The Direction of Causation: Ramsey's Ultimate Contingency

Huw Price

The University of Sydney

1. Introduction

Our present concern originates with two uncontroversial observations about causation: the causal relation is asymmetric, so that if A is a cause of B then B is not a cause of A; and effects never (or almost never) occur *before* their causes. Uncontroversial as they may be, these features of causation are far from unproblematic. A philosophical theory of causation thus has these two non-trivial tasks, among others: to explicate the difference between cause and effect—to reveal the true content of the "arrow" of causation, so to speak—and to explain why the arrow of causation is so well aligned with the arrow of time.

Note that the latter task permits two readings, depending on whether the temporal reference is read rigidly. On the stronger rigid or de re reading, the question is why the causal arrow points in *this* particular temporal direction, thought of as fixed independently of our disposition to treat the direction in question as that of the future rather than the past. On the weaker non-rigid or de dicto reading, the issue is that as to why we take the cause–effect arrow to point towards what we think of as later times; this allows that in a world in which we ourselves had the opposite temporal orientation, we might take both arrows to point in the opposite direction. (We might also consider rigid and non-rigid readings of the reference to the causal relation, which would yield four possible variants in all.)

An extreme form of a de dicto answer is that offered by Hume, who takes the two arrows to be related by definition. Thus the fundamental causal relation is held to be symmetric in time, and the claim is that we simply use the different terms "cause" and "effect" to distinguish the earlier and later members of a pair of events so related—the phrase "is a cause of" is simply shorthand for something like "is earlier than and causally related to".

It has often been noted that there is a heavy price to be paid for this convenience, however. For example it makes simultaneous causation and backward causation logical impossibilities, which seems too strong, and it precludes the project, attractive to many, of explicating temporal order in terms of causal order. (These arguments and

PSA 1992, Volume 2, pp. 253-267

https://doi.org/10.1017/S0270864700009346 Published online by Cambridge University Press

others may be found for example in Horwich 1987, p. 8, and Papineau 1985, pp. 273-4.) A more attractive suggestion is therefore that the asymmetry of the causal relation is not merely an analytic image of earlier–later ordering, but rests on some intermediate asymmetry—on a temporally asymmetric feature of the world which is itself aligned (typically, but perhaps not invariably) with the temporal ordering. This is what I shall call the *third arrow strategy*.

Our concern here is with a particular form of the third arrow strategy. A number of recent writers have sought a third arrow in statistical terms, a project which dovetails with the probabilistic approach to causation itself. If causation in general is to be understood in probabilistic terms, then the twin problems of the causal arrow—its constitution and its temporal alignment—may be expected to be themselves recast in probabilistic terms. Hence the project has been to find some objective statistical temporal asymmetry in the world to play the role of the third arrow.

My main aim in this paper is to offer a pessimistic view of the prospects for this popular project, at least in its standard objectivist form. My own view is that causal asymmetry is anthropocentric, being linked to our perspective as agents. This is a position originally advocated by Frank Ramsey. In the later and less well known of his two papers on laws (Ramsey 1978) Ramsey makes a number of remarks about causation and its direction. He links our notions of cause and effect to the agent's point of view, saying that "from the situation when we are deliberating seems to ... arise the general difference of cause and effect." (1978, p. 146) As we shall see, however, this view too seems to have an asymmetry of probabilistic inference at its heart (the crucial feature again being identified by Ramsey). A subsidiary aim of the paper is to draw attention to this anthropocentric version of an appeal to probabilistic asymmetry in the service of an understanding of causal asymmetry. I want to indicate how it avoids some problems which beset the usual non-anthropocentric or objectivist versions of the strategy, and also to suggest that it provides a particularly attractive explanation of some of the intuitions that have motivated these alternative approaches.

In particular, Ramsey's view helps to explain the intuitive force of a group of principles about time, statistical correlation and causation which together comprise what is known as the *fork asymmetry*. In this way it partially reverses the order of explanation characteristic of some of the views I am criticising. Roughly, these views take the fork asymmetry to provide what constitutes causal asymmetry—the objective third arrow, in other words. I shall argue that the fork asymmetry is not adequate to that task, and argue instead that its plausibility derives in part from the conceptual structure associated with the agent perspective.

As I say, this is the view I want to support. However, our first project is to identify some difficulties for objectivist versions of the third arrow strategy, particularly in its popular statistical form, and in the process to clarify the status of some of the statistical principles on which such accounts rely. I want to begin at an abstract level, by noting two pitfalls to which such accounts are liable to fall victim.

2. Two Dangers: Disguised Conventionalism and Passing the Buck

In renouncing Hume's easy way out—the suggestion that the connection between the causal and temporal arrows is analytic—advocates of the third arrow strategy must be careful not to endorse it by accident, as it were, by failing to notice some subtle definitional appeal to temporal asymmetry. This would amount to what I shall call *disguised conventionalism*, viz. the mistake of making the connection between the causal and temporal arrows analytic, albeit by a more indirect route than that suggested by Hume.

The second danger is of a kind very familiar in analytic philosophy. It is what might be called *conceptual buck-passing*. In the present context, in other words, it is the mistake of appealing to some notion whose own temporal asymmetry and orientation is no less problematic than that of causation itself. Of course, passing the buck is not always a mortal sin in philosophical analysis—a problem may be more tractable in a new form than in the old, for example—but it is always important to be clear as to what an analysis achieves and what it fails to achieve. I want to suggest that some statistical versions of the third arrow strategy have been claiming more than they are entitled to.

In the context of a probabilistic approach to causality, the notion of probability is itself a potential source of both mistakes. The risk depends in part on what notion of probability such an account employs. For example, if the notion employed is evidential then we are entitled to ask in what the evidential base consists. If the answer is temporally asymmetric, as for example if we are told that what is relevant to the probability of an event at a time is what happens *earlier* but not what happens *later*, then we should suspect conventionalism—for there would be an alternative account which simply took things the other way round. And although more objectivist accounts of probability may avoid this first difficulty, it is likely to be at the cost of another. For the usual metaphysics of chance is itself thoroughly asymmetric in time. Chances are normally taken to be "forward-looking", and dependent on the past but not the future, for example. There is clearly a danger that this asymmetry will turn out to be no less mysterious than that of causation itself. (Indeed, there is a natural tendency to appeal to causation in order to account for the asymmetry of chance. Why do the chances of an event at a time depend on the history of the world up to that time? Because there lie its potential causes.) So here is a trap of a different kind.

Both these pitfalls could be avoided by using a temporally symmetric notion of probability. A frequency account seems the best candidate. There need be no in-built temporal asymmetry in referring to correlations or patterns of association between events in the world; though even here we need to be careful with conditional probabilities: if the probability that P given Q is thought of in terms of the frequency with which P *succeeds* Q, then again we have an asymmetry unaccounted for, an arrow which appears to be oriented only by convention.

It seems to me an approach to causal asymmetry suggested by Frank Arntzenius involves disguised conventionalism of a similar kind. Arntzenius suggests that we "relate causation to the existence of transition probabilities which are invariant under changes of *initial* distributions." (1990 p. 95, my italics) He is careful to allow transition probabilities in both temporal directions, and thus avoids conventionalism on that score. The problem concerns the reference to *initial* rather than *final* distributions. A simple example will provide an illustration. Suppose we have 100 identical fair coins, each of which is to be tossed once at a time t. Then the probability that an arbitrary coin will be heads after t given that it was (say) heads before t is independent of the initial distribution—i.e., of the number of coins that initially show heads. Not so the reverse "transition probability", such as the probability of heads before t given heads after t: if 99 of the coins initially showed heads then this latter probability is 0.99 (assuming fair coins); if 99 showed tails then it is 0.01; and so on. Thus there is an asymmetry between forward and reverse conditional probabilities, in that only the former are independent of the initial distribution.

But what happens if we specify the final distribution instead of the initial distribution? In purely evidential terms the situation is precisely reversed. For example if we are told that after t 99 of the coins show heads then the (evidential) probability that an arbitrary coin will be heads after t given that it was heads before t is not 0.5 but 0.99.

Whereas the (evidential) probability of heads before *t* given heads after *t* is now 0.5. Thus far the direction of the probabilistic asymmetry depends on nothing more than the choice of initial rather than final boundary conditions. If Arntzenius's analogous account (which is formulated in terms of transition probabilities in Markov processes) is to avoid the same charge we need to be told why it cannot likewise be formulated in reverse, as it were—if it could be, the asymmetry Arntzenius claims to find would be shown to be simply conventional. The difficulty is that any way of meeting this objection seems bound to pass the buck. In effect, it would be an objective justification either for preferring initial to final conditions, or for invoking a notion of probability which was sufficiently asymmetric not to embody the symmetry just noted for evidential probabilities. In either case the asymmetry of causation would then rest not on Arntzenius's asymmetry in transition probabilities as such, but on whatever it was that sustained this asymmetry in the face of the above objection.

The general point here is that if we wish to refer to boundary conditions in an account of causal asymmetry then we must do so symmetrically. The lesson is nicely illustrated by the account offered by David Papineau. At first sight it may seem that Papineau's (1985) account falls into the trap just mentioned, for doesn't it treat causes as earlier *INUS* conditions and effects as later *INUS* conditions, and only then find an asymmetry in the fact that in the former case but not the later the associated background conditions exhibit statistical independence? If so, then the difference between initial and final conditions is again constitutive of that between cause and effect, and it is analytic that causes precede their effects. However, I think Papineau need only reply that his suggestion is that causes be thought of not as *earlier INUS* conditions, but simply as *INUS* conditions exhibiting the required independence feature. It is then a discovery about the world that these are generally initial conditions rather than final conditions. By beginning symmetrically the account thus avoids conventionalism.

Enough by way of abstract cautions. Let us now turn to what has been the most popular third arrow for probabilistic accounts of causal asymmetry, namely the fork asymmetry and its correlates.

3. The Fork Asymmetry

The term "fork asymmetry" has come to be used as an umbrella notation for a group of principles relating time, cause and effect, and statistical correlations between remote events (principles whose discussion in the recent philosophical literature stems largely from Reichenbach 1956). For present purposes technical precision is not crucial, and we may formulate these principles in more informal terms than is usual. At the same time, however, I want to draw attention to a logical point which is easily overlooked. This is that there are two importantly different sorts of principle in play in the fork asymmetry. Both relate the occurrence of correlations between pairs of events which are not themselves causal connected (i.e., such that neither is a cause of the other) to the presence of a common cause; and both draw a contrast between cause and effect, in noting that the analogous principle does not hold with respect to common effects. But the net relation thus described between correlations and common causes is (at least loosely) biconditional in form, and hence involves the following two ingredients (each loosely the converse of the other):

1. *Correlation productivity*. (a) Common causes produce or are typically associated with correlations between their joint effects, whereas common effects are not generally associated with correlations between their joint causes. Or to put it in what is effectively the contrapositive form, (b) joint causes are probabilistically independent of one another, whereas joint effects tend to be correlated.

https://doi.org/10.1017/S0270864700009346 Published online by Cambridge University Press

2. *Correlation explicativity.* A remote non-causal correlation between a pair of events is typically associated with a joint correlation with a third event which (a) is earlier than the two events in question, (b) is their common cause, (c) screens off the original correlation. The corresponding kind of correlation with a later event and/or a common effect is much more rare.

I am going to argue that these two kinds of principle call for different kinds of explanation: roughly, to the extent that (1) holds at all it is a consequence of the thermodynamic asymmetry, and thus contingent and a posteriori; while (2) is in part a consequence of the conceptual links between causation and the agent's perspective, and thus a priori.[1]

There are a number of ways to take such principles to provide what is required of the third arrow. One popular approach is to say that explaining remote correlations is what is constitutive of causes as against effects, so that (2b) is analytic; and then say that (2a) embodies the de facto temporal asymmetry in the world, in virtue of which causes typically precede rather than succeed their effects. A somewhat different route is taken by Papineau, who as we saw takes the difference between cause and effect noted in (1b) to be constitutive of that between cause and effect (and again takes it to be a contingent matter that we find such a temporal asymmetry in the world as we know it).

Abstracting however from particular accounts, let us focus on the issue as to how the talk of correlation and statistical dependence in these principles is to be interpreted. Bearing in mind the cautions of the previous section, an appealing suggestion is that it be read in an actualist and frequentist sense, so the principles refer to actual frequency correlations. But I want to argue that the resulting third arrow is then insufficiently general to ground causal asymmetry, in two senses. To illustrate the sort of point I have in mind, take any common case of a common cause producing joint effects—fire producing both heat and smoke, for example—and consider principle (1). It is no doubt true that there is actually a significant correlation between heat and smoke in the world, but the fact that fire causes heat and smoke surely does not depend on this being so. Had there only ever been one fire in the history of the universe (and lots of uncorrelated heat and smoke due to other causes) it would still have been true that that fire caused heat and smoke. Moreover, once we see this we see that there must be many common causes which are simply too infrequent to give rise to *actual* correlations between their joint effects. We only find *actual* correlations where the causes in question are big and/or frequent enough for their effects to stand out against the background "noise".

Construed in actualist terms, then, (1) appears to provide little handle on the cause–effect asymmetry as such: many common causes don't in fact give rise to correlations between their joint effects, and no common cause does so necessarily. So (1) provides no sharp distinction between cause and effect. This is symptomatic of a general dilemma facing the attempt to ground causal asymmetry on an objective statistical asymmetry. If the latter asymmetry is characterised in actualist terms, it seems likely to lack sufficient scope, in two senses: in failing to apply to some actual cases of causal dependence, and in being such that it might have failed to apply to any actual case (without it thereby being inappropriate to say that the causal dependence in question would still have obtained). If this charge of scope-insufficiency can be sustained, it will follow that the statistical asymmetry concerned needs to be construed in modal rather than merely actualist terms. However, the required temporal asymmetry of this modal notion will then be as problematic as that of causation in the first place: unless we resort to Humean conventionality, the crucial temporal asymmetry will again be left unexplained.

4. The Scope-insufficiency of the *Actual* Fork Asymmetry

How serious is the difficulty just described? The best way to address this question is to investigate what might be called the *actual* status of the fork asymmetry. What asymmetry of this kind do we *actually* find in the world? I propose to approach these questions indirectly, by beginning with a phenomenon which has seemed to many to provide a paradigm case of the fork asymmetry, namely the apparent temporal asymmetry of radiation. In a recent book on the physics of time asymmetry Dieter Zeh describes this asymmetry as follows:

> *After* a stone has been dropped into a pond one observes concentrically *out*going waves. Similarly, *after* an electric current has been switched on, one finds a retarded electromagnetic field. Since the laws of nature which successfully describe these events are invariant under time-reversal, they are equally compatible with the reversed phenomena in which, for example, concentrically focussing waves would eject a stone out of the water. Such solutions of the dynamical laws have however never been observed in nature. (1989, p. 12)

Accordingly, it is said that in nature radiation is always *retarded* rather than *advanced*. The example of the stone in the water is particularly associated with Popper, who used it in an influential (1956) note to argue that temporal irreversibility is not simply a matter of thermodynamics. But like many other writers, Popper failed to recognise the importance of the boundary conditions, and *their* connection with thermodynamics. The reason that outgoing waves are common is that the initial conditions that give rise to them are common. Solid objects are often so placed that they fall into ponds, to use Popper's example. And this can only be the case because our region of the universe is not in a state of thermodynamic equilibrium.

In other words, we need thermodynamic disequilibrium in order to generate the conditions that make radiation appear to be asymmetric in time. The asymmetry depends on the fact that we have big disturbances (such as flying stones) in the initial conditions but not in the final conditions. When there are no big disturbances at either end, the situation is entirely symmetric. This applies to other sorts of radiation, as much as to those on water surfaces. The reason that electromagnetic radiation appears temporally asymmetric is that we have concentrated transmitters or sources of radiation—such things as stars and radio transmitters—but no corresponding receivers or sinks. Again, we only have such transmitters because the universe is very far from thermodynamic equilibrium.[2]

So the asymmetry of radiation would disappear in a world in thermodynamic equilibrium. More importantly, it actually *does* disappear on the micro scale. This is because it is *essentially* macroscopic: just as the pictorial characteristics of a printed picture disappear if we focus on the individual dots of ink that make it up, so the asymmetry of radiation disappears if we concentrate on the microstructure of the processes in which it is manifest. It depends on the ordered alignment of vast numbers of microscopic events, and hence simply isn't the sort of feature of the world which can be manifest when the numbers involved are too small. Again, sources need to stand out against the noise, and this requires the cooperation of many individual events.

I want to suggest that these observations about radiation provide a model for the fork asymmetry in general, or rather for its *actual* manifestations in the physical world. In a world in thermodynamic equilibrium there would be no large coherent causes, no macroscopic beach walkers to leave their footprints in the sand. And in the actual world the asymmetric correlational structure provided by the thermodynamic asymmetry disappears at the micro level. If we are talking sub-statistically, of events too small to be concentrated by whatever it is that is responsible for the fact that our region of the universe is not in thermodynamic equilibrium, then there is no asymmetry of correlation productivity. The asymmetry described in (1) is therefore not only contingent but also essentially macroscopic: it depends on the fact that (due to the thermodynamic asymmetry) there are macroscopic concentrations in initial conditions but not (as far as we know) in final conditions.[3]

In sum, if cast in actualist terms an account based on the fork asymmetry is scope-deficient in two senses. It fails to apply to many *actual* cause–effect pairs, particularly microscopic ones; and might fail to apply to *any* cause–effect pair (and so fails to respect modal intuitions concerning causality in counterfactual circumstances). This scope-deficiency may be traced to the fact that the thermodynamic asymmetry is itself insufficiently "global", in both actual and modal terms.

It seems to me that this difficulty afflicts all attempts to ground causal asymmetry on the fork asymmetry, interpreted in actualist terms. Consider for example the appeal to the independence of initial conditions (e.g. by Ehring 1982, 1987, as well as by Papineau 1985). The lack of actual correlations between initial conditions depends on the absence of concentrated future sinks, or of anti-thermodynamic behaviour generally. If the future were like the past, in containing regions whose entropy was much lower than it is at present, then initial conditions would exhibit precisely the same kinds of correlation as final conditions do in fact. Alternatively if entropy were higher in the past, there would be fewer correlations in final conditions. So we only have an asymmetry of the sort described in (1b) because and in so far as we have the thermodynamic asymmetry. And this is insufficiently general, both modally and actually, to ground the asymmetry of causation. Actually speaking, there isn't enough asymmetry to go round.[4]

A slight digression at this point: it has sometimes been claimed that the thermodynamic asymmetry may be *explained* by the independence of initial conditions early in the history of the universe. For example, Horwich (1987) suggests that it is the fact that the initial micro state of the universe is highly random that explains why entropy generally increases. He points out that in contrast the final micro state must be highly correlated, reflecting the fact that it is the deterministic product of a very highly ordered early macro state. This proposed account of entropy increase seems to me to involve a serious confusion concerning explanatory priority, however. For it seems quite inappropriate to say that the universe has a highly ordered *early* macro state because it is *later* in a highly correlated micro state. In this case if there is an adequate explanation in either temporal direction it goes in the other direction: the ordered early macro state explains the later correlated micro state. Alternatively we might say that there is no substantial explanation in either direction—that we simply have two ways of describing the same fact about the universe, in effect. Either way, we don't take the final micro state to explain the initial macro state.

But then by what right do we take the initial micro state to explain the final macro state? In practice of course we are inclined simply to help ourselves to the principle that the past explains the future; but what could possibly justify that inclination here, where the temporal asymmetry of the universe is what we are seeking to explain? In sum, it seems to me that we have no more right to take initial micro chaos to explain the later macro state than we do to take the final micro state to explain the initial macro state. Again, either it is the later macro state which explains the earlier micro state, or (and this seems to me the more appropriate conclusion) there is no substantial explanation in either direction, and we simply have two ways of describing the same phenomenon.

5. Beyond the Actual—Manipulability as a Route to Modality

We have seen that if the principle of correlation productivity is to provide a generally applicable distinction between cause and effect, it needs to be given modal "bite". In practice this is commonly provided by a formulation going something like this:

3. Increasing the frequency of a common cause increases the frequency of its effects, thereby yielding an increased correlation between the latter; not so for increasing the frequency of a common effect.

Admittedly this is not unambiguously modal as it stands, for the phrase "increasing the frequency" might be taken to refer to changes in actual frequencies over time. However, if we are escape the insufficiency of actual frequencies in the service of an understanding of causal asymmetry, the reading must be modal. One option is then to read the formulation in agency terms; to see the implicit reference to manipulation of frequencies as essential, in other words. This is the approach I want to recommend, of course. It reflects the agency approach to causation for which I wish to argue. Is there any alternative? Can the notion of change of frequency the formulation embodies be understood in terms neither anthropocentric or conventional, but rather such as to depend on some objectively asymmetric feature of the dependence-structure of the world?

Against this possibility the most effective argument seems to be to point out that there are conceivable agent perspectives from which the asymmetry simply does not hold. Consider for example the perspective available to God, as She ponders possible histories for the universe. For all we presently know, God may have originally had a preference for a world in which the beginning of the Third Millennium in January 2001 is marked in spectacular fashion by the occurrence of many millions of tiny fires around the globe. Among the possible histories of the world are some in which the number of individually accidental fires at that time is several orders of magnitude higher than normal. In those histories there are simply many more "accidental" conjunctions of combustibles, oxygen and sources of ignition just prior to the given date than we would normally expect (reflecting the fact that combustibles and oxygen are among the joint causes of fires). In opting for such a history over others, God would have increased the frequency of a common effect—namely fire—and hence produced a correlation between its joint causes.[5]

The moral of this example is that if we view matters from a sufficiently atemporal perspective, the temporal asymmetry (3) attempts to capture will elude us. So if the notion of change of frequency is to yield an asymmetry, it has to be taken in such a way that the asymmetry is effectively imported from somewhere else. We might do this by stipulation, thus reintroducing Humean conventionalism. We might import some asymmetric modal notion (temporally asymmetric counterfactuals, for example) to do the trick, thus passing the buck. Or we might read the notion in agency terms, so that the imported asymmetry in (3) is that of the agent's perspective. This brings us to Ramsey's approach to causal asymmetry. In taking this approach we shall of course want to be convinced that it does not commit either of the sins just mentioned: it doesn't amount to Humean conventionalism, and it doesn't pass the buck. But first let me say some more about what the suggestion is, and how it avoids other difficulties.

6. Ramsey's Ultimate Contingency

In (1978) Ramsey extends to the topics of law and causality the subjectivist or pragmatic approach he had earlier taken to probability. Most importantly for our present purposes, he links the asymmetry of cause and effect to our perspective as agents, saying that "from the situation when we are deliberating seems to ... arise the general difference of cause and effect." (1978, p. 146) The suggestion is thus that we should look for the origins of causal asymmetry not in terms of some objective asymmetry in the world, but in terms of those features of our perspective as agents which lead us to conceptualise the world in these asymmetric terms. Note that Ramsey is not suggesting that human agency is beyond the reach of science in general—this is not Taylor's (1966) agency causation—but merely that some of our concepts originate in the fact that we are agents, and reflect the distinctive perspective with which an agent regards the world. For Ramsey our notion of causality depends on the fact that we are agents in much the same way that our notion of probability depends on the fact that we are creatures capable of partial belief.

Ramsey goes on to identify what he takes to be the crux of the agent's perspective, namely the fact that from the agent's point of view contemplated actions are always considered to be *sui generis*, uncaused by external factors. As he puts it, "my present action is an ultimate and the only ultimate contingency." (1978, p. 146) I think this amounts to the view that free actions are treated as probabilistically independent of everything except their effects; and I think that the appropriate move is to read this in reverse, saying that the effects of an event A are those events to which A would provide a means. That is, we consider a given event B as an end, as something whose occurrence we wish to bring about. If in the context of means–end deliberation to realise A as the immediate product of a free action would be to raise the probability of B, then B is thought of as an effect of A.

In (1991b) I have argued that this approach provides the most promising basis for a probabilistic theory of causality. Among its virtues is the fact that it avoids the problem of spurious causes: correlations due to common causes don't translate into probabilistic dependencies from the agent's point of view, because the presence of the common cause is incompatible with the assumption of *sui generis* origins. The argument turns on a defence of evidential decision theory against Newcomb-style objections. Indeed I think the viability of the approach in general depends on this defence, for it is this that ensures that the probabilities in question need only be evidential, and hence not dependent on a prior modal notion (as invoked in causal decision theories). Essentially what needs to be shown is that correlations between prior causal states and actions do not translate into evidential dependencies from the agent's perspective; and the crucial point is that in the means–end context any such subjective dependency would itself be a causal factor, so that the principle of total evidence would immediately undermine the judgement on which it was based. (For the details see my 1991b and 1986.)

This argument shows that Ramsey's suggestion is coherent in purely evidential terms, I think. As such, it is then available to ground our causal concepts in the way that Ramsey suggests. There are other potential objections to such an account of causality, of course, and in (Menzies & Price, forthcoming) Peter Menzies and I respond to a number of standard arguments of this kind. It would take me too far afield to go into these issues here. Want I want to do here is to point out that Ramsey's approach provides a source of asymmetry which avoids the scope problem, conventionalism and buck passing; and also, as promised, to argue that it accounts for some of the intuitive plausibility of the fork asymmetry—viz., it explains why we find it implausible that there should be correlations between non-causally related events which are not explicable in terms of a joint correlation with an earlier common cause.

First then, how does Ramsey's suggestion avoid buck-passing and disguised Humean conventionalism? The former is straightforward: there is simply no asymmetric modal notion left unaccounted for. Ramsey explains our use of the (asymmetric) modal notions in terms of something actual, namely our own constitutions.[6] As for Humean conventionalism, Ramsey relates our use of the notions of cause and effect to our perspective as agents. One consequence of this is that the concepts of cause and effect are conceptually distinct from those of past and future, earlier and later. Of course, it might well turn out that these latter notions themselves depend on related features of our constitution. Our distinction between past and future may well turn out to be importantly anthropocentric, for example. But as long as there is some significant distinction between whatever aspect of us supports the past–future distinction on the one hand, and whatever supports the cause–effect distinction on the other, the two distinctions will rest on distinct grounds and not be analytically related.

Earlier I distinguished de re and de dicto senses of the question why the cause–effect arrow aligns with the earlier–later arrow; in fact four different senses, given that each distinction might be read in either a de re or de dicto sense. For Ramsey there is presumably no de re sense of the causal distinction; nothing in the world, as it were. The two remaining issues are (i) Why do we impose the causal arrow in this particular temporal direction (taken de re)? And (ii) Why does the causal arrow align with the earlier–later arrow, taken de dicto? These questions will receive significantly different answers. The first is likely to appeal to what we know about the dependence of agents on the thermodynamic arrow. Roughly, the existence of agents appears to depend on the entropy gradient; thus the reason our agent's perspective (and thus our imposed causal arrow) is oriented this way rather than that is that the entropy gradient goes this way, at least in this region of the universe. But whichever way the gradient went we would align the causal arrow with what *we took to be* the past–future arrow; so that in the de dicto sense, it seems to be a priori that the causal arrow aligns with the past–future arrow—a priori not because the connection is analytic, but because of a relationship between what it is *in us* that supports each of the concepts concerned. (This is something like a Kantian synthetic a priori, in other words.)

7. Agent-grounded Causation and the Existence of Common Causes

It appears to be contingent that the world contains the kind of correlative forks that we think of as involving a common cause and its joint effects; for it depends on the thermodynamic asymmetry, and this is surely a contingent matter. All the same, I think there is an important sense in which the principle I have called correlation explicativity is a priori, given Ramsey's understanding of the connection between causality and the agent's perspective. This claim may seem to conflict with the contingency of the fork asymmetry, but as we shall see, there is a nice resolution of the apparent tension.

First to the argument. Suppose events of kind A are found to be correlated in practice with events of kind B, so that Freq(AB)>Freq(A).Freq(B); and let us assume that the correlation is not simply accidental. It follows from Ramsey's view that if this actual correlation is taken to support the conclusion that P(B/A)>P(B), under the assumption that an A event is produced as result of a free action, then A will be taken to be a cause of B. So if it is specified that the correlation between A and B is non-causal—i.e., that A is not a cause of B and vice versa—then apparently this can only be because the correlation requires a particular history for A, incompatible with its being produced by a free action. (It vanishes if we try to do it ourselves, so to speak.) In this case let us think of C as the disjunction of all possible histories for A which maintain the correlation between A and B. There may be many such histories, of course. The crucial point is that if the (non-accidental) correlation is not to support P(B/A)>P(B) under the assumption that an A event is produced by a free action, it must depend on A's being produced in some other way.

We now wish to show that C may chosen so as to provide a common cause for A and B. The situation is symmetric with respect to A and B, so that under the assumption that B is not a cause of A and that the correlation is non-accidental the agent is entitled to conclude that there is some C* such that the correlation only holds when B is produced by C*. Moreover if C and C* were non-identical then an agent might consider producing one without the other—let us say C* without C, for example. Given that the correlation is robust under these conditions this would provide a means of producing A by producing B, without producing A directly (for by assumption C is the disjunction of all possible causes of A which preserve the correlation, so that if we have C* without C, C* is not a cause of A directly). Given the assumption that A and B are not directly causally connected, in other words, it follows that C* and C are the same condition, and comprise a common cause for A and B.

It appears that C may also be chosen so that its occurrence screens off A from B. For if we had P(AB/C)>P(B/C).P(A/C) we could run same argument again: either A would be a cause of B (given presence of conditions C), again contradicting the initial assumption, or there would be some D such that only when A is produced by CD is it correlated with B. In the latter case we conclude that our initial choice of C was insufficiently inclusive; we should have chosen what we are now calling CD.

This is rather informal, and I am not sure that it is watertight as it stands. It seems to me to rest on a sound intuition, however. Given the way in which the notions of cause and effect are a product of our perspective as agents, to say that a correlation between remote events does not rest on a direct causal connection is to say that it is not a correlation that an agent might exploit by producing one event as a means to the other. This in turn can only be because there are preconditions of the correlation which prevent its being exploited by an agent in this way. (Why *pre*conditions rather than *post*-conditions? Because as agents we are oriented this way—we deliberate from past to future. Thus again the asymmetry is not merely conventional, but rather reflects our own constitutional asymmetry in time.) And these preconditions will comprise a common cause for the correlated events. Thus correlation explicativity is a priori, and at least in this respect the fork asymmetry is a *product* rather than a *constitutive part* of our notion of causation. This is what I meant earlier in saying that Ramsey's view partially reverses the order of conceptual priority between causal asymmetry and the fork asymmetry.

What of the apparent tension between this conclusion and the contingent character of the fork asymmetry? Suppose for example that we were to encounter a region of the world in which the thermodynamic asymmetry was partially reversed, so that macroscopic correlative forks occurred with the reverse of the normal orientation. Wouldn't correlative explicativity be expected to fail in this case? The resolution of this tension lies in the fact that if events A and B were correlated via a joint correlation with some later event E, agents could exploit this correlation to use A as a means to B, or vice versa. Hence it would not be true that A and B were not causally connected, or at least so the Ramsey view seems committed to saying. (Thus for example if my movements and those of Death are jointly correlated in virtue of our future meeting at some prearranged time (but not place!) I can affect Death's movements in the interim by choosing to journey say to Samarra rather than Baghdad on the day in question.) So although it is contingent that the world does not contain such reverse correlative forks, it is nevertheless a priori (given the orientation in time we human agents actually have, and given what we mean by causation) that a *non-causal* remote correlation will not turn out to be of this kind.

A number of people have objected that the above argument seems to fly in the face of apparent counterexamples to Reichenbach's common cause principle suggested by

Elliott Sober. Sober (1988a, 1988b) describes cases in which we appear to have remote non-causal correlations without common causes. His favourite example concerns the correlation between rising bread prices in Britain and rising sea levels in Venice. Given that both processes are (let us assume) monotonic, the event of a rise in the bread price is positively relevant to that of a rise in sea level. Clearly there need be no common cause at work here. So it would seem that the common cause principle is not even true in general, much less a priori. I want to conclude the section with a brief response to this objection.

It seems to me that the main lesson of Sober's examples is that we need to pay more careful attention than is usual to the distinction between token correlations and type correlations. As all parties recognise, the principle of the common cause is not intended to apply to token or "one-off" correlations, for these may simply be accidental. Suppose for example that I find an eggplant whose profile resembles that of Ronald Reagan. The fact that these similar profiles have no common cause provides no counterexample to Reichenbach's principle. It is not that the principle is completely inapplicable in the single case—clearly we regard some single-case similarities as more likely to be coincidental than others, and in thinking this way we are giving implicit recognition to the existence of a common cause constraint of some kind. But it is in the case of type correlations that the principle applies with full force, and with respect to which it has some claim to be accepted as a universal constraint.

Do Sober's examples concern type correlations or token correlations? Let us turn first to a simpler case. Think again of the Reagan-like eggplant. Suppose that its photograph is reproduced in *Newsweek*, beside a picture of Reagan himself. There now exist thousands of individual pairs of images, each pair displaying the same remarkable internal similarity. Is this multiple correlation any more in need of a common cause explanation than the original accident? Obviously not, for the original similarity has simply been replicated. A full explanation need only refer to the single original accident, and to the process of replication. The Reagan pictures trace their causal roots to Reagan, and the eggplant pictures trace theirs to the original eggplant, but the two chains have no common ancestor. Thus we need to distinguish between *genuine* type correlations—those between multiple independent instances of events of two given types—and multiple correlations which simply involve the bilateral replication of a single original pair of events. (Let us call these *replicated token correlations*.) The common cause principle need be no more concerned with replicated token correlations than it is with single-case token correlations.

Moreover, the term "replication" is to be read very broadly here. Almost any deterministic process will count, in particular. Thus suppose we have two deterministic processes G and H, responsible for the values of physical variables g and h, respectively. The initial values g_0 and h_0 thus determine later values g_t and h_t. Except in certain trivial cases (e.g. if one of the processes produces a constant value), the two variables will be correlated in much the same way as Sober's bread prices and sea levels. But clearly what is involved here is simply a deterministic "replication" of a one-off relationship between the initial values g_0 and h_0. No common cause need be involved.

In some cases the processes of replication themselves may be the main source of the observed similarity. In Sober's example it is that fact that each process is monotonic that does the work. But note that it is essential to the example that the two variables (bread price and sea level) each be governed by *some* (more-or-less) deterministic process. (After all, why isn't the monotonic increase in the bread price over time itself an outstanding coincidence, as it would seem to us to be if we thought of it as a random variable? Because we take it for granted that there is an underlying process or mechanism at work.) The progress of the bread price and the sea level in Venice over time are thus both in an important sense single prolonged events, rather than a series of independent events. The case is one of replicated or extended token correlation, rather than type correlation; and hence provides no counterexample to the common cause principle in its intended type-correlation form.

In practice it may not always be easy to determine whether we are dealing with a genuine type correlation or a replicated token correlation. Roughly, what needs to be decided is whether all the individual events on each side of the correlation trace their existence (or relevant characteristics) to a common ancestor, or whether instead they are the several products of multiple independent processes. It is in the latter case that we expect a common cause, given that we take there to be no more direct causal relationship to explain the remote correlation.[7]

8. Conclusions

To summarize: attempts to account for causal asymmetry and directedness in terms of objective statistical asymmetries come to grief on the scope problem—or at least they do so if they rely on *actual* correlations, for in that case their actual and counterfactual scope is essentially that of the thermodynamic asymmetry, which isn't wide enough. This scope-deficiency is commonly obscured either by disguised Humean conventionalism or by an appeal to some further modal asymmetry, but neither of these options is satisfactory. The most plausible way to account for causal asymmetry is to regard it as "put in by hand", that is as a feature that agents project onto the world. The crucial statistical asymmetry is then the anthropocentric one identified by Ramsey, namely that as agents we take our actions to be statistically independent of everything except (what we come to call) their effects; and its temporal orientation stems from that of ourselves as agents. Finally, the right story about the relevance of thermodynamics seems to be that the entropy gradient makes possible the existence of agents, but that once in place, as it were, these agents project their perspective beyond the confines of the gradient on which they depend. In that sense causal asymmetry is indeed more widespread than any physical temporal asymmetry—or so it properly seems to us, as we regard the world from the agent's point of view.

Notes

[1]Following Reichenbach, principle (2) is often called the principle of the common cause. My otherwise regrettable neologism above is intended to draw attention to the important distinction between (2) and (1).

[2]These points are discussed in more detail in my (1991a) and (1991c).

[3]I would very much like to be able to put the above discussion on a more formal footing, and in particular to be able to demonstrate more decisively the parallel between the case of radiation and the fork asymmetry in general. Failing that, I am aware that the suggested analogy may be found less than conclusive. All the same, an opponent who wishes to maintain that there is more to the fork asymmetry (in actualist terms) than is given to us by the thermodynamic asymmetry is committed to the view that the world exhibits a very significant temporal asymmetry which is not traceable to the same roots as the thermodynamic asymmetry; and this should seem a very strong claim indeed, given our current understanding of temporal asymmetry in

physics. True, it is not an unusual view that there is a *modal* temporal asymmetry distinct from those of physics; but the status and origins of modal temporal asymmetries such as those of causation are precisely the questions at issue.

[4]David Papineau's response to this argument seems to be to concede that (asymmetric) causation is simply very much more "patchy" than we have always thought. At the time of writing I haven't seen this response developed on paper, and in any case I don't have the space to attempt to do it justice here. My main concerns have to do with the devastation it appears to wreak on modal intuitions about causation, and its apparent inability to make sense of causal reductionism (i.e., the intuition that big bits of causation are made up of lots of little bits). On the latter point, see (Price 1992, pp. 511-12).

[5]Note that from our point of view this world would seem to exhibit fire-seeking teleology before January 2001, revealing non-independent initial conditions.

[6]True, we may employ modal notions in describing such things as our own constitutions. We may say for example that we are *disposed* to behave in this way rather than that. This raises some interesting issues for the pragmatist's project, but it would take us too far afield to explore them here.

[7]Many commentators take the EPR cases to provide a more substantial (if more esoteric) counterexample to the common cause principle. However, the argument depends on the assumption that the EPR correlations are not causal, itself a hotly disputed point. Most of those who argue that the EPR correlations are causal are thinking in terms of spacelike non-local effects. An alternative which might particularly appeal in the present context (where the direction of causation is the point at issue, as it were) is that they be thought of as mediated by joint correlations with *future* events. For more on this see (Price, forthcoming).

References

Arntzenius, F. (1990), "Physics and Common Causes", *Synthese* 82: 77-96.

Ehring, D. (1982), "Causal Asymmetry", *The Journal of Philosophy* 79: 761-74.

Ehring, D. (1987), "Papineau on Causal Asymmetry", *British Journal for the Philosophy of Science* 37: 81-7.

Horwich, P. (1987), *Asymmetries in Time*. Cambridge, MA: MIT Press.

Lewis, D. (1979), "Counterfactual Dependence and Time's Arrow", *Nôus* 13: 455-76.

Menzies, P. & Price, H. (forthcoming), "Causation as a Secondary Quality", *British Journal for the Philosophy of Science*.

Papineau, D. (1985), "Causal Asymmetry", *British Journal for the Philosophy of Science* 36: 273-89.

Popper, K. (1956), "The Arrow of Time", *Nature* 177: 538

Price, H. (1986), "Against Causal Decision Theory", *Synthese* 67: 195-212.

_ _ _ _ _. (1991a), Article review of K. G. & J. S. Denbigh, *Entropy in Relation to Incomplete Knowledge* (Cambridge University Press, 1985) and H. D. Zeh, *The Physical Basis of the Direction of Time* (Springer-Verlag, Berlin, 1989), *British Journal for the Philosophy of Science* 42: 111-144.

_ _ _ _ _. (1991b), "Agency and Probabilistic Causality", *British Journal for the Philosophy of Science* 42: 157-76.

_ _ _ _ _. (1991c), "The Asymmetry of Radiation: Reinterpreting the Wheeler-Feynman Argument", *Foundations of Physics* 21: 959-75.

_ _ _ _ _. (1992), "Agency and Causal Asymmetry", *Mind* 101: 501-20.

_ _ _ _ _. (forthcoming), "Neglected Route to Realism About Quantum Mechanics", *Mind*.

Ramsey, F.P. (1978), "General Propositions and Causality", in *Foundations: Essays in Philosophy, Logic, Mathematics and Economics*, D.H. Mellor (ed.). London: Routledge & Kegan Paul, pp. 133-51.

Reichenbach, H. (1956), *The Direction of Time*. Berkeley: University of California Press.

Sober, E. (1988a), *Reconstructing the Past: Parsimony, Phylogeny and Inference*. Cambridge, MA: Bradford/MIT Press.

_ _ _ _ . (1988b), "The Principle of the Common Cause", in *Probability and Causality: Essays in Honor of Wesley Salmon*, J.H. Fetzer (ed.). Dordrecht: Reidel, pp. 211-28.

Taylor, R. (1966), *Action and Purpose*, Englewood Cliffs: Prentice Hall.

Zeh, H.D. (1989), *The Physical Basis of the Direction of Time*, Berlin: Springer-Verlag.