

Causal Asymmetry

by DAVID PAPINEAU

- 1 *Introduction*
- 2 *A Probabilistic Asymmetry*
- 3 *The Connection between Probabilities and Causes*
- 4 *Causal Asymmetry*
- 5 *Lewis's Account of Causal Asymmetry*
- 6 *Indeterminism and Screening-Off*
- 7 *Physical Asymmetry*

I INTRODUCTION

It would be nice to be able to give some explanation of the asymmetry of causation. Not least because the standard philosophical analyses of causation make it mysterious that causation should have a direction at all. Hume said that one event being the cause of another is a matter of events like the first being constantly conjoined with events like the second. But constant conjunction is manifestly a symmetrical relationship. If *A* is constantly conjoined with *B*, then *B* is likewise constantly conjoined with *A*. Whence the asymmetry?

Of course Hume also said that of two constantly conjoined events the earlier is the cause and the later the effect. But there are good reasons for resisting this Humean suggestion.

- (1) There seem to be cases of simultaneous cause-effect pairs. The causal asymmetry in such cases obviously can't be accounted for in terms of any temporal asymmetry.
- (2) The Humean suggestion implies that backwards causation is an immediately obvious contradiction. But surely this is wrong. Even if there are in the end deep conceptual reasons why effects can't precede their causes, they aren't immediately obvious. We can certainly make some initial sense of precognition (where present mental states are affected by future events) and of travel back in time (where the traveller's past behaviour is affected by his or her present-day experience). We ought to allow that backwards causation is at least conceivable.
- (3) The Humean suggestion would also undermine any attempt to account for the asymmetry of *time* in terms of the asymmetry of *causation*. This reverse reduction has a number of attractions. Central

Received 17 February 1984

T

to the arrow of time, to the asymmetry of earlier and later, are such basic features of human life as memory and action. One can say that the past, for a given person, just is that side on which events they can remember lie. And the future is the side on which events they can influence lie. But what are memories but mental states *caused* by the events they are of? And, even more obviously, what is it to influence an event but to *cause* it to happen? So it seems plausible to conclude that 'earlier' is the direction of those events that can affect us, and 'later' those events that we can affect.¹ Of course this argument for reducing temporal asymmetry to causal asymmetry needs a lot of tightening. But it certainly seems a promising line of thought. (See in particular Mellor [1981], ch. 9.) Yet if we were to rest with the Humean suggestion any such reduction would be rendered pointlessly circular, for talk of causal asymmetry would already be presupposing the notion of causal asymmetry.

2 A PROBABILISTIC ASYMMETRY

The problem, then, is to produce an account of causal asymmetry which does not rely on the temporal relationships between causes and effects.

I shall approach this problem via recent work on probabilistic accounts of causation. (See in particular Suppes [1970] and Salmon [1971].) The basic idea of such probabilistic accounts of causation is that one event's causing another does not require that the former determine the latter, but only that it make it more probable than it would otherwise have been. For smoking (*S*) to cause cancer (*C*) it doesn't have to be the case that all smokers get cancer, only that the conditional probability, of cancer given smoking, is greater than the probability of cancer in general. That is, $\text{Prob}(C/S) > \text{Prob}(C)$. (Note that on this account smoking does not have to make the probability of cancer *high*. It only has to increase it. It would be enough if the proportion of smokers who got cancer was 5%, provided the proportion of non-smokers who got it was less.)

Of course this suggestion in itself does not show causation to be asymmetrical. If smoking makes cancer more likely ($\text{Prob}(C/S) > \text{Prob}(C)$) then it immediately follows that, speaking probabilistically, cancer makes smoking more likely ($\text{Prob}(S/C) > \text{Prob}(S)$). The two conditions on the probabilities are trivially equivalent, both indeed being different ways of

¹ Wouldn't the reduction of temporal to causal asymmetry have the above-mentioned undesirable consequence of making backwards causation immediately contradictory? But nothing in the initial argument for the reduction immediately ensures that causes always lie on the same temporal side of their effects; nothing automatically ensures that the causal arrow points in the same direction in all regions of space-time. And so there is certainly initial conceptual room for causal arrows that go 'against the general flow'; and these could plausibly be counted as 'backwards causes' by the causal theory of temporal direction. For further discussion of such variation in causal orientation and of the complications it raises for the causal theory of temporal direction see Mellor, *op. cit.* ch. 10, and Healey [1983].

expressing the overtly symmetrical claim that cancer and smoking are probabilistically associated ($\text{Prob}(C \& S) > \text{Prob}(C)\text{Prob}(S)$).

But there is a further twist to the probabilistic approach to causation. After all, it is a familiar thought that a probabilistic association between two events does not in itself establish a causal connection: 'Correlation doesn't prove causation.' Thus in our example it is possible (as the cigarette companies sometimes half-heartedly suggest) that some further factor—a genetic predisposition, say, or a high degree of psychological stress—is on the one hand leading people to smoke, and on the other giving them cancer. If this really were the case then the correlation between smoking and cancer would be 'spurious'. Smoking wouldn't 'really' make cancer more likely—it would just appear to do so because it was associated with the genetic predisposition (or stress, or whatever, depending on the precise form of the story).

In probabilistic terms the question the cigarette companies are raising is this—is there some further factor X such that when we take it into account smoking doesn't make any further difference to the probability of cancer? That is, is there an X such that $\text{Prob}(C/S \& X) = \text{Prob}(C/X)$, and $\text{Prob}(C/S \& \neg X) = \text{Prob}(C/\neg X)$? As it is sometimes put, is there an X which 'screens' the cancer 'off' from the smoking? If there isn't such an X , then the cigarette companies are wrong and there is a real causal connection between smoking and cancer. But if there were such an X , then it seems that they would be right— X would be affecting S and C separately ($S \leftarrow X \Rightarrow C$) and there would be no direct causal connection.

But note now that this does give us an asymmetry. When we find an X that screens off an association between S and C we conclude that X is a common cause of the joint effects S and C . We don't conclude that X is a common effect and that S and C are its joint causes ($S \Rightarrow X \leftarrow C$).¹

Let us focus then on this probabilistic asymmetry. In general terms it can be put like this. Common causes screen off associations between their joint effects. But common effects don't screen off associations between their joint causes.

To illustrate, let us forget about the cigarette companies and their forlorn hope that smoking and cancer will turn out to be joint results of some hidden predisposition, and let us instead consider smoking itself as a common cause of the joint effects, cancer and, say, yellowish fingers (Y). ($Y \leftarrow S \Rightarrow C$). Here we will find an association between the joint effects C and Y —more

¹ As stated this 'asymmetry of screening-off' is an oversimplification. For not only do common causes screen off associations amongst their effects (as in: effect \leftarrow cause \Rightarrow effect), but so also do causal intermediaries screen off associations between their original causes and their subsequent effects (as in: prior cause \Rightarrow intermediate cause \Rightarrow effect). But this still gives us an asymmetry, for these two cases are still different from that of: cause \Rightarrow effect \leftarrow cause, where there aren't any associations to be screened off. The existence of screening-off by causal intermediaries is by and large irrelevant to my argument, and it would complicate the exposition unnecessarily to bring it in explicitly. (But see n. 1, p. 281.)

people with yellow fingers get cancer than in the population at large ($\text{Prob}(C/Y) > \text{Prob}(C)$). And this association disappears when we control for smoking—since, presumably, non-smoking yellow-fingered people (jaundiced people, chalk manufacturers, . . .) aren't more likely to get cancer than the population at large.

But now consider cancer in turn as a common effect of both smoking and, say, working in an asbestos factory (A). ($A \Rightarrow C \Leftarrow S$). Here we won't find any association between the joint causes. It won't be the case (at least not because of our story) that cigarette smoking is more common amongst asbestos factory workers than in the population at large. And so, *a fortiori*, it won't be the case that an association between the joint causes is screened off by their common effect.

3 THE CONNECTION BETWEEN PROBABILITIES AND CAUSES

What we make of the probabilistic asymmetry just identified (the 'screening-off' asymmetry henceforth) depends on what view we take of the connection between probabilities and causation.

Most writers in this area seem to adopt what might be called a *definitional* view: the view that what it is for one event to cause another just *is* for the appropriate probabilistic relationships to obtain between events of those types. On this view causal relationships are, so to speak, *constituted* by probabilistic relationships, analogously to the way that on the old Humean constant conjunction view causal relationships were held to be constituted by deterministic relationships.

I have a different view of the connection between probabilities and causation, a view that could appropriately be called an *evidential* view. On my view the relevant probabilistic relationships don't constitute the causal relationship as such: rather they provide indirect evidence for the existence of 'underlying' causal connections.

This is not the place to argue for the evidential view against the alternative definitional one. I have another paper that does that (Papineau [1985]). Here I shall merely present my view and content myself with a few brief comments on its virtues. My concern in this paper is not to defend the evidential view as such, but rather to explore certain consequences that it has for the way the screening-off asymmetry should be interpreted.

It is easiest to explain the evidential view if we assume that causation is deterministic, that the real, 'underlying' causal relationships are deterministic ones. (This is not essential to the evidential view—see Papineau, *op. cit.* section 4(b)—but, saving the comments in Section 6 below, I shall assume it henceforth.) So I am going to take it that if S causes C , there is some law which says that S , in conjunction with certain other background conditions X , is sufficient for (determines) C . ($S \& X \rightarrow C$).

Suppose we also take it that the world (or at least the bit of it we are considering) is deterministic, in the sense that everything that happens has a deterministic cause. (Again, this is not essential, but it is helpful to assume it.) So the idea is that if cancer isn't caused by smoking (in conjunction with some X) then one of its other causes must be present. Let us further simplify by assuming that asbestos factories (plus other background conditions) are the only other possible cause of cancer, and let us symbolise this other possible conjunction of conditions sufficient for cancer by A . Then we can write $S \& X \vee A \leftrightarrow C$. Cancer occurs if and only if a conjunction of factors which determines it does. (In effect I am requiring here that S counts as a cause of C iff it is an 'inus' condition for it. Cf. Mackie [1974], ch. 3.)

Once we have $S \& X \vee A \leftrightarrow C$, it is easy to see why causal connections should show up in such statistics as $\text{Prob}(C/S) > \text{Prob}(C)$. Given that you are a smoker, you will get cancer if either X or A . Whereas the non-smokers get cancer only if A . Since X or A can't help being more likely than A alone, the smokers are more likely to get cancer than the rest.

Of course one has to be a bit careful about making inferences here—it may, as before, be that smoking is correlated with cancer only because it is associated with A (more smokers in the asbestos business) and not because it is really part of a sufficient condition itself. But then, again as before, there is the screening-off check—is S still associated with C within the class of asbestos workers? If not, then it wasn't part of an independent sufficient condition; but if it is, then it must be.

I hope this suffices to give a general idea of the evidential view. In effect what I am doing is supposing that the substance of causation consists in the existence of underlying deterministic processes. Probabilities other than zero and one only come in because we relate the outcome (C) not to the full set of relevant initial conditions (S, X, A), but to S alone. The probability of C given S is thus a reflection of the probability of getting one of the other relevant initial conditions (X, A) when we have S . Obviously, these probabilities might, for all that has been said so far, be equal to any number between zero and one. The important point for the evidential view is that, whatever the exact probability of the initial condition $X \vee A$, it is going to be greater than the probability of the initial condition A .

These last remarks should mollify those who feel that serious probabilities are incompatible with determinism, those who feel that we ought not to be talking about probabilities other than zero and one unless we are dealing with properly indeterministic chance set-ups. I have no argument with this feeling. My probabilities derive from the probabilities of my initial conditions, and I have said nothing at all about where they come from. It is perfectly compatible with the determinism of the process *between* initial conditions S, X, A and the outcome C , that the initial conditions themselves be generated by indeterministic chance processes. (Or that they be generated by deterministic processes which in turn have initial conditions generated by chance mechanisms. . . .)

I said I would content myself with some brief comments about the merits of the evidential view. Let me make just two.

Firstly, note that in the context of our deterministic assumptions there is a strong reason for preferring the evidential view to the definitional alternative. The evidential view, of course, is at home in the deterministic situation, in the way I have outlined in this section. But at first sight it might seem that the definitional view can also cope satisfactorily with deterministic situations. After all, as I have just pointed out, the determinism doesn't mean that the probabilities $\text{Prob}(C/S)$, $\text{Prob}(C)$, *etc.* cease to exist. So can't the definitional view simply forget about the underlying deterministic process and continue to say that it is because of the probabilities as such that cancer causes smoking? No it can't. Because if the process is deterministic then we can have a case of this form: somebody who smokes, and then gets cancer, but who does so not because of the smoking in conjunction with X , but because of A . Clearly such a person is *not* caused to get cancer by smoking. But if the definitional view is going to ignore the underlying processes and rest with the surface probabilities it will be unable to avoid this conclusion.¹

Secondly, and not unrelatedly, note that the evidential view yields a natural explanation of the somewhat puzzling principle that an increase in probability is sufficient for causation, even if the probability remains well below 50 per cent. There is certainly something intuitively appealing about this principle. But nevertheless one feels uneasy with the thought that somebody's smoking *caused* them to get cancer if it only made it 5 per cent probable that they would get it. Nineteen times in twenty such a person wouldn't get cancer. Do we really have an account of what *made* them get it? The evidential view reconciles the conflicting intuitions. An increase in probability, however small, does suffice to 'establish causation'. But this is in the sense that it shows that one factor (S) *on occasion* operates in conjunction with other factors (X) to cause the result. The low increase in probability doesn't in itself constitute the causal connection: on those occasions where the causal connection is actually instantiated the probability of the result (C) given all the relevant circumstances ($S\&X$) is one. Or, to put it the other way round, the fact that smoking is on occasion part of a fully sufficient condition for cancer doesn't mean that the smoking as such makes the probability of cancer high. If the probability of X is itself low, smoking will be followed by cancer only relatively rarely.

4 CAUSAL ASYMMETRY

It might seem that I have now argued myself out of the position I wanted to get to. I have argued that probabilities don't give us the real substance of causation, that they are only superficial indications of underlying deter-

¹ This form of argument can be mobilised in defence of the evidential view even in the context of indeterministic assumptions. See Papineau, *op. cit.* section 4(b).

ministic relationships. But it was the probabilities that seemed to give us a hold on the asymmetry of causation. By returning to determinism haven't I let go of the asymmetry?

Perhaps not. At least I've progressed beyond Hume's trivially symmetrical 'constant conjunction' view of causation. I now have it that S is a cause of C just in case we have an equivalence of the form $S \& X \vee A \leftrightarrow C$. As I said above, this is roughly the requirement that S be an 'inus' condition of C . Perhaps this will give us an asymmetry after all. Perhaps 'being an inus condition' is itself an asymmetrical relationship.

But unfortunately it isn't. It certainly isn't asymmetrical; and given a fairly natural assumption it is downright symmetrical. (Cf. Mackie, *op. cit.*, p. 161.) Thus, suppose that S is indeed an 'inus' condition of C : $S \& X \vee A \leftrightarrow C$. Then obviously $C \& \neg A \rightarrow S$. And if we assume also that whenever S occurs it leaves some 'trace' (that if smoking doesn't cause cancer it will at least show up in something—smelly clothes, tired skin, . . . , the disjunction of which I shall signify by T , for 'trace') then we will have $C \& \neg A \vee T \leftrightarrow S$. That is, C is an 'inus' condition of S . (The 'fairly natural assumption' is the one about traces. It is the mirror image of the deterministic assumption that every event has a determining cause. Instead of assuming cancer is always determined, we are assuming smoking is always ineliminable.)

So we seem to be stuck with the problem raised at the beginning of this section. By switching from probabilities to 'inus' conditions, haven't I let go of the asymmetry of causation?

But let us focus once more on the probabilistic asymmetry that seemed to offer us some grip on the asymmetry of causation—namely, the screening-off asymmetry that common causes screen off associations amongst their joint effects, but common effects don't screen off associations amongst their joint causes. How should this be accounted for, given the evidential view of the connection between probabilities and causes I have been defending?

Let us suppose, as before, that smoking is a common cause of two effects, cancer and yellow fingers. ($Y \Leftarrow S \Rightarrow C$.) I am now taking this to be a matter of smoking being an inus condition both of cancer ($S \& X \vee A \leftrightarrow C$) and of yellow fingers ($S \& M \vee N \leftrightarrow Y$, for some M and N along the lines of M = not scrubbing your fingers and N = being jaundiced, or a chalk worker, or . . .). Now it is easy enough to see why this pair of equivalences should give rise to a probabilistic association between C and Y , the common effects. C is equivalent to $S \& X \vee A$, and Y is equivalent to $S \& M \vee N$. Given that the 'background conditions' are suitably independent of each other, given that the chance of X or A occurring with (or without) S isn't affected by whether M or N occurs, the presence of S in both these necessary and sufficient conditions makes it inevitable that C and Y will turn up together more often than one would expect given their separate probabilities of occurrence. Intuitively, cancer makes smoking more likely; and smoking makes yellow fingers more likely; so cancer makes yellow fingers more likely.

But why then doesn't it work the other way round? If, as I argued a moment ago, effects are as much 'inus' conditions of their causes as *vice versa*, why can't we argue analogously that we will find probabilistic associations amongst the joint causes of a given effect?

Well, let us try it. Let us take cancer, as before, as the common effect of two alternative causes, smoking and asbestos. Then we have the 'inverted' inus conditions: $C \& -A \vee T \leftrightarrow S$, and $C \& -S \vee R \leftrightarrow A$. (R here signifying the 'traces' of asbestos.) So why can't we argue that the common appearance of C in both $C \& -A \vee T$ and in $C \& -S \vee R$ means that smoking must be more likely in asbestos factories, as above we argued that cancer must be more likely amongst yellow-fingered people?

But now the reason should be obvious. The reason we don't get an association between smoking and asbestos factories is that the 'background factors' in the 'inverted' inus conditions (the $-A$, T , $-S$, R) are *not* probabilistically independent of each other. In the common-cause-joint-effects case above I took it that the analogous background conditions were so independent. If the 'inverted' background conditions in the present case were similarly independent, then a statistical association between smoking and asbestos factories would have to follow. But of course the 'inverted' background conditions aren't so independent. T stands for traces of smoking, and that is clearly going to be negatively correlated with $-S$ (namely, non-smoking); and similarly R stands for traces of asbestos, which is negatively correlated with $-A$. Indeed I take it that the negative correlations here are just what are required to 'cancel out' the appearance of C in both $C \& -A \vee T$ and $C \& -S \vee R$ and leave us with a null correlation between smoking and asbestos factories.

This then is my account of the asymmetry of causation. I want to claim that the asymmetry of causation derives from the fact that the background conditions together with which causes determine their effects are independent of each other, whereas the same does not hold of the background conditions together with which effects 'determine' their causes. If we focus solely on the general laws involved then the situation seems symmetrical: the effects (the final conditions) fix their causes (the initial conditions) as much as *vice versa*. But if we look at the probabilistic conditions satisfied at the different ends we get our asymmetry back: for the initial conditions satisfy independence requirements which do not hold of the final conditions.

Note that this is not intended as anything like a full analysis of causation as such. All I want to do is give an account of causal direction. I want to be able to take it for granted that there are causal connections amongst events: my claim is only that, given a set of such events, the causes are the ones that operate in conjunction with independent background conditions.

One immediate caveat about the suggested analysis of causal asymmetry. Contrary to my suggestion, isn't it perfectly possible that there should be correlations amongst the background conditions at the causal origin of some process? Mightn't it be that asbestos working (A) is associated with finger-

scrubbing ($-M$), thus, so to speak, obscuring the correlation between cancer and yellow fingers? Of course. But I would argue that, in order for there to be such a correlation between asbestos working and finger scrubbing, there would then have to be a common cause of these two factors (perhaps meticulous people are attracted to asbestos factory work), and I would argue that the initial conditions in conjunction with which *that* common cause gave rise to its joint effects would then be independent of each other. And I would deny that associations amongst 'background conditions' at the effect ends of processes can similarly be explained away as 'resulting' from common effects which turn out to be operating in conjunction with their own independent 'background conditions'.¹

But even given this caveat my analysis of the asymmetry of causation no doubt seems implausible. Why ever *should* it depend on the probabilistic independence of background conditions? But let us take stock. If we accept that the asymmetry of screening-off is central to the asymmetry of causation, and if we accept the above evidential account of the connection between probabilities and causes, then the 'independence of background conditions' account of causal asymmetry is pretty much forced on us—for it is just what we need, *modulo* the evidential view, to account for the screening-off asymmetry.

It would be nice to offer independent motivations for the independence account of causal asymmetry. It would be nice, for instance, to show that it fitted with such principles as that we can only remember things that are causally 'upwind' of us, and only influence things that are causally 'downwind'. But I must confess that I am not at all clear how to show this.

But I do have three lines of argument that offer some reinforcement of my analysis. The first relates to the importance of the screening-off asymmetry; the second suggests that an evidential approach to the explanation of this screening-off asymmetry will remain appropriate even when we relax our deterministic assumptions; and the third connects my analysis of causal asymmetry with the explanation of such physical asymmetries as the increase of entropy and the expansion of radiation.

¹ A further complication for the analysis of causal asymmetry arises from the screening-off by intermediate causes mentioned in n. 1, p. 275. For this means that not only do causes operate in conjunction with independent background conditions in 'fixing' their joint effects, but so also do intermediate causes operate in conjunction with independent conditions when 'fixing' their prior causes and their subsequent effects. Still, we can at least say the following. If *A*, *B*, and *C* are event-types related by 'inus' conditionship, and there is probabilistic independence amongst the 'background' conditions in conjunction with which *B* 'fixes' *A* and *C* respectively, then *B* is either the cause of both *A* and *C*, or it is causally intermediate between them, but it is not an effect of both. A question which then arises is how far such information will serve to determine the directions of the causal arrows between the directly related pairs in a set of causally related event-types. The simple case of three event-types where one screens off a correlation between the other two itself shows immediately that such information won't always give a complete determination. But one can imagine more complex cases where it will. What general conditions will suffice for such a complete determination is an interesting topic for further investigation.

5 LEWIS'S ACCOUNT OF CAUSAL ASYMMETRY

David Lewis has argued (Lewis [1973]) that c causes e just in case e counterfactually depends on (something which counterfactually depends on . . .) c . For David Lewis the counterfactual dependence of e on c requires (putting it at its simplest) that in the nearest possible world in which c is absent, e is absent too. In his [1979] Lewis addresses the question of whether this kind of analysis can account for the asymmetry of causation. The difficulty is this: given that in general the nearest world in which a cause is absent will be one where the effect is absent, what stops us symmetrically taking it that the nearest world in which an effect is absent will be one where the cause is absent, thus ending up with effects causing their causes as much as *vice versa*?

Lewis's answer is roughly along these lines. A given particular cause will characteristically produce a whole throng of independent chains of particular effects. (Consider the way in which an explosion, say, will leave traces all over the place.) But a given particular effect will very rarely be produced by more than one chain of particular causes. (Only one lighted match, or bolt of lightning, or finger on the button, or whatever, will normally be responsible.) As Lewis puts it, effects are rarely overdetermined by their causes; but the overdetermination of causes by effects is absolutely normal. And then he argues that the nearest possible world without a given effect will still contain the cause, because there will be lots of other effects left to fix the cause; but take away the cause and there will be no alternatives left to fix the effect.

Lewis of course has his work cut out to establish the requisite comparisons of inter-worldly distances. But those are his problems, not mine. What I am interested in is the connection between the phenomenon of differential overdetermination that Lewis appeals to and the asymmetry of screening-off. It is not difficult to see that these are closely related. The asymmetry of screening-off hinges on the fact that joint effects are probabilistically associated, whereas joint causes aren't. Joint effects are found together more often than one would expect given their separate probabilities of occurrence, but not so for joint causes. And this happens because any particular instance of a given cause will as a rule be followed by a number of its effects, but a particular instance of an effect will normally be preceded by just one of its causes. Which (given that we take into account a wide enough range of effects of any given cause) is just the phenomenon that Lewis invokes.

6 INDETERMINISM AND SCREENING-OFF

Let me now say a bit about indeterministic processes. Everything I have said in the last three sections has presupposed that we are working with deterministic processes. I have been supposing that probabilities other than zero and one only come in because the occurrence of initial conditions is

probabilistic; given the initial conditions everything else is determined. But of course we don't nowadays believe that all physical processes are like that. What then becomes of my analysis if we relax the assumption of determinism?

This is a somewhat complicated matter. The first thing to note is that allowing that the *world* is indeterministic doesn't force us to conclude that *causation* can be indeterministic. We can simply maintain that in so far as there is indeterminism there isn't causation. I defend a version of this view in my [1985]. I also show in that paper that even those who *do* believe there is genuinely indeterministic causation will need to recognise the possibility of an evidential attitude to the relationship between 'surface' probabilities and 'underlying' causal connections. But I don't want to go into all this here. Let me assume that, whatever view we take on these matters, the asymmetry of screening-off will still be central to the asymmetry of causation. I want to suggest that, even in the presence of indeterminism, the 'independence of background conditions' analysis remains the right approach to the asymmetry of screening-off, and hence to the asymmetry of causation.

What I have in mind is this. Even if the processes by which smoking leads on the one hand to cancer, and on the other to yellow fingers, are indeterministic, they might nevertheless require a specific sequence of determinate stages. It might be that whenever smoking leads to cancer there will have been a specific sequence of, say, fracture of the genetic material, rapid cell division, breakdown of the immune system, *etc.* Whether any stage is followed by the next one will be indeterministic, but whenever smoking does give rise to cancer it will be in virtue of that sequence of actual stages having been gone through. (There may of course be more than one possible 'route' from smoking to cancer, or even an interlinked collection of 'branching' routes. But that won't matter for the argument.) Again, one can imagine a similar chain for the smoking-yellow finger link, involving, say, successive stages in the typical mental processes leading up to decisions as to whether to scrub one's hands or not.

Now, if the chances of one stage following the next in one chain are appropriately *independent* of the corresponding chances in the other chain, then we will still have an explanation, analogous to the one given earlier for deterministic situations, for the existence of a cancer-yellow finger association, and indeed an explanation that shows why that association is screened-off by smoking. For we will still have equivalences of the form: $C \leftrightarrow S \& X \vee A$ and $Y \leftrightarrow S \& M \vee N$, though with X now standing for the conjunction of the stages needed for smoking to 'go through' to cancer, and M standing similarly for the stages between smoking and yellow fingers. And as long as we are still given, as seems plausible, that the probability of finding $X \vee A$ with S is independent of the probability of finding $M \vee N$ with S (and similarly without S), these equivalences will as before imply an association between C and Y which disappears when we control for S .

It is my conjecture that any attempt to construct 'inverted' chains of

stages leading 'back' from cancer to smoking on one hand, and from cancer to asbestos on the other, will fail to generate any association between smoking and cancer, because the 'intervening' stages on the two chains *won't* then be appropriately independent. (The stages required to 'get back' from cancer to smoking will be breakdown of the immune system, rapid cell division, *etc.*, *plus* the absence of the corresponding asbestos-cancer stages—the latter to ensure the cancer did come from smoking and not from asbestos. And similarly the stages needed to get back from cancer to asbestos will include the absence of the smoking-cancer intermediaries. And thus the occurrence of the relevant stages in the two 'backwards' chains will not be independent, but negatively correlated.)

But won't some common causes generate their joint effects straight off, without any determinate intermediate stages being realised along the way? And won't this then give us some screened-off correlations amongst joint effects (and thus a manifestation of the asymmetry of causation) which *can't* be accounted for in terms of independence requirements on intermediate conditions? I'm not so sure.

Consider the situation in these terms. If we *do* have determinate intermediate stages, then there is a sense in which the probability of cancer given smoking ($\text{Prob}(C/S)$) does not give the real chance of cancer given all prior circumstances on any particular occasion. Rather $\text{Prob}(C/S)$ reflects a 'mixture' of situations, in some of which the *DNA* has fractured, and in some of which it hasn't, and in which the eventual chance of cancer is therefore different. It is because we are thus dealing with mixtures, and the probabilities in a sense reflect our ignorance of all relevant factors, that the general lines of the evidential approach get a grip, and we are still able to explain the correlation between joint effects as due to the common cause operating in concert with other probabilistically independent conditions.

We are now being asked to consider a situation where $\text{Prob}(C/S)$, and $\text{Prob}(Y/S)$, do not reflect mixtures, but are the real chances, all things considered, of the relevant results. Prior to their 'measurements' there is nothing beyond *S* which is determinate and which is relevant to the occurrence of *C* or *Y*. In the terminology of quantum mechanics we have 'superpositions' of *C* and not-*C* (*Y* and not-*Y*), not mixtures.

I agree that in such a case we won't be able to invoke the 'independence of intermediate conditions' analysis to account for the correlation between *C* and *Y*, and the fact it is screened off by *S*, for by hypothesis there won't be any such intermediate conditions. What I am not sure about is whether any such cases actually occur.

It is because of quantum mechanics that we believe there are irreducible, 'unmixed' probabilities in the world. But quantum mechanics does not force us to accept that there are common causes producing 'unmixed' correlations amongst their joint effects. Indeed, as the experimental violation of Bell's inequality confirms, when we *do* have correlated quantum mechanical observables which prior to their measurements are in superpositions and not

mixtures (as we are now supposing C and Y to be) the nature of the correlation will as a rule be such as to preclude an explanation in terms of a prior screening-off cause. (Cf. in particular Mermin [1983].) How to interpret the correlation between such observables in causal terms is of course a puzzle. But what the impossibility of a screener-off seems definitely to show is that it is not due to the existence of a common cause of which the observations are two joint effects.

Nothing *a priori* rules out a given cause leading 'unmixedly' to two joint effects in such a way as to produce a screened-off correlation between them. And if this did in fact happen then it would discredit my claim that such screened-off correlations (and hence the asymmetry of causation) are always to be accounted for by reference to the independence of background/intermediate conditions. But both quantum theory and experiment suggest that such 'pure' cases of screenable-off correlations don't in fact happen. While there are 'unmixed' productions of correlations, they aren't ones that can be screened-off by common causes. So whatever else they may be, they aren't relevant to the asymmetry of causation.

7 PHYSICAL ASYMMETRY

In this last section I want to bring out some interesting analogies between my account of causal asymmetry and the explanation of such 'irreversible' physical processes as the increase of entropy and the expansion of radiation.

Let me begin with the familiar case of entropy increase in a closed system such as a collection of molecules in a box. One approach to explaining why such a system will normally move from a low entropy state to a higher one is in terms of Gibbs phase-space. For a closed system S of N molecules this is a $6N$ -dimensional space, in which each point represents a possible arrangement of positions and momenta for all the molecules in the system (3 position co-ordinates, and 3 momentum co-ordinates, for each molecule). Given this space, one can represent the *probability* that S is in a given microstate at a given time by a probabilistic distribution over the points of the space. Suppose S has total energy E . Since a closed system has constant energy, it is natural to assign probability zero to points representing microstates with total energies other than E . The function which distributes the probability equally over the remaining points is called the 'microcanonical distribution'. There are general reasons for taking the microcanonical distribution to be the appropriate distribution for equilibrium systems, for supposing that an equilibrium system has an equal chance of being in any of the microstates compatible with its total energy. (Cf. Malament and Zabell [1980], especially pp. 345–6.)

Correspondingly, if we have a non-equilibrium system, say one characterised by a certain degree of low entropy, then it seems natural to suppose that the probability distribution of its microstates should be given by a uniform distribution over all the points which are compatible with its total

energy *and* with that degree of low entropy. If we do suppose this, then there are plausible arguments for concluding that any system in a low entropy state will evolve into a state of higher entropy (see for instance Davies [1974] Sections 2.5 and 3.2, and Balescu [1975] pp. 695–729).

However, despite its power as a tool in mathematical physics, the Gibbs-space analysis needs to be approached with some caution by the philosopher. Thus note that the notion of ‘entropy increase’ used by the Gibbs approach is not the intuitive idea of an increase in internal disorder *within* a given system, but rather the idea of the initially ‘allowable’ microstates getting more disordered in *phase-space*. That is, ‘entropy’ is not defined directly in terms of the homogeneity of the positions and momenta of all the molecules inside the system, but rather in terms of the extent to which the initial ‘ensemble’ of phase-space points assigned to the system spread themselves out over the phase-space.¹ One striking result of this lack of direct interest in the internal molecular arrangements of thermodynamic systems is that on the Gibbs approach the ‘entropy’ of *any* system with initially low entropy will increase: even if we did have a freak low entropy system whose initial conditions were such as to make it even *less* disordered, the Gibbs analysis would still deem its ‘entropy’ to have increased, since the ‘entropy’ does not depend on the *actual* system’s microstate and subsequent internal evolution, but rather on the evolution in phase-space of the whole ‘ensemble’ of points compatible with the system’s initial *macrostate*. And a corollary is that the Gibbs approach gives us no insight into the mechanical processes by which particular ordered systems evolve into disordered ones—it tells us nothing about the kind of interactions which bring it about that the molecules within a particular system get ‘mixed up’.

There is an alternative approach to entropy increase, stemming from Boltzmann’s work in the second half of the nineteenth century. Boltzmann defined entropy in terms of the uniformity of the distribution of the *N* molecules in the system over a *six*-dimensional space of position and momentum co-ordinates. And Boltzmann then proceeded to show, by direct dynamical analysis of the possible collisions between the molecules, that a non-equilibrium starting point would always be followed by an entropy increase.

In fact Boltzmann originally proved rather too much. The simple ‘reversibility’ thought-experiment (imagine the velocities of all the molecules to be reversed) showed that there must be something wrong with the conclusion that entropy *always* increases. As a result of this objection (together with Poincaré’s ‘recurrence’ objection—see Davies, *op. cit.* pp. 56–7) it was recognised that Boltzmann’s analysis contained a dubious premise, the ‘Stosszahlansatz’, or, as it is now known, the ‘assumption of molecular chaos’. This states that in any spatial volume element of the gas

¹ I am referring here to the ‘coarse-grained’ Gibbs entropy, which can increase with time, rather than the ‘fine-grained’ entropy, which can’t. Cf. Davies, *op. cit.* section 2.5.

the number of pairs of molecules such that the first molecule has velocity v_1 and the second velocity v_2 will be proportional to the product of the total number of molecules in the gas with velocity v_1 by the total number of molecules with velocity v_2 —in short, that the number of molecules in a given volume with one velocity will be independent of the numbers of molecules with other velocities.

Since this ‘molecular chaos’ assumption does indeed imply that an entropy increase is inevitable, the ‘reversibility’ and ‘recurrence’ objections show it can’t be universally true. But a version of Boltzmann’s analysis will still go through if we read the ‘molecular chaos assumption’ not as an assumption about the actual, finite *number* of pairs of molecules with velocities $\langle v_1, v_2 \rangle$ in a particular volume of gas, but rather as an assumption about the *probability* of getting such a pair. The claim is then that the joint probability of getting such a pair will be the product of the separate probabilities of two single molecules having v_1 and v_2 respectively: that is, that the velocities of nearby molecules are *probabilistically* independent of each other. And the conclusion of the argument is then, not that entropy *always* increases, but that it very probably does so.¹

I have been dwelling on these details of the explanation of entropy increase, and in particular of the original Boltzmann approach, because of the close analogy they seem to me to offer to my explanation of causal asymmetry. Compare the molecular chaos assumption, the assumption that nearby molecules in a gas should have probabilistically independent velocities, with my appeal to the probabilistic independence of the background conditions in connection with which causes act. In both cases the way a given system evolves away from the ‘tail’ of the arrow of (physical, causal) asymmetry depends on the disposition of various unknown conditions. And in both cases it is precisely in so far as those background conditions satisfy certain probabilistic independence conditions that we get the characteristic development that makes us think of such processes as asymmetric ones.

Of course, this is at best only an analogy. It would be nice to show that causal and physical asymmetry are really both instances of the same phenomenon. But I don’t claim to have done this. Indeed there are obvious points at which the analogy runs into difficulties. What corresponds, in the

¹ It is worth being clear about the reasons for turning from the Gibbs approach to the Boltzmann analysis. This is not simply because the Gibbs approach uses probabilities. The Boltzmann strategy depends just as much on assumptions about probabilistic distributions over microstates (namely, such assumptions as that the probability of getting a microstate in which molecule x has v_1 is independent of the probability of getting a microstate in which molecule y has v_2). Nor is it because the Gibbs probabilities are somehow especially subjective. There are some delicate issues involved here, but there is no immediately obvious reason why Gibbs theorists shouldn’t take just the same attitude towards their probabilities as the Boltzmann theorists take towards theirs. The difficulty is not the Gibbs probabilities as such, but rather the way they get built into the Gibbs notion of *entropy*, and the consequent fact that the Gibbs approach gives an account of ‘entropy’ increase which fails to cast any light on the structure of the mechanical interactions involved.

analysis of entropy increase, to the positive correlation amongst the *different* effects of a common cause? The fact that the gas mixes up in all the different parts of the container? But it is not clear that this really fits the bill.

In some ways the asymmetry of the expansion of radiation (the fact that we don't find waves 'collapsing inwards' as the 'reversal' of normal expanding waves) seems a rather better analogy for causal asymmetry. This asymmetry too can be given an explanation in terms of probabilistic independence of background conditions (see Davies, *op. cit.* ch. 5, especially pp. 119–20). And in this case we would have the correlations between the different portions of the wavefront to correspond to the correlations between the different effects of a common cause. But there are points where this analogy breaks down too, and it would take me too far afield to pursue this line of thought here.

Let me conclude by noting one consequence which would follow from even a weak analogy between causal and physical asymmetry. In the first section I alluded to the plausibility of the view that the direction of time can be reduced to the direction of causation. What about the corresponding reduction, of the direction of time to the direction of entropy increase and the expansion of radiation? To most commentators (see in particular Mehlberg [1961] and Sklar [1974], ch. V) this latter reduction has seemed inherently implausible, because of the statistical, '*de facto*' basis of physical asymmetry. (Surely a mere change in the distribution of background conditions wouldn't change the direction of *time*?) But if I am right about my analysis of causal asymmetry, exactly the same objection could be made to the reduction of the direction of time to the direction of causation.¹ And so those who find this latter reduction plausible ought not to reject out of hand all attempts to reduce temporal asymmetry to a physical basis.²

University of Cambridge

REFERENCES

- BALESCU, RADU [1975]: *Equilibrium and Nonequilibrium Statistical Mechanics*. Wiley.
 DAVIES, P. C. W. [1974]: *The Physics of Time Asymmetry*. University of California Press.
 HEALEY, RICHARD A. [1983]: 'Temporal and Causal Asymmetry', in Swinburne, R. (ed.), *Space, Time and Causality*. Reidel.
 HUANG, KERSON [1963]: *Statistical Mechanics*. Wiley.
 LEWIS, DAVID [1973]: 'Causation', *Journal of Philosophy*, 70, pp. 556–7.
 LEWIS, DAVID [1979]: 'Counterfactual Dependence and Time's Arrow', *Nous*, 13, pp. 455–76.
 MACKIE, J. L. [1974]: *The Cement of the Universe*. Oxford University Press.
 MALAMENT, DAVID B. and ZABELL, SANDY L. [1980]: 'Why Gibbs Phase Averages Work—The Role of Ergodic Theory', *Philosophy of Science*, 47, pp. 339–49.

¹ Though it does seem to me that '*de facto*' is an inappropriate term for the statistical conditions behind either the causal or the physical asymmetry: without going into the metaphysics of probability, I take it that the probabilities in question are serious ones with counterfactual import, and not merely sampling artefacts.

² I would like to thank Jeremy Butterfield, Peter Clark, Richard Healey, and David Lewis for help with this paper.

- MEHLBERG, H. [1961]: 'Physical Laws and Time's Arrow', in Feigl, H. and Maxwell, G. (eds.), *Current Issues in the Philosophy of Science*. Holt, Rinehart and Winston.
- MELLOR, D. H. [1981]: *Real Time*. Cambridge University Press.
- MERMIN, N. D. [1983]: 'Pair Distributions and Conditional Independence: Some Hints about the Structure of Strange Quantum Correlations', *Philosophy of Science*, 50, pp. 359-73.
- PAPINEAU, DAVID [1985]: 'Probabilities and Causes', *Journal of Philosophy*, 82.
- SALMON, WESLEY *et al.* [1971]: *Statistical Explanation and Statistical Relevance*. University of Pittsburgh Press.
- SKLAR, L. [1974]: *Space, Time and Spacetime*. University of California Press.
- SUPPES, P. [1970]: *A Probabilistic Theory of Causality*. North-Holland.