Time and Causation

Mathias Frisch

To appear in:
The Blackwell Companion to the Philosophy of Time,
edited by Heather Dyke and Adrian Bardon

Abstract

In this chapter I discuss the relation between the temporal and causal asymmetries. I critically examine temporal accounts of the causal asymmetry (such as Hume's) and attempts to reduce both the causal and temporal asymmetries to a third asymmetry. These attempts include physicalist reductions that see the asymmetry of thermodynamics as fundamental as well as subjectivist accounts that take a deliberative asymmetry of agents as basic. Existing versions of both kinds of account face significant problems. The prospects for a causal theory of the temporal asymmetry, by contrast, are better than is often assumed in the literature.

Keywords: Hume; Reichenbach; common cause; thermodynamics; relativity; physicalism; causal models; randomness; reduction; intervention.

1. Introduction

What, if anything, is the relationship between the arrow of time and the asymmetry between cause and effect? One of the central characteristics of the causal relation is that it is asymmetric: If *c* is a cause of *e*, then it is not the case that *e* is a cause of *c*.[1] This asymmetry seems to line up with a temporal asymmetry: at least in cases with which we are familiar causes do not occur after their effects. Thus, it is tempting to look for possible relations between the direction of time and the causal asymmetry. *Prima facie* there are three options for possible connections. First, the causal asymmetry might supervene on the temporal asymmetry; second, the temporal asymmetry might supervene on the causal asymmetry; or, third, the two asymmetries might have their origins in some third asymmetry on which they both supervene. In this paper I will examine these three options and try to offer a partial defense of the second option: a causal theory of the temporal asymmetry.

2. Reducing the causal to the temporal asymmetry

That the causal asymmetry can be reduced to the temporal asymmetry was famously claimed by David Hume, who takes the time symmetric notion of constant conjunction to be at the core of the causal relation and defines the cause as the earlier of the two tokens of the two types of object that are constantly conjoined :

> We may define a CAUSE to be 'An object precedent and contiguous to another, and where all the objects resembling the former are plac'd in like relations of precedency and contiguity to those objects that resemble the latter. (Hume, Bk. 1, Part III, sec. XIV))

Thus on Hume's regularity account, which proposes a conceptual reduction of the causal to the temporal asymmetry, it is simply a matter of definition that causes precede their effects.

But, as has often been argued, Hume's account seems to be both too strong and too weak (see Price and Weslake 2009, 414–416). On the one hand, the connection postulated by Hume between the two asymmetries appears to be too tight. First, Hume explicitly disallows simultaneous causation. Many philosophers, however, take Newton's second law $\mathbf{F}=m\mathbf{a}$ to be a paradigm example of a causal law and, if interpreted causally, the law seems to postulate a simultaneous causal connection between forces and acceleration. Second, it seems to be an interesting and substantive question, whether backward causation is physically or conceptually possible. Indeed, that backward causation must at least be conceptually possible is for many one of the conditions of adequacy that any adequate account of the direction of causation must satisfy (see, e.g., Healey 1983). Yet if Hume were right, then it would follow immediately from the very concept of cause that backward causation is impossible and investigations into possible backward causal structures would simply be confused.

While positing that causes precede their effects as a conceptual truth arguably posits too tight a connection between the causal and temporal asymmetries, it is worth stressing that our common sense conception of causation nevertheless does see a rather strong connection between the two asymmetries: effects never precede their causes in familiar circumstances. Thus, it also is a condition of adequacy that an account of the causal asymmetry not make the connection so loose as to allow for widespread backward causation even in mundane circumstances.

On the other hand, Hume's account appears to be too weak, as Price and Weslake argue. The causal asymmetry is intimately related to a number of other temporal asymmetries:

(i) We deliberate with future goals in mind, but it does not seem to make sense to deliberate with past goals in mind.

(ii) We take ourselves to be able to intervene into the world and be able to affect the future (at least to some extent), but we do not believe that our actions can influence the past.

(iii) We can know more about the past than about the future and our access to knowledge about the past seems to be different from that about the future;

(iv) In particular, we have records and memories of the past but not of the future.

All these asymmetries are closely linked to the causal asymmetry and on some views are simply special cases of the asymmetry of causation. Yet these asymmetries appear to be more substantive than a Humean account would allow. Our sense that we can affect the future and not the past seems to call for a deeper explanation than insisting that, whatever regularities involving our actions there may be, only those events that lie to the future of our actions are to be called 'effects' of these actions, while events in the past are labeled 'causes' of our actions.

To the extent that Price and Weslake's worry presents a genuine problem, it arises not only for Hume's account but more broadly for any account of causation that, like Hume's, begins with a time-*symmetric* notion of causal connectedness and then introduces the distinction between cause and effect by appealing to the notion of temporal priority. Conserved quantity accounts of causation, for example, also have this feature. And, despite its slightly different structure, Mackie's INUS condition account[2] also seems to face this problem, at least if we assume determinism. For then effects are INUS conditions of their causes just as causes are INUS conditions of their effects and the asymmetry of the causal relation has to be added through an additional step. W. H. Newton-Smith, for instance, attempts to arrive at a notion of directed causation within a Mackie-style account by introducing the notion that an event *has occurred by time* t—an explicitly time-asymmetric notion (see Newton-Smith 1983).

Price and Weslake have correctly identified a *potential* problem for accounts of causation that tag on a notion of temporal directedness to a prior account of symmetric causal connectedness. But whether the problem *actually* arises also depends on the conception of the temporal asymmetry we presuppose. Price and Weslake's problem arises most starkly, if we take the past to future direction itself to be a mere matter of stipulation. If we assume a four-dimensional block universe and take it to be merely a matter of stipulation which temporal direction in the block we identify as the past direction, then it seems indeed doubtful that this by itself can explain the intimate connections between the causal asymmetry and the other asymmetries we distinguished above.

But we might presuppose a richer conception of what the arrow of time consists in. If, for example, we were to posit a rich conception of temporal becoming, then this conception might be able to explain why it does not make sense to deliberate with past goals in mind, since the past unlike the future already has already occurred. Alternatively, if, with Tim Maudlin (2007, 133–134), we took the direction of time to consist in the fact that earlier states *generate* or *produce* later states, then the temporal asymmetry might be able to explain the cluster of asymmetries even under the assumption of a four-dimensional block universe. In the case of Maudlin's view, however, it may no longer be obvious that the account is one that reduces the causal asymmetry to a more fundamental temporal asymmetry rather than the other way around, since the fundamental temporal facts are themselves characterized in explicitly causally loaded terms.

2. Causal Theories of Spacetime

As we saw in the last section, temporal accounts of the causal asymmetry seem to face a dilemma: merely tacking on a temporal asymmetry to a symmetric notion of causal connection either does not allow us to explain various other asymmetries closely associated with the causal asymmetry, unless we presuppose a rich and perhaps already causally characterized notion of temporal becoming or temporal production. Thus, the opposite strategy of trying to reduce temporal notions to causal notions might appear to be more promising. Here we can distinguish two different kinds of project—one more ambitious and the other less so. The more ambitious project aims to reduce *all* temporal relations or even all spatiotemporal relations to causal relations. The aim of this project is not only to ground the temporal *asymmetry* in the causal *asymmetry* but also to recover the topological structure of time or spacetime from more fundamental causal relations. The less ambitious project takes a four-dimensional spacetime manifold to be given and merely tries to account for the temporal asymmetry—that is the difference between the past and future directions—in terms of the causal asymmetry. The more ambitious project aims to provide an account of the asymmetry *of* time: all temporal relations, including their asymmetry, on this account are ultimately reducible to causal relations. The less ambitious theory might only be interested in offering an account of asymmetries *in* time: the theory presupposes a spacetime manifold with symmetric spatiotemporal relations between events and then argues that what fundamentally distinguishes the past and future directions from each other is a causal asymmetry.[3]

The more ambitious project has two planks: an account of the topological structure of time or spacetime in terms of a *symmetric* notion of causal connectibility and a causal theory of the temporal asymmetry. In this section I want to briefly discuss the first component of this project, which in the literature is often referred to as 'the causal theory of time', before returning to a discussion of the temporal asymmetry in the next section.

Causal theories of spacetime have in the last century been defended by Hans Reichenbach (1958; 1956), Adolf Grünbaum (1963) and Bas van Fraassen (1992). Defenders of causal theories point to special relativity as providing one of the main sources of inspiration and support for the theory. The Lorentzian metric of special relativity defines a so-called 'lightcone structure' on the spacetime manifold, which allows us to distinguish pairs of events that can be connected by a signal travelling at or slower than the speed of light from those pairs of events that cannot be so connected. Pairs of events that can be connected by a signal traveling at most at the speed of light are said to be causally connectible. Exactly those events that are causally connectible exhibit a frame-independent temporal ordering. Thus, one can correlate the notion of causal connectibility with spatiotemporal notions, as for example van Fraassen does in the central theorem of his version of a causal theory of spacetime: "*X* is causally connectible with *Y* if and only if *X* and *Y* are either spatiotemporally coincident or temporally separated." (1992, 193)

Van Fraassen's theory takes the terms *events*, *causal connectibility* (and *genidentity*) as its primitive terms and then defines temporal and spatiotemporal relations in terms of the causal structure. But one might also read the biconditional as a definition of the notion of causal connectibility in terms of the metrical structure of the four-dimensional Minkowski spacetime of special relativity. Thus, John Earman worries that the causal theory of time is trivial and that 'causal connectibility' might be "just another name for a spatiotemporal relation, a relation which must be understood in terms of space-time structure." (1972, 74)

While a causal theory of time can be most successfully defended within the context of the special theory of relativity, extending the theory to general relativity is more problematic, since Einstein's field equations have solutions for which the spacetime topology cannot be recovered from the relation of causal connectibility. If we restrict possible spacetimes to those that are "strongly causal", then the construction is still possible.[4] But general relativity also allows for causally pathological spacetimes in which the causal structure is not sufficient to determine the topology. The existence of such solutions to the field equations might invite two different responses: on the one hand, one might maintain that these 'pathological' spacetimes are physically possible precisely because they can be represented by solutions to the field equations and, hence, that any adequate reductive account of spatiotemporal relations would have to be able to generate these spacetimes as well. Thus one might take the existence of the problematic solutions as a reason to reject the causal theory. On the other hand one can insist that the class of solutions to a physical equation can outrun the class of physically possible worlds or situations countenanced by that theory: some solutions may not represent genuine physical possibilities at all and hence, the causal theory need not be able to account for these particular solutions. Instead, the causal theory may be taken to provide an additional constraint on the physically possible solutions to the field equations.

Many philosophers of physics strongly favor the first option and simply take it for granted that the class of mathematical models defined by a theory's basic equations represent the class of physically possible worlds, according to that theory. But physicists themselves often reject certain solutions as 'unphysical', invoking considerations relevant within a given theoretical context beyond what is embodied in the theory's equations. Thus, it is far from clear that we are committed to accepting every model of a theory, as defined by the theory's basic equations, as representing a physically possible world according to that theory. At the very least the view that each solution to a theory's dynamical equations represents a possible world allowed by that theory is not as self-evident as it is often taken to be and requires an additional argument defense.

Against the view that causal considerations might provide an additional constraint on the class of physically possible worlds Lawrence Sklar has argued that causally pathological spacetimes seem to be perfectly intelligible to us and that this already shows that our actual topological notions cannot be defined in terms of the notion of causal connectibility (Sklar 1977).[5] Sklar's own response to the problem is to propose a richer causal reduction basis—the set of continuous causal curves—which allows us to recover causally pathological spacetime topologies as well. Yet it is not clear how far the argument from intelligibility can really take us. Even if spatiotemporal relations are physically reducible to causal relations, it seems that we can come to understand spatiotemporal notions independently of causal notions. And even once we have learned that spatiotemporal notions are physically reducible to causal notions, we might be able to expand the use of the former notions beyond their proper domain and interpret abstract mathematical manifolds as representing spatiotemporal relations, even in cases in which a spatiotemporal interpretation of the manifold cannot be introduced via causal reduction relations.

3. Toward a Causal Theory of the Asymmetry of Time

The second plank of a complete causal theory of time consists in an account of the arrow of time in terms of the causal asymmetry. Broadly two options present themselves at this point: one might appeal to the causal asymmetry as a basic asymmetry grounding the temporal asymmetry; or one could argue that the causal asymmetry is itself reducible to a further, non-causal asymmetry. Reichenbach proposed accounts of both kinds: while his initial causal theory of time invoked a basic causal asymmetry, which he tried to explicate in terms of his mark method (see 1958), he later argued that the causal asymmetry is further reducible to the thermodynamic asymmetry (1956). In this section I want to examine the prospects of an account of the first kind—an account that aims to ground the temporal asymmetry in basic asymmetric causal relations. While this account can be understood as part of an overall causal theory of time, it can also stand on its own, as part of what I have called 'the less ambitious project' that takes a four-dimensional spacetime manifold for granted and is interested only in offering a causal account of the arrow of time.

At first sight, a causal account of the asymmetry of time seems to face a formidable obstacle: it seems to conflict with physicalism—the view, as Price and Weslake put it, "that the abilities the world grants us, and restrictions it imposes on us, are determined ultimately by physics" (Price and Weslake 2009, 416). Basic asymmetric causal relations, many philosophers of physics believe, are incompatible with the time-reversal invariant dynamical laws of our theories of physics. Whatever legitimate role asymmetric causal notions might play in our common sense conception of the world, such notions cannot be part of a physicalist conception of the world in terms of time-symmetric basic laws, or so it is argued.

Often this worry is not expressed in the form of an explicit argument, but it seems to be motivated in part by the following conception of how physical theories represent the world. Theories of physics, it is thought, present us with mathematically precise dynamical equations that define the class of models of a theory, or the possible worlds allowed by the theory. Different models are distinguished by different initial and boundary conditions. Thus, once we are given the dynamical equations in conjunction with appropriate initial conditions, there is no work to be done by putatively causal principles: the laws plus initial conditions tell us everything there is to know about the system in question. Moreover, if the dynamical equations of our well-established theories are time-symmetric, these equations cannot themselves be understood as embodying asymmetric causal relations. Thus, in a discussion that is representative of the view of many philosophers of physics, Price and Weslake conclude from the premise that "fundamental physics seems to be time-symmetric" (416) that if time asymmetric causal relations were to be real they would have to be something "over and above physics" (417). They call the latter view "the hyperrealist view of causation" and maintain that "the main difficulty with hyperrealism is that in putting causation beyond physics, it threatens to make it both *epistemologically inaccessible* and *practically irrelevant*." (417, italics in original) Price and Weslake conclude that more promising than to postulate such extra-physical relations is to reject the view that causal relations are a real aspect of the inventory of the world.

We can represent this argument in explicit premise-conclusion form as follows:

1) Reasoning and inferences in physics can be exhaustively characterized in terms of a theory's dynamical models together with choices of particular initial and boundary conditions.
2) Time-symmetric equations cannot provide evidence for asymmetric causal assumptions.
3) Asymmetric causal notions could play a legitimate and substantive role in a physicalist conception of the world only if either they played a substantive role in explanations or inferences in addition to the purely dynamical models or their use was justified by the character of our theories' dynamical laws.
4) Therefore, asymmetric causal notions can play no legitimate role in a physicalist conception of the world. (1, 2, 3)

This argument is an Ockham's razor argument: physics has no need for asymmetric causal relations and therefore one should not posit such relations. Yet both premises (1) and (3) can be challenged. Consider (3) first. In order to establish that causal relations would have to be extra-physical relations that are epistemologically inaccessible and practically irrelevant it is not enough to point to the temporal symmetry of the laws, but one would in addition have to show that time-asymmetric causal assumptions do not play a role in our treatment of initial or boundary conditions.[6] Second, contrary to what premise (1) asserts, many (and perhaps even most) inferences in physics do not proceed from fully specified initial conditions fed into the appropriate dynamical equations. Rather, in many cases our observational data severely underdetermine which purely dynamical model of a given theory adequately represents the phenomena. This underdetermination problem is (at least in some cases) solved by embedding the dynamical models into richer causal structures.

Picture yourself looking up at the night sky.[7] How can we scientifically justify the belief that the focused disturbances in the radiation field, which we observe coming in from a specific direction, were emitted by a star as their source? We cannot justify this belief by solving a final value problem, for that would require knowledge of the state of the radiation field on a complete final value surface, which in this case would be a complete cross section of the forward lightcone centered on the putative source point.[8] Rather our inference is paradigmatically causal: We infer the existence of a source as the best explanation of the strong correlations among our observations at different spacetime points (which may be either simultaneous observations at different locations or observations at different temporal instants, for example if we look at a star through a period of time).

This inference can be reconstructed as comparing the likelihood of a common cause explanation with that of a separate cause explanation: the correlations in the observed radiation fields are far more probable, given the hypothesis that the disturbances were produced by one and the same source, than if we assumed source-free fields coming in from spatial infinity as separate causes. An implicit assumption in this assessment is that the initial fields coming in from distant regions of space are randomly distributed—or rather as random as possible, given our observational evidence: assuming initial randomness makes it radically improbable that the observed correlations could be due to correlations in the free incoming fields.

Given the assumption of an asymmetric causal interactions between sources and fields and that of initial randomness we can represent the interaction between the radiation field and its putative source in terms of the kind of causal structures developed in the causal modeling literature, for example by Judea Pearl(2009). According to Pearl, a causal model is a triple $M = <U, V, F>$, where U is a set of background variables, determined by factors outside the model, V is a set of variables determined by other variables in the model, and F is a set of functions that assigns a values to each $V_i \in V$ that depends on its causal parents. Each causal model can be associated with a directed graph $G(M)$, which in our case looks as in Figure 1.

[Insert figure 1 approximately here]
Figure 1: Directed acyclic graph for a radiation fields associated with a source.

The associated causal equations specify the observed fields as functions of the earlier state of the source and any free incoming fields: $O^i = f^i(S, F^i_{in})$. Adding the initial randomness assumption amounts to offering a probabilistic causal model $<M, P(u)>$, where M is a causal model and $P(u)$ is a probability function defined over the domain of U. As Pearl shows, the initial randomness assumption is equivalent to the causal Markov condition, according to which each variable is independent of all its nondescendants, given its parents.

While the *purely dynamical* equations for the fields O^i are time-symmetric and the O^i can be expressed *both* as functions of earlier fields and an earlier state of the source *and* as functions of later fields and a later state of the source, the *causal equations* do not exhibit a corresponding symmetry. In particular, while the earlier incoming fields can be assumed to be distributed randomly, later fields will generally not satisfy the randomness assumption.

Examples like this can be easily multiplied. Indeed, the underlying pattern of reasoning appears to be extremely wide-spread: very often do we not know the state of a system on a full initial value surface and therefore need to use partial knowledge of the dynamical state of a system together with asymmetric causal assumptions and the dynamical laws to draw inferences about the history of the system.[9]

Once we have introduced causal models, we can also see that, contrary to what (3) asserts, causal models can play an explanatory role even if we assume full knowledge of the initial conditions. Consider the explanation of the so-called 'wave asymmetry' that is standardly offered in physics textbooks: the question as to why we observe coherently diverging radiation fields in nature, but not coherently converging fields, is usually answered by singling out the solution representing the field in terms of the earlier state of any sources (and, thus, in terms of disturbances *diverging* from the sources) together with any free fields as the physically correct solution. A mathematically equivalent solution in terms of the later state of the sources (and, thus, *converging* field disturbances) is rejected as unphysical. Considered *purely dynamically*, this textbook answer does not make much sense, for purely dynamically the two representations (in terms of an initial or final value problem) are simply equivalent representations of one and the same field. But the answer does make sense if we understand it as proposing an embedding of the dynamical models into causal models. In terms of the causal models, the explanation of the asymmetry is that, given an initial randomness assumption, coherently diverging waves can be explained by the earlier action of a source acting as common cause, while coherently converging waves are radically improbable. In particular, the *later* action of the source with which the converging disturbance is correlated cannot explain the correlations in the field, since it is not a cause of the correlations in the field and final fields are *not* distributed randomly.

We find a similar pattern of explanation in the case of thermodynamic phenomena (see Maudlin 2007, 131–133). The fact that the entropy of a closed system tends to increase is explained by appealing to the assumption that the system originated in a state that is macroscopically 'atypical'—a state of low entropy corresponding to an extremely small region of the phase space accessible to the system—but microscopically 'typical' or random. By contrast, the final state of a thermodynamic system is macroscopically typical in that it has high entropy, but is microscopically highly atypical, in that there are delicate correlations among the system's microscopic constituents, such that if the system's state is evolved *backward* in time, these correlations result in an evolution toward the low entropy initial state.[10] Again, the assumption that the initial but not the final states are random or typical plays a crucial role in the explanation.

The connection between an asymmetrically applied randomness assumption and a time-asymmetric evolution is also evident, if we consider a thermodynamic system that at time t_0 is in some intermediate state, with an entropy lower than the equilibrium value, but higher than the system's initial entropy. If we are interested in the system's *backward* evolution for times $t<t_0$, then we cannot assume that the microstate of the system is random, since the state contains correlations accounting for its backward evolution toward a macroscopically less typical state of lower entropy. If, however, we are interested in the system's *future* evolution, then we can generally assume that the microstate at t_0 is effectively random. (Formally, this is due to the fact that the very small phase space region corresponding to the low entropy initial state of the system evolves at t_0 into a region of equal size, which is, however, highly 'fibrillated' and smeared out through the much larger phase space region corresponding to the macrostate at t_0.) That is, any correlations that exist among the microscopic subsystems of the system are in general irrelevant for the future evolution of the system but not for its past evolution. This is what we would expect, given a causal representation of the system: we cannot treat the state at t_0 as microscopically random with respect to the system's earlier states, because the state at t_0 was *caused* by the system's earlier state.

We now have all the ingredients in place to sketch a causal theory of the asymmetry of time. In response to the eliminativist argument I argued that asymmetric causal reasoning plays an important role in cases where the full initial state of the system is not known and that causal structures allow us to represent the asymmetric treatment between initial and final conditions. For ease of exposition I took the temporal asymmetry to be given and, thus, simply assumed that the direction of explanation goes from earlier to later states and that the randomness assumption holds at initial but not at final times. A causal theory of the asymmetry of time, however, would have to show that we can characterize the relevant asymmetry in strictly non-temporal terms and that the temporal asymmetry can then be defined in terms of the causal asymmetry.

It is easy to see, for example in terms of Pearl's causal modeling framework, how this can be done. The orientation of causal models is given, independently of a temporal orientation, in terms of a directed graph. The randomness assumption holds for all exogenous variables in the model—that is variables in the *causal* past of endogenous variables in the model. If the variables of the model represent events that also have spatiotemporal coordinates, one can use the orientation of a causal model to introduce a temporal orientation: If an event *e* is an effect of an event *c*, and c and *e* do not occur simultaneously, then *e* is in the future of *c*. We can strengthen this to a necessary and sufficient condition as follows:

> An event *B* is in the future of an event *A* if and only if: (i) *B* and *A* are not spatiotemporally coincident and (ii) either *B* is an effect of *A* or *B* is in the same half lightcone centered on *A* as is an effect of *A*.

The direction of time can be represented mathematically by defining an orientation on the spacetime manifold in terms of an equivalence class of timelike vector fields.[11] At issue is whether this direction can be defined in a non-arbitrary manner, since for every time-like vector field we can define a different field with the reversed orientation. The claim of a causal theory of the asymmetry of time is that an orientation is uniquely picked out by the direction given by our causal models.

The causal asymmetry itself is not imposed arbitrarily but is underwritten by our inferential and explanatory practices and asymmetric causal representations can be tested and confirmed empirically. Moreover the asymmetric treatment of initial and final states tracks an explanatory asymmetry. We explain (facts about) the state of a system by appealing to earlier states of the system (and interactions of the system with its environment), but not by appealing to later states of the system. Indeed, the physicist Eugene Wigner thought that the central characteristic of our physical understanding of the complexity of nature consisted in distinguishing a domain of regularities, the laws of nature, from one of randomness, the initial conditions. The aim of physics, according to Wigner, is to explain order and correlations that may exist within the state of a system at a time in terms of an *earlier* random state and lawful regularities (see Wigner 1979). Far from being an element foreign to physics, as Price and Weslake claim, an explanatory asymmetry is built into the very foundations of physics, on Wigner's conception.

How does the causal theory of the temporal asymmetry fare with respect to the problems we identified for Hume's temporal theory of the causal asymmetry above? As I formulated it, the theory allows for the possibility of simultaneous causation but not for causal orientations that are temporally directed opposite to each other, and hence cannot allow for putative cases of backward causation. But one could adjust the account slightly by allowing for isolated divergences from the dominant causal orientation and taking the direction of time to be given by the orientation of the majority of causal relations within a certain spacetime region. Isolated causal relations that are temporally directed opposite to the majority of causal relations (such as journeys in a time machine) are then backward causal.

Finally, the causal theory seems to be able to account for the other asymmetries closely related to the temporal and causal asymmetries. Exploiting the close connection between the notions of causation and intervention is at the very core of the recent causal modeling literature. The temporal asymmetry of deliberation and of action can, thus, simply be seen as an instance of the causal asymmetry: Our deliberations concern events that we can affect causally and an action is a cause of the action's future consequences. Moreover, recording interactions exhibit a common cause structure—that is causal forks that are open toward the causal future. Indeed, one can derive the asymmetry of records—that is, the fact that there are records of the past but not the future—from the initial randomness assumption and time-symmetric assumptions about the nature of recording interactions.

Many philosophers of physics seem to worry that granting a role to causal structures within physics commits us to a metaphysically rich theory of causal production. But positing asymmetric causal relations in the first instance amounts only to postulating that a certain type of structure can play an explanatory role within a scientific account of the world. These structures are causal, in that they are directed and allow what are paradigmatically causal inferences. Yet as such the causal theory is metaphysically largely neutral. It is a separate question how metaphysically to interpret these structures.[12]

Of course it is possible to interpret causal models as representing metaphysically 'thick' relations of causal production. Earlier I said that Price and Weslake's challenge to a temporal theory of the causal asymmetry does not arise for a rich theory of temporal production, as the one proposed by Maudlin, for example. What, then, is the relation between a causal theory of time of the kind I proposed here, but interpreted metaphysically richly in a way going beyond what I want to advocate here, and Maudlin's theory of temporal generation or production? One might think that there must be a significant amount of disagreement between the two views, since Maudlin has explicitly argued that the notion of causation is less fundamental than that of causal laws that drive the dynamical evolution of a system (2007, 149–154), but it seems to me that the disagreement is to a large extent merely apparent.

Maudlin's argument for the priority of dynamical laws of production over causal relations appeals to a variant of Conway's game of life, in which the dynamical laws of the future evolution of the system consisting of a two-dimensional grid of spaces that are either occupied or empty are well-defined. The laws of evolution fully determine how the system would evolve under different counterfactual suppositions. Yet, Maudlin argues, fixing the truth values for all counterfactuals, in this case does not fix the truth values of all causal claims: whether the fact that a certain space is occupied at one time is a *cause* of it being occupied at a later time can be open to dispute even when the truth values of the relevant counterfactuals are not. Maudlin concludes that the notion of causation cannot be more fundamental than the notion of counterfactual or that of a law of dynamical generation. Since he also argues that the notion of counterfactual dependency is not more fundamental than that of causation, this leaves the notion of dynamical laws, which account for how states at one time produce or generate states at later times as the most fundamental notion.

Yet in Maudlin's example a causal model of the system, such as a structural model as defined by Pearl, is perfectly well-defined. The indeterminacy concerns only the question which, of a range of all clearly causally relevant interacting factors, we are willing to single out as causes of an event. Thus, one could maintain that the indeterminacy that Maudlin identifies concerns a pragmatic dimension of our concept of cause—when to bestow the label 'cause' on a factor that is causally relevant—but not the directed causal structure itself. Moreover, since Maudlin's laws of evolution are explicitly and purposefully characterized in causal terms—they are, after all, laws of production and generation—his account arguably does not offer a purely temporal account of the causal asymmetry. Rather, like the account I have discussed here, it proposes that there is a fundamental asymmetry that has to be expressed in causal terms.

4. A subjectivist account

In the last section I sketched a causal theory of the temporal asymmetry that takes the causal asymmetry to be basic. In this section and the next I want to examine two kinds of account that take asymmetric causal relations to be further reducible. The first kind are subjectivist accounts of causation that argue that the fact that we describe the world in asymmetric causal terms has its origins in a more fundamental asymmetry of deliberation. Accounts along these lines have been proposed by (Healey 1983) and (Price 1997). I here want to focus on the more recent defense of such an account in (Price and Weslake 2009).

Price and Weslake's starting point is an evidentiary asymmetry of deliberation consisting in the fact that we deliberate with future end and not past ends in mind. They argue, first, that our deliberations exhibit a temporal asymmetry that underwrites an asymmetry of influence: "we can't use evidence as a 'causal handle' to influence the earlier states of affairs for which it provides evidence." (433) This asymmetry, they claim, can then be extended to cover causal claims in general, since it provides us with a time-asymmetric perspective on the world, which underwrites a time-asymmetric prescription as to how to evaluate counterfactual interventions into systems not involving human agents. Causal and counterfactual claims can, thus, be treated in terms of interventionist accounts of causation, such as the ones developed by Pearl or James Woodward. But unlike Pearl or Woodward, Price and Weslake do not introduce asymmetric causal relations as basic, but rather take them to be a projection of a more fundamental asymmetry of deliberation.

The fact that we are beings with a certain temporal bias that manifests itself in the deliberative asymmetry, Price and Weslake suggest, might ultimately be explicable in terms of the statistical asymmetry underlying thermodynamic phenomena. But they believe that the latter asymmetry cannot itself directly ground the asymmetry of causation, but only via a detour through the deliberative asymmetry characterizing agents like us. The deliberative asymmetry locally defines a direction of causation, which we then 'spread over the objects' and extend into a global asymmetry of causation (as long as the universe we live in is temporally orientable).

How do we evaluate Price and Weslake's proposal in comparison with a causal account like the one proposed above? Does the deliberative asymmetry reflect a more basic causal asymmetry, since actions are causes, or is the deliberative asymmetry the more fundamental of the two? To the extent that our universe is causally 'normal', the two proposals will not merely agree locally but also globally on the direction of causation: judged from the perspective of a causal theory, Price and Weslake's subjectivist theory inherits the deliberative arrow from the *causal* arrow of deliberation and then spreads that arrow over the universe as a whole. One way to distinguish the two competing proposals, however, is to examine what they would say about hypothetical universes in which there seem to be oppositely directed causal arrows.[13]

(Price 1997) has considered the possibility that the universe may exhibit a global time symmetry and not only evolved from an initial low entropy macro-state together with a condition of initial randomness but also evolves towards a future low entropy state that satisfies a condition of final randomness. If we posit that the age of such a universe is more than twice its relaxation time, then the two periods of entropy increase and decrease are separated by a time period in which the universe is in a maximum entropy equilibrium state. The thermodynamic arrow points in opposite directions during the two temporal halves of the universe. (Which of these two directions is the future direction? Price's suggestion is that beings living in either half would disagree and would identify the direction of entropy increase within their local environment as the future direction.) Let us assume that both epochs can be characterized in terms of a causal asymmetry. In which direction does the causal arrow point during the two epochs?

On the one hand, Price and Weslake maintain that the direction of the causal arrow globally is simply the same as that of our local arrow of deliberation. Yet on the other hand they suggest that one can extend the deliberative asymmetry to causal claims not involving human agency by using recent interventionist accounts of causation: "Ideally, the subjectivist will want to step into the interventionists' shoes—all the more so, now that Pearl, Woodward and others have shown us how far those shoes may take us!" (438) But for regions of the universe with an opposite entropic arrow the two prescriptions give conflicting results. According to Price and Weslake's first prescription, the arrow of causation, as *projected by us*, globally points in the same direction as our deliberative arrow. This means that *for us* the causal arrow during the entropically reversed epoch points from states with delicate microscopic correlations to states of increasing micro-randomness: during that epoch we would take higher entropy states exhibiting microscopic correlations to be the cause of later lower entropy states which are microscopically more random.

Yet, according to the second, interventionist prescription (as given by the causal discovery framework formulated by Pearl and others) the randomness assumption holds in the causal *past* of the models and the causal arrow will point in opposite directions during the two epochs of our hypothetical universe. The interventionist models satisfy the causal Markov condition and, therefore, a screening-off condition. The causal models resulting from Price and Weslake's first prescription, by contrast, are not Markovian. In fact, the prescription implies not only an isolated failure of the Markov condition but a general failure of the condition during the entire (from our perspective) anti-entropic epoch of the universe. Yet as Pearl maintains, it is "not clear how one would predict the effect of interventions from such a model, save for explicitly listing the effect of every conceivable intervention in advance" (Pearl 2009, 61).

If Price and Weslake allow the two causal arrows to come apart, the question arises as to why we should identify the two arrows during our own epoch of the universe. Price and Weslake, thus, seems to face the following dilemma: Either they follow Pearl and others in positing as default (but in isolated special cases perhaps defeasible) constraint on causal models that disturbances in the causal past are distributed randomly. Then the causal and entropic arrows will be aligned during both epochs of a symmetric universe, but Price and Weslake's 'detour' through our own perspective of deliberation in assigning a global causal arrow does no work. Or they take the direction of causation everywhere to be that given by the local arrow of deliberation and action. But this has the consequence that we are now faced with two distinct arrows that might be called 'causal': the perspectival causal arrow on which Price and Weslake focus and the interventionist inferential arrow. The need to posit two distinct causal arrows, where other accounts require only a single arrow, might itself be taken to be a disadvantage of the account. At the very least Price and Weslake owe us an explanation as to why the two arrows should coincide for us locally.

5. An entropy account

A crucial component of causal models is the causal Markov condition which is equivalent to an assumption of initial randomness. What is the status of the randomness assumption? One view is to take the assumption as an integral part of the causal models and read it as itself a causal assumption, as for example Pearl does. Another option, however, is to see it as part of a reductive account that aims to derive the causal asymmetry from a set of non-causal and, in particular, thermodynamic assumptions.

After having posited the causal asymmetry as basic in earlier work, Reichenbach defends an entropy account of the causal asymmetry in (Reichenbach 1956). I here want to focus on a more recent version of such an account, the account developed by David Albert and Barry Loewer. Albert and Loewer's strategy is to argue that a certain probabilistic counterfactual with truth conditions given in terms of the thermodynamic account exhibits a strong temporal asymmetry and that the temporally asymmetric counterfactuals can be used to ground causal claims, along broadly Lewisian lines. Thus, the explanatory order between causes and counterfactuals is opposite in Albert and Loewer's account from that in Pearl's. While Pearl and Woodward give a semantics for counterfactuals in terms of Markovian causal models, Albert and Loewer introduce a certain type of counterfactual that is meant to ground causal claims. Here the latter account also differs from Price and Weslake's. Yet there are also several similarities between this project and that defended by Price and Weslake. Both take the core assumptions underlying the thermodynamic asymmetry to be basic and both give a role to deliberative agents in their derivation of the thermodynamic asymmetry.

(Albert 2000) proposes that the thermodynamic asymmetry that the entropy of a closed system tends to increase can be explained in terms of the following assumptions: (i) the assumption of an extremely low entropy past (the 'past hypothesis'); (ii) a time-symmetric micro-dynamics; (iii) an equiprobability distribution over all microstates compatible with the initial low entropy macrostates. These posits, Albert and Loewer argue, can also ground the direction of time and the causal asymmetry: the past direction simply is that temporal direction in which there is a low entropy constraint and the asymmetry of causation follows from the purported fact that certain counterfactuals are time-asymmetric.

David Lewis thought that a temporal asymmetry of counterfactuals was a consequence of an asymmetry of overdetermination, according to which multiple localized facts about the present are nomological sufficient for the occurrence of events in the past, but that the future is not similarly overdetermined by the present (Lewis 1979). But Lewis's overdetermination thesis is false. No separate local facts at one time are individually nomologically sufficient for events at other times—rather the laws require the state on a complete initial or final value surface as input to determine the state of a system at other times; and in the case of time-symmetric laws there is no *asymmetry* of determination. (see Elga 2001; and Frisch 2005 for different arguments to that conclusion) An overdetermination thesis similar to Lewis's is defended by Daniel Hausman in (Hausman 1998). Hausman's argument is criticized in (Frisch 2005, 185–187) and in (Schurz 2001), where Gerhard Schurz shows that the purely probabilistic relations governing intervention variables do not exhibit a temporal asymmetry.

Albert and Loewer try to rescue a Lewisian project by appealing to the thermodynamic asymmetry in the evaluation of counterfactuals. The counterfactuals on which they focus have the form "If a small, localized non-actual present macro-event *C* occurred, then the probability of some past or future macro-event *E* would be *p*." This probability is calculated by conditionalizing *E* on the occurrence of *C*, the assumption of an extremely low entropy past ('the past-hypothesis'), the dynamical micro-laws, which are assumed to be deterministic and by assuming an equiprobability distribution over all microstates compatible with the low entropy past. On one version of the account events *C* are primarily decision events—events in the brain corresponding to the act of deciding to perform a certain act (Loewer 2007).

While this account has the advantage over Lewis's account that the counterfactuals clearly exhibit *a* temporal asymmetry, since it builds an asymmetric constraint, the past-hypothesis, directly into the semantics for the counterfactuals, it is unclear whether the asymmetry is sharp enough to underwrite the sharp asymmetry that we associate with common sense causal claims: it seems to be part of our common sense notion of causation that in standard circumstances causes have absolutely *no* effects that occur earlier than their cause and Albert and Loewer's account cannot give that result.[14]

In response to criticism Albert and Loewer have conceded that there will in fact be many non-actual past events that have high probabilities conditional on small changes to the present. Thus, their proposed semantics implies that there will be large-scale backward causation. Yet they maintain that we could not exploit such backward causal influence, since we could only influence past events that have no records in the present. This explains, they maintain, how our concept of cause can exhibit a sharp temporal asymmetry even though the counterfactuals grounding the concept do not exhibit a similarly sharp asymmetry. But it seems possible for us to come to learn of strong correlations between actions of ours of a certain kind and past events that trigger them. Thus, that we have performed such an action should provide us with good evidence for the occurrence of its past trigger, which according to Albert and Loewer's counterfactual account, would come out as an *effect* of the action rather than as one of its *causes*.[15]

One of the advantages of Albert and Loewer's entropy account appears to be that it promises a unified account of the various asymmetries I distinguished in section 2 above and, in particular, promises to derive the epistemic asymmetries together with the causal asymmetry from the entropic account. But, as can be shown, the entire work in the derivation is done by the probability postulate or the condition of initial randomness and this assumption, as we have seen above, is also an integral part of the causal modeling framework.

I want to mention two additional possible problems for Albert and Loewer's account. First, the account posits a *cosmological* asymmetric constraint, the past hypothesis, and a corresponding probability postulate, yet paradigmatic causal counterfactuals concern small, 'human-sized' systems. It is a controversial question whether the cosmological account allows us to recover a localized 'past hypothesis' and initial randomness assumption also for small quasi-isolated thermodynamic systems (Earman 2006; Winsberg 2004). As Lawrence Sklar has argued, small systems like a body of gas in a container can, following Reichenbach, be represented as branch systems that are relatively isolated from the rest of the universe for a short period of time. In our spatiotemporal region of the universe both temporal 'ends' of the branch system occur at times of overall very low entropy. The question, thus, is what can account for the fact that the randomness assumption can be made at one end of a branch system's history but not at the other end (to retrodict the system's evolution), and it is not clear whether a cosmological account, like Albert and Loewer's, can adequately answer that question.

In a thought-provoking essay on the relation between conceptual and physical reductions (Sklar 1981), Sklar raises a second concern, drawing on a discussion by Eddington (Eddington 1953, ch. 5 "Becoming"). Sklar and Eddington point to the fact that a direction of time and the notion of *becoming* are intimate features *both* of our experience *and* of our conception of the physical world (as is evident in our explanatory practices). As Eddington puts it, "we have direct insight into 'becoming' which sweeps aside all symbolic knowledge as on an inferior plane." (1953, 97) and imagining a universe that evolved backward in time would require imagining reversing the "the trend in world-texture" (1953, 93–4) underlying our explanatory inferences. According to Sklar, the problem is this. In the case of other physical reductions, such as that of color to something involving electromagnetic wavelength, we learn to distinguish an item of our subjective experience—an experience of color—from a physical correlate that gives rise to our experience. That is, what the account 'reduces away' is color as existing in the objects. But in the case of the asymmetry of time it is the character of our experience itself—our immediate and intimate experience of the directedness of time—that is to be reduced away. There is no room here to try to fully unpack Eddington's suggestive remarks and Sklar's arguments, but the fact that a temporal directedness is an intimate feature of all experiences presents a challenge that existing attempts to reduce the temporal asymmetry to entropic considerations have not yet adequately addressed.

6. Conclusion

In this chapter I have examined possible connections between the causal and temporal asymmetries. According to what is perhaps the most popular view today, both the causal and temporal asymmetries ultimately supervene on thermodynamic features of the universe (either along the lines of a subjectivist account, like Price's, or a 'pure' entropy account, like Albert and Loewer's). I have argued that such attempts to ground the temporal and causal arrows in a third arrow are far from unproblematic and face significant challenges. The prospects for a causal account of the temporal asymmetry, by contrast, are brighter than many currently seem to believe.

References

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

Earman, John. 1972. "Notes on the causal theory of time." *Synthese* 24 (1-2): 74 - 86.

———. 2006. "The 'Past Hypothesis': Not even false." *Studies in History and Philosophy of Science Part B* 37 (3): 399-430.

Eddington, Arthur Stanley. 1953. *The nature of the physical world*. Reprinted. Cambridge: Univ. Press.

Elga, Adam. 2001. "Statistical Mechanics and the Asymmetry of Counterfactual Dependence." *Philosophy of Science* 68 (S1): S313-.

Faye, Jan. 2010. Backward Causation. In *The Stanford Encyclopedia of Philosophy*, ed. Edward N. Zalta. 2010th ed. http://plato.stanford.edu/archives/spr2010/entries/causation-backwards/.

Fraassen, Bas C. Van. 1992. *An Introduction to the Philosophy of Time and Space*. Columbia University Press Morningside Ed. Columbia Univ Pr, August.

Frisch, Mathias. forthcoming. "No place for causes? Causal skepticism in physics." *European Journal for Philosophy of Science*

———. 2005. *Inconsistency, Asymmetry, and Non-Locality: A Philosophical Investigation of Classical Electrodynamics*. Oxford University Press.

———. 2007. Causation, Counterfactuals and Entropy. In *Russell's Republic: The Place of Causation in the Constitution of Reality*, ed. Huw Price and Richard Corry, 351-395. Oxford: Oxford University Press.

———. 2010a. Causal Models and the Asymmetry of State Preparation. In *EPSA: Philosophical Issues in the Sciences. Launch of the European Philosophy of Science Association*, ed. M. Rédei, M. Dorato, and M. Suárez. Springer.

———. 2010b. Does a Low-Entropy Constraint Prevent Us from Influencing the Past? In *Time, Chance, and Reduction: Philosophical Aspects of Statistical Mechanics*, ed. Andreas Hüttemann and Gerhard Ernst, 13--33. Cambridge University Press.

Grünbaum, Adolf. 1963. *Philosophical problems of space and time*. Knopf.

Hausman, Daniel M. 1998. *Causal asymmetries*. 1st ed. Cambridge studies in probability, induction, and decision theory. Cambridge: Cambridge Univ. Press.

Healey, Richard. 1983. Temporal and Causal Asymmetry. In *Space, Time, and Causality*, ed. Richard Swinburne, 79-105. D. Reidel Publishing Co.

Hume, David. 2012 *A Treatise of Human Nature*. accessed January 15. http://www.gutenberg.org/ebooks/4705.

Lewis, David. 1979. "Counterfactual dependence and time's arrow." *Noûs* 13 (4): 455-476.

Loewer, Barry. 2007. Counterfactuals and the Second Law. In *Causation, Physics, and the Constitution of Reality: Russell's Republic Revisited*, ed. Huw Price and Richard Corry, 293-326. New York: Oxford University Press.

Maudlin, Tim. 2007. *The Metaphysics Within Physics*. Oxford ; New York: Oxford University Press.

Newton-Smith, W.H. 1983. Temporal and Causal Asymmetry. In *Space, Time, and Causality*, ed. Richard Swinburne, 105-121. D. Reidel Publishing Co.

Pearl, Judea. 2009. *Causality : models, reasoning, and inference*. 2nd ed. Cambridge U.K. ;;New York: Cambridge University Press.

Price, Huw. 1997. *Time's arrow & Archimedes' point : new directions for the physics of time*. New York: Oxford University Press.

Price, Huw, and Brad Weslake. 2009. The time-asymmetry of causation. In *The Oxford Handbook of Causation*, ed. Helen Beebee, Peter Menzies, and Christopher Hitchcock, 414-443. Oxford University Press.

Reichenbach, Hans. 1956. *The Direction of Time*. Berkeley: University of California Press.

———. 1958. *The Philosophy of Space & Time*. New York: Dover Publications.

Schurz, Gerhard. 2001. Causal Asymmetry, Independent Versus Dependent Variables, and The Direction of Time. In *Current Issues in Causation*, ed. Wolfgang Spohn, 47-67. Paderborn: Mentis Verlag.

Sklar, Lawrence. 1977. "What might be right about the causal theory of time." *Synthese* 35 (2): 155 - 171.

———. 1981. "Up and down, left and right, past and future." *Noûs* 15 (2): 111-129.

———. 1985. *Philosophy and Spacetime Physics*. Berkeley: University of California Press.

Wigner, Eugene Paul. 1979. *Symmetries and Reflections*. Ox Bow Press, June.

Winsberg, Eric. 2004. "Can conditioning on the 'past hypothesis' militate against the reversibility objections?" *Philosophy of Science* 71 (4): 489-504.

---

[1] Or perhaps weaker: If *c* is a cause of *e* through a causal chain *K*, then *e* is not a cause of *c* through chain *K*. This formulation allows for the possibility of causal loops.

[2] According to Mackie's account, a cause is an *i*nsufficient yet *n*ecessary part of an *u*nnecessary yet *s*ufficient condition for its effect. Hence the acronym 'INUS'.

[3] See (Price 1997) for a discussion of the distinction between the asymmetry of time and asymmetries in time.

[4] A Lorentzian manifold (M,g) is strongly causal exactly if for any point $p \in M$ on a manifold M there exists a neighborhood $U$of p such that there exists no timelike curve that passes through U more than once. Given the weaker notion of past- and future-distinguishability one can implicitly define the topology in terms of causal connectibility (see Sklar 1977).

[5] This paper is reprinted in (Sklar 1985).

[6] See (Maudlin 2007, 120) for an argument along these lines.

[7] For a more extended discussion of the following argument, see (Frisch forthcoming).

[8] Put less technically (and almost correctly), we would have to know the present state of the electromagnetic radiation field in a three-dimensional sphere centered on putative source and with a radius equal to the distance between the source and Earth.

[9] The asymmetry in this case is purely observational. For an argument that our experimental interactions with physical systems exhibit an asymmetry that support a causal interpretation of the system's behavior see (Frisch 2010a). See also (Faye 2010).

[10] Imagine a video of the mixing of milk in coffee shown backward: for the unmixing process to occur, the microscopic milk and coffee particles have to exhibit extremely delicate correlations that ensure that all the milk particles end up being separated from the coffee particles.

[11] If a general relativistic spacetime is not temporally orientable then one can move to a so-called "covering spacetime" which is everywhere locally the same as the given spacetime and is temporally orientable.

[12] This was also Reichenbach's view: "the conception of a directed causality does not necessarily entail any metaphysics." (1956, 28)

[13] (Newton-Smith 1983) makes an argument similar to the one that follows against the subjectivist account proposed in (Healey 1983).

[14] For detailed criticisms of Albert and Loewer's account (or accounts) see (Frisch 2007; Frisch 2010b)

[15] Albert and Loewer might want to insist that *during the process of deliberation* we have to think of our actions as being completely free and uncaused. But this does not undermine my criticism. As long as we remember having performed the action at time *t*, and have good evidence that at times before *t* this kind of action was reliably associated with a certain kind of past trigger, we are in a position to believe with high probability that the past trigger also occurred this time, immediately preceding *t*, even if there are no other records of the trigger's occurrence and even if *during our deliberation* we ignored any knowledge we have of correlations between our actions and past triggers.