sciendo

Kairos. Journal of Philosophy & Science 22, 2019
Center for the Philosophy of Sciences of Lisbon University

Should Special Science Laws Be Written into the Semantics of Counterfactuals?

Daniel Dohrn
Università degli Studi di Milano,
Dipartimento di Filosofia
daniel.dohrn@unimi.it

DOI 10.2478/kjps-2019-0010

Abstract Adam Elga has presented an anti-thermodynamic process as a counterexample to Lewis's default semantics for counterfactuals. The outstanding reaction of Jonathan Schaffer and Boris Kment is revisionary. It sacrifices Lewis's aim of defining causation in terms of counterfactual dependence. Lewis himself suggested an alternative: «counter-entropic funnybusiness» should make for dissimilarity. But how is this alternative to be spelled out? I discuss a recent proposal: include special science laws, among them the laws of thermodynamics. Although the proposal fails, it serves to uncover the limits of Elga's example.

Keywords: Counterfactuals; Lewis; Similarity; Special Science Laws; Elga.

1. Introduction

Adam Elga has presented an intriguing counterexample to Lewis's default semantics for counterfactuals.[1] I discuss a recent proposal how to deal with Elga's example. To summarize Lewis's semantics: his truth-condition for counterfactuals is:

1 Elga, 2001.

A counterfactual p□→q is non-vacuously true iff some p and q - world is more similar to the actual one than any p and not-q-world.[2]

Following Elga, I will talk as if there were a set of closest p-worlds. For the deterministic case, Lewis presents a default similarity ordering of worlds:

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

(2) It is of second importance to maximize the spatio-temporal region throughout which perfect match of particular fact prevails.

(3) It is of third importance to avoid even small, localized simple violations of law [small miracles].

(4) It is of little or no importance to secure approximate similarity of particular fact, even in matters that concern us greatly.[3]

2. Elga's Counterexample to Lewis

I sketch Elga's counterexample to these criteria. He considers

> «At 8:00, Gretta cracked open an egg onto a hot frying pan. According to the analysis, are the following counterfactuals true?
>
> [C1] If Gretta hadn't cracked the egg, then at 8:05 there wouldn't have been a cooked egg on the pan.
>
> [C2] If Gretta hadn't cracked the egg, then at 7:55 she wouldn't have taken an egg out of her refrigerator».[4]

Consider two competitors for closest non-cracking worlds: Lewis's favourite W_2 perfectly matches the actual world W_1 until shortly before 8:00. At that point, a small miracle occurs such as to prevent the egg from being cracked. Then W_2 develops according to the laws such as to never again perfectly match the actual world. W_3, in contrast, differs from

2 cf. Lewis, 1986, 41.

3 Lewis, 1986, 47–48.

4 Elga, 2001, S314.

the actual world before 8:00. The egg is not cracked and not taken out of the refrigerator at 7:55. Some time after 8:00 W_3 converges to the actual world by dint of a small miracle. Lewis insists that the W_3 - strategy is not feasible. For any normal event leaves many and varied traces. It would need many and varied unlawful events, a big miracle, to suppress these traces.[5] So the asymmetry is this: certain divergence worlds are better candidates for closeness than any convergence worlds.

Elga sets out to show that pace Lewis there are candidate worlds for W_3 that need nothing but a small miracle.[6] He confines his discussion to the framework of deterministic statistical mechanics. A physical state can be completely described by the positions and momenta of particles. Starting from the egg in the pan at 8:05, Elga contrasts a normal *future* development to a thermodynamically atypical future development. The latter reverses the actual forward-directed development of the egg between 8:00 and 8:05: the egg uncooks and jumps back into the shell. The reversed development is extremely unstable. Consider its initial state: the closest neighbours to this state in phase space are states which differ from it just in a small local group of molecules. It needs only a tiny change at 8:05, a small miracle, to proceed to one of these states. But the tiny difference very quickly spreads such as to give rise to a completely different, thermodynamically normal development. A tiny change leads to a completely different result. Now an analogous situation can be achieved by running the symmetrical laws *backwards*: the actual past development of the egg is a precise mirror image of the atypical future development. From the viewpoint of statistical mechanics, both are equally atypical. Judging the situation of the cooked egg by statistical mechanics alone, we should expect entropy to increase towards the past. Take again the imagined *future*-directed development of the egg jumping back into the shell. Now imagine a slightly varied initial situation. A completely different, thermodynamically normal future development ensues. Entropy increases towards the future. The same goes for the past. Take a possible situation which is slightly different from the actual one immediately before 8:05. Instead of the actual *past* of the egg (cracking and cooking),

5 Lewis, 1986, 47.

6 Elga, 2001, S318.

we get a completely different, thermodynamically atypical past development. Elga envisages a process of «reversed rotting»: a normal future development of the egg is reversed and projected into the past. A coniform «infected region» comprises this complete development from the very distant past up to the point of convergence at 8:05. In the distant past, the infected region was huge.[7] But it rapidly shrank and gave way to thermodynamically normal developments up to the small miracle immediately before 8:05.[8] The egg has never been taken out of the refrigerator, never been cracked. Nevertheless, the thermodynamically reversed development eventually approximates the actual development. A small miracle administered immediately before 8:05 yields perfect match. Since this match is perfect, it comprises all the alleged traces of the cracking.

As it seems, we cannot rule out that the Elga world is closer than all competing non-cracking worlds. It perfectly abides by the actual fundamental laws of nature except for a small miracle, and it counters the perfect pre-antecedent match Lewis's candidate worlds exhibit by perfect post-antecedent match. If W_3 is closer than its competitors, for instance because the end of the world is not far away, Lewis's criteria have the wrong counterfactuals come out true, for instance the intuitively false

(C2) If Gretta hadn't cracked the egg, then at 7:55 she wouldn't have taken an egg out of her refrigerator.

If W_3 is as close as W_1, neither C1 nor C2 are true. So Lewis's similarity metrics grossly misses our common counterfactual verdicts.

3. The Revisionary Reply: Writing Causal Dependency into the Similarity Metrics

Jonathan Schaffer provides the following strategy against Elga: one should only maximize match in facts *from regions causally independent of whether or not the antecedent obtains*:

7 Why is the infected region coniform? Imagine a thermodynamically abnormal region with *future* developments of the egg-jumps-back-into-the-shell-sort. Varying a few molecules at some point within this region initiates an «infection» by thermodynamically normal developments which rapidly grows as the molecules begin to interact with others. Elga's region which grows towards the past is a mirror-image of such an infection.

8 Elga, 2000, S323.

(1c) It is of the first importance to avoid big miracles.

(2c) It is of the second importance to maximize the region of perfect match, *from those regions causally independent of whether or not the antecedent obtains*.

(3c) It is of the third importance to avoid small miracles.

(4c) It is of the fourth importance to maximize the spatiotemporal region of approximate match, *from those regions causally independent of whether or not the antecedent obtains*.[9]

Boris Kment has recently presented a version of this strategy in terms of sameness of explanatory history.[10] On balance, in spite of criticism,[11] I deem the Schaffer-Kment strategy the most promising way of dealing with Elga's counterexample and other recent revivals of the future similarity objection.[12]

However, from an *orthodox* Lewisian viewpoint, the Schaffer-Kment strategy seems unattractive. It forestalls Lewis's aim of defining causation in terms of counterfactual dependence. Such a definition is the boldest and most straightforward version of a strand in the philosophy of causation which has grown ever more important in recent years.[13] But if causal dependence is written into the very metrics which determines the truth of counterfactuals, this is irreconcilable with the aim of defining it without circularity.[14]

9 Schaffer, 2004, 305.

10 Kment, 2006.

11 cf. Noordhof, 2005.

12 At least within a nearness semantics as Lewis presents it, cf. Schaffer, 2004, Wasserman, 2006.

13 cf. the literature in Hitchcock, 2011, 230–231. To be sure, not anyone working on this project aims at a non-circular definition of causality.

14 Admittedly, Schaffer's approach is Lewisian in sharing Lewis' truth-condition for counterfactuals and most of Lewis' four-part lexical similarity ordering. Moreover, far from announcing his ideas as un-Lewisian, Schaffer takes great pains at reconciling the use of causal terms with Lewis' metaphysical ambitions:
«Might one adopt *both* a causal independence account of counterfactuals, *and* a counterfactual account of causation? Is the resulting circularity *problematic*? *Ontologically* speaking, I see nothing problematic here. The truth about both counterfactuals and causality still *supervenes* on the arrangement of events. Or at least, nothing here contradicts that. The causal and counterfactual facts can still, for instance, be regarded as «co-supervenient» upon a Humean base. If there were a problem, it could be a *conceptual* problem.

So if one looks for a Lewisian strategy of answering Elga's challenge (not taking into account other new variants of the future similarity objection), it might seem worthwile to look for one which does not write outright causal relationships into the similarity metrics.

4. A Conservative Proposal: Writing Special Science Laws into the Similarity Metrics

4.1 Summary of the Proposal

Lewis himself hints at such a strategy. Bennett reports his reaction to Elga's counterexample:

> «The worlds that converge onto worlds like ours are worlds with counter-entropic funnybusiness. 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».[15]

Lewis suggests that counter-entropic regions should play a role comparable to the role of miracles: Worlds like Elga's W_3 are disqualified as candidates for being closest antecedent worlds because they exhibit counter-entropic funny-business. But how is «funnybusiness» to be spelled out as a general criterion? Jeffrey Dunn has recently presented such a criterion which pays due respect to the intuitive weirdness of counter-entropic processes.[16] While Lewis' original metrics draws on fundamental laws, Dunn also takes into account the laws of the special sciences,

---

One would lose *linear definability* – no ordered chain of definitions could wind from the Humean base up through the conceptual superstructure. But perhaps linear definability was never in the offing. *Because concepts do not have definitions».* (Schaffer, 2004, 307–308.)

Schaffer intimates that preserving Lewis' conception of supervenience matters more than the non-circular definability of causality by counterfactuals: the truth about both causality and counterfactuals supervenes on the mosaic of events; but there is no non-circular definition of concepts like causality.

15 Quoted in Bennett, 2003, 296.

16 I note in passing that Dunn misconstrues Elga's argument. He thinks that Elga just takes the velocity reverse of the actual scenario at 8:05: the cooked egg sitting in the pan (Dunn, 2011, 80). But it is clear that the reversal of all velocities constituting the scenario cannot be reached by a small miracle.

among them «lawlike relations that are not entailed by the fundamental laws».[17] In his Lewisian view, «a special science law is a theorem in the best systemization of … phenomena when we limit ourselves to the vocabulary of a special science».[18] He gives laws of the special sciences fourth importance, demoting Lewis's fourth criterion (concerning the role of approximate match) to fifth importance:

(4´) It is of the fourth importance to avoid violation of the special science laws.[19]

The Elga world is disqualified as a candidate for closeness. It violates the laws of thermodynamics, construed as a special science. Briefly, Dunn invokes the second law of classical thermodynamics: heat cannot spontaneously flow from a hotter location to a cooler location.[20] Yet that is what Elga's reverse process would amount to. The backwards rotting egg would have to absorb heat from its relatively cool surroundings in order to end in the pan at cooking temperature. Lewis's candidate worlds abide by the second law of thermodynamics while the Elga world violates it. So the former are closer according to Dunn's criteria.

I think that Dunn's proposal deserves closer attention independently of dissolving Elga's problem. Counterfactuals play an important role in the special sciences (for instance in functional biology.[21] One may wonder whether the role of special science laws in such counterfactuals is exhausted by Lewis's criteria:

> «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:
>
> [C3] If the apple farmers' crop yield had outrun demand, the price of apples would have gone down.

17 Dunn, 2011, 73.

18 Dunn, 2011, 88.

19 Dunn, 2011, 84.

20 For a more detailed modern formulation cf. Dunn, 2011, 82.

21 cf. Lange, 2004.

> 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».[22]

In my critical discussion of Dunn's approach, I will proceed as follows: in the next section, I will point out several problems of integrating special science laws into Lewis's metrics. In section (4), I draw a lesson from my criticism: Dunn's proposal ultimately fails. Yet as long as there might be fundamental laws which underlie the thermodynamic asymmetry, there is hope for Lewis's *original* metrics.

4.2 Criticism: Why Special Science Laws Should Not Be Written into the Similarity Metrics

My first three queries concern the special science solution in general, the fourth its application to Elga's counterexample.

(4.2.1.) I begin with Dunn's own independent motivation for integrating the special science laws, the intuitively true apple price counterfactual:

(C3) If the apple farmers' crop yield had outrun demand, the price of apples would have gone down.

I take Dunn's basic argument to be the following: Lewis maintains that fundamental laws do not allow for exceptions.[23] Thus, the closest miracle-worlds (relative to the actual fundamental laws) are worlds where at least some actual fundamental laws are no laws. But as witnessed by the apple counterfactual, the special laws of economics obtain.

22 Dunn, 2011, 81 ann. 9.

23 Lewis, 1986, 45.

Assume the varied fundamental laws in the closest situation where the antecedent of (C3) is true entail the laws of economics. Then (C3) is ensured by Lewis's original metrics and does not motivate Dunn's amendment. Assume that the special science laws at stake are not entailed by the fundamental laws. Still we do not need (4´) to account for (C3). In a Lewisian account, the special science laws must somehow supervene on fundamental laws and particular matters of fact. In dealing with the counterfactual supposition, we minimally revise the fundamental laws. And we hold onto perfect match in particular matters of fact as far as reconcilable with the antecedent and the revised laws. Provided the revised laws and particular matters of fact support the special science laws, the latter are preserved. This is why (C3) seems acceptable. The revised fundamental laws and particular matters of fact are minimally modified by the everyday situation of the crop yield outrunning demand. The result also supports the usual functional relationship of supply and demand. Just imagine an increase in apple supply. We hold fixed (the fundamental facts underlying) the rational behaviour of apple traders. They will react in their usual way: sellers compete in prices, there are few buyers, the price will go down. Things would be different if the antecedent required a *substantial* change in actual matters of fact: a change such that actual matters of fact do not any longer support certain special science laws. Dunn does nothing to show that in such a situation, we would hold onto the special science laws in question. In contrast, the arguments to come point in the opposite direction.

However, an argument by Frigg and Hoefer seems to support Dunn's claim. They reject the counterfactual

[C4] If the physics of our world was vastly different from what it is, then the chance for heads should be different too![24]

Frigg and Hoefer are silent on the issue of special science laws. Yet we may suspect that there is a comparable regularity involved. For coin-like set-ups, the general chance of a particular outcome is 1/n where n is the number of sides of the coin-like object. To Frigg and Hoefer, the

24 Hoefer and Frigg, 2010, 365.

sort of general patterns which underpin objective chances might obtain in worlds which are microphysically very different from ours:

> «…these sorts of patterns may obtain even in worlds with radically different micro-laws. Imagine a universe in which matter is a continuum and obeys something like the laws of Cartesian physics; imagine that coins exist in this universe and are tossed repeatedly. Despite the basic physics being very different, suppose it turns out that the overall pattern of outcomes of rolls/tosses of such n-sided objects in the continuum universe's HM [Humean mosaic] is very similar to the pattern in our universe. What would the chance of heads be in the continuum universe? Clearly it would be given by the 1/n rule, since this is the best rule relative to the HM, irrespective of the microconstitution of matter».[25]

However, we must distinguish two readings of the consequent. In one reading, «heads» means «heads within the undisturbed social practice of tossing a coin». It is a necessary feature of the undisturbed social practice that the coin is fair, i.e. the 1/n rule holds. Then we may read the counterfactual as subject to the following implicit qualification:

(C5) If the physics of our world was vastly different from what it is, then, if there were a chance of a coin being tossed at all, the chance for heads would not be different.

Assuming import-export [(p□→(q□→r) ⊃ (p&q)□→r)], this boils down to (C5´) If the physics of our world was vastly different from what it is and there was a chance of a coin being tossed at all, the chance for heads would not be different.

(C5´) seems true as the consequent involves Ch(heads/a coin is tossed within our undisturbed social practice).

If, in contrast, we just consider the coin as a physical object which is subject to a certain kinematics, I see no reason to accept

(C6) If the physics of our world was vastly different from what it is, then the chance for heads would be the same.

25 Hoefer and Frigg, 2010, 368.

And (C6) is what Dunn would need for Frigg and Hoefer's observations to support his case. Still I would join Frigg and Hoefer in rejecting (C4). The reason will become obvious in a moment.

I do not deny that, in order to evaluate counterfactuals within a special science, the laws of this science must be upheld. But I do not see why, say, biology should require to uphold the laws of economics. Rather any special science may create its own non-default context. But this has no impact on Lewis's default similarity ordering. If one eschews non-default contexts for special sciences, alternatively one may impose criteria of selection. These criteria select precisely those counterfactuals which are relevant to a special science among all the counterfactuals which are true by default. For instance, biologists may be interested in selecting antecedent events which *could* have formed part of normal biological functioning (although the closest world is one where they have come about by a small miracle).

(4.2.2) The Elga world is described in terms of statistical mechanics. Dunn confesses his uncertainty how the latter relates to thermodynamics:

> «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».[26]

Dunn's treatment raises a question: assume certain laws of the special sciences can be reduced to more fundamental laws and facts. What should the place of these laws in the similarity ordering be?

One may deny that this is a problem, arguing as follows: assume a special science law is *entailed* by the fundamental laws. For the special science law to be violated, the fundamental laws must be violated as well but not vice versa. So in case of reducibility, Dunn's criterion yields the following result: a world where the fundamental laws are violated but

26 Dunn, 2011, 82.

the special science laws are not is closer, other things being equal, than a world where the special science laws are also violated.

To elaborate the problem, I begin with an example of Kment's. He discusses the counterfactual dependence of particular facts on fundamental laws:

> «[C7] If (Law of Gravitation) had not been a law, then events would still have at least approximately conformed to it.
>
> No one I asked believed that this counterfactual was true…
>
> [C8] If the master law [comprising all fundamental laws] had not been a law, the history of the world would still have been very similar to what it was actually like».[27]

Kment suggests that there is an explanatory relationship between laws and particular matters of fact. This relationship makes us reject (C7) and (C8). In the closest counterfactual situation where the actual *explanans* does not obtain, we do not hold onto the *explanandum* either. I do not commit myself to this view of explanation. Yet I maintain that there is a parallel counterfactual dependence of special science laws on laws they can be reduced to. I remain neutral as to whether this dependence ultimately enforces the Schaffer-Kment amendment. There is a close parallel to the example of Frigg and Hoefer. They accept an explanatory relationship as envisaged by Kment: «…there is the pervasive intuition that if a macro result can be derived from a more fundamental theory, there is explanation».[28]

According to Frigg and Hoefer, there is an explanatory relationship between microphysics and the macroscopic chance of getting heads in a coin toss. Now by Kment's lights, in a situation where the microphysics is different, we do not hold onto what is explained by it. So on the one hand, we do not have any reason to hold onto the chance of *heads*.

27 Kment, 2006, 280, 281.

28 Frigg and Hoefer, 2010, 366.

On the other hand, however, there is no reason why it must be *different* from the way they actually are. Consider again:

(C4) If the physics of our world was vastly different from what it is, then the chance for heads should be different too.

We do not hold onto facts explained by the actual microphysics. It is simply open whether the different microphysics in the closest (C4)-antecedent worlds would support the chance of heads as it actually is or not. The chances *might* be the same, but they might also be different.

Coming back to problem (4.2.2), consider an outstanding candidate for reducing the laws of thermodynamics, the Albert-Loewer recipe.[29] It comprises:

(i) the Newtonian dynamical law: $F = ma$; (ii) the Past Hypothesis: the initial conditions are low entropy; and (iii) the Statistical Postulate: there is a probability distribution uniform on the standard measure over those regions of phase space compatible with our empirical information.[30]

Just for the sake of argument, I assume two things: the laws of thermodynamics can be reduced à la Albert-Loewer. And the Past Hypothesis is a fundamental law. Then the following should be rejected:

(C9) If the Past Hypothesis had been violated, still the laws of thermodynamics would not have been violated.

For if we do not perfectly hold onto the Past Hypothesis, we do not have to hold onto what is explained by it. But Dunn seems committed to (C9). One concern will be discussed in appendix I. Now Dunn is critical of the Albert-Loewer recipe. He has a strategic motive for his criticism: if the Past Hypothesis is a fundamental law, the Elga world can be rejected for violating it. We do not need Dunn's amendment. To counter this threat, Dunn voices doubts that the Past Hypothesis is a fundamental law. I will come back to these arguments below. Fundamental laws à la Lewis are confined to perfectly natural properties. *Entropy* is no such property. And fundamental laws are usually regarded as regularities. The Past Hypothesis is no regularity.[31] Be that as it may, I here use the Albert-Loewer

29 cf. Albert, 2000, Loewer, 2012.

30 Schaffer, 2007, 122, cf. Dunn, 2011, 83.

31 cf. Dunn, 2011, 83–84, Callender, 2011.

recipe only *by way of an example* as an outstanding model of reduction. So if we do not principally eschew reduction of special science laws, any reductive efford could be used to underpin my argument.

There are two further problems. Both are mentioned but not thoroughly solved by Dunn.

(4.2.3) I come to what could be the most grievous difficulty. For it is easily fleshed out in terms of individual counterfactuals. Contingent matters of fact might interfere with laws of the special sciences:

> «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 led 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».[32]

The problem (L1) poses is that in the standard Lewisian construal, it entails.

(L3) It is not the case that if lightning had struck in region R, then the laws of biology would have been the same.

This contradicts Dunn's (4´). (4´) prescribes to hold onto the laws of biology. Dunn tries to avoid this result. He follows a suggestion of Lewis: sometimes read «might» as «it would be that: different laws are possible».[33] To judge this proposal, we have to distinguish two cases. In one case (Case I), say where lightning directly strikes the critical event, the fundamental laws and the facts in the scenario entail different biological laws.

32 Dunn, 2011, 92.

33 Dunn, 2011, 93.

Then, Dunn concedes, we accept:

(L4) If lightning had struck the critical event, then the biological laws would have been different.

Thus, the laws might have been different, «might» understood in the standard way. There is another, more problematic case (Case II): the fundamental laws and the facts modified by lightning do not ensure which biological laws will come to obtain. How can such a situation be reconciled with determinism? The antecedent is vague. Nothing precludes that two different small miracles lead to two different antecedent worlds, one with our biological laws, one with alien ones.[34] In that situation, doctoring the «might»-conditional won't help. For we pace Dunn reject that the laws *would* have been the same. This clashes with Dunn's explicit commitment to:

(L2) If lightning had struck in region R, then the laws of biology would have been just as they actually are.[35]

I discuss a possible reaction on Dunn's behalf in appendix II.

(4.2.4) I come to a difficulty of Dunn's argument against Elga: we cannot dispel the Elga world. It might exhibit more match in particular facts than Lewis's candidate, divergence world W_2. This can happen when the world has a beginning and stretches further into the future than into the past (relative to the time of convergence). As a remedy, Dunn gerrymanders a reading of the second criterion: «we do not quantitatively compare a region of past match with a region of future match».[36]

There are two ways of fleshing out Dunn's reading. The first is: assume that world A exhibits more perfect match with actuality in pre-antecedent facts but B exhibits more perfect match with actuality in post-antecedent facts. Then, other things being equal, both are *equally* similar to actuality. The unfortunate consequence is this: for some A and B, there will be a world C. C fares even worse than A in perfect post-antecedent match but equals A in pre-antecedent match (assume a different small miracle leads to larger regions of post-antecedent mismatch than the small

34 Dunn presents a structurally analogous case (cf. Dunn, 2011, 95).

35 Dunn, 2011, 92.

36 Dunn, 2011, 86.

miracle by which A departs from actuality). As a consequence, C is as similar to actuality as B but less similar than A. To Lewis, the relations of overall similarity among worlds form a total preorder. A total preorder is transitive and complete.[37] Since A and B are equally similar and so are B and C, by transitivity A and C must be equally similar. But C is less similar than A.

The second way of putting Dunn's interpretation is that A and B are *incommensurable* with regard to overall similarity.[38] This conflicts with the completeness of the similarity ordering. If A and B qualify for overall similarity, either one is more similar or both are equally similar. Even if an account must ultimately be judged by its dealing with exemplary counterfactuals, the basic formal properties of the similarity relation are crucial to Lewis's standard analysis of counterfactuals.[39] Dunn's proposal conflicts with these properties.

5. The Limits of Elga's Counterexample

I think that Dunn's strategy fails. But *malgré lui,* he teaches us an important lesson. Most people involved in the debate grant that Elga has devised a successful counterexample to Lewis's metrics. But there are reasons for doubt Elga does not attend to. The success or failure of Elga's argument depends on our explanation of the thermodynamic asymmetry.[40] Many philosophers feel the need of an account in terms of fundamental physical laws.

Let us consider an example:[41] Dunn dismisses the possibility of using the Albert-Loewer recipe against Elga. Following the Albert-Loewer recipe, the Past Hypothesis (the initial conditions of the universe are low entropy) is treated as a law. Let us assume it is a fundamental law. Elga worlds would gravely violate that law and thus be ruled out from being

37 cf. Lewis, 1973, 14.

38 cf. Morreau, 2010.

39 cf. Lewis, 1986, 41.

40 cf. Callender, 2011.

41 Another example would be a suitable version of the «General-Causal» approach. There is a fundamental mechanism which ensures that our world exhibits the thermodynamic asymmetry (cf. North, 2002, 123).

closest by Lewis's first criterion (no big miracles). For we have seen that the atypical region of reversed rotting rapidly grows the further we go back. So it leads to relatively widespread initial conditions which are not low entropy.

However, Dunn presents two reasons why the Past Hypothesis cannot be a fundamental law: (i) the Past Hypothesis is no *regularity*. Thus it cannot be a law at all. (ii) Entropy is no perfectly natural but a high-level organizational property. Thus the Past Hypothesis cannot be a *fundamental* law.[42] In reply to (i), the initial conditions of the universe are special. For instance, there is an intense debate on whether such conditions allow for explanation at all.[43] Their peculiarity may be spelled out in fundamental terms. The Past Hypothesis may quantify over any situation, at any time, which displays those general fundamental properties which as a matter of fact only the *initial* conditions of the universe display. In quantifying over all times, it is a regularity, but a regularity whose only actual instantiation are the initial conditions of the universe.

In reply to (ii): assume for the sake of argument that the Past Hypothesis is a law. Then, says Dunn, it must be a special science law as it is not formulated in terms of perfectly natural properties. Take a disjunction of all fundamental ways of realising the property of low entropy. Given Dunn's own account of special science laws, the fundamental law behind the Past Hypothesis may be expressed in terms of this disjunction. To Dunn, fundamental terms relate as follows to the less fundamental ones used in the special sciences:

> «We have a special science law:
>
> SS-Law : $Sx \supset S^*x$
>
> «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.

42 Dunn, 2011, 83–84.

43 cf. Price, 2002.

> 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».[44]

In this passage, Dunn wants to account for exceptions to the special science laws. But his template also fits exceptionless laws. The Past Hypothesis may be put as an exceptionless disjunctive law in terms of fundamental properties: let S be a disjunction of initial states of the universe which fulfil a general microphysical description (which only initial states satisfy). And let S* be the disjunction which spells out all fundamental ways of realizing low entropy.[45]

Given the doubts about the status of the Past Hypothesis, I close with a moderate claim: the success or failure of Elga's example depends on whether there are fundamental natural laws which underlie the thermodynamic asymmetry. If there are such laws, they can probably be used to rule out the Elga world as a competitor for closeness. So what concerns Elga's egg, there is hope for Lewisian orthodoxy. Just as Lewis intimates, «counter-entropic funny-business makes for dissimilarity in the same way a big miracle does»: by violating the fundamental laws of nature on a large scale. I mention but cannot discuss a move that would allow us to simulate the effects of the assumption that the Past Hypothesis is a fundamental law without endorsing this hypothesis. The exclusion of counterentropical funny-business can be written into the semantics of counterfactuals.[46] This leaves us with the following options in the neighbourhood of Lewis's original criteria: first, amend the criteria by a clause which precludes post antecedent match from counting as far as it explanatorily depends on the antecedent. The cost is that the semantics depends on explanatory relations like causation. Second, accept that the Past Hypothesis is a fundamental law. The cost is that we lose the

44 Dunn, 2011, 90.

45 For further discussion of objections to the Albert-Loewer recipe cf. the essays in Wilson, 2014, ch. 7–12). In particular, Albert tries to rebut the objection that the recipe allows for a localized backwards influence and thus cannot account for the sharpness of the opposition between past and future (Wilson, 2014, ch. 8). Albert replies that such an extremely local influence could not be of practical avail. Moreover, there are further substantial differences between our access to the past and our access to the future.

46 Bennett, 2003, 296, Williams, 2008.

symmetry of laws and have to accept a somewhat strange law. Third, write it into the similarity criteria that counterentropical funny-business detracts from similarity. The consequence is that the criteria look more gerrymandered. I leave it to future debate which option will prevail.

Appendix I to 4.2.2

Dunn distinguishes between violating the laws of thermodynamics and these laws not being laws at all.[47] In the latter case, they are not violated. Lewis insists that a fundamental law which has an exception is not a law. *Provided we follow Dunn* that a law which is no law is not violated, the Past Hypothesis cannot be violated.

(C9) If the Past Hypothesis had been violated, still the laws of thermodynamics would not have been violated.

(C9) is vacuously true. In this case, consider instead of (C9):

(C9´) If the Past Hypothesis had not been a law, still the laws of thermodynamics would not have been violated.

Again I think that (C9´) should be rejected; Dunn must accept it. For criterion (4´) requires us to hold onto the special science laws even if the fundamental laws are no laws.

However, consider any situation where the Past Hypothesis is not a law. Shouldn't we expect it to be a situation where the laws of thermodynamics are no laws either (and thus are not violated)? I use the Elga world to conjure up a counterexample. By Dunn's lights, it violates the laws of thermodynamics. To be violated these laws must be laws in the first place, says Dunn. In Elga's vision, the infected region grows the further back we go in time. So we may construe the Elga world as follows: the Past Hypothesis is not a law at that world. In the infected region, initial conditions are not low entropy. However, in the regions surrounding the infected region, initial conditions are low entropy. And that is sufficient to ensure that the laws of thermodynamics are laws. Otherwise Dunn could not rule out the Elga world. Thus, the Elga world is a world where

47 cf. my extensive discussion in appendix II, and Dunn, 2011, 94–95 footnote 33.

the Past Hypothesis is not a law but the laws of thermodynamics are violated, contrary to (C9´). Of course, there is no reason to deem the Elga world the closest world where the Past Hypothesis is no law. But some world like it is a good candidate. We cannot accept (C9´) as long as we cannot decide between two candidates for the closest worlds where the Past Hypothesis is not a law: a world where the laws of thermodynamics principally are laws but are violated; and a world where the laws of thermodynamics are no laws.

Appendix II to 4.2.3

In light of Dunn's comments in his footnote 33, it is not clear that his account really commits him to (L2).

(L2) If lightning had struck in region R, then the laws of biology would have been just as they actually are.[48]

Footnote 33 presents a situation where the actual biological properties fail to be instantiated. As a consequence, the actual biological laws are no laws.[49] Dunn insists that his criterion (4´) does not rule out such a world from being closest. Dunn is not committed to (L2) if the following holds: consider any lightning situation where the actual facts and the lightning together with the fundamental laws do not entail the actual biological laws. In such a situation, either the laws of biology are laws or they are no laws.

This is literally right as far as (4´) demands that the laws of the special sciences should *not be violated,* not that they should be laws at all – but only given a specific reading of the criterion: a law is not violated if the properties to which it applies are never instantiated. Dunn must distinguish two kinds of situation: situations where the laws are not violated because they are no laws and situations where they are violated. Otherwise laws, at least those of the special sciences, *could* not be violated at all. All possible situations whatsoever would fare equal with respect to Dunn's (4´). The Elga world could not be ruled out.

48 Dunn, 2011, 92.

49 cf. Dunn, 2011, 94–95.

Consequently we can further develop my second lightning scenario (Case II) into two subcases: in one lightning scenario (Case II.I), the laws and the facts together only ensure that one of two alternatives will come to pass: either the actual laws of biology are laws and unviolated. Or they are no laws (and thus unviolated) because our biological properties are not instantiated. For that subcase, Dunn gets the right result: if the lightning situation had obtained, the actual laws would not have been violated but might have been no laws. Yet there is another subcase (Case II.II): in a different lightning scenario, the fundamental laws and the actual facts modified by lightning do not ensure that our biological laws *will not be violated.* For in that situation, our biological laws perfectly hold or they will be violated.[50] Dunn must falsely maintain that if this situation had come to obtain, the laws of biology would not have been violated.

To assess the possibility of (Case II.II), I take a closer look what it could mean to violate a law. By Lewis's lights, there is no distinction between a *fundamental* law being violated and being no law. What concerns laws of the special sciences, things are more intricate. These laws permit exceptions. To fit Dunn's needs, a violation must steer between an exception permitted by the laws and the laws being no laws. For instance, Dunn must prevent Elga from retorting: in the Elga world W_3, the laws of thermodynamics are not violated because they are no laws. So Dunn must insist that in W_3, the actual thermodynamic properties are instantiated although the infected region violates the laws of thermodynamics.

In a parallel vein, Dunn must accept that the lightning scenario can be further developed along the following lines: let there be several regions R_1–R_n. Each is sufficient to bring about our biological properties and the concomitant laws. But assume that had lightning struck region R_1, that region might have been infected by biological systems behaving deviantly. Deviant behaviour would have violated the actual laws of biology.[51] More precisely, the fundamental laws and the facts modified by lightning do not ensure that R_1 would not have been infected. In one lightning sit-

50 We could also integrate the third alternative of the laws being no laws.

51 One may doubt that such deviant biological systems are microphysically possible. But there will be other examples; one is my scenario of the Elga world conflicting with the Past Hypothesis such as to violate the laws of thermodynamics.

uation R_1 would have been infected. In another, it would not have been. Dunn must accept.

(L5) If lightning had occurred in region R_1, the actual laws of biology would have gone unviolated.

We should reject (L5). In sum, it does not help to introduce the possibility that laws are not violated because the respective properties are not instantiated at all. Note that my argument does not commit me to accepting the distinction of laws being violated or being no laws. I just point out what Dunn is committed to.

References

Albert, David 2000. *Time and Chance*. Cambridge, MA: Harvard University Press.

Albert, David 2014. The Sharpness of the Distinction between the Past and the Future, in: Alasdair Wilson (ed.), *Chance and Temporal Asymmetry*. Oxford: Oxford University Press: 159–174.

Bennett, Jonathan 2003. *A Philosophical Guide to Conditionals*. Oxford: Oxford University Press.

Callender, Craig 2011. Thermodynamic Asymmetry in Time, *The Stanford Encyclopedia of Philosophy* (Fall 2011 Edition), Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/fall2011/entries/time-thermo/>.

Dunn, Jeffrey (2011) "Fried Eggs, Thermodynamics, and the Special Sciences", *The British Journal for the Philosophy of Science*, 62, pp. 71–98.

Elga, Adam (2001) "Statistical Mechanics and the Asymmetry of Counterfactual Dependence", *Philosophy of Science*, 68, pp. S313-S324.

Frigg, Roman, Hoefer, Carl (2010) "Determinism and Chance from a Humean Perspective", in: Friedrich Stadler et al. (ed.), *The Present Situation in the Philosophy of Science*, F. Stadler et al (Ed.), Dordrecht: Springer, pp. 351–371.

Hitchcock, Christopher (2011) "Trumping and Contrastive Causation", *Synthese*, 181, pp. 227–240.

Kment, Boris (2006) "Counterfactuals and Explanation", *Mind*, 115, pp. 261–310.

Kutach, Douglas (2002) "The Entropy Theory of Counterfactuals", *Philosophy of Science*, 69, pp. 82–104.

Lange, Marc (2004) "The Autonomy of Functional Biology: Reply to Rosenberg", Biology and Philosophy, 19, pp. 93–109.

Lewis, David (1973) *Counterfactuals*, Oxford: Blackwell.

Lewis, David (1986) "Counterfactual Dependence and Time's Arrow, Postscripts to «Counterfactual Dependence and Time's Arrow»", in *Philosophical Papers II, Oxford: Oxford University Press*, pp. 32–66.

Loewer, Barry (2012) "Two Accounts of Laws and Time", *Philosophical Studies*, 160, p. 115–37.

Morreau, Michael (2010) "It Simply Does Not Add Up: Trouble with Overall Similarity", *Journal of Philosophy*, 107, pp. 469–490.

Noordhof, Paul (2005) "Morgenbesser's Coin, Counterfactuals and Independence", *Analysis*, 65, pp. 261–263.

North, Jill (2002) "What is the Problem about the Time-Asymmetry of Thermodynamics? – A Reply to Price", *The British Journal for the Philosophy of Science*, 53, pp. 121–136.

Northcott, Robert (2009) "On Lewis, Schaffer and the Non-Reductive Evaluation of Counterfactuals", *Theoria*, 75, pp. 336–343.

Price, Huw (2002) "Boltzmann's Time Bomb", *The British Journal for the Philosophy of Science*, 53, pp. 83–119.

Schaffer, Jonathan (2004) "Counterfactuals, Causal Independence and Conceptual Circularity", *Analysis*, 64, pp. 299–309.

Schaffer, Jonathan (2007) "Deterministic Chance?", The British Journal for the Philosophy of Science, 58, pp. 113–140.

Wasserman, Ryan (2006) "The Future Similarity Objection Revisited", *Synthese*, 150, pp. 57–67.

Williams, Robert G. (2008) "Chances, Counterfactuals, and Similarity", *Philosophy and Phenomenological Research*, 78, 385–420.

Wilson, Alasdair (ed.) (2014) *Chance and Temporal Asymmetry*. Oxford: Oxford University Press.