BACKTRACKING INFLUENCE

DOUGLAS KUTACH

Abstract. Backtracking influence is influence that zigzags in time. For example, backtracking influence exists when an event E_1 makes an event E_2 more likely by way of a nomic connection that goes from E_1 back in time to an event C and then forward in time to E_2. I contend that backtracking influence is redundant in the sense that any existing backtracking influence exerted by E_1 on E_2 is equivalent to E_1's temporally direct influence on E_2. I prove the redundancy of backtracking influence using several plausible physical principles without assuming any fundamental temporal or causal asymmetry. This explanation can play a prominent role in an account of why causation appears to be objectively asymmetric regardless of whether the fundamental laws are temporally symmetric.

1. Introduction

One principle that appears to play an important role in causation is the following: an event E_1 never makes an event E_2 more or less likely by way of a nomic connection that goes from E_1 back in time to an event C and then forward in time to E_2. In our neighbourhood of the universe, apparently, such backtracking influence does not exist. I will attempt to demonstrate the truth of this principle by showing that backtracking influence in general fails to augment the relations of temporally direct influence that backtracking presupposes.

One simple way to explain the lack of backtracking influence is to rule out the existence of all past-directed influence by incorporating some fundamental temporal asymmetry into one's theory of influence. Theories using a fundamental asymmetry, e.g., Hume (1739), Kant (1781), Broad (1923), Beauchamp and Rosenberg (1981), Ellis (2006), Maudlin (2007), can do so in various forms including assuming that causation can only be directed towards the future (where the future direction is an aspect of reality independent of the universe's material layout) or that the counterfactual conditionals relevant to causal dependence

---

Date: October 14, 2010.

presuppose the fixity of the past. Such fundamental causal asymmetries provide a straight-forward explanation for the non-existence of backtracking influence, albeit at a theoretical cost. A minor drawback is ontological bloat, for one should not postulate fundamental asymmetries beyond necessity. A more serious concern is the coordination problem: ensuring that any explanation of influence in terms of the primitive asymmetry coordinate properly with explanations given in terms of the microphysical evolution. A theory where a fundamental direction of influence can flip back and forth without consequence for the motions of particles and fields would undermine its explanatory value. Although these problems can be overcome, I am among the many who prefer a framework eschewing primitive asymmetries, e.g., Gasking (1955), Reichenbach (1956), von Wright (1975), Papineau (1985), Mackie (1973), Lewis (1979), Price (1992), Mellor (1995), Hausman (1998), Albert (2000), Dowe (2000), Loewer (2007).

My goal here is to explain the absence of backtracking influence without postulating any fundamental temporal asymmetry. The framework I employ interprets influence as a kind of difference-making based entirely on what follows from fundamental laws of nature. The underlying metaphysical picture is one where the evolution of the world is governed fundamentally by laws that resemble the kind of laws appearing in paradigm theories of fundamental physics. The totality of facts about causation and influence, I will assume, hold in virtue of such laws and the complete spatiotemporal layout of fundamental properties and relations.

With the leaner metaphysics that postulates only fundamental physics and refuses to cook the books in favour of unidirectional causation, it is prima facie mysterious what forbids the existence of backtracking influence. This is clearest when the laws are deterministic in both temporal directions. Suppose (1) that event C determines the occurrence of future events E_1 and E_2, and (2) that E_1 could fail to occur only in the absence of C, and (3) that each of the two events can occur only if C occurs. By (3), E_1 implies C, which by (1) implies E_2. By (2), a non-occurrence of E_1 implies a non-occurrence of C, which by (3) implies a non-occurrence of E_2. Thus, E_1 makes a difference as to whether E_2 occurs, raising its probability from zero to one.

My explanation does not require any particular account of the matter asymmetries, e.g., no assumption of a fork asymmetry or entropy asymmetry. The end result is a boon to anyone who wants to derive the direction of causation merely from fundamental laws plus some sort of asymmetry in the historical layout of the universe's material contents because my account helps to explain why it is useful for people to conceive of causation as objectively asymmetric even if the fundamental laws turn out not to possess the right kind of temporal asymmetry to ground a fundamental causal direction.

2. Prob-Influence

Among the many formalizations of influence one could construct, there exists one I call 'prob-influence' that is especially useful for understanding causation. 'Prob-influence' stands for 'probabilistic influence that exists in virtue of fundamental laws.' It quantifies how much the probability of some chosen effect counterfactually depends on the occurrence of some chosen candidate cause. In order to define this notion more precisely, several preliminaries are needed.

Causation, I assume, exists in virtue of fundamental laws that resemble (near enough) the kind of laws found in the most famous theories proposed in physics: classical mechanics with gravitation, relativistic electromagnetism, general relativity, and non-relativistic quantum mechanics. These four paradigm fundamental theories all describe reality using spacetime structure with material contents like corpuscles and fields[1] that can be described using the language of events. There are four kinds of events I will distinguish.

Fine-grained events, labelled with lowercase variables, are maximally precise specifications of all the fundamental stuff in a determinate region of the arena. Coarse-grained events, labelled with uppercase variables, allow us to fuzz the precise microscopic details of a fine-grained event. Formally, a coarse-grained event C is just a set of possible fine-grained events, $\{c_1, c_2, \dots\}$. It is convenient (and does no harm) to restrict attention to coarse-grained events whose elements all have the same size and shape. That allows us to think of any coarse-grained

---

[1] In the case of quantum mechanics, most fundamental properties are represented as part of the universal wave function, Ψ, which encapsulates the entanglements among position, momentum, spin, etc. for all the particles in the universe. Ψ's degrees of freedom require a mathematical space of very high dimension called 'configuration space' for adequate representation. One way to interpret Ψ is to say that it represents a holistic relation among all the particles in the universe at some time t and thus occupies the infinitely extended time slice at t. However, one might wish to interpret this configuration space as a part of (or perhaps all of) the fundamental space (or arena) inhabited by all fundamental entities and properties. In that case, one would need to translate my talk of spacetime locations into talk of regions in this more general arena that includes spacetime and configuration space or perhaps just configuration space alone. Everything I say in this article is compatible with such interpretational manoeuvres, though for rhetorical convenience, I will presume spacetime as the arena of fundamental reality.

event C as occupying a determinate spacetime region with each of its elements representing a possible instantiation of everything that happens in that region. Thus, coarse-grained events serve (in effect) as *types* of fine-grained events. I will introduce the two remaining kinds of events shortly.

All four paradigm theories include fundamental laws of temporal evolution whereby some fine-grained events fix an objective probability distribution over what can happen at other locations. For instance, some theories include laws of determination, where the right kind of fine-grained event c (in conjunction with the fundamental laws) entails a certain fine-grained effect, e. When an event c determines that e occurs, we say that c is a determinant of e. These relations of determination have a built-in relativization to the spacetime location of c. For example, when c determines a later collision between an electron and positron, it determines that they collide at a specific spacetime location relative to c, not just that an electron and positron eventually collide somewhere in the universe. Throughout the rest of the discussion, this spatiotemporal relativization will be left implicit.

As far as we know, the fundamental laws might be chancy. In order to accommodate this possibility, determination can be generalized as follows. Whenever an event c fixes, by way of the fundamental laws, a probability distribution over a set of fine-grained events, at least one of which is e or includes e as a part, we can say that c termines e and that c is a terminant of e. A terminant is just a fine-grained probability-fixer, and determination is a special case of this broader relation, terminance.

Notice that when a terminant, c, fixes a uniform probability distribution over a continuum of possible chance outcomes, as happens in most interpretations of quantum mechanics, the probability that c fixes for any given e is typically zero. Such laws, however, still provide nontrivial probabilities for numerous coarse-grained events. For example, suppose c is the complete state of the universe at some time t, instantiating a lone neutron that decays exactly three minutes later. Let e be an event located in a small neighbourhood around the neutron decay, instantiating its last moments and the very precise way it explodes into decay products. The laws are such that c fixes a probability for e of zero, which makes the probabilistic relation between c and e appears to be not terribly informative. However, c does fix a positive probability for the neutron's decaying sometime between $t + 2$ minutes and $t + 4$ minutes, and for the neutron not decaying before $t + 7$ minutes, and for many other events of interest.

A remarkable feature of the paradigm fundamental theories is that none of them permit probability-fixing relations among the ordinary-sized events we humans typically care about. For example, we tend to think that throwing an ordinary rock over the surface of a large, calm, liquid lake will cause ripples ten seconds later, but the fundamental laws say nothing about the causal impact that the rock alone has on what happens ten seconds later. In order for the fundamental laws to be at all informative, they require a full specification of all the microscopic details of some vast event that includes the rock. If the laws are relativistically local like the theory of relativistic electromagnetism, the relevant throwing event needs to span at least ten light-seconds in radius and must specify everything about that region, including where it instantiates empty space. If the laws are non-relativistic, as in classical gravitation, the event needs to encompass the entire breadth of the universe, again specifying every last detail about what exists at every location. The upshot is that influence among ordinary human-sized events only exists insofar as they are parts of events that are fully specified in microscopic detail and are vast enough.

The notion of probability-fixing that arises from fundamental laws found in physics should not be confused with the kinds of probabilistic relations employed in standard probabilistic theories of causation, e.g. Reichenbach (1956), Suppes (1970). On their views, probabilistic relations can hold among localized events (or event types) without holding in virtue of the detailed structure of vast physical events. For example, they may refer to the probability of person S acquiring heart disease given that S smokes regularly. On the view I am presenting, although there are certainly statistical correlations among such localized happenings, there are no underlying probabilistic relations among them that play a role in grounding relations of influence.

The notion of probability-fixing can easily be extended to allow coarse-grained events to fix probabilities. Let us temporarily focus on theories where the fundamental laws are chancy. According to such theories, if an event c termines an event e, that terminance holds in virtue of c's fixing a probability distribution over a set of events, one of which includes e, at least as a part. Let us say that this set of events together with the assigned probability distribution constitutes a contextualized event, $\overline{E}$. In general, a contextualized event is just a coarse-grained event with a probability measure over all its elements. (The motivation for the label ‘contextualized’ will become apparent later.) I will signify contextualized events by placing a bar over the corresponding capital letter. The bar is not a function or an operator; it is just part of the label for a contextualized event. So, in addition to saying that c fixes a probability for an ordinary coarse-grained event E and that it therefore termines E's instantiation, e, we can also say that c fixes $\overline{E}$, which means that c entails $\overline{E}$ (when conjoined with the fundamental laws). This fixing is similar to determination in the sense that $\overline{E}$ is guaranteed by the laws to occur when c occurs, but it is different from determination in that only one element of $\overline{E}$ is instantiated and c does not determine which element that is.

Now that we have made sense of how fine-grained events can fix contextualized events, it is easy to understand how contextualized events can fix probabilities for ordinary coarse-grained events as well as contextualized events. One simply defines the probability that the contextualized event $\overline{C}$ fixes for E to be the probability that each element of $\overline{C}$ fixes for E, weighted by $\overline{C}$'s built-in probability distribution. Also, we can say that the contextualized event that $\overline{C}$ fixes for spacetime region r is just the union of all the contextualized events fixed for r by the elements of $\overline{C}$, weighted by $\overline{C}$'s built-in probability distribution.

Contextualized events were introduced for three reasons. First, they allow us to more richly represent coarse-grained events to take into account that not all ways of instantiating an event are equally likely to be instantiated. For example, by representing an archer's shot with a contextualized event, we can incorporate the fact that the archer is generally very accurate while also allowing that a badly aimed shot is still possible. Second, contextualized events overcome the following deficiency of plain coarse-grained events. We cannot in general say that a coarse-grained event, C, fixes a determinate probability for E even if all C's elements are filled in with full microscopic detail and are large enough to fix a probability for E because the elements of C might fix different probabilities for E. By construing such a happening as a contextualized event, $\overline{C}$, it is guaranteed that $\overline{C}$ will fix a determinate probability for E. Third, contextualized events are able to fix non-trivial probabilities for events even when determinism holds. That allows them to represent, for example, that the roll of a die fixes a 1/6 probability for each outcome even though every fine-grained instantiation of a die roll (together with its broader environment) fixes a probability of either one or zero. Contextualized events thus help to insulate our claims about macroscopic causation from the controversial question of whether our world is fundamentally chancy.

We are now in a position to represent the relation of probability-fixing using a function, $p_{\overline{C}}(E)$. This value quantifies how likely the event $\overline{C}$ makes E. It is the objective probability (of the coarse-grained event E) generated by assuming the existence of the contextualized event $\overline{C}$ as a hypothetical starting point and letting the fundamental laws alone dictate everything else that happens in the universe. It is not a conditional probability.

Although we ordinarily evaluate the nomic consequences of hypothetical situations by reckoning how they would likely evolve towards the future, the probability-fixing I am countenancing does not presuppose that it only works towards the future. For example, if the laws are deterministic and $\overline{C}$ represents a fully specified time slice, then $\overline{C}$ will fix probabilities for earlier events. The fundamental laws may turn out not to fix any probabilities for past events, but we should not forbid past-directed terminance because that would unnecessarily restrict the pool of candidate explanations for the causal asymmetry.

We now have two sources of probability being countenanced: that which arises from fundamentally chancy laws of evolution and that which is just stipulated as part of one's choice of contextualized event. Presumably, if contextualized events are to be useful for understanding causation, there needs to be some account of why some probability measures are better than others for abstracting away from the microscopic details of fundamental physics. However, this issue can be bracketed for present purposes because backtracking influence will turn out to be redundant regardless of a contextualized event's probability measure.

The fourth and final conception of an event is a contrastive event. A contrastive event, formally speaking, is an ordered pair of contextualized events. I will signify contrastive events with a tilde on top of a capital letter. It again proves to be convenient (and does no harm) to restrict attention to contrastive events whose two contextualized events are the same size and shape, so that we can think of the contrastive event as representing two contrasting possibilities for what happens in a given spacetime region.

The purpose of a contrastive event is to allow us to represent localized, human-scale happenings in a way that is compatible with the fact that the fundamental laws do not provide any lawful relationships between such events. Suppose Guy throws a rock at the surface of a calm lake at time t, the rock strikes the surface of the lake, and ten seconds after t, there are expanding ripples. Intuitively, we think of the throw as a cause of the ripples. The fundamental physics, however, tells us that insofar as terminance goes, the fish in the lake, the sunlight on the clouds, and the craters on the Moon are just as much partial causes of the ripples as the throw because any fine-grained event that omitted any one of these components would fix no probability whatsoever for the existence of ripples and would not be in any other way informative about the ripples. Contrastive events, though, allow us to make sense of the privileged causal status of the throw.

To see how this works, let c be a localized representation of Guy's throw, an event at t occupying a sphere one meter in radius instantiating the precise way Guy threw the rock. Let us now contextualize c as $\overline{C}$ by defining $\overline{C}$ to occur at t and occupy a region centred around c and stretching a few light-minutes in all spatial directions. Let one element of $\overline{C}$ be the fine-grained event that instantiates the complete actual state of the world at time t, which includes c as one small part. One can then add as many other elements to $\overline{C}$ as one likes by taking the one actual fine-grained event and forming new events by shifting an electron here or a proton there until there are enough fine-grained events in $\overline{C}$ to represent a fuzzy version of what actually happened. One is free to include whatever background physics one likes into each element, but every element needs to instantiate Guy throwing a rock in region c. Finally, one should stipulate some reasonable probability measure over all these elements so that the resulting $\overline{C}$ captures the intended range of possibilities, including that Guy's throw is very probably one that is directed at the lake with enough strength and without any obstacles in the way.

To form the second contextualized event, simply take every element of $\overline{C}$ and replace the physics in the region that instantiates Guy throwing a rock with some alternative physical arrangement that instantiates Guy not throwing the rock. This could be Guy holding on to the rock, or Guy dropping the rock, or Guy not having any rock at all, among many other possibilities. The chosen contextualized event, $\overline{\neg C}$, constitutes a fuzzy way of representing Guy not throwing the rock, embedded in the same vast environment used for $\overline{C}$.

The resulting contrastive event, $\tilde{C}$, defined as $(\overline{C}, \overline{\neg C})$, represents Guy throwing the rock rather than not throwing the rock. Notice that because of the way the two contextualized events were constructed, they disagree about what happens in the small region where Guy exists, but they agree on everything happening elsewhere. We can say that any spacetime region where the two contextualized events disagree counts as the foreground, and everywhere they agree counts as the background. Thus, $\tilde{C}$ can be said to represent the localized event of Guy's throwing the rock (rather than not) within a constant background field.

Now we are finally in a position to define prob-influence. Let us say that the degree of prob-influence that $\tilde{C}$ exerts on some coarse-grained event E is equal to the difference in the probability of E fixed by its two elements. For example, Guy's throw (as represented by $\tilde{C}$) prob-influences the existence of ripples, E, to the degree $p_{\overline{C}}(E) - p_{\overline{\neg C}}(E)$. When this quantity is positive, $\tilde{C}$ makes E more likely; it promotes E. When this quantity is negative, $\tilde{C}$ makes E less likely; it inhibits E.

We can also construe prob-influence as another form of event fixing. If a contrastive event, $\tilde{C} \equiv (\overline{C}, \overline{\neg C})$, is large enough to be informative of what will happen throughout a spacetime region r, then each of its contextualized events individually fixes a contextualized event occupying r. The resulting ordered pair $\tilde{E} \equiv (\overline{E}, \overline{\neg E})$ can thus be thought of as the unique contrastive event that $\tilde{C}$ fixes for region r. This contrastive effect represents the existence of ripples and the rock in the lake and Guy's memory of having thrown the rock, etc., rather than the non-existence of ripples and the rock being in Guy's hand (or on the ground) and Guy's memory of not having thrown the rock, etc.

It should now be easy to see how Guy's throw counts as an important causal factor for the ripples whereas virtually every other salient human-sized event taking place at t does not. First, it presumably follows from the fundamental laws and our chosen background conditions that $p_{\overline{C}}(E)$ is very high and that $p_{\overline{\neg C}}(E)$ is nearly zero, vindicating the intuitively correct claim that throwing the rock promotes ripples. Second, almost any representation of the actual state at t as a contrastive event with a foreground that is some ordinary object in the environment does not result in significant prob-influence. For example, consider the causal impact of a nearby tree by formulating a contrastive event that uses $\overline{C}$ as the first element in the ordered pair and uses as its other element, a contextualized event that differs from $\overline{C}$ only by replacing the tree with some ordinary, empty patch of land. Because both contextualized events instantiate Guy's throwing the rock with the only difference being the presence of the tree, they will both fix nearly the same high probability of ripples. Hence, the tree will count as not significantly prob-influencing the existence of ripples. Consideration of a wide range of other examples leads to the conclusion that although one can find narrowly crafted contrastive events that significantly prob-influence the existence of ripples, the contrastive events that represent our ordinary ways of construing events (by using contrast classes that strike us as natural) vindicate our intuitions about the relative causal importance of various localized events.[2]

---

[2] A more thorough discussion of this approach towards causation is provided in Kutach (2011b).

3. The Redundancy of Backtracking Prob-influence

Backtracking prob-influence occurs whenever an event at one time prob-influences an event in the past (future) that in turn prob-influences an event in the future (past). One might think that backtracking prob-influence is possible in virtue of the nomic connections that we know exist in the form of a common cause of two distinct effects. In Fig. 1, three events are located in ways that exemplify a paradigmatic common cause pattern, a situation we would justifiably describe as the event C making two distinct later events, E_1 and E_2, more likely than they would be in the absence of C. For concreteness, one could imagine that C is the activation of a school's fire alarm; E_1 is some particular student's being frightened; and E_2 is a fire truck arriving at the school. The intuitive claim to be explained is that the student's being frightened does not make the appearance of a fire truck more likely in virtue of a nomic connection that goes from the student's fright back in time to the alarm and then forward in time to the truck's arrival. The disutility of such backtracking promotion holds even if the student's fright backwardly makes the sounding of the alarm more likely (by way of the fundamental laws) than it would have been if the student had been bored or sleepy. Of course, everyone can agree that the student's fright can promote a fire truck's arrival in virtue of a purely future-directed nomic connection. For purposes of illustration, I will assume E_1 temporally precedes E_2. By the end of the discussion of the proof, it should be obvious how the argument applies to cases where E_2 happens earlier or at the same time.

According to the model of probabilistic influence I have presented, the claim that E_1 affects the probability of E_2 by way of C must be understood in terms of contrastive events because a localized, human-scale event like E_1 implies nothing by way of the fundamental laws about other localized, human-scale events. So, the claim to be considered is whether any contrastive event $\tilde{E}_1 \equiv (\overline{E_1}, \overline{\neg E_1})$ whose foreground part instantiates E_1 rather than a non-existence of E_1 is able to prob-influence E_2 by way of some causal intermediate $\tilde{C} \equiv (\overline{C}, \overline{\neg C})$ that lies to the past of both $\tilde{E}_1$ and E_2. The claim I intend to prove is that the prob-influence $\tilde{E}_1$ exerts on E_2 in virtue of backtracking through any past event $\tilde{C}$ is redundant in the sense that it is equal to the prob-influence that $\tilde{E}_1$ exerts directly towards the future on E_2.

Although the definition of a contrastive event permits the background part of $\tilde{E}_1$ to occupy arbitrary spacetime regions, I will only consider contrastive events whose foreground part is located inside a

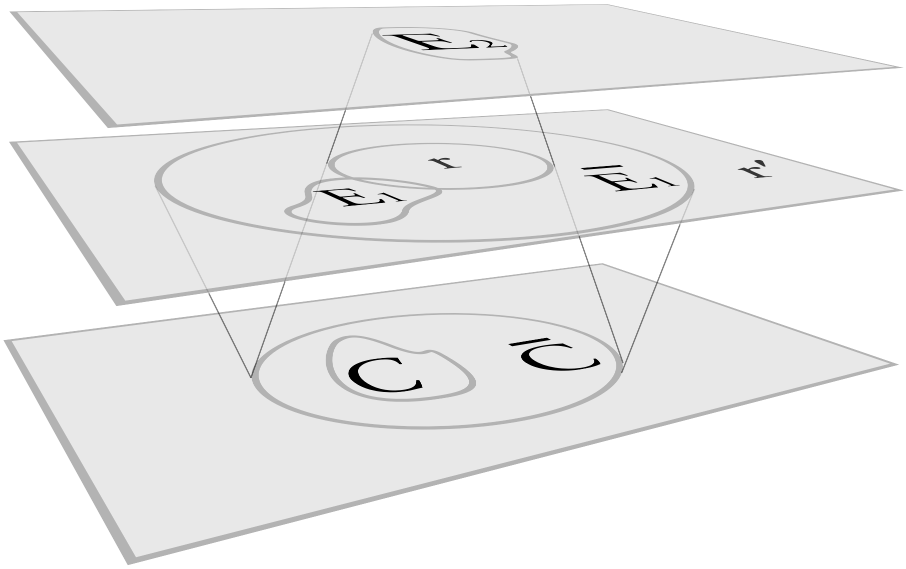

FIGURE 1. $\overline{E_1}$ shields $\overline{C}$'s probability fixing of E_2 because $\overline{C}$ can only impact E_2 through the region r.

compact region of spacetime and whose background part occurs at the same time as the foreground part.[3]

Let us first consider the contextualized event $\overline{E_1}$ in order to ascertain what it implies about the probability of E_2. It is possible that $\overline{E_1}$ extends far enough out into space to fix a probability for E_2 purely in virtue of the fundamental laws of nature dictate the future-directed evolution of any instance of $\overline{E_1}$. Let us say that $p_{\overline{E_1}}(E_2)$ is the temporally direct probability of E_2 fixed by $\overline{E_1}$, assuming it is defined. It is also possible that there is a probability fixed for E_2 by way of a nomic connection that starts with $\overline{E_1}$, goes towards the past to some intermediate event $\tilde{C}$, and then goes back towards the future to E_2. My aim is to prove that when any such probability is defined, it is equal to $p_{\overline{E_1}}(E_2)$.

I will assume the structure of spacetime is either Galilean, which is appropriate for the non-relativistic theories, or is semi-Riemannian,

---

[3] In particular, I will not consider the prob-influence exerted by spatiotemporally disconnected contrastive events where the foreground part is located inside a compact region that lies to the future of a vast time slice. The prob-influence exerted by such events does not count as backtracking influence because it exists only in virtue of the temporally direct probability-fixing of the vast time slice. The conclusion I am defending here only concerns genuine backtracking influence. I discuss the prob-influence exerted by such disconnected events in Kutach (2011b), but in any case they do not constitute exceptions to any of the claims I make here.

which is appropriate for relativistic theories. Both kinds of spacetime allow one to define space-like tangent vectors. For relativistic theories, a vector at p is space-like iff it points outside of p's light cone and for the non-relativistic theories, a vector at p is space-like iff it points inside the unique plane of simultaneity that p inhabits. For convenience, let us say that any path that is everywhere smooth and whose tangent is nowhere space-like is a c-path. Then, we can say that an event c is c-connected to an event e iff every point of e can be reached from some point of c by way of a c-path. To get an intuitive grasp of a c-connection, it helps to think of relativistic electromagnetism; in that case, c is c-connected to all and only events that are at least partially in c's light cone. The reason to care about c-connections is that, as far as paradigm fundamental laws are concerned, only c-connected events are related to each other via fundamental laws. In short, according to paradigm theories of fundamental physics, events only termine events to which they are c-connected. Fig. 1 depicts events as if the spacetime is Minkowski spacetime, but my proof nowhere depends on this. In non-relativistic theories, the contextualized events extend infinitely far out into space.

For the sake of a simpler proof, I will assume the event $\tilde{E}_1$ is instantaneous. It is not technically challenging to generalize the proof to handle temporally thick $\tilde{E}_1$'s, but accommodating arbitrarily shaped events unnecessarily obscures the important philosophical issues. In any case, it should be intuitively plausible that if the backtracking prob-influence issuing from every instantaneous slice of a temporally thick $\tilde{E}_1$ is guaranteed to be redundant, one does not acquire non-redundant backtracking prob-influence merely by considering $\tilde{E}_1$'s time slices in aggregate.

I will also assume the situation under consideration does not involve any spacetime structure with wormholes. The right kind of wormhole might permit non-redundant backtracking influence by allowing some influence to first go back in time and then go forward into the future by sneaking past $\tilde{E}_1$ through the wormhole. I will only be arguing that if such wormholes are unavailable for exploitation, then any backtracking prob-influence is redundant.

The central idea behind the proof is simply that the dynamical laws of our world propagate events continuously through time. This idea can be broken down into several rules governing terminance. Because all four paradigm fundamental theories obey the following principles and because there is no empirical evidence that disconfirms any of them, we have some reason to believe they are true.

- Transitivity: Any terminant of e termines whatever e termines. Furthermore, this transitivity extends to probability-fixing more generally; if an event occupying region r fixes an event in region r' which in turn fixes an event in region r'', then the event in region r fixes a unique event in r'', though not necessarily the same event fixed by the event in r'.
- Density: For any terminant c of an event e, if there exists a spacetime region r such that (1) every point of r lies on a c-path going from c to e, and (2) no point of c can be c-connected to e without intersecting r, then there exists an event i fully occupying r such that c termines i and i termines e. Density merely generalizes the claim that physical states evolve continuously through spacetime without ever hopping over intervening states. Density extends to probability-fixing among coarse-grained events; if an event occupying region c fixes an event occupying region e and there is a region meeting conditions (1) and (2), then there is a unique event occupying that region that fixes the event at e and is fixed by the event at c.
- Non-spatiality: All events that fix a probability for e are c-connected to e. Non-spatiality just expresses that events never fix probabilities for what happens in regions that are space-like separated from them, i.e. for other distinct events that occur at the same time. I will comment further on this principle after the presentation of the proof.
- Shielding: The probability that a terminant c fixes for e is unaltered by augmenting c with events that cannot be c-connected to e without passing through c. Shielding in effect claims that the state of the world at any one time incorporates all relevant information from its past (future) when it fixes probabilities for what happens in its future (past).

Because we do not know of any good candidates for the true, complete set of fundamental laws, we cannot be sure that all actual laws obey these principles, but they are reasonably uncontroversial as far as physical principles go.

Here is the proof. The contextualized event, $\overline{E_1}$, can only fix a probability for E_2 in virtue of the probabilities fixed by its elements. So let us consider an arbitrary element, e_1, of $\overline{E_1}$. In order to fix a probability for E_2 by backtracking though what happens in the past, e_1 needs to fix some contextualized event $\overline{C}$ that lies wholly to the past of e_1, which in turn fixes a probability for E_2. Let region r' include all the spacetime points occupied by e_1 but extended out in space (as a space-like surface) as far as the universe's spacetime structure allows. Let region r be defined as all the spacetime points of r' that are c-connected to E_2. This definition ensures that r satisfies conditions (1) and (2) of the density principle, taking into consideration that the possibility of wormholes and similar topological pathologies have been set aside. In virtue of the density principle, because $\overline{C}$ fixes a probability for E_2, it also fixes a unique event occupying r. Then, because e_1 fixes $\overline{C}$ and $\overline{C}$ fixes an event at r, by transitivity it follows that e_1 fixes an event occupying r.

Furthermore, the region r must be a subset of the region occupied by e_1 because if r were to include any point lying outside e_1, that point would be located at space-like separation from e_1 and so e_1 would fix a probability for what happens at space-like separation, in violation of non-spatiality.

Because region r is already fully occupied by e_1, the event that e_1 fixes for r must be just the event e_1 itself, restricted to the region r.

Because all the probability-fixing that is routed through $\overline{C}$ can only get to E_2 through r, it can only get to E_2 through e_1. Thus, e_1 forms a shield, which according to the shielding principle means that the probability e_1 fixes for E_2 is unaltered by adding any events, like $\overline{C}$, that occur to the past of e_1. Thus, any probability fixed for E_2 by backtracking from e_1 through some $\overline{C}$ must equal $p_{e_1}(E_2)$, which already holds purely in virtue of future-directed terminance. Since this equality holds for all the elements of $\overline{E_1}$, it holds for $\overline{E_1}$ itself. Thus, any backtracking probability-fixing issuing from $\overline{E_1}$ and arriving at E_2 is equal to $p_{\overline{E_1}}(E_2)$.

The argument can be repeated to show that the probability fixed by any nomic connection that backtracks from $\overline{\neg E_1}$ through some $\overline{\neg C}$ to E_2 must be the same as the probability fixed by the direct connection from $\overline{\neg E_1}$ to E_2. Thus the degree of prob-influence we get from backtracking from $\tilde{E}_1$ is always $p_{\overline{E_1}}(E_2) - p_{\overline{\neg E_1}}(E_2)$. Thus, no matter how the original E_1 event is represented as a contrastive event, its backtracking prob-influence of E_2 amounts to nothing more than its purely future-directed prob-influence of E_2. Q.E.D.

The proof assumed that E_2 happened entirely after $\tilde{E}_1$, but it should be clear enough what to say when E_2 precedes $\tilde{E}_1$. $\tilde{E}_1$ can exert non-trivial past-directed prob-influence on E_2 directly, but any probability-fixing that goes from $\overline{E_1}$ into the past through E_2 to $\overline{C}$ and then backtracks towards the future to E_2 will be the same as the directly fixed probability, $p_{\overline{E_1}}(E_2)$. Essentially, propagating $\overline{E_1}$ or $\overline{\neg E_1}$ back to E_2

fixes a determinate probability for E_2 that is not altered or augmented by further propagation of the state into the past. Furthermore, whatever parts of E_2 happen at exactly the same time as $\tilde{E}_1$ must themselves be a part of $\tilde{E}_1$, which ensures that the probability of E_2 is trivially fixed by $\tilde{E}_1$ regardless of any backtracking. When E_2 occurs partially before $\tilde{E}_1$, partially at the same time, and partially after, then the above considerations apply individually to each of these three temporal parts.

Although non-spatiality is obeyed by all four paradigm fundamental theories, it may not have overwhelming prima facie appeal. Given that there is a form of instantaneous influence in the arbitrarily fast action of the classical gravitational force[4] and that quantum mechanics reliably exhibits non-local correlations that are arguably causal, it is not too much of a stretch to posit fundamental laws that violate non-spatiality. However, I think there are good reasons for doubting that violations of non-spatiality actually occur.

First, if there were laws enforcing constraints on the arrangement of matter in a single time slice t where an event located in region p fixes a probability distribution over what happens in region q, then it would be puzzling how we are often seemingly able to control (within reasonable bounds) what happens at a given time in different locations. For example, I can routinely place a book on the right half of some table and if there were a law under which the book fixes a probability for the existence of a lamp on the left half of the table, it is unclear why I could not disconfirm the law by placing or not placing a lamp on the left half at will to contravene the probability assigned by the law. At least, it is a challenge to formulate laws with space-like terminance that could withstand a determined attempt to flout them.

Second, the function of non-spatiality in the proof is merely to rule out future-directed nomic connections coming from $\tilde{C}$ that zigzag spatially around the edge of $\tilde{E}_1$ to reach E_2 without going entirely through

---

[4] In classical gravitation, the events that play the key role in causation are determinants that span a full time slice of Galilean spacetime, specifying the relative location and speeds of every particle in the time slice and also specifying where the time slice is just a vacuum. Events that are smaller than a full time slice, even if they only omit specification of what happens at a single point p, will typically fail to determine (or fix probabilities) for anything else because what happens at other times depends on whether there is a massive particle at p. And such events do not determine or fix a probability for whether there is a particle at p. In some models of classical gravitation, one can postulate that the gravitational field is an ontologically independent quantity, which could lead to determination at space-like separation, but I am assuming a model that omits the gravitational field from the fundamental ontology.

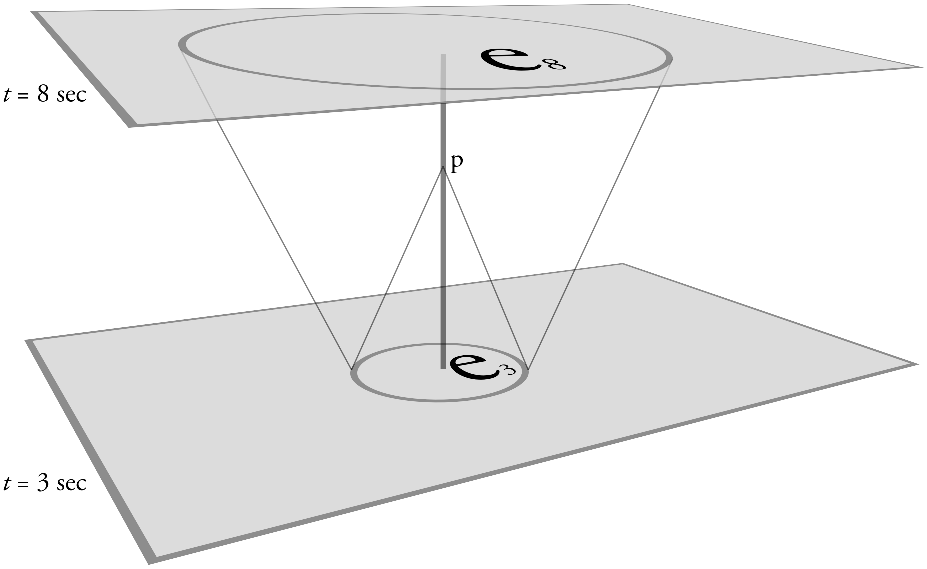

FIGURE 2. As e_3's influence expands towards the future, its terminance contracts until it vanishes at p. Also, e_8's terminance contracts towards the past at the same rate that e_3's influence expands towards the future.

$\tilde{E}_1$, and such models appear to be implausible because they violate a reasonable relationship between influence and terminance. Stated simply, as time goes by, influence spreads, and as influence spreads, terminance shrinks. For illustration, consider the rest frame of a single inertial particle in Minkowski spacetime, as depicted in Fig. 2. Let e_3 be an instantaneous spherical event centred on the particle at time $t = 3$ sec that is three light-seconds in radius, and let e_8 be an instantaneous spherical event centred on the particle at time $t = 8$ sec that is eight light seconds in radius. It is uncontroversial that every part of e_3 influences e_8 in the sense that hypothetical alterations to the magnitude of fundamental fields or particles instantiated by e_3 can make a difference as to what happens (or to the probability of what happens) at e_8. There are also physical quantities at $t = 3$ instantiated outside of e_3 that influence what happens in their future light cone, and together they influence everything in the future except what happens in the cone whose base is e_3 and whose tip is p, the particle's location at $t = 6$. This conic region also includes everything that e_3 termines towards the future. Notice that because the influence from events at

$t = 3$ outside e_3 spreads at the speed of light, the region that e_3 termines shrinks at the speed of light. Thus, there is a sort of reciprocal relationship between influence and terminance.

In non-relativistic theories, this relationship is obscured because it exists in the limit where influence is infinitely fast. Because non-relativistic interactions such as the gravitational force are transmitted immediately, terminance of what happens at any given point p requires a completely specified state s occupying a global time slice that does not intersect p, except for some rare special cases that can be ignored. If the reciprocal relation held in full generality, the terminance issuing from s would immediately shrink to a single point, but because every point located off the global time slice is equally preceded by (or followed by) the global time slice, s termines every point and thus termines the entire history of the universe.

The reciprocal relation between influence and terminance also holds with regard to past-directed influence and terminance. In relativistically local theories, for example, any event at $t = 8$ that does not include e_8 will fail to termine e_3, and any event at $t = 8$ that termines e_3 does so only in virtue of its including e_8 as a part; the quantities instantiated outside e_8 are nomologically irrelevant. If the fundamental laws include past-directed terminance at all, as one goes back in time from $t = 8$ to $t = 3$, what is *termined* by e_8 at t is an ever smaller sphere until e_3 is reached. Correspondingly, as one goes forward in time from $t = 3$ to $t = 8$, what is *influenced* by e_3 at t is an ever larger sphere until e_8 is reached. Assuming this relation between influence and terminance holds of the actual world, we can expect $\tilde{C}$ to fail to exert prob-influence on E_2 by zigzagging around $\tilde{E}_1$ because in order to do so, the terminance issuing from the elements of $\overline{E_1}$ and $\overline{\neg E_1}$ would have to go towards the past and then back to the future and expand at some stage during this backtracking. This expansion cannot occur because (1) any influence $\tilde{E}_1$ has on the past must spread out as time goes by (towards the past), so that its past-directed terminance must correspondingly shrink, and (2) any influence $\tilde{C}$ has on the future must spread out as time goes by (towards the future), so that its future-directed terminance must shrink. In the non-relativistic case, $\tilde{E}_1$ cannot be circumvented because it must already span the entire universe in order to exert any prob-influence at all.

4. Causal Asymmetry

It might be puzzling why one should ever bother to demonstrate the redundancy of backtracking influence in a temporally neutral way.

After all, we need to account for why causation is temporally asymmetric, and it presumably follows from the asymmetry of causation that the past-directed portion of backtracking influence does not exist and thus that backtracking influence itself does not exist. The reason I presented the argument is that I think it is better to reverse the order of explanation so that instead of using the asymmetry of causation to rule out backtracking influence, one can use the redundancy of backtracking influence to provide a superior analysis of the asymmetry of causation. Because the topic of causal asymmetry is too complex to address adequately here, I will just briefly discuss how the redundancy of backtracking prob-influence helps to explain one crucial part of the asymmetry of causation. I provide a more thorough discussion elsewhere (Kutach 2011a, 2011b).

A key component of what we ordinarily describe as the asymmetry of causation is the advancement asymmetry. The advancement asymmetry is the demonstrable fact that agents are sometimes able to advance goals they have for the future, but they are always unable to advance any goals they have for the past. You can test this claim yourself by choosing some kind of event E to represent a possible goal like writing a poem or eating a sandwich. Then, randomly assign a zillion people either the task of trying to accomplish E or the task of trying to prevent E. I think you will find that when the task is located in the agents' past, E will occur at very nearly the same frequency for both groups. For example, assign to one randomly selected group of people the task of chewing gum yesterday and assign everyone else the task of not chewing gum yesterday. You should find that the fraction of gum chewers in the first group is very nearly the same as the fraction of gum chewers in the second group.

One common explanation of the advancement asymmetry is simply to presume some sort of fatalism about the past. Agents cannot advance goals for the past, so the story goes, because it is not up to them what went on beforehand, or because effects never precede their causes, or because the past is essentially immutable, or because the past does not counterfactually depend on the future, etc. However, anyone who has (rightly) adopted a framework where influence is construed as some sort of difference-making and has accepted that such difference-making ought to be evaluated in terms of what the laws of nature imply about hypothetically postulated situations is in a position to provide an alternative answer that does not take for granted the fixity of the past. A variety of explanations for the advancement asymmetry are potentially available because there are many ways influence can fail to be exploitable for furthering goals. For example, many instances of influence are too weak, such as our personal gravitational influence on the planets. Other instances of influence are too chaotic, such as our present influence on the number of insects that will be alive three hundred years from now. If it could be shown (e.g. Kutach 2011a, 2011b) that we have influence over the past but that it is always unexploitable for advancing goals, we would have a good explanation for why the past appears to be immune from influence, or at least why it is safe to treat it as such.

The redundancy of backtracking prob-influence can be employed to great effect in such an explanation. The goal is to explain why an agent who, at time t, is randomly assigned the task of bringing about an event E before t is no more likely to be paired with E's previous occurrence than an agent who is assigned the task of preventing E. In the language of contrastive events, this becomes the claim that E is not prob-influenced by the later contrastive event $\tilde{C}$, whose foreground constitutes an assignment of the task to bring about E rather than prevent E and whose background includes a suitably competent agent who is capable of understanding the task assignment and is motivated to accomplish the task.

The explanation of why $\tilde{C}$ does not prob-influence E can be divided into two tasks. The first is to explain why $\tilde{C}$ does not prob-influence E in virtue of what it implies for what happens after t, and the second is to explain why $\tilde{C}$ does not prob-influence E in virtue of what it implies directly towards the past. The above argument for the redundancy of backtracking prob-influence ensures that if one can accomplish the second task, the first task is automatically accomplished. That is, the redundancy of backtracking guarantees that nothing happening after the agent finds out she is supposed to accomplish task E will affect the probability of E beyond what is already encoded in the event at t, when the agent has not yet found out what she is supposed to do. That goes a long way towards showing that the activity of the agent will not make E any more or less likely because it is not very plausible that $\tilde{C}$ itself directly (towards the past) prob-influences E. To get an intuitive feel for this claim, let E be "The agent carved her initials into a tree last week," so that $\tilde{C}$ at time t is the event of a message being delivered to the agent which has ink patterns in the form, "Carve your initials into a tree last week" rather than in the form of "Don't carve your initials into a tree last week." Intuitively, it would be puzzling how, when propagating the complete physical state of the world at t towards the past according to the fundamental laws, the precise text on the paper alone—holding everything else at t fixed—could make a difference as to whether the person who will soon receive the note carved her initials into a tree recently. Although caution about such intuitions is advised owing to the subtly of past-directed prob-influence, I believe one can adequately explain why $\tilde{C}$ does not directly prob-influence E. However, to identify the best explanation and to close all the loopholes comprehensively, though, requires an assessment of the relative explanatory merit of various asymmetries in the historical layout of matter, i.e. an investigation of the details of the fork asymmetry, the asymmetry of entropy, and other candidate explainers of causal asymmetry. The importance of the redundancy of backtracking prob-influence is that it significantly reduces the problem of accounting for the advancement asymmetry without making any assumptions whatsoever about which aspects of the universe's material layout ultimately account for the asymmetrical character of causation.

Douglas Kutach is at the Department of Philosophy, Brown University. Correspondence to: Box 1918, Brown University, Providence, RI 02912. Email: Douglas_Kutach@brown.edu.

References

Albert, D. 2000. *Time and chance.* Cambridge: Harvard University Press.

Beauchamp, T. and R. Rosenberg 1981. *Hume and the problem of causation.* Oxford: Oxford University Press.

Broad, C. D. 1923. *Scientific thought.* New York: Harcourt, Brace and Co.

Dowe, P. 2000. *Physical causation.* Cambridge: Cambridge University Press.

Ellis, G. 2006. Physics in the real universe: Time and spacetime. *General relativity and gravitation* 38, no. 12: 1797–1824.

Gasking, D. 1955. Causation and recipes. *Mind* 64: 479–487.

Hausman, D. 1998. *Causal asymmetries.* Cambridge: Cambridge University Press.

Hume, D. 1739. *A treatise of human nature.* London.

Kant, I. 1781. *Critique of pure reason.* English translation by N. K. Smith, 1933. London: Macmillan.

Kutach, D. 2011a. The asymmetry of influence. In *The Oxford handbook of philosophy of time*, edited by C. Callender. Oxford: Oxford University Press.

Kutach, D. 2011b. *Causation and its basis in fundamental physics.* Forthcoming.

Lewis, D. 1979. Counterfactual dependence and time's arrow. *Noûs* 13: 455–476.

Loewer, B. 2007. Counterfactuals and the second law. In *Causation, physics, and the constitution of reality*, edited by H. Price and R. Corry. Oxford: Oxford University Press.

Mackie, J. L. 1973. *The cement of the universe.* Oxford: Oxford University Press.

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

Mellor, D. H. 1995. *The facts of causation.* New York: Routledge.

Price, H. 1992. The direction of causation: Ramsey's ultimate contingency. *PSA 1992, Volume 2*, edited by K. Okruhlik, D. Hull, and M. Forbes. East Lansing, MI: Philosophy of Science Association, 253–267.

Reichenbach, H. 1956. *The direction of time.* Berkeley: University of California Press.

Suppes, P. 1970. *A probabilistic theory of causality.* Amsterdam: North-Holland.

von Wright, G. 1975. *Causality and determinism.* New York: Columbia University Press.