Causality and Relativity

Steven Weinstein*
Dept. of Philosophy, U. Waterloo, Waterloo, ON N2L 3G1
Perimeter Inst. for Theoretical Physics, Waterloo, ON N2L 2Y5

Draft of November 1, 2007.
Not for general distribution. Please do not quote!

Abstract

Causality is analyzed in the light of modern relativistic theories. It is shown that such theories provide a natural home for causation, and provide a framework for understanding both the time-asymmetry of causation and single-event causation without introducing any further, contingent features of the world.

1 Introduction

Modern physical theory makes frequent reference to principles of causality – ‘relativistic’ or ‘Einstein’ causality in classical contexts, and ‘microcausality’ in the context of relativistic quantum field theory. Yet many philosophers, frustrated by difficulties encountered in making the notion of causality itself precise, have recently inclined toward what a recent collection of papers on the subject calls ‘causal republicanism’, the view that “although the notion of causation is useful, perhaps indispensable, in our dealings with the world, it is a category provided neither by God nor by physics, but rather constructed by us.” ([PC07], 2). Eagle argues that "our best physics don’t provide a determination relation that meshes at all well with our folk theory of causation" and that "the real fault lies in the conception of causation as requiring determination" ([Eag07], 161). Similarly, Kutach claims,

The idea that a central organizing principle of nature is a causal relation between events is not motivated by a serious examination of fundamental physics. What we do find in our best fundamental theories are equations expressing relationships between physical quantities at different times and places, equations that have no obvious connection with the notion of causation. ([Kut07], 327).

*Email: sw@uwaterloo.ca, sweinstein@perimeterinstitute.ca
This attitude toward causation finds its inspiration in Russell [Rus13], though the idea that the physics of the day does not incorporate causation at a fundamental level also finds expression in the logical positivist movement [Car28].

Skepticism about finding causation in fundamental physics is echoed on a larger scale by what sometimes seems to be kind of despair about finding a philosophically sound and robust conception of causation, period. In his review of Phil Dowe’s book *Physical Causation* [Dow00], Daniel Hausman, author of his own tome on causality [Hau98], begins by saying:

Causation is a frustrating subject. Suppose one begins with some promising idea such as that causation is counterfactual dependence or statistical relevance. One then develops this idea with care and intelligence, revises and improves it to cope with criticisms, and by the time one is finished, sane people will be looking elsewhere... Causation continues to stump us all.[Hau02]

In this paper I shall argue that there is a very robust notion of causation associated with relativistic theories, one which both accords with and illuminates our intuitive, “folk” notions of causation. In short, the idea is that the lawlike determination characteristic of relativistic theories is causation.

This is in itself not a particularly new idea (see e.g. [Mac65]). But it seems to have been thought that such a view does not allow us to make sense of the time-asymmetry of causation, and that it does not accommodate our intuition that causes can be single events. I will argue that, correctly viewed, it does both. And in the process I will show that the apparent physical asymmetries in the world described by the second law of thermodynamics have nothing to do with the time-asymmetry of causation. Furthermore, we shall see that this view has the advantage of providing a framework within which one can see both the power and the limitations of counterfactual [CHP04] and interventionist [Car79] [Hit07] [Woo03] accounts of causation.

2 Relativity

For our purposes, the feature of relativistic theories which is most important is that the physical properties of a given point in spacetime $E$ – the physical values of the fields (including matter) at $E$ – are determined by the values of the fields at a compact set of spacetime points at some other time. For ordinary relativistic theories, the field at $E$ is determined by the field on (and in general within) a “slice” of the past or future lightcone.1 [See Figure 1] For example, the properties of the electromagnetic field at the point where you sit reading this paper now are fully determined by the values of the field one second ago at the points

---

1 One can also write down relativistic (i.e., Poincare-invariant) theories for which the domains of influence and dependence are generically wider than the lightcone. Several examples are discussed in [Wei06].
in space 300,000,000 meters distant from you. (Distances and times are relativistically meaningful only with respect to a given reference frame, but it is characteristic of relativistic theories that any frame will do.) Thus this set of physical facts – I will be provocative and collectively call them \( C \) – determine \( E \). And I propose that this sort of determination grounds the correct and proper analysis of the concept of causation.

Two problems with this analysis are likely to occur immediately to the devoted student of causation:

1. The time-reversal invariance of Maxwell’s theory of electromagnetism (characteristic of many, if not quite all, physically meaningful theories), implies that \( E \) is also determined by the physical facts \( C' \) which correspond to the values of the electromagnetic field at the same points but at a time one second to the future of \( E \). This implies that the “causal arrow” may be understood as running from future to past, as well as, or in addition to, running from past to future.\(^2\) This seems highly counterintuitive, and offers no explanation for the seeming fact that we cannot affect the past.

2. There is a counterintuitive asymmetry between causes and effects – not the usual one whereby causes precede their effects but effects never precede their causes, but one whereby causes are sets of events, whereas effects are single events.\(^3\)

These will now be addressed in turn.

### 3 Time-asymmetry of causation

Consider two dominoes \( L \) and \( R \) standing on edge, \( L \) on the left and \( R \) on the right. If \( L \) is perturbed so as to fall to the right toward \( R \), \( R \) will be perturbed and fall in turn (see Figure 2). We say that \( L \)’s fall is the cause of \( R \)’s fall, and that \( R \)’s fall is an “effect”. We are not, however, inclined to say the opposite, that the fall of \( R \) causes the fall of \( L \). Clearly we are inclined to identify the earlier of two causally connected events as the *cause*, and the later as the *effect*. But why? How can we ground this in a physical theory (such as classical or relativistic mechanics) which is time-reversal invariant at the fundamental level?

The usual answer to this question involves invoking, in some way or other, the time-asymmetry encoded in the second law of thermodynamics. The idea behind this line of response seems to be that the second law tells us that entropy (a quantity which, roughly speaking, represents the amount of disorder in the world) increases with time, that we identify the direction of time with the direction of increase in entropy, and that interactions which

\(^2\)It can also be understood as running “sideways” (up to down, right to left, etc.), at least in the case of the fields themselves (as opposed to the charged matter coupled to the fields). See [CW07] for discussion.

\(^3\)I note in passing that this asymmetry is more problematic in Newtonian (i.e., non-relativistic) gravitation.
merit the name ‘causal’ are ones which lead to an increase in entropy, such that the production of an effect by a cause results in an increase in entropy. In our example, there is indeed an entropy increase due to the fact that energy is dissipated as the dominoes fall through the air and collide, the collision converting some (eventually all) of the kinetic energy of the falling domino into sound and heat. Thus the interaction is not reversible, in the sense that, in the absence of a near-miraculous orchestration of air and floor molecules, we will never observe R to rise up from the floor and push L back so that they are both standing upright at the end of the interaction.

This sort of story explains the causal asymmetry insofar as one is prepared to define the distinction between cause and effect in terms of the thermodynamic arrow. Under such an approach, effects are never observed to precede their causes because entropy almost never decreases. One problem with this approach is that the distinction between cause and effect requires coarse-graining. On a fine-grained view, where one keeps track of the motion of all relevant molecules, there simply is no change in entropy. Thus on a fine-grained view, one must say that the distinction between cause and effect vanishes, leaving us without an unambiguous distinction between cause and effect. Many who advocate this view would readily accept such a conclusion, arguing that causation is perspectival, and (more importantly) that part of one’s perspective is a certain coarse-graining built into one’s description. Alternatively, one might argue in a related way that this type of event (the time-reversed situation) is rarely observed, and that it is in virtue of the fact that the token event is a member of a rare type that we reject it as a ‘cause’. Thus one could either say that the division of the actual world into causes and effects is meaningful only at a macroscopic level, or that it is meaningful on a microscopic level as well in virtue of the fact (such as it is) that the microstate in question is a member of a certain kind of macrostate.

Yet there is another potential problem, for one might object that the question of time-asymmetry is not the question of whether the time-reversed events can happen, but why we don’t say that R’s fall causes L’s fall. And there are even further variations which I shall not essay here. Luckily, the situation is not nearly so complicated, for the second law is a red-herring. We are both willing and able to tell unambiguous causal stories in cases which do not involve dissipation and entropy increase in an essential way. Consider Price’s beautiful “stargate doughnut” example:

Imagine a photon, p, which spends billions of years in intergalactic space as it travels from one distant galaxy, G_{past}, to another, G_{future}. At some point in between, at a time t, it passes through the central aperture in a tiny doughnut-shaped object, which happens to be spinning on a transverse axis, somewhere in deep space. As this doughnut spins, it periodically occludes the path the photon takes from G_{past} to G_{future}, and hence acts as what we could call a ‘stargate’. At t, however, the gate is open, and the
path is unobstructed.

Consider the following counterfactuals:

Proposition 1: If Stargate Doughnut had been closed at time t, the photon p would not have been absorbed at $G_{\text{future}}$.

Proposition 2: If Stargate Doughnut had been closed at time t, the photon p would not have been emitted at $G_{\text{past}}$.

We take Proposition 1 to be true, and Proposition 2 to be false. Accordingly, we take the orientation of the stargate to be a cause of the later position of the photon, but not the earlier position of the photon. What is the source of this time-asymmetry?

On the face of it, the example seems completely independent of any thermodynamic details of the systems in question. It is far from clear that any sense can be made of the notion of the entropy of a discrete microscopic system of this kind. Even if it could, however, the example doesn’t depend on the existence of a thermodynamic gradient. After all, we could quite well imagine the same kind of situation, in a universe in thermodynamic equilibrium. (We’d need to replace galaxies with some photon source at the same temperature as everything else, but this makes no significant difference to the example.)

In some sense, of course, our intuitions about the case are sensitive to temporal direction. What else could we rely on in distinguishing Proposition 1 from Proposition 2, after all, given their apparent symmetry in other respects? But as we saw, however, it is no use trying to rely simply on temporal direction, to explain the difference between the two cases. If we say it is merely a conventional matter—a matter of the meaning of ‘cause’ and ‘effect’—that causes occur before their effects, this will certainly imply that Proposition 2 (or a causal variant of it) must be false; but it doesn’t explain why that is relevant to our decision behaviour, in the way that it is. (We think that we could use the stargate to influence future photon positions but not past photon positions, for example.)

([Pri07], 266-267)

Price understands this example to show that our intuitions about the direction of causation have no basis in fact, but rather in the temporal asymmetry of the perspective we bring to the world (a perspective which he thinks is not imposed upon us, hence can be opted out of). He writes,

[W]hat we bring to the case, in imagination, is the typical perspective we have as deliberating agents – the perspective we bring to the situation, quite unconsciously, when we think about manipulating the stargate. This deliberative perspective displays a very marked temporal bias, of course. Roughly, it treats the past as fixed, and only the future (or some subset of the future) as under our control. ([Pri07], 267)
In effect, then, Price is arguing that in cases of causal connection, our inclination to hold the past fixed rather than the future in evaluating the truth of counterfactuals is a choice we make which reflects our perspective, a choice grounded neither in the second-law of thermodynamics nor in any other physical facts. Notably, this leaves the door open for backward causation, which Price elsewhere suggests is a possible explanation for the curious phenomena displayed in EPR-type experiments, phenomena which quantum theory regards as arising from entanglement. ([Pri96], 233-260).

The perspective I suggest on perspectivalism is different. It is a perspective from which one may view the perspectival aspects of causation as objective. And why? Because describing a physical situation – e.g. a set of field values at some spacetime point or in some compact spacetime region – as an event presupposes the selection of a temporal orientation. Without a temporal orientation, there is nothing one can call an event at all.

For example, Price’s stargate example involves two events, which may or may not be “caused” by the fact that the stargate doughnut is closed. What are these two events? In Price’s description, they are the absorption of the photon by some galaxy $G_{\text{future}}$ in the future and the emission of the photon from some other galaxy $G_{\text{past}}$ in the past. Consider the former, and ask, what is the time-reversed description of this situation? It is not the absorption of a photon, it is the emission of a photon from $G_{\text{future}}$. Similarly, the time-reversed description of the latter event is the absorption of a photon at $G_{\text{past}}$.

Causally, what’s going on in Price’s example is as follows. The world consists of a photon passing through a spinning doughnut and matter in two locations which can act as a source or a sink for the photon. A “cause” of any event occurring at a point $x$ on the photon’s trajectory is, from the nomological standpoint urged here, simply the state of the world (the state of the photon, the doughnut, the absorbers/emitters, plus the trivial state of the otherwise empty space) at any given instant in time over all points in or on the forward or past lightcone of $x$. But which, forward or past? That depends on the event, which in turn depends on the temporal perspective. Describing the physical situation at $G_{\text{future}}$ as the absorption of a photon involves adopting a perspective, a choice of temporal orientation, whereby the photon’s velocity $v$ is directed toward $G_{\text{future}}$. If we time-reverse the description, adopting the opposite temporal orientation, we find that we have $G_{\text{future}}$ emitting a photon with opposite velocity $-v$, which takes it away from $G_{\text{future}}$ as it travels into the past. We have two different events corresponding to the same physical situation, just as we would have two electric fields corresponding to the same physical situation described in different inertial frames. If we agree that we want to ascertain the cause of the absorption of the photon at $G_{\text{future}}$, then we will look to the past lightcone for a set of events corresponding to the cause. These events, too, presuppose a temporal orientation, though the only salient one here is the velocity of the photon. Just as in the domino case it would make no sense to say that domino $L$’s reverse-fall (from the time-reverse perspective) causes $R$’s fall, it makes no sense
in the Stargate case to say that the photon with velocity $-v$ causes the photon to arrive at $G_{future}$ with velocity $v$. There is no privileged temporal orientation, but one must keep the temporal orientation the same everywhere.\footnote{It is probably not surprising that spacetimes which are not temporally-orientable (a Möbius strip is an example of a non-orientable surface) present problems for causal analyses. However, it is worth noting that causality worries for spacetimes with closed timelike curves, spacetimes which may well be temporally-orientable, arise only at the level of single-event causation, discussed below.}

One might argue that viewing causation in this way makes it perspectival after all, but this is not so important. What is important to note, in any case, is that perspective enters at a very different place than it does for Price.\footnote{An analogous point arises for quantum mechanics, where the time-reversed description involves complex conjugation.} Price tends to think of events as spacetime points (thus ‘events’ in the technical sense of relativity theory), but in doing so misses their temporal orientation, an orientation having to do with the fact that both Newtonian and relativistic classical theories insist that initial conditions, and states in general, are characterized in terms of properties (positions, field strengths) and their time-derivatives.\footnote{An analogous point arises for quantum mechanics, where the time-reversed description involves complex conjugation.} Thus Price is correct in identifying the role of perspective in yielding the time-asymmetry of causation, but incorrectly associates the perspective with the prejudices introduced by observers, who after all do have an apparent temporal orientation. Because of this, Price is inclined to view backward causation as not only a real possibility, but a possibility distinct from forward causation. He thinks that our inclination to regard future events as caused by past events might undergo revision, and in particular that quantum phenomena might be just the sort of phenomena that encourage us to adopt a different perspective. From the point of view argued for here, this makes no sense, as the time-reversed phenomena are different phenomena. A causal explanation of our observation of the cosmic microwave background radiation, for instance, looks to the past toward the big bang, not because the radiation has an intrinsic direction (it doesn’t), but because the very thing we are trying to explain, our observation, our absorption of radiation of a certain kind, is temporally oriented.

\section{4 Single-event causation}

I have so far discussed examples in which we understand one event to be the cause of another. Yet the model of causation I am using, the model coming from relativity theory, is one in which causes are sets of events, not single events. The true, full cause of domino $R$’s falling includes not only the prior fall of domino $L$, but also other conditions simultaneous (in some/any frame) with the fall of $L$, for it is only in conjunction with those conditions that one has any way of using the laws to predict the fall of $L$. For instance, we assume that no one stands at the ready to prop up $R$ in the event that it is struck by $L$. How, then, does it make sense to say that $L$’s fall causes $R$’s fall?
Strictly speaking, it does not make sense, because the fall of $L$ only causes the fall of $R$ if conditions allow. Specifying the conditions — spelling out \textit{ceteris paribus} clauses — is the bane of counterfactual analyses of causation, analyses whereby “$A$ causes $B$” is understood to mean “If $A$ had not happened, then $B$ would not have happened.” [Lew73] Such analyses tend to be forward-looking, holding the past fixed and pondering what effect locally different initial data would have on future behavior, but aside from the issue of time-asymmetry (addressed above), they suffer from two major problems. One, evaluation of the future evolution of the system requires one to know what happened \textit{instead} of $A$. Two, one needs to make peace with the possibility that other events at the same time, events which are held fixed in evaluating the counterfactual, may have the “effect” (I use this word advisedly!) of somehow interfering with the production of $B$. The causal \textit{relevance} of these other events is what gives rise to most, if not all, of the other problems which plague counterfactual analyses, problems involving multiple causation, trumping, preemption, negative causation, and so on [Lew04]. Single events do not “cause” anything in any precise sense. It is the fact that the determinacy relation in relativistic physics relates \textit{sets} of events to single events that explains the failure of counterfactual accounts to provide a globally satisfactory analysis of causation.

However, what is absolutely crucial to the \textit{possibility} of evaluating counterfactual claims \textit{at all} is the fact that data at any two points $x$ and $y$ spacelike points are \textit{independently specifiable} (modulo differentiability considerations). More specifically, it means that there is a well-defined set of allowed initial data for the partial differential equations governing the behavior of the system, and that knowing the values of the data at one spacetime point does not limit one to a proper subset of possible values at another point. So for example data for the free Maxwell equations are independent; for distinct points $x$ and $y$ on some initial (Cauchy) surface, $\vec{E}(x)$, $\vec{E}(y)$, $\vec{B}(x)$, and $\vec{B}(y)$ are mutually independent in this sense, despite being subject to constraints $\vec{\nabla} \cdot \vec{E}(x) = 0$ and $\vec{\nabla} \cdot \vec{B}(x) = 0$ for every spacetime point $x$. On the other hand, if the constraints were nonlocal, as they are in the timelike initial value problem [CW07], then the data at a given point restricts the data at neighboring points.\footnote{More precisely, given a uniform probability distribution with respect to the standard Lebesgue measure on the space of allowed initial data (i.e., data satisfying the constraints), $P(\phi(x)|\phi(y)) < P(\phi(x))$, in general, for a given field $\phi$ and points $x$ and $y$ on the initial data surface.}

The point, then, is that our common-sense causal claims, claims which identify single events as causes, are founded on a highly non-trivial feature of the fundamental physics, a feature which is independent of the determinacy properties. This goes a long way toward making sense of the meaning of “causality” as it appears in physics in the vague formulations of “Einstein causality” or “relativistic causality” and in the more precise quantum theoretic constraint of “microcausality”. I will discuss this further below, but first, let us sum up the significance of this discussion for the philosophical project of explicating causality.

I have suggested that single events are not “causes”, strictly speaking, but rather partial
(even if particularly salient) causes. This in effect takes the pressure off counterfactuals to provide a robust account of causality. From this perspective, counterfactuals can’t hope to eliminate ambiguity, because there is no objective sense in which single events bring about other single events. If one insists that a proper philosophical analysis of causation must allow single events to unambiguously cause other single events, then one will likely despair of finding a unified account of causation.

5 Conclusion

For our understanding of the world in general, and causation in particular, it matters little whether one accepts the idea that causation is a relation between sets of events and individual events, or whether one rejects this idea in preference for a set of somewhat vague yet perhaps more commonsensical notions, in which the associated idea that causes “bring about” their effects is more problematic. The difference is semantic. What is most important is that two central aspects of causation are elucidated.

The time-asymmetry of causation, as we have seen, is derivable directly from relativistic physics, once one accepts that lawlike determination is required in order for two events to be related as cause and effect. (That is, one accepts either that only sets of events are candidate causes, or that individual events are candidates only if they belong to sets of events which fully determine the effect.) The description of the effect $E$ picks out a temporal orientation, and it is this which determines whether future events $C'$ or past events $C$ (with appropriately oriented description) are salient. Since our event descriptions invariably share our own temporal orientation from past to future, so do our causal attributions. But they need not.

The second point that remains is that our ability to single out individual events as having any particular causal efficacy of their own depends on a very particular feature of the physical description, yet one so common as to be almost invisible: the independence of initial data at spacelike points. This independence is a feature of a theory, rather than the world: it is a property of the kinds of initial data allowed by a theory. The world, considered as a whole, simply has some initial data. Or more precisely, there are data at any given point in time which (assuming the standard cosmological model\(^7\)) are determined by data at prior times, and which determine data at subsequent times. The world is, from this perspective, simply one solution to the equations. On the other hand, our notions of signaling, and more generally of causation as manipulability are meaningful because in many situations, our choices and their “effects” are themselves part of initial conditions. The fact that our theories regard local events as “freely specifiable” allows us at least the possibility of regarding ourselves

\(^7\)Here I am referring to the hot big-bang model, though for our purposes most, perhaps all, inflationary cosmological models share this property.
as agents able locally to adjust or prepare physical objects and systems in a way that is independent of other things going on elsewhere at the same time. (Whether they are in any sense "independent" of our own past experience is arguably the nub of the free will question.)

What I have done here is to ground causation in the framework provided by modern relativistic physics, the framework of special and general relativity\(^8\). The laws of such theories tell us unambiguously how sets of events determine other, individual events. The time-asymmetry inherent in the description of events enables us to reconcile the time-reversal symmetry of the laws with the apparent asymmetry of the causal relation, while the fact that these theories allow one to freely specify data on spacelike Cauchy surfaces enables us to recover our common-sense notions of causation via counterfactual analyses, with all their potential for fuzziness and ambiguity. Sharpening the boundaries and removing the ambiguities of such accounts remains a fascinating and important task.

6 Appendix: “Einstein causality”

There is a variously formulated, somewhat vague principle of “Einstein causality” in the physics literature. (It is also known as “relativistic causality” or just “causality”.) It is generally understood as a prohibition on faster-than-light transmission of signals. It is sometimes pointed out that faster-than-light signaling would correspond to signaling “into the past” with respect to some reference frame, and often this is further understood (incorrectly, in my opinion; see [Wei06] for examples) as opening the door to causality paradoxes such as the grandfather paradox. Whatever the case with spacelike signal transmission, Einstein causality seems universally to be understood as precluding signaling into the past lightcone.

What is the origin of this principle? It is not too hard to see. Essentially, signals are understood as disturbances, physically manifested for example by a light pulse in an otherwise light-free background. Or consider a radio broadcast, a continuous stream of electromagnetic radiation which can be thought of as a stream of signals, the individual signal (say, the opening drum hit of “Like a Rolling Stone” [Dyl66]) being a particular property of the continuous stream of radiation. In either case, one can show that, in the case of signals qua electromagnetic radiation, signals cannot travel faster than light.\(^9\)

Perhaps the only puzzling aspect of Einstein causality is its evident time-asymmetry. I suggest that the time-asymmetry is justifiable in just the way shown above. That is, a given

---

\(^8\)More precisely, the present analysis applies straightforwardly to globally hyperbolic spacetimes, and certainly illuminates the situation with respect to non-globally hyperbolic spacetimes.

\(^9\)For a more precise discussion, see [Bri60]. Note that the claim is not that certain properties of radiation cannot travel faster than light – it is straightforward to direct a pulse of light against a distant object such that the image moves superluminally, and it is also well-known that the group velocity of a wave-pulse (as well as the so-called “signal velocity”) can move superluminally when passing through media with the appropriate properties.
phenomenon is a signal only when viewed in a particular temporal orientation, and so the “effects” of the signal are properly sought in the future with respect to that orientation. Pressing “send” to send an email is initiating a signal, but the time-reversed description of this same process is not a description of sending a signal.

References


