

David Lewis's awkward cases of redundant causation

HUGH RICE

The main line of Lewis's account of causation is in terms of chains of counterfactual dependence. According to his original account (1973), a *causal chain* is a sequence of two or more events, with counterfactual dependence at each step; and one event is a *cause* of another iff there is a causal chain from one to the other. But some awkward cases involving redundant causation lead him to introduce the notion of quasi-dependence (1986: 205–7). Laurie Paul (1998) has suggested a way of dealing with one important class of these cases in terms of dependence proper. I shall suggest a different (and, I argue, preferable) way of dealing with this class of cases; I shall also suggest that it is possible to deal with other awkward cases too in terms of dependence proper.

1. Cases of late preemption

Let us consider first a type of case which poses a major problem for Lewis's original account – a major problem because it is not far fetched, and we have clear intuitions about what the verdict should be. (Lewis 1986: 203–5) We have redundant causes c_1 and c_2 of e : that is to say e would have happened if c_1 had not occurred, and would have happened if c_2 had not occurred, but would not have happened if neither had occurred.¹ The difficult case at issue is one where we have (apparently) a process which starts with c_1 (say) and leads to e , and the occurrence of e preempts the process which began with c_2 . (Lewis (1986: 200) describes preemption of this sort as *late preemption*.) We want to say that c_1 is a cause of e , but that c_2 is not. But the original account does not allow us to.

Paul (1998) gives the following example. Quickdraw McGraw and Slow Joe face off against Billy the Kid. Both draw their guns virtually simultaneously a few seconds before noon. But McGraw fires his gun (c_1) a little before Joe fires his (c_2), his bullet hits Billy first, and Billy dies instantly from the bullet wound at noon precisely. However, but for McGraw's act, Joe's bullet would have killed Billy at a second past noon. Assuming that the event of the death at noon is the same event (e) as the death a second later, c_1 and c_2 are redundant causes of e . It is clear that we want to say

¹ More generally we have a set of n redundant causes, c_1, \dots, c_n , of e iff e would not have occurred if none of the c_i s had occurred, but would have occurred if any of them had occurred.

that c_1 is a cause of e , but the original account does not allow us to. For, not only is e not counterfactually dependent on c_1 , there is also no causal chain connecting it with c_1 . Lewis says that he must have thought that the solution lay in the *fragility*² of the effect: if c_1 had not occurred, the death that occurred would have been different in a number of ways – in particular it would have been later. That is why c_1 is the cause of the death. The same cannot be said of c_2 . But then, he says, this must be wrong – because then we should have to regard *any* event which had some effect on the death as a cause of the death; because, just as in the case of c_1 , if such an event had not occurred, the death would have been different. So he proposes a solution in terms of quasi-dependence. Is there an alternative solution?

Paul herself offers an elegant alternative solution, which I shall discuss later. But first I shall offer a different solution.

It certainly will not do, for the reason Lewis gives, to say that, if something is a cause of the very fragile death, it is also a cause of the unfragile death. What I want to suggest is that the converse does hold. If something is a cause of the unfragile death it must also be a cause of the very fragile death. It must be a cause of the death with all its details – though it need not, of course, be a cause of the details. In fact I think that it must be true that, if c is a cause of e , it must also be a cause of (to adopt Lewis's analysis of events) every actual subset of e .

To return to the example. How does this idea enable us to say that c_1 is a cause of e ? Well, c_1 and c_2 are members of a pair of redundant causes. So, if we can rule out one as irrelevant to what actually happened, the other ought to be a cause. And we can. The very fragile death is not dependent on c_2 ; but it *is* on c_1 .

The following principle seems to be true:

If c is a redundant cause of e , c is a cause of e *simpliciter* only if e would not have occurred (more or less)³ just as it did if c had not occurred.

Can we substitute 'if and only if' for 'only if'? Unfortunately not. For consider a variation of Paul's example. Suppose that Joe fired a little before

² 'Call an event *fragile* if, or to the extent that, it could not have occurred at a different time, or in a different manner.' (Lewis 1986: 196)

³ I have added 'more or less' because otherwise perhaps *any* redundant cause c would satisfy this condition, on the ground (for instance) that the bare fact that e had not been preceded by c would constitute a difference in the way e occurred. This imports some regrettable vagueness. Though there is, of course, already a good deal of vagueness on Lewis's account over the question of whether event e_1 is the same event as event e_2 .

McGraw, and that Billy saw him fire and (having super fast reflexes) ducked. Suppose also that Joe's bullet travelled much more slowly than McGraw's, and that the outcome was that Billy was hit in the head by McGraw's bullet. If he had not ducked, he would have been hit in the heart. The ducking would, however, have been ineffective against Joe's bullet. If McGraw's bullet had not killed him, Joe's would have. Now it is clear that we want to say that c_1 (McGraw's firing his gun) is a cause *simpliciter* of e , but that c_2 is not. But, if we were to substitute 'if and only if' in our principle, we would have to say that both c_1 and c_2 are causes. But equally, if we leave the principle as it stands, we cannot say that c_1 is a cause on the grounds that c_2 is ruled out by the principle; because it is not. Suppose, though, that we look later in the course of events, after Billy has ducked – at t_2 , let us say. We now have a new pair of redundant causes: the position and behaviour of McGraw's bullet at t_2 (d_1 , say) and the position and behaviour of Joe's bullet at t_2 (d_2). Now the first of these meets the condition laid down by the principle; for, if d_1 had not occurred, the death would have been rather different. But the second does not; for, if d_2 had not occurred, the death would have been more or less as it was. The ban on backtracking forbids us to say that, if Joe's bullet had not been where it was at t_2 , it would not have been fired, and so the death would have been different. So it seems that we can rule out d_2 as a cause of e , and can say that d_1 is a cause. And so we can go on to say that c_1 is a cause of the death, because d_1 is dependent on it, and d_1 is a cause of the death. And we cannot say the same about c_2 .

There are two things to notice in this line of thought. The first is that we have again appealed to the idea that, if we have a pair of redundant causes of e , and one of them meets the necessary condition laid down by the principle while the second does not, then the first is a cause of e and the second is not. The second thing to notice is that we have appealed to Lewis's idea that a cause may be indirectly connected, by way of a causal chain, with an event of which it is a cause. Both of these ideas seem plausible. So, with them in mind, let us first give an account of what we will call *direct dependence*:

Event e is directly dependent on event c iff either (i) it is counterfactually dependent on c or (ii) c is a member of a set of redundant causes of e , and e would not have occurred (more or less) just as it did if c had not occurred, and would have occurred (more or less) as it did if the remaining members had not occurred.

Then, adapting Lewis, we can say that a causal chain is a sequence of two or more events, with direct dependence at each step; and one event is a cause of another iff there is a causal chain from one to the other.

Paul's account, as I have said, is different. She suggests that we amend the account (for deterministic causation) to read:

e depends causally on *c* iff *c* occurs, *e* occurs, and if *c* had not occurred, then *e* would not have occurred at all, or would have occurred later than the time that it actually did occur.

This is evidently a simpler account. Should it be preferred? I think that it should not for two reasons. The first is the reason that Paul herself gives, that on her account any event that hastens another event will count as one of its causes. Well, we could certainly rule out some of these cases by amending her account thus:

e depends causally on *c* iff *c* occurs, *e* occurs, and either (i) if *c* had not occurred, then *e* would not have occurred at all, or (ii) *c* is a member of a set of redundant causes of *e*, and, if *c* had not occurred, then *e* would have occurred later than the time that it actually did occur.

But even this, arguably, lets in too much. For suppose again that Joe fired earlier than McGraw, and that his bullet travelled more slowly. But suppose also that McGraw (blest also with super fast reflexes) was aware of Joe's firing and as a result (wishing to have the glory of killing Billy for himself) fired a little earlier than he would otherwise have done. Again it is his bullet that killed Billy. Now it seems that c_1 (McGraw's firing) was a cause of *e*, but that c_2 (Joe's firing) was not. But the amended version of Paul's account (like her unamended account) would mean that both were causes. To avoid having to say this, we could make a further amendment, so that the new account paralleled my own account:

e depends causally on *c* iff *c* occurs, *e* occurs, and either (i) if *c* had not occurred, then *e* would not have occurred at all, or (ii) *c* is a member of a set of redundant causes of *e*, and *e* would have occurred later than the time that it actually did occur if *c* had not occurred, but would not have occurred later than it did if the remaining members had not occurred.

And then we could establish (as we did earlier) that c_1 was a cause of *e* by finding a d_1 to link c_1 with *e* by a chain of causal dependence.

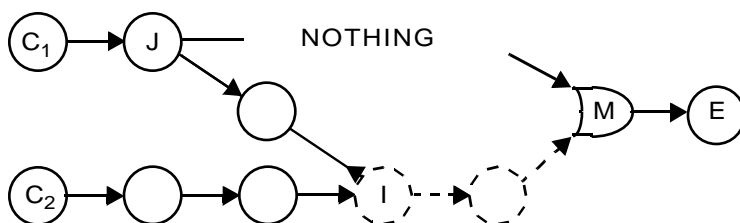
I think that this account would still be better than my account, if there were not another reason to prefer mine – better because the question of whether an event would have occurred later than it did is easier to settle than whether it would have occurred more or less as it did. The other reason is that, if we appeal only to time, it seems that we cannot deal satisfactorily with the following sort of case. Suppose that we have the firing of three neurons, C_1 , C_2 and E . E would not have fired if neither C_1 nor C_2 had fired. It would, however, have fired if just one of them had. However

let us further suppose that, if C_2 alone had fired, E 's firing would have been slightly weaker than it actually was, but not so much weaker as to constitute a different event; but that, if C_1 alone had fired, E would have fired exactly as strongly as it actually did. And let us suppose that there were no (relevant) intermediate events between the firing of C_1 and C_2 and the firing of E . Then, if the only detail of E 's firing that we take into account is its timing, we cannot discriminate between the firing of C_1 and the firing of C_2 .⁴ On my account, however, we can say that the firing of C_1 was a cause of the firing of E and that the firing of C_2 was not. And this seems the right thing to say.

If my account is right, we can deal with the relatively commonplace type of problem case without introducing quasi-dependence. But Lewis also describes two other sorts of case, which he describes as 'far-fetched'. Can we also deal with them?

2. *Far-fetched cases*

The first far-fetched case involves action at temporal distance. Lewis (1986: 202) gives an example involving the firing of neurons:



The ordinary arrows represent the stimulation of the neuron at the arrow head by the firing of the neuron at the tail; and the arrows with reversed heads represent the inhibition of the neuron at the reversed arrow head. In the diagram there is a process beginning with J which inhibits I , and J also stimulates M – but at a temporal distance; there is no process in between. We want to say that the firing of C_1 is a cause of the firing of M , but that the firing of C_2 is not. But, does the account I am proposing allow us to?

It seems at first sight that it does not. If we assume, as we may for the sake of the example, that M 's firing would have been much as it was if C_1

⁴ The account in Ganeri, Noordhof and Ramachandran 1998 will not discriminate between them either, since there is no event missing from the process by which C_2 's firing would have caused E 's firing if C_1 had not fired.

had not fired, and also much as it was if C_2 had not fired, there seems to be no way of arguing that C_1 's firing, unlike C_2 's firing, is connected to M's firing by a chain of direct dependence.

But perhaps appearances are deceptive. I want to suggest that we can in fact connect the firing of M to the firing of C_1 by a chain of dependence – a chain of counterfactual dependence, in fact. Suppose that one takes a time, t_2 , between the firing of J and the firing of M. It will be true, of course, at that time that C_1 fired when it did – let us say at t_0 . It will also be true that, if C_1 had not fired at t_0 , it would not have been true at t_2 that C_1 had fired at t_0 . But the crucial fact is that, if at t_2 it had not been true that C_1 had fired, then M would not have fired; because by t_2 J had already fired and the chain from C_2 to M was already doomed. What makes this dependence possible is again the ban on backtracking. When considering what would have happened, if at t_2 it had not been true that C_1 had fired, we are not allowed to say that in that case J would not have fired, and so I would have fired, and so M would have fired. In fact the gap between the firing of J and the firing of M will be densely packed with what one might at a pinch call events of its being true at the time that C_1 had fired; and there will be *counterfactual* dependence linking these events. One might protest that these 'events' are pretty odd events. But then action at a temporal distance is pretty odd. One might protest that the application of the ban on backtracking to such odd events is a little odd. But no odder than the events themselves. It is also worth noticing that in some quite ordinary cases we distinguish between 'if e had happened' and 'if it had been true that e happened'. Suppose, for instance, that John tells me that Oswald did not kill Kennedy. I might reasonably reply, 'If it had been true that Oswald did not kill him, it would have been true that someone else did.' This would be quite different from my saying, 'If Oswald had not killed him it would have been true that someone else did.' And the difference seems to stem from the fact that in assessing the first comment we are not allowed to backtrack to a time before John's utterance. Finally, one might, I suppose, protest that, if things are as I say, we do not after all have a case of causation at a temporal distance, as was supposed to be true in the example. The answer to this is that we have all we could possibly expect. Any world which contains both the firing of C_1 at t_0 and the time t_2 , is bound to contain also the 'event' of its being true at t_2 that C_1 fired at t_0 .⁵ So the solution to the

⁵ Perhaps, though conceding that these 'events' can enter into chains of counterfactual dependence, one might deny that they could enter into causal relations. This raises the general question of whether counterfactual dependence is enough for causal dependence. But that question is beyond the scope of this paper.

problem posed by the example is that the firing of M is linked by a chain of dependence with the firing of C_1 . So the firing of C_1 is a cause of the firing of M . And the firing of C_2 is not.

Lewis's second far-fetched case (1986: 202–3) involves infinite multiple preemption. We have a main process starting, let us say, with c_1 and culminating in e . We also have an infinite number of alternative processes starting with, say, c_2 , c_3 , c_4 , and so on. The process beginning with c_2 is doomed 128 seconds before e , that beginning with c_3 is doomed 64 seconds before e , that beginning with c_4 is doomed 32 seconds before, and so on. So, however close we get to e , although some of the alternative processes will be doomed, there will be others which are not yet doomed. So, however close we get to e , we cannot find an event upon which e is dependent, and which itself is (stepwise) dependent on c_1 . The solution to this problem, I suggest, is not to look for an event which establishes c_1 as indeed a cause of e . That is doomed to failure. Rather we should look for an argument for saying that all the competing redundant causes which initiate the alternative processes can be ruled out, while c_1 remains. If that can be argued, then we can say that c_1 is a cause, but the other c_i s are not. But it can. 128 seconds before e , the process beginning with c_2 is doomed. So there will be an event then of which c_1 is a cause, and which is in its turn a redundant cause of e . The same will be true of c_3 and each c_i ($i > 2$). But it will *not* be true of c_2 . And that seems enough to rule out c_2 as a cause of e . Similarly, if we consider what happens 64 seconds before e we can rule out c_3 , and so on. In short *each* the c_i s ($i > 1$) will be ruled out at some time before e . So c_1 triumphs. We can put it like this. However far one goes towards e , there will be an event of which c_1 is a cause, and which belongs to a set of redundant causes of e ; and this is true of none of the other c_i s. So c_1 is a cause of e . Alternatively we could think of things like this. The mereological sum of the c_i s is a cause of e . (Lewis 1986: 212) Call this sum C . Now consider what happens 128 seconds before e . There will be an event, d_1 , belonging to the process beginning with c_1 , which is dependent on c_1 , and is a member of a set of redundant causes of e . Also there will be an event, d_3 , belonging to the process beginning with c_3 , which is dependent on c_3 , and is a member of a set of redundant causes of e . And so on for all the c_i s apart from c_2 . Call the mereological sum of these d_i s D . Call the mereological sum of the c_i s apart from c_2 $C-c_2$. Then $C-c_2$ is a cause of D , which is in its turn a cause of e . So $C-c_2$ is also a cause of e . By the same token, if we consider what happens 64 seconds before e , there will be an event which is a cause of e and of which the mereological sum of the c_i s apart from c_2 and c_3 is a cause. So *that* mereological sum, $C-c_2-c_3$, is a cause of e . And in general, in the case of each of the c_i s ($i > 1$),

$C-c_2-c_3, \dots, -c_i$ is a cause of e . But, since this is true of each of the c_i s ($i > 1$), the mereological sum of c_1 and nothing else is a cause of e . That is to say, c_1 is a cause of e .⁶

Christ Church, Oxford OX1 1DP, UK
hugh.rice@chch.oxford.ac.uk

References

- Ganeri, J., P. Noordhof and M. Ramachandran. 1998. For a (revised) PCA-analysis. *Analysis* 58: 45–47.
- Lewis, D. 1973. Causation. *Journal of Philosophy* 70: 556–67. Reprinted in Lewis 1986: 159–72.
- Lewis, D. 1986. *Philosophical Papers*, vol. II. New York: Oxford University Press.
- Paul, L. A. 1998. Keeping track of time: emending the counterfactual analysis of causation. *Analysis* 58: 191–98.

⁶ An ancestor of this paper was read at a meeting of the Philosophical Society at Oxford in 1984 as a reply to an ancestor of Lewis's discussion of redundant causation in Lewis (1986: 193–212). I am grateful to David Lewis for his comments.