Law. Specifically, consider the (approximately) closed system of the kitchen containing the frying pan and the egg. As we move forward in time before the small miracle, the pan spontaneously heats up while the egg spontaneously gets warmer and less decayed. For this to happen, energy is dispersed from the rest of the kitchen and into the frying pan and egg. As the frying pan and egg grow warmer there is spontaneous energy dispersal from the cooler surroundings to the warmer frying pan and egg. This is in direct conflict with Clausius's formulation of the Second Law, but also with the entropy-version of the Second Law. The dispersal of energy from warmer bodies to cooler bodies in a closed system results in a net increase of entropy. Thus, in our kitchen system, there is a net decrease of entropy since there is dispersal of energy from the cooler bodies (the rest of the kitchen) to the warmer bodies (the pan and egg). Now, the kitchen is not an entirely closed system, but as long as there is no offsetting increase in

---

unambiguous as possible, and to relate these quantities by means of universally valid equations (the *equations of state* and the *laws of thermodynamics*).' (Greiner, *et al* [1995], p. 3). Callender ([2006]) offers us the following: 'Thermodynamics is a 'phenomenal' science, in the sense that the variables of the science range over macroscopic parameters such as temperature and volume. Whether the microphysics underlying these variables are motive atoms in the void or an imponderable fluid is largely irrelevant to this science.'

[11] In thermodynamics, the entropy of a system A is usually defined as the integral of dQ/T from some (arbitrarily picked) state B to A of a reversible process, where Q is the energy of the system and T is the temperature. It is important to note that it is a contentious point whether or not the concept of entropy used in classic thermodynamics is the same as the concept of entropy used in statistical mechanics. The relation between statistical mechanics and thermodynamics is interesting, contentious, and difficult to determine. I do not make claims about that relation here. An extremely tentative view about the relation between the two is that statistical mechanics is an attempt to explain how we get the special science laws of thermodynamics, given certain fundamental physical laws. It is important to note that I am attempting to construe *classical thermodynamics* as a special science, not *statistical mechanics*.

[12] See Callender ([2001]); Buckingham ([1964], p. 28); Craig ([1992], p. 39); Lieb & Yngvason ([2000]).

entropy outside the kitchen (as there would be if, say, we were considering a refrigerator as our approximately closed system), then we will have a violation of the Second Law. Further, I follow Elga in assuming that not only will the small miracle put the system onto a backwards-directed entropic path, but this backwards-directed increase in entropy will spread out infecting a larger and larger area as time goes backwards. Here is how Elga describes w_3: 'In the distant past, the infected region is huge. Within that region are events that look thermodynamically reversed. Events outside of the infected region look thermodynamically typical.' (Elga [2000], p. S323) Since the entire world *is* a closed system, we have here a violation of the Second Law. Thus, if the Second Law is somehow worked into the similarity metric, w_3 is less similar to @ than w_1.

One could attempt to construe the Second Law as a fundamental law and work it into the similarity metric in this way. This option is problematic.[13] A Lewisian fundamental law is a theorem in the best deductive systematization of the phenomena. The best system has the best balance of simplicity and strength. Importantly, however, the theorems must reference only perfectly natural properties. If there is no such restriction, then the constraint of simplicity is rendered trivial. The problem with construing the Second Law as fundamental is that the predicate 'entropy' does not seem to refer to a perfectly natural property. [14] The sharing of a certain amount of entropy between two systems doesn't make for similarity in the way that the sharing of perfectly natural properties is supposed to make for similarity. The Second Law would be very complex if formulated by referring only to

[13] This problem with candidate Lewisian fundamental laws is given in Schaffer ([2007]).
[14] The predicate 'entropy' refers to an organizational property that complex systems can have. Take two systems with parts that are intrinsic duplicates of each other and that both instantiate the property of having low entropy. These systems need not be perfect duplicates of each other. According to Lewis, though, things that have exactly the same perfectly natural properties are perfect duplicates of one another. See for instance, Lewis [1986a], p. 61.

perfectly natural properties. Thus it seems unlikely the Second Law is eligible to be a fundamental law of nature.[15]

One might think that there is a different way of getting thermodynamics into the similarity metric. David Albert ([2000]) and Barry Loewer ([2007], [*forthcoming*]), for example, give a different kind of picture where the fundamental laws *entail* the special science laws and the thermodynamic regularities that we observe. However, on this picture, the fundamental laws are somewhat different than expected, consisting of Newton's Laws, the Past Hypothesis—which posits that the universe began in a state of low entropy—and a statistical mechanical probability distribution over initial conditions. This is very interesting proposal, and would allow for a different response to Elga's example. However, let me note four reservations with such an account. First, the Past Hypothesis and the probability distribution are not *regularities*. Lewis ([1999], p. 41) claims that only the regularities in a best system will be laws. Second, if the predicate 'low entropy' does not refer to a perfectly natural property, then the Past Hypothesis does not qualify as a fundamental law. Third, it is an article of faith that these three fundamental laws will *entail* all the special science laws. Certainly we'd expect the special science laws to be roughly consistent with the fundamental laws, but it is a bold claim that they are *entailed* by the fundamental laws. Finally, it is important to note that on a natural way of extending the Albert/Loewer account so that it gives truth-conditions for counterfactuals (see Loewer [2007]), counterfactuals are never fully true or false, but instead have probabilities associated with them. The solution I will advocate does not have this consequence. Although there are unlikely possibilities where a small miracle puts us on an anti-entropic course in the *future* (and so even though Nixon *did* press the button, no nuclear

[15] To avoid this problem, one could argue that the predicates referenced by thermodynamics in fact *are* perfectly natural. A full treatment of this is beyond the scope of this paper. Nevertheless, note that the suggestion moves away from Lewis's conception of perfectly natural properties and so presents a *prima facie* difficulty.

holocaust resulted), such worlds do not compete with the well-behaved worlds because they violate the Second Law. I do not think that any of these points motivate a rejection of the Albert/Loewer picture. However, I think they do motivate an exploration of a different Lewisian solution.

Here, then, is the structure of such a solution. The basic idea is that we work the Second Law into the similarity metric by adding an extra criterion. If Lewis's first three criteria leave us with worlds that are tied for similarity, we turn to the new criterion:

(3.5) *It is of the fourth importance to avoid violation of special science laws.*

Lewis's criterion 4, which does little work for Lewis, is then moved to fifth importance.[16] Worlds that violate special science laws are thus less similar to the actual world than those that do not. For this special science solution to be successful, several things must be shown: (i) it must be shown that thermodynamics can be thought of as a special science, and (ii) it must be shown how special science laws could sit with Lewis's view of lawhood. I will address both of these shortly. But first note how this new criterion in the similarity metric would block Elga's counterexample: w_1 and w_3 are tied for similarity with respect to criteria 1-3, so we turn to criterion 3.5. If thermodynamics is a special science, then the Second Law is a special science law, so the fact that w_3 violates it, but w_1 doesn't counts against w_3. Thus, w_1 is more similar to @ than w_3 and the counterexample is avoided.

But note that even before addressing points (i) and (ii) things are not so simple as this.[17] For the proposed solution assumes that worlds like w_3 and more well-behaved worlds like w_1 will always be tied for similarity up through the first three criteria. But this may not be the case. Say that Gretta's egg-cracking occurs very early in @. Thus, there is a relatively short spatiotemporal region before Gretta cracks the egg and relatively long spatiotemporal

[16] In order of importance, then, the criteria are ordered: (1), (2), (3), (3.5), (4).
[17] Thanks to an anonymous referee for suggesting this objection.

region after she cracks the egg. Now imagine the w_1 and w_3 worlds relative to such an @. Both will be tied with respect to criterion 1 since neither w_1 nor w_3 exhibit big violations of fundamental law. However, w_3 would seem to be favored over w_1 with respect to criterion 2. For w_1 has a relatively short region of perfect match with @ while w_3 has a relatively long region of perfect match. Given this, w_3 is more similar to @ than w_1, and we never have a chance to appeal to criterion 3.5. So, it would seem, the special science solution doesn't solve all Elga-style counterexamples, only counterfactuals with antecedents that aren't relatively early in @'s history. This seems like a problem.

It is useful in working towards a solution to note Lewis's primary purpose for criterion 2. He uses criterion 2 to rule out worlds like w_2 (see fig. 1). But to rule out such worlds, we do not need to quantitatively *maximize* the region of perfect match. Rather, we only need to require that there *is* some region of perfect match. w_2 fails in this respect, so it is ruled out of contention for being most similar to @. So, it seems that we could reinterpret criterion 2. Instead of instructing us to quantitatively maximize the region of perfect match, it is to be modified so that any future perfect match counts for the same as any past perfect match no matter the size of the region of perfect match.

On its own, such a modification is probably inadequate. For criterion 2 may play a secondary role with regard to how early or late a miracle occurs to bring about the truth of the antecedent of the counterfactual. Consider a counterfactual with the antecedent, 'If my alarm hadn't gone off this morning…'. Now imagine two worlds, world *m* where there is a small miracle that causes the clock to malfunction, and world *s* where there is a small miracle the night before that keeps me from setting the alarm. One world has a slightly longer region of perfect match than the other, but according to the suggestion above this does not render *m* more similar than *s*. Addressing the general question of which world *should* be closer to get the right truth conditions for various counterfactuals would go beyond the scope of this paper.[18] But it may be desirable to appeal to maximization of spatiotemporal match in such situations to decide the issue of which world is closer to the actual world. The solution is to note that the second criterion can be formulated so that quantitative comparisons *do* matter when we're considering worlds with different amount of past match, or different amounts of future match, but does not matter when comparing a world with one amount of future match and a different world with a different amount of past match. So, quantitative comparison between two worlds with differing regions of perfect past match or between two worlds with different perfect future match *does* matter for similarity. But we do not quantitatively compare a region of past match with a region of future match. Importantly, this does not stipulate any time asymmetry. It treats the past as something different than the future, but it does not treat the past differently from the way it treats the future. So, no asymmetry of time is put in by fiat in a way that would be unacceptable. Adopting this modification will have the effect that even an @ with an early egg-cracking is a world where w_1 and w_3 are tied for similarity for criteria 1-3. This then allows criterion 3.5 to decide all such Elga counterexamples as desired.[19]

6 The Special Science Solution

---

[18] For some thoughts about the complexity of such considerations, see Bennett ([2003], pp. 209-221).

[19] One might be worried that there is no intuitive support for taking this modification of criterion 2 to be getting at similarity any better than the original criterion 2. There is a sense in which this is correct. However, in this paper I am explicitly adopting Lewis's methodology with respect to similarity metrics. Such metrics, Lewis notes, need not be intuitive in themselves. Rather, the intuitiveness of a similarity metric is to be judged in what it says about the truth-conditions of counterfactuals. As Lewis writes in his ([1986b]):

> […] we must use what we know about the truth and falsity of counterfactuals to see if we can find some sort of similarity relation—not necessarily the first one that springs to mind—that combines with Analysis 2 to yield the proper truth conditions. …we must use what we know about counterfactuals to find out the appropriate similarity relation—not the other way around. (p. 43)

'Analysis 2' refers to Lewis's well-known account of the truth-conditions of counterfactuals that I set out at the start of the paper.

With the structure of the special science solution explained, I will address points (i) and (ii). First, I will show that thermodynamics can be thought of as a special science, and then I will show how special science laws could be accounted for given Lewis's view of lawhood. After this I will consider several objections to the proposal.

6.1 Thermodynamics as a Special Science

Thermodynamics isn't among the paradigm examples of special sciences. But it seems that any plausible way of articulating which features make for a special science will put thermodynamics in with the others.

One key feature of a special science is the following: special science laws hold specifically at their own "level" of inquiry. This idea can be made more perspicuous. Fundamental laws hold between the entities of fundamental physics, even when we consider these entities singly.[20] For instance, the fundamental laws of motion hold between the entities of fundamental physics. These laws of motion, of course, *also* hold between entities of the special sciences – entities like molecules, cells, brain states, and monetary units. A feature of the special sciences is that their laws do not have this wide applicability. Laws of supply and demand simply do not apply to the entities of fundamental physics taken singly. Instead, we need massive pluralities of specially arranged fundamental entities before the laws of supply and demand apply. This is one feature of special science laws. If we consider this feature, thermodynamics looks like a special science. Thermodynamic laws invoke properties like pressure, temperature, and entropy, and macroscopic entities like adiabatic

[20] What is a fundamental entity of physics? Following Lewis, we can think of the fundamental entities as those things—whatever they may be—that form the ultimate parts of the Humean mosaic. There is a worry here, that the entities of fundamental physics might not be single particles, but actually massive pluralities of quantum-entangled particles. This possibility is set aside for two reasons. First, we are considering classical correlates of the actual world. Second, it is not clear that if this possibility damages any of what is said.

systems. Laws invoking such entities and properties do not apply to fundamental entities taken singly.

A second key feature of special sciences is that they are blind, in a certain sense, to what takes place at the fundamental level. In economics, for example, we have laws about supply and demand and as long as we have something that plays the monetary role (and several others), these laws hold. This is true even if fundamental physics were such that collections of different fundamental entities, related in different ways, combined to play that role. Imagine we find that fundamental physics is terribly mistaken. There is a clear sense in which it doesn't matter to economic laws *what* our fundamental science looks like, as long as the monetary role (among others) is filled by *something*. Similar comments could be made about psychological laws and biological laws. If this is a mark of the special sciences, then we again have reason to think that thermodynamics is a special science. The laws of classical thermodynamics are relations between things like pressure, temperature, and entropy. It doesn't matter to thermodynamics whether the systems we are discussing are composed of the fundamental particles of physics so long as the systems have something that play the proper roles. In fact, Lieb & Yngvason ([2000]) derive the Second Law in a way that is completely blind to what the various states are like at the fundamental physical level.[21] Thus, though thermodynamics is not traditionally listed as a special science, it seems that it should be.[22]

---

[21] Also, Greiner, *et al* write: '…classical macroscopic thermodynamics…is of great importance: the concepts of thermodynamics are very general and to a great extent independent of special physical models, so that they are applicable in many fields of physics and the technical sciences.' ([1995], p. 3).

[22] Callender ([1997]) also claims that thermodynamics should be thought of as a special science, though his reasons are different than mine.

6.2 Lewisian Special Science Laws

The next task is to show how special sciences can sit within a Lewisian framework of lawhood. Recall that for Lewis a fundamental law is a theorem (referencing only perfectly natural properties) in the deductive system that is the best balance of simplicity and strength. I propose that a Lewisian special science be thought of as a science that studies a certain subset of all the phenomena and a special science law is a theorem in the best systemization of these phenomena *when we limit ourselves to the vocabulary of a special science*. For example, special science laws in psychology are the regularities in the best system that is formulated using psychological vocabulary. These terms do not refer to perfectly natural properties, and so psychological laws are not fundamental. But they are kinds of laws nonetheless.

There is an immediate worry: according to this account we describe the world using a certain vocabulary. Then we systematize these descriptions in the best way possible. The regularities are our special science laws. But where does this vocabulary come from? It is tempting to claim that any vocabulary is on the table. Describe the world in some vocabulary. If that description can be strongly and simply systematized, then we have some special science laws. This, however, will not work. If we choose bizarre enough vocabularies, then we can get *any* kinds of laws we want. This is not only a problem because such a move would allow us to construct bizarre laws. The primary problem is that by constructing the predicates in the right way, we would be able to make any world violate or agree on as many special science laws of any other world as we want. This would render criterion (3.5) of the similarity metric irrelevant.

The solution to this difficulty is to restrict the vocabularies that can be used. In the fundamental case, Lewis limits the vocabulary to those that reference only perfectly natural properties. In this case, one should say that the vocabularies can only reference *imperfectly* natural properties, those properties the having of which make for similarity. Note that there are two conceptions of imperfectly natural properties. One corresponds to those properties that can be built up out of logical constructions of perfectly natural properties. On this conception, *every* non-perfectly-natural property is imperfectly natural. The other corresponds to what Lewis needs in his response to Putnam's paradox:

> […] the realism that recognizes a nontrivial enterprise of discovering the truth about the world needs the traditional realism that recognizes objective sameness and difference, joints in the world, discriminatory classifications not of our own making […]. What it takes to solve Putnam's paradox is an objective inegalitarianism of classifications, in which grue things (or worse) are not all of a kind in the same way that bosons, or spheres, or bits of gold, or books are all of a kind. (Lewis [1999], p. 67)

This conception is distinct from the first conception. We can see this by noting that the property of being grue is more natural than the property of being a book when we rate things in terms of constructions from perfectly natural properties, and yet Lewis says that grueness is less natural. It is this second notion of imperfect naturalness that is needed to ground special science laws. Lewis already needs and uses such a notion of imperfect naturalness for his philosophy of language (Lewis ([1986a], p. 61, [1999], pp. 13, 49, and 66). Further, this notion of naturalness is not in any way subjective. Lewis appeals to such a notion of naturalness to solve problems of indeterminacy of reference. On pain of circularity, then, it is *not* the fact that we refer to these properties that make them imperfectly natural. Rather, it is their imperfect naturalness that make them eligible referents.

In a recent paper, John Hawthorne speaks favorably about this conception of imperfect naturalness. He writes: 'We should thus be willing to give relative naturalness a life of its own, one that allows properties that are of equal definitional distance from the microphysical ground floor to be of radically unequal naturalness.' ([2007], p. 434) And later, using this notion of imperfect naturalness to get a handle on what he calls 'semantic properties'—the eligible referents of our words—he writes:

> […] once we have relinquished the idea that relative naturalness is to be tied to ease of definability from the ground floor, we are free to take a more elevated view of semantic properties and relations themselves. In particular, we should take seriously the idea that while semantic properties and relations don't occur at the ground floor, they are very natural, not gerrymandered. […] Granted, there is no recipe for generating the semantic properties from the fundamental ones. But given that naturalness does not have to be tied to ease of definability, that does not indict the naturalness and importance of semantic properties. ([2007], p. 435)[23]

This notion of imperfect naturalness is a respectable one, and it is one that is needed elsewhere in Lewis's system. The virtue of such a notion rests in the uses to which it can be put. It can and should be used here.[24]

So, with our vocabulary restricted to referencing the imperfectly natural properties we go about describing the world and trying to uncover regularities. Certain subsets of the phenomena will permit of simple and strong systematization. The regularities of such systems are the special science laws. In practice, of course, the process of coming up with special science laws is intimately related with the process of correctly formulating the

---

[23] For more on this distinction, see Schaffer ([2004]). Lange ([2000b]) also discusses something like this notion of imperfectly natural properties (see, especially, pp. 215-216), but he seems to tie this notion of naturalness to *our* practices in a way that I do not.

[24] It is worth noting that we could rank the imperfectly natural properties in terms of their degree of naturalness, if we desired, via their definability in terms of the perfectly natural properties. In this way, for example, the properties of chemistry might be more natural than the properties of biology. Importantly, however, the properties of both these sciences are members of the privileged class of imperfectly natural properties.

vocabulary to be used. Certain choices of vocabulary simply won't yield a best system that is good enough: the best system that results is one with no regularities, or one that is terribly uninformative. The actual process is one of mutual refinement, in which the formulation of laws and the vocabulary used affect each other. Nevertheless, as a characterization of what a special science law *is*, we can say that a special science law is a regularity in the best system formulated when restricting oneself to the vocabulary of a special science, where a vocabulary of a special science is one that refers to imperfectly natural properties.

There is a remaining worry, however, concerning the fact that special science laws usually have exceptions. The picture I have sketched doesn't seem to allow for such exceptions. Though *incredibly* unlikely, it might be true that in the actual world there are never exceptions to the Second Law.[25] Nevertheless, a general account of special science laws, should show how such laws *could* have exceptions.

One way to think about exceptions follows Jerry Fodor ([1974]). Here is his familiar picture. We have a special science law:

SS-Law: $Sx \rightarrow S^*x$[26]

'S' and 'S*' are predicates of a special science, so they refer to properties that are not perfectly natural. Thus, we can formulate all the ways to be S and all the ways to be S* in terms of perfectly natural properties:

Sx: P_1x or P_2x or … or P_nx

S*x: P^*_1x or P^*_2x or … or P^*_mx

Each of the P_is and P^*_is can be thought of as an abbreviation for a longer specification of the pattern of instantiation of perfectly natural properties. One might think that there are

[25] Of course, this feature might make one doubt that the Second Law is a special science law. I don't think this is right. Exceptions are typical of special science laws, but they need not be necessary.

[26] Fodor, it seems, means to read the '$\rightarrow$' in a metaphysically robust sense. I mean no such thing. The '$\rightarrow$' is just shorthand for 'It is a law that if…then…'.

fundamental laws entailing that each P_ix evolves to some P^*_ix. However, if the SS-Law is not *entailed* by the fundamental laws, then there will be some P_ix that does not evolve to some P^*_ix. This would not necessarily result in an *exception* to the SS-Law unless there actually is some x that instantiates some P_i that does not evolve to P^*_i. But when there is, we get an exception to the SS-Law. Perhaps there is some instantiation of perfectly natural properties by a, P_na, which is also an instantiation of S-ness. However, P_na evolves by fundamental law to some state $P'a$, where $P'a$ is not identical to any P^*_ia. Given this, something that is S does not become S* and we have an exception to our SS-Law.

According to this picture of exceptions, the special science laws need not be consistent with the fundamental laws. The special science law says that S will become S*, but it does not. How, one might ask, can this be? The right thing to say is that it might be that *no* imperfectly natural vocabulary yields perfect regularities, but *some* such vocabulary gives us approximate regularities. Thus, the best systematization of the phenomena described in some imperfectly natural vocabulary is one that doesn't give us perfect regularities, but only approximate regularities. But this is not catastrophic. A certain special science law might come out as a theorem in the best system for whatever vocabulary the special science is conducted in, even if the special science law is not *strictly* consistent with the fundamental laws and so has exceptions. It is a virtue of best system view of laws that there can be laws with exceptions.[27] The gains in simplicity could make it worth adopting a special science law not strictly consistent with the fundamental laws.[28]

[27] In his presentation of the best system account of fundamental laws, Lewis holds that fundamental laws are exceptionless. This is not, however, a consequence of the best system view of laws, but rather an additional constraint on fundamental laws adopted by Lewis.

[28] Obviously this will be a balancing act. If there are too many exceptions no matter how we formulate things, we simply reject the idea that there are any special science laws for the phenomena picked out by the vocabularies in question.

This concludes my sketch of how special science laws can fit within a Lewisian picture of laws. I have also argued that thermodynamics should be treated like a special science, just as we treat chemistry, biology, psychology, or economics. Given all this, the hope is that one special science law that is true of the actual world is the Second Law of Thermodynamics. It unifies a *vast* amount of information about the behavior of *many, many* different macroscopically described systems. Because of the vast number of systems the law applies to (in part due to the imperfectly-natural predicate 'entropy'), the law is extremely simple but also very strong. It serves to unify a myriad of different kinds of time-asymmetric behavior.

It is instructive to pause here and consider some recent work by physicists Lieb & Yngvason ([2000]) on the foundations of thermodynamics. Their work supports the view that the Second Law is a special science law in the Lewisian sense just described. In their paper, they give an axiomatic treatment of entropy and, in turn, are able to formulate the Second Law in what they believe is its most general form. They start with the relation: adiabatic accessibility. A state Y is said to be adiabatically accessible from state X iff Y can be reached from X 'without leaving an imprint on the rest of the universe, apart from the displacement of a weight.' (Lieb & Yngvason, [2000], p. 33) They then give six axioms that restrict how this relation must behave. Finally, they make the assumption they call the Comparison Hypothesis. This hypothesis stipulates that we have enough pairs of states that bear the adiabatic accessibility relation to each other, ensuring that we have enough states that are related to each other in the right sorts of ways so as to be able to derive a lawlike relation. Given, this, they show how to derive a unique entropy function for every state, and a corresponding Second Law that unifies this behavior. Two points deserve note. First, it does not matter to the entropy function or to the Second Law, what the various states are like at the fundamental physical level. All that is required is that a certain accessibility relation holds between the states. This supports the claim that thermodynamics should be seen as a special science. Second, note that the Comparison Hypothesis is important because it ensures that there are enough of the right sorts of regularities to derive the entropy function and the Second Law. So, on this account, the Second Law is straightforwardly a systematization of all the data when we restrict ourselves (via the Comparison Hypothesis) to certain states and systems. This fits in well with the Lewisian conception of special sciences just given. A special science law is a law in the best system, when we restrict ourselves to a certain class of phenomena described in the vocabulary of a special science.

7 Objections

So far I have shown how Thermodynamics may be seen as a special science, and how a Lewisian might offer an account of special science laws. Together with the new similarity metric, this proposal blocks Elga's counterexample. However, one might worry that the new similarity metric will lead to trouble in other areas.

Consider the following objection to the proposed new similarity metric.[29] Grant that biology is a special science, and imagine that there was some critical event that occurred in the past, say a crucial step in the move toward DNA, in spacetime region R, that lead biology on its current course. Let's assume that had this particular critical event not occurred, then biology would have been very different. Now, consider the counterfactual:

(*L1*) *If lightning had struck in region* R*, then the laws of biology might have been very different.*

(*L1*) strikes us as true. But one might wonder how the proposed similarity metric can issue this result. For imagine two worlds, *c* and *d*, tied for similarity according to criteria 1-3. In *c*,

[29] This objection was suggested by an anonymous referee.

lightning strikes in just the particular region of R to prevent the critical event, and biology is radically different. In *d*, lightning strikes in a different part of region R, the critical event occurs, and biology is the same. It looks as if criterion 3.5 will issue the verdict that *d* is closer to the actual world than *c* in which case the following is true:

> (*L2*) *If lightning had struck in region* R*, then the laws of biology would have been just as they actually are.*

But (*L1*) and (*L2*) appear to conflict.

The solution to this difficulty is to say that (*L2*) is true, but that there is a reading of (*L1*) on which it is consistent with (*L2*). There are two ways of reading a counterfactual like (*L1*). Consider:

> (*L1-nwn*) It is not the case that: if lightning had struck in R, then the biological laws would have been just as they actually are.

This does indeed conflict with (*L2*). However, consider:

> (*L1-wbp*) If lightning had struck in R, it would be that: different biological laws are possible.

(*L1-wbp*) says that all the most-similar lightning strike worlds, are worlds where different biological laws are possible. But this is compatible with (*L2*). So, I opt for (*L1-wbp*) as a way of understanding why (*L1*) strikes us as true.[30] One might object that the (*L1-nwn*) reading of

[30] This kind of response is not novel. Lewis makes essentially this sort of move in the postscript of 'Counterfactual Dependence and Time's Arrow' ([1986b]). Lewis is worried about the following sorts of cases. Assume the world is indeterministic. Then there are all sorts of chancy events happening. But now consider the Nixon case. Lewis ruled out the convergence world (different-past world) by saying that it would take a big miracle to get perfect convergence to the actual world where there is no nuclear holocaust. But if there are chancy events, then these chancy events might do the job of the big miracle on their own, converging different past world onto the actual world, without any miracle at all. Lewis's solution is to appeal to what he calls 'quasi-miracles', which are to be chancy lawful events that are very much like miracles, and detract from similarity in the same way. Never mind what one thinks of quasi-miracles. What is important is that Lewis notices a problem:

> For if quasi-miracles make enough of a dissimilarity to outweigh perfect match throughout the future, and if I am right that counterfactuals work by similarity, then we can flatly say that if Nixon had pressed the button there would have been no quasi-miracle. ([1986b], 61)

the might-counterfactual is much more plausible than the (*L1-wpb*) reading. If so, then this is indeed a cost of this view. But it is not a devastating cost as one need not flatly reject the truth of the might-counterfactual in (*L1*).

That is the structure of the response. There is, however, one worry to be addressed. I just said that (*L1*) could be interpreted as

> (*L1-wbp*) If lightning had struck in R, it would be that: different biological laws are possible,

which is compatible with (*L2*). (*L1-wbp*) says that all the most-similar lightning strike worlds (call these the L-worlds) are worlds where different biological laws are possible. What does this actually mean? Well, the L-worlds are all worlds where the biological laws are the same as they are in the actual world. This is what makes (*L2*) true. But, I claim, the L-worlds are also worlds where different biological laws are possible. That is, there are worlds accessible to the L-worlds where there are different biological laws. Now, whether or not this is true depends on what is meant by 'accessible' here. I think we can get the right result if we take the worlds accessible relative to the L-worlds to be the set of worlds, L*, that are most similar to the actual world under all precisifications of the antecedent of the counterfactual we're considering. So, for instance, we can precisify the antecedent by saying that lightning strikes exactly at location x in region R. Then we take all the most-similar worlds relative to that precisification and put them in L*. Then precisify the antecedent by saying that lightning strikes exactly at location y in region R, and take all the most-similar world relative

---

But, of course, this seems wrong, and it seems wrong in just the way that (*L1*) and (*L2*) seemed wrong. Lewis's response is essentially the one I offered above.

There is, however, a slight difference. Lewis's case concerns objective chance, whereas the sort of counterexample we are considering does not deal with situations concerning objective chance. This difference, however, appears inessential to the kind of response. The heart of the response is that such pairs of 'would' and 'might' counterfactuals need not be treated as contradictory. Further, there is something very chance-like going on in the purported counterexample. For why do we think that (*L1*) is true? One might say: 'If lightning were to strike in R, then there'd be some chance that it hits the critical molecule. And if the lightning were to hit the critical molecule, then the laws of biology would be different.' So I think the cases are very analogous.

to that precisification and put them in L*. The set of worlds in L* are then the worlds accessible relative to the L-worlds. Since one way of precisifying the antecedent is a way in which the biological laws end up different (when the strike hits the critical event)[31] it is true to say that relative to the L- worlds, different biological laws are possible. This renders (*L1-wbp*) true as desired.[32, 33]

---

[31] Notice that the counterfactual, 'If lightning struck the critical event, then the biological laws would have been different,' comes out true on my account. If lightning hits the critical molecule, then (presumably) the fundamental laws entail that biology won't go as it actually does. This is, after all, what drives the intuition that (*L1*) is true. So, given such a lightning strike, the only way to get biology to go as it actually does is to have another fundamental miracle. But any world where this is the case is less similar than the non-biology world in virtue of the second fundamental miracle.

[32] A somewhat related worry concerns the following counterfactual, which may seem to come out true on my proposed account:

If it were that A, then there would have been no exceptions to any special science laws.

This sounds bad, and it isn't obviously solved by the solution given in the text. What should we say about this? The answer is to note that the most-similar A-worlds will be worlds where the fundamental laws are obeyed (apart from the small miracle that brings about A). Since we are not assuming that the fundamental laws *entail* the special science laws, there is no reason to think that there will be *any* most-similar A-worlds where there are no exceptions to the special science laws. So we have no reason to think that the counterfactual above will come out true. What criterion 3.5 does is to favor worlds where the special science laws hold with some exceptions to those in which the special science laws are egregiously violated.

[33] The lightning counterexample considered in the text concerns a very unlikely unlawful world that is ruled out of consideration when it shouldn't be. One might wonder if we don't get a counterexample going the other way, where very likely unlawful worlds are ruled out because of the presence of an unlikely lawful world. That is, consider some event, E, that we stipulate will *almost* certainly result in different biological laws. However, we also stipulate, there is one particular way in which E can happen so that the biological laws are the same. Now consider the counterfactual:

(*E1*) *If E had happened, then the biological laws would have been different.*

This counterfactual might strike one as true. But my account would say that it is false, since the one world where the biological laws are the same seems to be ruled as closer to the actual world than the rest. There are three main things to say in response to this. First, it is not obvious that (*E1*) is true. Imagine that I describe a way that I've soaked a match in water so that it *almost* certainly won't light. However, I stipulate that I've soaked it in such a way that there is still one way in which it could light. Now consider the counterfactual:

(*M1*) *If I had struck this match, then it wouldn't have lit.*

In much the same way as the counterfactual above, this counterfactual strikes me as false.

The second thing to say is that E is a very odd sort of event, in that it is stipulated to almost certainly result in different biological laws. In virtue of uttering this sort of counterfactual, it is likely that we are moving farther away from the standard context of evaluation, which the modified similarity metric is to account for.

Third, it is important to note that worlds with different biological laws need not transgress criterion 3.5. Consider a world with no biological properties. The biological laws would have been different in such a world, since there would be no such laws. However, such a world does not violate any biological laws, and so is not ruled out. Alternatively, consider worlds where there are none of the actual biological properties, but different kinds of properties that are similar in certain ways to the actual biological properties. Further, imagine that these properties behave in lawlike ways. One might want to call these the biological laws of these worlds. These, then, will be worlds that have different biological laws. But these worlds do not violate the actual biological laws, and so these worlds will not be ruled out of contention by criterion 3.5. This shows that the following counterfactual will come out true, even if we stipulate that there is one way in which E can happen so that the biological laws are the same:

(*E2*) *If E had happened, then the biological laws might have been different.*

A second kind of worry concerns what to say in situations where special sciences come into conflict.[34] For instance, imagine that we have a counterfactual of the form:

If it were that A, then it would be that C.

Further, imagine that there are two A-worlds, *b* and *p* tied for similarity on criteria 1-3, and that in world *b* C is true and in world *p* C is false. Further, *b* and *p* differ in the following respect: in world *b*, in addition to the small miracle to bring about A, there is also a violation of biological law; in world *p*, in addition to the small miracle to bring about A, there is a violation of psychological law. One might wonder: does the counterfactual come out true or false?

I don't think there is a unique answer to this question. Sometimes context will make clear that one world or the other is the one on which we should be focusing. If we're at a psychology conference, for instance, it is plausible to think that context fixes the psychological laws and so world *b* is preferred. However, if there is no such context to fix things in this way, there is a natural way to break ties. Either the biological law or the psychological law will have more exceptions in the actual world than the other. If the psychological law has more exceptions, then *p* is closer than *b*. If the other way, then *b* is closer than *p*. If they have equal exceptions, then the counterfactual goes indeterminate.

This, then, leaves us with the following modified similarity metric:

(1) It is of the first importance to avoid big, widespread, diverse violations of fundamental law.

(2) It is of the second importance for there to be a spatio-temporal region

---

This seems to be the correct verdict, in just the same way that the following is correct, when considering the match case:

(*M2*) *If I had struck this match, then it might not have lit.*

Thanks to an anonymous referee for mentioning this kind of scenario.

[34] Thanks to an anonymous referee for pushing this worry.

throughout which perfect match of particular fact prevails (with quantitative comparisons for worlds with different-sized past match or different-sized future match).

(3) It is of the third importance to avoid small, localized, simple violations of fundamental law.

(3.5) *It is of the fourth importance to avoid violation of special science laws (with more exceptionless laws preferred in case of ties).*

(4) It is of little or no importance to secure approximate similarity of particular fact.

8 Conclusion

This proposal, if successful, blocks Elga's counterexample. Not only that, but it does so in a satisfying way. First, it seems to accord with Lewis's own views about what a good response to the problem would look like: 'I think the remedy […] is to say that such funny business, though not miraculous, makes for dissimilarity in the same way that miracles do.' (quoted in Bennett [2003], p. 296) This solution shows that the funny business is not *fundamentally* miraculous, but that it does make for dissimilarity in the same way that fundamental miracles do.

Second, this solution seems to correctly locate what is odd about Elga's w_3. Given Lewis's original similarity metric, it is true that w_3 is just as similar to the actual world as w_1. But this seems wrong to us. The reason that it seems wrong, I conjecture, is the region of anti-entropic behavior. This solution identifies that region as the problem.

Third, this solution does not simply *stipulate* the Second Law or the asymmetric nature of it. Rather, this solution tells us to take account of any special science laws in the evaluation of similarity. Since one true special science law of the actual world is the asymmetric Second Law, *this* fact gets us the asymmetry.

Fourth, and importantly, the holding of the Second Law at our world can plausibly provide the physical basis for the epistemic asymmetry of overdetermination mentioned in section 2. At least one true epistemic asymmetry of overdetermination is that there are many macroscopic approximate determinates after an event (the smoking gun, the bloody glove, the footprint, etc.) and very few (if any) macroscopic approximate determinates before an event. A physical world that has such an epistemic asymmetry of overdetermination will have it in virtue of some physical asymmetric macroscopic regularities. These asymmetric macroscopic regularities are plausibly the regularities that special science laws in general, and the Second Law in particular, capture. Thus, worlds that exhibit small-miracle convergence to the actual world *do* lack certain macroscopic asymmetric regularities of the actual world. And a similarity metric that takes into account the Second Law is able to give a principled reason for ruling out such easy convergence worlds.

Acknowledgements

Special thanks to Phil Bricker and Chris Meacham for very helpful discussion on issues related to this paper. Thanks also to the participants of the 2008 Central APA session where a version of this paper was read, especially Valia Allori who provided written comments. I also appreciate the helpful comments from the participants of the UMass Philosophy Graduate Student Workshop, and two anonymous referees for this journal.

References

Albert, D. Z. [2000]: *Time & Chance*, Cambridge, MA: Harvard University Press.

Albert, D. Z. [1994]: 'The Foundations of Quantum Mechanics and the Approach to Thermodynamic Equilibrium', *British Journal for the Philosophy of Science*, **45**, pp. 669-677.

Bennett, J. [2003]: *A Philosophical Guide to Conditionals*, New York: Oxford University Press.

Bennett, J. [1984]: 'Counterfactuals and Temporal Direction', *Philosophical*

Review, **93**, pp. 57-92.
Buckingham, A. D. [1964]: *The Laws and Applications of Thermodynamics*, Oxford: Pergamon Press.
Burshtein, A. I. [1996]: *Introduction to Thermodynamics and Kinetic Theory of Matter*, New York: John Wiley & Sons, Inc.
Callender, C. [2006]: 'Thermodynamic Asymmetry in Time', *The Stanford Encyclopedia of Philosophy (Fall 2006 Edition)*, Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/fall2006/entries/time-thermo/>.
Callender, C. [2001]: 'Taking Thermodynamics Too Seriously', *Studies in the History and Philosophy of Science Part B*, **32**, pp. 539-553.
Callender, C. [1997]: 'What is 'The Problem of the Direction of Time'?', *Philosophy of Science*, **64**, pp. S223-S234.
Craig, N. C. [1992]: *Entropy Analysis: An Introduction to Chemical Thermodynamics*, New York: Wiley-VCH.
Earman, J. [1986]: *A Primer on Determinism*, Dordrecht: D. Reidel Publishing Company.
Elga, A. [2000]: 'Statistical Mechanics and the Asymmetry of Counterfactual Dependence', *Philosophy of Science*, **68**, pp. S313-S324.
Fodor, J. [1974]: 'Special Sciences', *Synthese*, **28**, pp. 97-112.
Gokcen, N. A. & Reddy, R. G. [1996]: *Thermodynamics (Second Edition)*. New York: Plenum Press.
Greiner, W. *et. al.* [1995]: *Thermodynamics and Statistical Mechanics*, Translated: Dirk Rischke, New York: Springer-Verlag.
Hawthorne, J. [2007]: 'Craziness and Metasemantics', *Philosophical Review*, **116**, pp. 427-440.
Lange, M. [2000a]: *Natural Laws in Scientific Practice*, New York: Oxford University Press.
Lange, M. [2000b]: 'Salience, Supervenience, and Layer Cakes in Sellars's Scientific Realism, McDowell's Moral Realism, and the Philosophy of Mind', *Philosophical Studies*, **101**, pp. 213-251.
Lewis, D. [1999]: *Papers in Metaphysics and Epistemology*, New York: Cambridge University Press.
Lewis, D. [1986a]: *On the Plurality of Worlds*, Oxford: Blackwell.
Lewis, D. [1986b]: *Philosophical Papers, vol. II*, New York: Oxford University Press.
Lewis, D. [1973]: *Counterfactuals*, Malden, MA: Blackwell.
Lieb, E. H. & Yngvason, J. [2000]: 'A Fresh Look at Entropy and the Second Law of Thermodynamics', *Physics Today*, **53**, pp. 32-37.
Loewer, B. [*forthcoming*]: 'Why There is Anything Except Physics', in *Being Reduced: New Essays on Reduction, Explanation, and Causation*, Hohwy, J. and Kallestrup, J. (*eds*), New York: Oxford University Press.
Loewer, B. [2007]: 'Counterfactuals and the Second Law', in *Causation, Physics, and the Constitution of Reality: Russell's Republic Revisited*, Price, H. and Corry, R. (*eds*), New York: Oxford University Press, pp. 293-326.
Loewer, B. [2001]: 'Determinism and Chance', *Studies in the History and Philosophy of Modern Physics*, **32**, pp. 609-620.
Sanford, D. H. [1989]: *If P, then Q: conditionals and the foundations of reasoning*, London: Routledge.
Schaffer, J. [2007]: 'Deterministic Chance', *British Journal for the Philosophy of Science*, **58**, pp. 113-140.

Schaffer, J. [2004]: 'Two Conceptions of Sparse Properties', *Pacific Philosophical Quarterly*, **85**, pp. 92-102.