Disturbances in the ether

by Tony Rex Smith

A thesis presented for the degree of Master of Arts

Department of Philosophy
Victoria University of Wellington
1999

Abstract

We criticise the current philosophical practice of invoking causation as a solution to various problems in various fields of philosophy. Our specific concern is that many of these solutions to problems rely on the intuition that causation is "the cement of the universe". We question whether several different analyses of causation which are supposed to substantiate this intuition (or at least are treated as if they substantiate this intuition) in fact substantiate this intuition.

We begin by establishing a basic desideratum for such an analysis of causation — that causal dependence ought to track physical dependence in this universe. We investigate in turn a Lewis-style counterfactual analysis of causation, the transference analysis developed by Aronson, Fair and Heathcote, and the process analyses developed by Salmon and Dowe. Rather to our surprise, none of the analyses fulfil our basic desideratum. Although this is not in itself conclusive grounds for scepticism about causation, our results speak against casually invoking analyses of causation in order to solve particular varieties of philosophical problems.

Contents

Acknowledgments

Academically, many acknowledgments are in order. Ken Perszyk for his course on time and space that showed me that I was going to have to say something about causation at some stage. Laurie Paul for a seminar on causal transitivity that taught me that I was a naive transference analyst, and that there were other alternatives that had to be considered. Ed Mares and Kim Sterelny for a tea room conversation on pre-Galilean physics, which led to the counterfactuals chapter (whether they realise it or not). Staff and students at Victoria University of Wellington who attended my 1998 seminar "Bonking Frank", an early version of the counterfactuals chapter (particularly Max Cresswell, for making me realise that causation needn't be the cement of the universe) and the audience for the performance at the 1998 University of Canterbury Graduate Student Conference. Helen Beebee at ANU for extensive comments on "Bonking Frank". Ed, again, for introducing to me several trees worth of causation articles, lending me *Causality and explanation* (which let me rationalise a multitude of photocopies into a single reference) and for some dogged supervision. My students from PHIL 403 in 1999 for both their well-judged scepticism and their positive contributions. Debbie Luyinda for the keys to the cupboards at important moments. Adele Donachie at the FHSS for help with increasingly odd requests over an extended length of time. Alison Munro at FHSS for helping arrange the money for the UC performance of "Bonking Frank". Maureen Penning at the Scholarships Office for support and smiles.

My wife Helen McCormick, for her forbearance and support again. (Sorry hon, but at least it wasn't dead German philosophers this time. Well, not too many of them.) Our cats Abu and Yasser. (Abu sadly didn't see the thesis handed in. We miss you.) All the friends who I've neglected so disgracefully over the past year or so, but who still seemed pleased to see me by accident. The umbra and the penumbra at the department and staff club on Thursday nights. Nick Agar who taught me bridge and kept me sane when Helen couldn't be in three places at once.

Guided by voices, Everclear, Mogwai, Slint and especially Pavement:

I need to sleep/Why don't you let me/I need to sleep/Why don't you

Introduction

> There are two ways of dealing with a difficulty — the metaphysical and the scientific way. The first is very simple and expeditious — it consists merely in giving the unknown a name whereby it may be classified and categorised. Thenceforward the unknown is regarded as having become part of knowledge. The scientific man goes further, and endeavours to find what lies concealed under the name. If it were possible for a metaphysician to be a golfer, he might perhaps occasionally notice that his ball, instead of moving forward in a vertical plane (like the generality of projectiles, such as brickbats and cricket balls), skewed away gradually to the right. If he did notice it, his methods would naturally lead him to content himself with his caddie's remark — "ye heeled that yin", or "Ye jist slicet it" (we here suppose the metaphysician to be right-handed ...). But a scientific man is not to be put off with such flimsy verbiage as this. He *must* know more. What is "heeling", what is "slicing," and why could either operation (if it could be thoroughly carried out) send a ball as if to cover-point, thence to long slip, and finally behind back-stop? These, as Falstaff said, are "questions to be asked."[1]

Such questions are also to be asked about causation. The notion of causation plays a central role in many areas of contemporary analytic philosophy. In the philosophy of language we find the causal theory of language. In metaphysics we find causal theories of personal identity. In the philosophy of science we find causal solutions to the problem of induction, causal decision theories, and causal theories of scientific explanation. There is something distinctly unsatisfactory about leaving causation as an unanalysed primitive if it is doing such important work for us.

Indeed, such a policy might even be extremely dangerous. What do we mean by "causation" when we use the term as an unanalysed primitive? Presumably, we use our common-sense understanding of the term as it is used in the course everyday talk — although there is a certain tension in claiming that a term is unanalysed when we apparently have some understanding of what it means. But as noted by Hilary Putnam, "causal talk arises from a variety of 'vulgar'

1 Tait, P G: "The unwritten chapter on golf", pp80–1, Gratzer, Walter, editor: *A bedside* Nature (London: Macmillan Magazines Ltd, 1996)

ways of speaking and has a variety of uses".[2] So if causation is to be treated as an unanalysed primitive, then philosophical positions such as the claim that words "hook on to the world" by virtue of causal chains linking utterances to their references may well be massively ambiguous and relentlessly *ad hoc*.

Such worries, or at least the spirit of such worries, underlie the perennial claim that talk of causation *belongs* in the philosophical wilderness, a view famously proposed by Bertrand Russell, and in more recent times by Willard Van Orman Quine. Talk of causation is a vestige of our pretheoretic understanding of the world, made redundant or about to be made redundant by our more advanced scientific understanding of the world, and the sooner we stop jabbering on about causes, the better off philosophy will be.

We have some initial sympathy for this position, which we call the Russell-Quine thesis. Yet it seems a little hasty. Even if our vulgar, pretheoretic and unanalysed primitive notion of cause needs to be abandoned, there are a lot of philosophical analyses of causation about. Perhaps one of these analyses would serve to provide a respectable foundation for various philosophical theories in the philosophy of science, philosophy of language and metaphysics. Our broad concern is whether or not several contemporary analyses of causation are up to the job.

This concern is probably *too* broad for practical purposes. The extensive variety of uses of causation in vulgar talk is reflected in the extensive variety of philosophical analyses of causation — the various different analyses seeking to capture and consolidate different sets of important pretheoretic intuitions. This diversity is also driven by the wide variety of philosophical problems for which causation is proposed as a therapy. The perceived importance of particular causal intuitions is largely a function of what needs to be established in a particular philosophical context. As often as not, we find perceived theoretical desiderata driving the development of the supporting analysis of causation, rather than due consideration of a particular analysis of causation dictating which theoretical positions should be taken. So no particular analysis of causation is likely to be the panacea for all philosophical ills.

---

2 Putnam, Hilary: "Is water necessarily H_2O?", pp54–79, Putnam, Hilary: *Realism with a human face* (Cambridge, MA: Harvard University Press, 1990)

So be it: but if this is the game, we insist that the players follow the rules. We shall pick a particular desideratum that an analysis of causation ought to fulfil in order to support some philosophical contentions. We shall examine several analyses of causation to see if they fulfil this desideratum. If they do not do so (and we shall argue that this is the case), then they *should not* be invoked in contexts where this desideratum needs to be fulfilled.

Our desideratum seems innocent enough at first glance. We are looking for an analysis of causation that gives a satisfactory account of physical dependence between things in the world — the vulgar intuition that we are trying to analyse is that causation is "the cement of the universe". Such a desideratum requires a few presuppositions: first, that there is a physical world with physical things in it, and second that the states of these physical things physically depend on the states of other physical things. We won't defend the first presupposition, but some explanation of what is meant by "physical dependence" is in order.

We have both a pretheoretic, common sense understanding and a more advanced theoretical understanding of the notion of physical dependence. The coffee cups on my desk do not fall through the desk and shatter on the floor. Why? At a fairly pretheoretic level of discourse we might say the desk is a strong solid object and the coffee cups are not particularly heavy, so the desk holds the cups up. At a more advanced theoretic level, we note that the cups and the desk are mostly composed of fermions (subatomic particles such as protons, neutrons and electrons) bonded together by various fundamental interactions. The cup cannot fall through the desk, because the Pauli exclusion principle dictates that no two fermions with the same spin state can occupy the same spatiotemporal position. Hence the predominant interaction between the cups and the desk is the repulsive component of the electromagnetic interaction, which provides a normal force balancing the force of the gravitational interaction between the cups and the Earth. So the table holds the cups up, and the cups do not float up into the air as a consequence of the interaction with the table. In either case, we are asserting that a particular set of physical circumstances (the state of the cups) is *thus*, because other sets of physical circumstances (the states of the desk and the Earth) are *so*. We will insist that an analysis of causation is consistent with our best possible theoretical understanding of physical dependence.

We do not require that all physical dependence reduces to causal dependence in a particular set of circumstances, or *vice versa*. It is not clear that all physical dependence *need* resolve to causal dependence. (We suspect that many forms of physical dependence are more comfortably subsumed to — or comprise — nomic dependence, but discussing the point is beyond the scope of this thesis.) But we will require that when there is no relevant physical distinction between sets of physical circumstances, and there is physical dependence in both cases, then the analysis of causation should either hold that there is causal dependence in both cases or there is no causal dependence in both cases. Our shorthand expression for this relation is that causal dependence should *track* physical dependence.

We do not require that causal dependence tracks non-physical dependence between entities (such as logical entailment) nor do we require that the analysis tracks dependence between non-physical entities (whatever they might happen to be). Nor do we require that the analysis of causation gives a satisfactory account of causation in *every possible world*. This one will do nicely. Moreover, we will not treat of issues of causal transitivity, nor causation by omission.

Although few analyses of causation have been produced with the deliberate intention of fulfilling our basic desideratum, such a desideratum *ought* to be fulfilled by analyses of causation underlying, for example, a causal theory of scientific explanation. At least one of the roles of scientific explanation is to explain physical dependence in this in this world. If all good explanations are supposed to be causal explanations, then it better be the case that causal dependence tracks physical dependence. Otherwise, the theory would not be a theory of *scientific* explanation. Indeed, most of the analyses of causation we shall discuss have been developed to support a theory of scientific explanation, or are at least invoked in the service of theories of explanation. Similarly, the causal chains linking utterances of terms and their referents required by a causal theory of reference also appear to be — or at least track — relations of physical dependence between physical things (physical objects in the world, brain states, utterance tokens and so forth).

We shall investigate three analyses of causation in detail: a basic version of David Lewis's counterfactual analysis, the transference analysis developed by Jerrold Aronson, David Fair and Adrian Heathcote, and three process analyses of causation developed by Wesley Salmon and Phil Dowe. Our conclusion is depressingly negative: none of these analyses fulfil our basic desideratum. Although this result does not constitute a proof of the Russell-Quine thesis, it must undermine the casual practise of invoking causation as a solution to a philosophical problems when what is required in order to solve those problems is an account of physical dependence.

Counterfactuals

Why investigate the counterfactual analysis?

The prevailing orthodoxy among analyses of causation is probably still the counterfactual analysis initially presented by David Lewis in the 1973 paper "Causation".[3] Part of the attraction of the analysis is the deceptive ease with which a basic formulation of the analysis can be expressed. Consider, for example, the commonly-cited formulation by Jaegwon Kim:

(1) An event *e* causally depends on an event *c* just in case if *c* had not occurred *e* would not have occurred

(2) An event *c* is a cause of an event *e* just in case there is a chain of events from *c* to *e*, each event in this chain being causally dependent on its predecessor[4]

Actually, this is too simple an expression. The case that "just in case" refers to is the closest possible world in which *c* does not occur. The closeness of possible worlds is cashed out by a relation of comparative similarity across possible worlds — the more similar two possible worlds are, the closer they are. Events are to be taken as properties of spatiotemporal regions of possible worlds. Even so, possible worlds, similarity, properties, and spatiotemporal regions are all comfortable furniture of philosophical discourse, so the more complicated expression of the basic analysis is still readily understandable.

So do Lewis-style counterfactual analyses fulfil our basic desideratum? Arguably, there is no point in even asking the question.[5] Lewis's basic desideratum is to produce an analysis which is consistent with both our advanced theoretical understanding of the way the world is and our common

3 Originally published as Lewis, David: "Causation", pp556–67, *Journal of philosophy* vol 70 (1973), reprinted with extensive postscripts in pp159–213, Lewis, David: *Philosophical papers volume II* (Oxford: Oxford University Press, 1986). All page references are to the latter source.

4 Kim, Jaegwon: "Causes and counterfactuals", pp205–7, Sosa, Ernest; Tooley, Michael; editors: *Causation* (Oxford: Oxford University Press, 1993)

5 Edwin Mares (incessant personal communications)

sense understanding of the notions of causation and events. Yet common sense, useful though it may be, is ever likely to be in conflict with our advanced theoretical understanding. If the counterfactual analysis truly respects common sense, then it could hardly be expected to fulfil our desideratum.

It is not even clear that the counterfactual analysis gives an account of causal dependence between physical things in the sense that causation is the cement of the universe. Although Lewis appears to be asserting that causal dependence is a relation between events (which we will allow to be some species of physical thing for the purposes of the present discussion), not all proponents of Lewis-style counterfactual analyses require this. For example, Philip Kitcher has championed the use of a Lewis-style counterfactual analysis of causation in the context of a theory of scientific explanation, but Kitcher suggests that the assignment of causal dependence is driven by explanatory dependence.[6] Counterfactual dependence is properly a relation between explanations of events (or at least statements about events) rather than events themselves.

Similarly, Lewis might deny that there is any such thing as "physical dependence". The Lewis-style counterfactual analysis is intended to be consistent with "Humean supervenience":

> It is the doctrine that all there is to the world is a vast mosaic of local matters of particular fact, just one little thing and then another … We have geometry: a system of external relations of spatiotemporal distance between points. Maybe points of spacetime itself, maybe point-sized bits of matter of aether or fields, maybe both. And at those points we have local qualities: perfectly natural intrinsic properties which need nothing bigger than a point at which to be instantiated. For short: we have an arrangement of qualities. And that is all. There is no difference without difference in the arrangement of qualities. All else supervenes on that.[7]

Under Humean supervenience it simply is not the case that the cups are thus because the desk and the Earth are so. The cup being on the table is just the cup being thus and the desk being thus. Why should causal dependence track physical dependence if there ain't no such thing?

---

6 See Kitcher, Philip: "Explanatory unification and the causal structure of the world', pp410–505, Kitcher, Philip; Salmon, Wesley; editors: *Minnesota studies in the philosophy of science vol 13: Scientific explanation* (Minneapolis: University of Minnesota Press, 1989)

7 *Philosophical papers volume II*, ppix–x

Well, fair enough. We have no quarrel with a Lewis-style counterfactual analysis in which causation is explicitly not intended to be fully consistent with our physical understanding of the world, or in which causal dependence is a relation between explanations, or in which causation is nothing to do with physical dependence because there is no such thing as physical dependence. (Although we really aren't sure what such an analysis could be useful *for*.) What we are concerned about is the common tendency to use Lewis-style analyses *as if* they are totally consistent with our best theoretical understanding of the world, *as if* causal dependence is the cement of the universe, and *as if* the analyses give an account of physical dependence. The use of a Lewis-style analysis in this way would only be warranted if counterfactual dependence tracked physical dependence. So there is some point to investigating whether such a counterfactual analysis fulfils our basic desideratum.

We will carry out this investigation by examining very simple examples of physical dependence — impacts. Any analysis of causation which fulfils our basic desideratum ought to be able to deal with collisions. We shall demonstrate that some of the classical failures of the basic Lewis-style counterfactual analysis to assign causal relations in these cases arise because counterfactual dependence does not track physical dependence. Furthermore, attempting to improve the Lewis-style analysis by trying to make causal relations more sensitive to physical dependence produces an analysis which is *less* successful in assigning causal relations in cases of impacts.

Overdetermination problems

We shall investigate instances of physical dependence associated with putative instances of causation in which cause and effect are directly rather than transitively linked — cases for which the first of Kim's characterisations of the Lewis-style analysis should hold:

> (1) An event *e* causally depends on an event *c* just in case if *c* had not occurred *e* would not have occurred

Such basic Lewis-style analyses are subject to problems of overdetermination. Consider the case of an ordinary glass window struck by a rock plenty large enough to break the window (we will call a rock this size a *standard* rock). The window breaks. The counterfactual analysis justifies our claim that the rock caused the window to break — *ceteris paribus* if the rock had not struck the window the window would not have broken.

But consider the situation where a shower of standard rocks strikes the window. No particular rock would then seem to have caused the window to break, because if that rock had not struck the window *ceteris paribus* the window would still have been broken by another rock — the effect would still have occurred. No subset of the rocks smaller than the entire shower of rocks would have caused the window to break, because if those rocks had not struck the window, any of the remaining rocks would still have broken the window — the effect would still have occurred. We can't even say that the entire shower of rocks caused the window to break, because in the counterfactual instance where *ceteris paribus* one less than the entire shower of rocks strike the window, the window is still broken. So rocks have hurtled through the window, there is glass all over the show, but we cannot say that a rock causes the window to break.

The failure of the analysis is ontological rather than epistemological. An epistemological failure would be that in the single rock case we *can* tell that a rock *caused* the window to break and in the overdetermined case we *can't* tell that a rock *caused* the window to break, even though the window breakage clearly physically depends on the impact of rocks in both cases. Rather, in both cases our commonsense knowledge claim is that "rocks caused the window break" but in the overdetermined case the counterfactual analysis cannot supply a truthmaking state of affairs (an appropriate counterfactual circumstance) for the knowledge claim.

If you feel there has been some illegitimate slide between "a rock" causing the window to break and "rocks" causing the window to break, the example can be adjusted to recreate the problem without such ambiguities. Consider a case where two *small* rocks strike a window. Neither rock can break the window individually — jointly they can break the window. Hence under the counterfactual analysis we can justifiably claim that "rocks cause the window to break". But when two standard rocks strike the window we have a case of overdetermination; we no longer have counterfactual support for the claim that "rocks cause the window to break".

We can also express the problem unambiguously in terms of "a rock". In the case where a standard rock strikes the window we can justifiably claim that "a rock caused the window to break". Yet in the case of a shower of rocks hitting the window in a slightly staggered formation — each rock strikes the window or passes through where the window would have been a fractionally different times — we still wish to claim that "a rock broke the window" yet we no longer have counterfactual support for the knowledge claim. This latter case of overdetermination is usually referred to as an instance of preemption — to be precise, a case of late preemption.[8]

Prima facie, in these cases the Lewis-style counterfactual analysis fails to assign relations of causal dependence in cases of physical dependence which seem extremely similar to cases of physical dependence in which relations of causal dependence *are* assigned. So do these various overdetermination problems conclusively demonstrate that causal dependence does not track physical dependence in a Lewis-style counterfactual analysis?

Well, perhaps not quite yet. To begin with, a philosophical "heavy" industry has grown up around devising these problems and defending and amending the counterfactual analysis in response to these problems.[9] (An unfortunate consequence of this feverish activity is that virtually every different overdetermination problem has a different proposed solution, which converts an attractively simple analysis into a ferociously complex and suspiciously *ad hoc* theoretical apparatus.) Some of these amendments would involve explicitly or implicitly denying that causal dependence tracks causal dependence. We have no problem with this, so long as the resulting counterfactual analysis was not used as if it was a satisfactory account of physical dependence.

Producing an amendment or defence of a Lewis-style analysis which *is* intended to give a satisfactory account of physical dependence seems much

---

8 A shower of rocks breaking a window is a case of *late* preemption, since the effect of the window being broken by *later* rocks in the shower is blocked by the final effect of the window being broken already taking place. *Early* preemption would involve an effect being blocked before the actual effect came about. For example, a standard rock is about to break a window, but it is knocked away by another standard rock which then goes on to break the window itself. See pp27–8, Ehring, Douglas: *Causation and persistence: a theory of causation* (New York: Oxford University Press, 1997) for a good description of Lewis's taxonomy of preemption.

9 A comprehensive bibliography of the literature on this topic would probably be longer than this thesis.

more problematic. *Prima facie* the overdetermination problems *do* show there is a tracking problem.

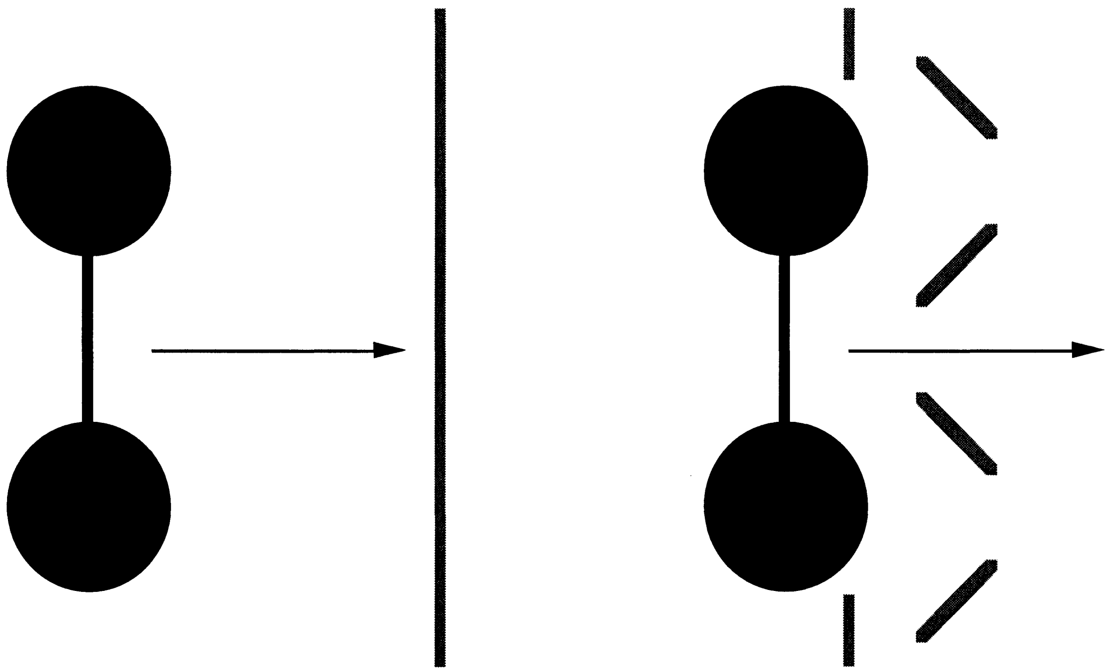
Presumably, the justification for such a defence programme would be that the *basic* Lewis-style analysis *does* track physical dependence, but it is only sensitive to a particular class of relevant physical considerations. The pairs of cases which appear to involve failure of tracking due to overdetermination problems involve differences in relevant physical circumstances between the cases. One standard rock; two standard rocks at once. Two small rocks; two standard rocks. One standard rock; several standard rocks one after another. These problems arise because one of the cases in each pair involve extra physical considerations which are not attended to by the basic analysis. The various proposed amendments to the basic Lewis-style analysis make the analysis sensitive to these extra physical considerations on a class-of-case by class-of-case basis.

But we shall present two problem cases which show that this programme is just wrongheaded. The case of the overdetermining dumbbell shows that the analysis can supply a truthmaker for causal claims about one physical system but not for another even though there are *no* relevant physical differences between the systems. The counterfactual analysis is sensitive to varying "common sense" assignments of object status within events, not to relevant differences of physical dependence in the systems. We can try and avoid this problem by keying assignment of object status to relevant physical differences, but then the problem of single-object overdetermination arises. We find the amended Lewis-style analysis suddenly fails to provide truthmakers to many more apparently straightforward causal claims than the basic analysis. In other words, trying to make the Lewis-style analysis *more* sensitive to physical dependence produces an analysis of causation which is *even less effective* at fulfilling our basic desideratum than the basic analysis. We conclude that some other analysis of causation is required.

The overdetermining dumbbell

Overdetermination problems can arise in Lewis-style analyses purely as a result of our common-sense individuation of objects within events. A *gendankenexperiment* can bring these object-oriented problems into sharp relief.

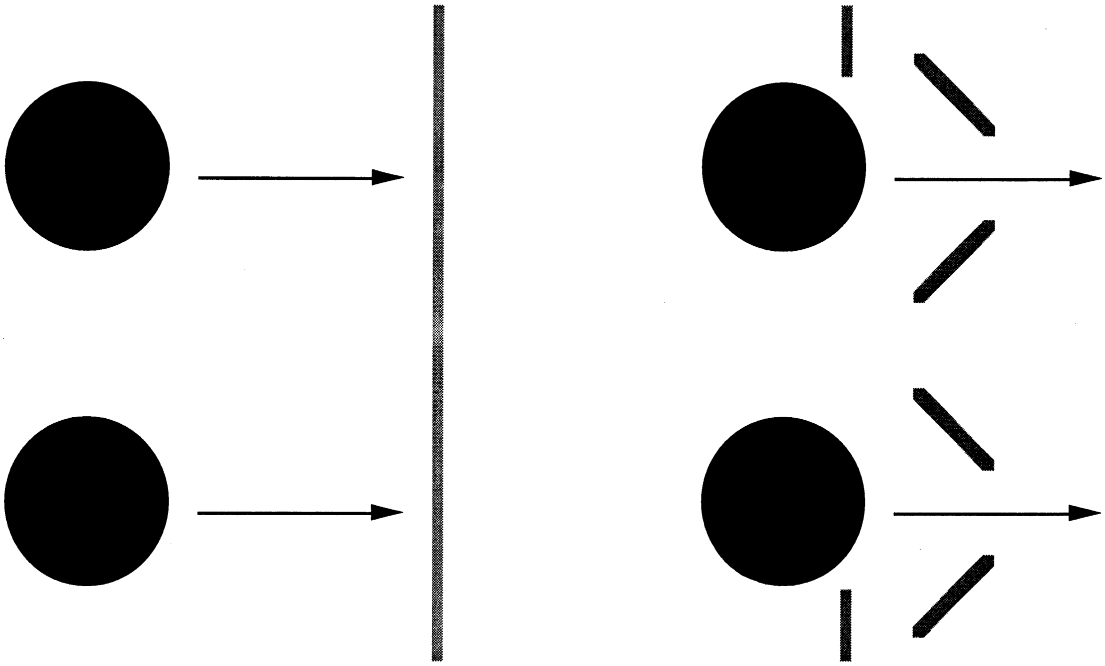
Consider a *standard dumbbell.* A standard dumbbell is a physical object made up of two standard rocks connected by a bar. The bar is solid but its mass is negligible by comparison to the two rocks. Consider a case where the dumbbell is hurled at a standard window. The trajectory of the dumbbell is perpendicular to the surface of the window. Both rocks strike the window simultaneously and the bar is perfectly parallel to the surface of the window at the moment of impact. A standard window has a nice homogenous structure, so any forces perpendicular to the path of the dumbbell generated by the impact with the window are balanced: thus the dumbbell is not deflected to the side at all by the impact with the window and the tension on the bar neither increases nor decreases as a result of the impact. (This can be approximated physically by making the standard window relatively weak or the standard rocks very large or heavy, so any perpendicular perturbing forces are negligible.) Under these circumstances, a basic Lewis-style analysis can justify the causal claim that "the dumbbell caused the window to break".



A standard dumbell strikes a standard window.
According to the counterfactual analysis, the dumbbell causes the window to break.

Figure 1 — an unproblematic example of causation for a Lewis-style analysis

Now consider the case where two standard rocks precisely equivalent to the weights on the dumbbell strike the window simultaneously. Under the basic Lewis-style analysis this is a case of overdetermination, and we can't say that any rock or collection of rocks caused the window to break.

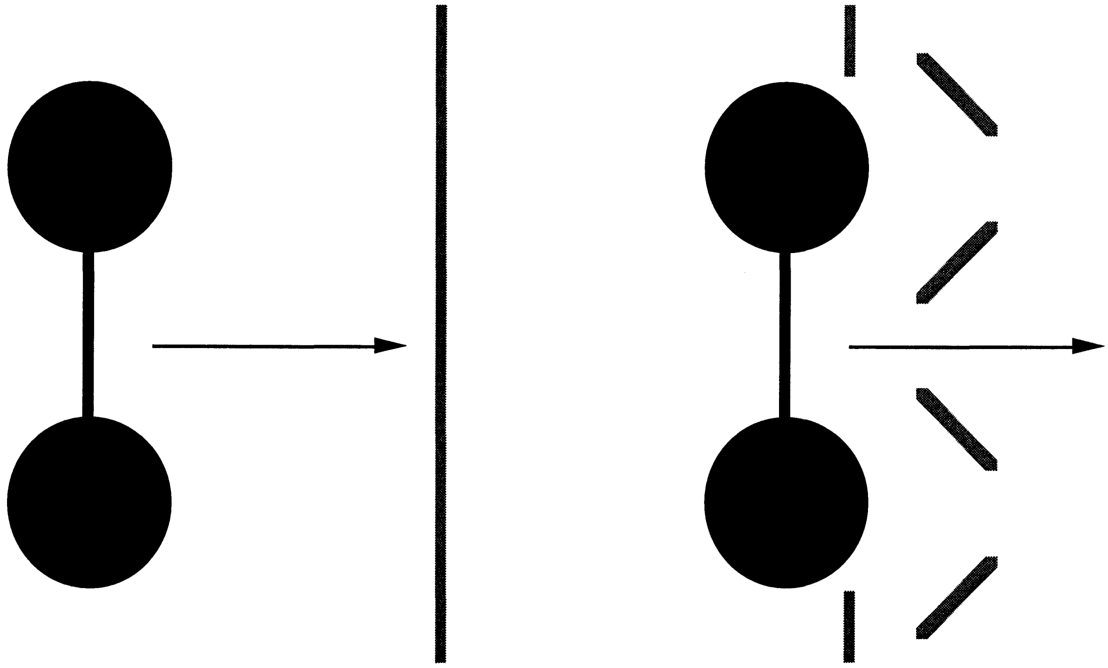


Two standard rocks strike a standard window, producing an instance of overdetermination. Neither rock can be said to break the window under the counterfactual analysis.

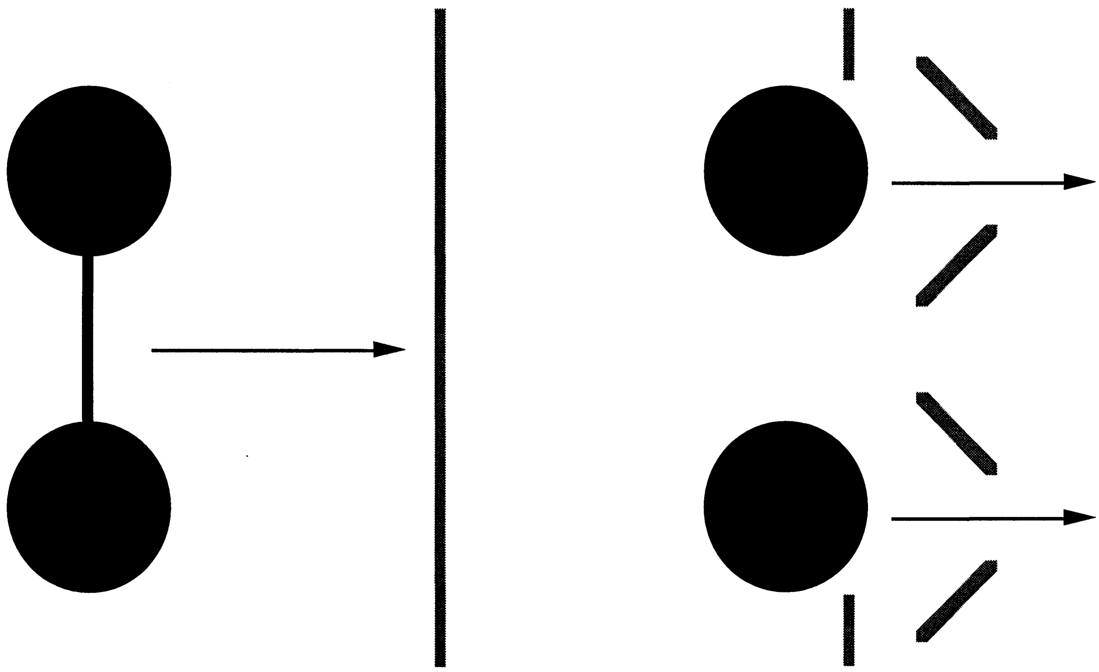
Figure 2 — an instance of overdetermination

Now consider the case of the *overdetermining* dumbbell. An overdetermining dumbbell has the same overall shape, mass and size as a standard dumbbell, but the bar contains a quick-release mechanism — it breaks in half in the middle and each half retracts into a rock. An ingenious recoil-damping system inside each rock ensures that the rocks do not move together or apart when the quick-release is triggered (or at least renders the effect negligible). We will assume energy loss in the system due to friction is also negligible.

Hurl the overdetermining dumbbell at the window, and let it strike the window with the bar still in place. As in the case of the standard dumbbell, the window is broken, and we can assign a cause to the effect of the window being broken. Now hurl the dumbbell at the window, but disengage the bar before the window is struck. The window is broken. Can we assign a cause to the breakage? The answer seems to be *no* — two standard rocks are striking the window, which is an instance of overdetermination.



Connected dumbell strikes the window: an instance of causation.



Disconnected dumbell strikes the window: an instance of overdetermination.

Figure 3 — the paradox of the overdetermining dumbbell

This is very odd indeed. The disconnected parts of the overdetermining dumbbell have the same mass as the connected dumbbell. They travel in the same trajectory as the connected dumbbell both before and after striking the window. The window is broken in precisely the same way. Yet we can no longer assign a cause to the effect of the window being broken. There is no relevant *physical* difference between the two systems. The only distinction between the connected and disconnected cases appears to be that in the connected case we normally say we are dealing with one object, and in the disconnected case, we normally say two objects are involved. Hence the overdetermination problems arise from our common sense assignment of object status.

Note that we can turn this *gedankenexperiment* into a story about preemption simply by considering a case where the overdetermining dumbbell strikes the window "end-on". The connected dumbbell is then a case of causation and the disconnected dumbbell is a case of late preemption.

Let's look at a possible objection to my conclusion about the overdetermining dumbbell *gedankenexperiment*. We assumed the influence of the mass of the connecting bar, the composition of the window, and recoil and friction in the quick-release were all effectively zero. Hence the things we want to say are causes in the two cases are effectively physically identical, and the two effects are effectively physically identical. So the counterfactual analysis tells conflicting causal stories about effectively identical physical systems. But perhaps this talk of "negligible influences" and "effective identity" is all epistemological sleight of hand. Even if we can't perceive or detect the changes in the effects and (putative) causes, there *will* be some slight physical differences between them. So maybe there is still room to defend the line that some relevant physical differences are driving the failure to assign causal dependence in the disconnected case.

But surely if we have to apply such harsh strictures to the qualitative identity of events in the *gedankenexperiment*, we have to apply these harsh strictures to other instances of causation. But this would conflict with part of *Lewis's* basic desideratum for development of the counterfactual analysis — that the analysis should be consistent with our common-sense understanding of the notions of causation and events. To be precise, such harsh strictures would make events too "fragile" to be consistent with our common-sense understanding of the identity of events.

An event for the Lewis-style analysis is a property of a spatiotemporal region: effectively, a feature of a place at a time. But an event need not necessarily, or essentially, be any particular feature at any particular place at any particular event. For example, consider the event of David taking a walk in the park one fine morning. If this event was essentially associated with David walking in that park at that time, then any slight difference in the circumstances of David's walking (for example, that it took place at a different speed) would render that other walking a completely different event. In Lewis's parlance, such an account would render events very fragile (it doesn't take much to change them into different events). But in common parlance events are nowhere near as fragile as this. If the weather had not been propitious *that very same walk* might have been taken in the afternoon. Or David could have taken that walk in the quad. Or if there had been a sudden cloudburst, half of that walking would have been a sprinting. Indeed, if David had gotten a job at a different university, that very same walk could have been taken by Ed instead.

If we intend to avoid the problem of the overdetermining dumbbell by making the analysis fantastically sensitive to physical differences between sets of circumstances, then how can we also consistently give an account of event identity which allows the degree of robustness displayed by events in everyday causal talk? Even if we think it is possible that the standards for fragility of events may differ between causal talk and causation itself, the fragility defence to the overdetermining dumbbell smacks of a double standard. As Lewis puts it:

> Extreme fragility of effects might get rid of all but some far-fetched cases of redundant causation, but it leads to trouble that we don't know how to control. Moderate fragility gets rid of some cases and casts doubt on others, but plenty are left. Our topic has not disappeared.[10]

Extreme fragility is certainly required to avoid the overdetermining dumbbell problem. The moral is that you can't think the overdetermining dumbbell case arises from physical differences between the connected and disconnected cases

---

10 "Causation", p199. See pp197–9 for rather a good discussion by Lewis why he thinks fragility is a bad defence to overdetermination problems.

and enjoy the other apparent advantages of a Lewis-style analysis at the same time.

So there is no relevant physical difference between the disconnected and connected cases of the overdetermining dumbbell which leads the counterfactual analysis to make a causal determination in the connected case and fail in the disconnected case. The counterfactual analysis tells conflicting causal stories about effectively identical physical systems due to its reliance on common-sense individuation of objects within events. The case of the overdetermining dumbbell provides good grounds for supposing that causal dependence does not track physical dependence in a Lewis-style counterfactual analysis.

Bonking Frank

The might be another way out of the strange case of the overdetermining dumbbell for a Lewis-style counterfactual analysis — give up common sense individuation of objects within events, and key individuation of objects within events to physical involvement. Since there are no relevant physical differences between the disconnected dumbbell and the connected dumbbell we should treat the disconnected dumbbell as if it were a single object, just like the connected dumbbell. If the disconnected dumbbell is a single object there is no overdetermination, and we have a truthmaker for the claim "the disconnected dumbbell caused the window to break".

This flight from common sense is not immediately implausible. We commonsensically think that the planets of the solar system are individual objects, and the solar system is thus a collection of distinct objects. Yet when considering the gravitational influence of the solar system on nearby stars, we might treat the solar system as a single object. At any rate, experience has often shown us that common sense can just be *wrong*.

The new story about objects certainly fits with the programme of trying to fit causal dependence to physical dependence. But this "sophisticated" interpretation of objects might lead to different problems for a Lewis-style counterfactual analysis. Think of cases of physical dependence involving a single common-sense object, like a standard rock hurtling through a standard window. The "naive" Lewis-style analysis (naive in the sense of appealing to common-sense objects) unproblematically supplies truthmakers to the causal claims involved. But if the analysis is keyed to other "objects" than the common sense variety, might it not turn out that many of these situations are actually instances of overdetermination?

In fact, this is precisely the case. Many cases of "single-object overdetermination" do arise. But note that these cases don't arise as cheaply as we might expect.

For example, a tempting but incorrect line of argument for single-object overdetermination in the case of the single standard rock striking a window is that the assignment of single-object status to the rock is purely arbitrary. A rock is made of trillions of independent objects: atoms and molecules. (It should not be too controversial in these times to claim that atoms are viewed as real objects by the woman in the street.) It seems plausible to suggest that only a subset of those independent atoms and molecules are sufficient to break the window. Let's suppose that subset amounts to less than half of the mass of the aggregate. So surely the independent action of trillions of separate objects striking the window amounts to a case of overdetermination. But then the counterfactual analysis is telling conflicting causal stories about identical physical systems again — when viewed as one object the rock causes the window to break, when viewed as an aggregate of independent objects striking the window we have a case of overdetermination.

This line of argument fails because "an aggregate of independent atomic and molecular objects striking a window" and "a rock striking a window" are not equivalent descriptions of the same physical system. The atoms and molecules in a rock are chemically bonded into larger structures — generally crystals — which are also physically tangled within together in such a way that (quantum effects aside) the smaller objects stand in fixed positions within the rock. These bonding and entanglement details dictate the physical properties of the rock — its density, its elasticity and so forth. These properties dictate the nature of the collision with a standard window and the way the window breaks. For example, a globe of liquid with the same size and mass as a standard rock would probably break a window in quite a significantly different way (splash, tinkle) to a standard rock (crash, tinkle), and a globe of gas of the size and mass of a standard rock (whoosh, tinkle). For the sake of charity, we'll accept that the counterfactual analysis can distinguish between these cases. So "a rock striking the window" is only identical to the description "an aggregate of *interdependent* objects striking a window", and it is not clear at all that this latter description picks out a case of overdetermination. So this line of criticism of the counterfactual analysis fails.

But the counterfactual analysis would be in trouble in cases of physical dependence involving single objects if it could be shown that *some* parts of an object were physically irrelevant to an interaction, even though all the parts of the object were interdependent in some way. Let's set up one such case.

In the George Alec Effinger novel[11] *The nick of time*. Effinger finds it necessary to crush his hero Frank by crashing the Moon into the Earth. The Moon has been thoroughly wrapped in duct tape so it will not break up when it passes through the Earth's Roche limit[12] and forms a single contiguous object when it strikes Frank (and the Earth very shortly after). We will assume there is no atmospheric involvement in this process.[13] Does the Moon striking Frank and the Earth cause Frank to be crushed?

The instant off-the-cuff response to that question is: "Of *course* it does! It is as certain as certain could be that the Moon striking the Earth crushes Frank. What else do you think could have done it?" The naive counterfactual analysis falls into step with the off-the-cuff response — but for the Moon striking the Earth, Frank would not have been crushed. Unfortunately, common sense puts us wrong, because most of the Moon is physically irrelevant to Frank's being crushed, producing an instance of single-object overdetermination.

---

11 Effinger, George Alec: *The nick of time* (Sevenoaks: New English Library, 1987)

12 The Roche limit is the orbital distance at which a satellite with no tensile strength (a "liquid" satellite) will begin to be tidally torn apart by the body it is orbiting. A real satellite can pass well within its Roche limit before being torn apart, particularly if it is wrapped in duct tape. See Wesstein, Eric: "Roche limit" (http://www.astro.virginia.edu/~eww6n/physics/RocheLimit.html 1996–8)

13 Partly this is to ensure that the moon strikes Frank in one piece — regardless of size, most objects entering the Earth's atmosphere are expected to break up unless they burn up first. See pp205–9, Melosh, H J: *Impact cratering: a geologic process* (New York: Oxford University Press, 1989). Partly this is out of charity to the counterfactual analysis. If we allow atmospheric effects to be considered, the massive overpressure produced when the atmosphere is squeezed between the Earth and the Moon would crush Frank well before the moon actually touched him. Under a Lewis-style analysis this is a case of early preemption, and Frank's crushing *would not be caused*. So we avoid the complication — perhaps the Earth has lost its atmosphere and Frank is standing around in a spacesuit waiting for the big event.

In fact, even informed common sense puts us wrong. When the educated layperson thinks about the collision of solid bodies at all, it tends to be in terms of the classical theory of impact, or stereomechanics. Stereomechanics amounts to a specification of initial and terminal velocities of colliding objects and the linear or angular impulses applied during the collisions.[14] The colliding bodies behave as if they were single mass points — all internal elements of the bodies are perfectly rigidly connected and are instantaneously subject to the same change of motion.[15] The momentum calculations exercises of high-school physics are applications of simple stereomechanics. Under such an analysis of physical dependence, when the Moon strikes Frank, the compressive force applied to Frank as he is crushed thinner than a sheet of paper and beyond will be a function of the entire mass of the Moon, and the entire mass of the Earth. So the entire Moon and the entire Earth are treated as if they are physically involved in the crushing of Frank, and common sense and the counterfactual analysis are supported. But a more advanced examination of the physics of impact tells a much different story about physical dependence and causal dependence under a Lewis-style analysis.

Consider a more realistic "instant-by-instant" picture of the dynamics of impact of the collision of two chunks of some solid substance. At the first instant of impact only the barest outside surfaces of the two objects will interact and be deformed. The "forces of collision" felt between the two objects at that instant will be a function of the masses in the "contact zone," not of the masses of the entire objects as in our simple stereomechanical model.[16] So at that instant, the portions of the objects outside the contact zone are *physically* irrelevant to the collision, and to preserve our basic desideratum we would want to say they are *causally* irrelevant to the collision process *at that instant*.

As we step through the collision process instant by instant, we observe the contact zone expanding further and further through the objects at a finite velocity. The boundary of the contact zone will be a stress disturbance within the material of the spheres.

---

14 See pp4–21, Goldsmith, Werner: *Impact: the theory and behaviour of colliding solids* (London: Edward Arnold, 1960)

15 *Impact*, p23

16 *Impact*, p22, see also p9, Zukas, Jonas; Nicholas, Theodore; Swift, Hallock; Greszczuk, Longin; Curran, Donald: *Impact dynamics* (New York: John Wiley & Sons, 1982)

Precisely how quickly the stress wave will propagate will be a complex matter. In a low velocity impact, where the stress is below the "yield point" of the material composing the objects, the objects will deform elastically. After deformation the objects will return to more or less their original shape — much like squeezing a tennis ball. At least two types of elastic disturbances will propagate through the objects.[17] A longitudinal stress pulse — where the motion of particles within the solid are parallel to the direction of propagation of the stress disturbance — will travel through the solid at, roughly speaking, the speed of sound in that particular solid. The longitudinal pulse will be followed by slower transverse disturbances, where the particles' motion is perpendicular to the propagation of the disturbance. (Residents of the Shaky Isles will recognise these effects as the two distinct P wave and S wave jolts of an earthquake.)

In higher energy impacts, where forces exceed the yield points of the solids, the objects will also deform plastically. The objects do not return to their original shapes — much like squeezing a ball of play dough. The disturbance front has a twin-peaked structure: an elastic disturbance is followed by a slower but much more intense plastic deformation. At still higher energies the double-pulse structure is eliminated by the production of shock waves (much like the sonic boom created by a supersonic jet). Shock pulses move supersonically through the material, outstripping and absorbing the energy of the elastic disturbance and forming an abrupt shock front — so abrupt that it is a good approximation to represent it as a discontinuous jump of pressure, particle velocity, density and internal energy.[18]

At high impact energies, solid materials will start to behave *hydrodynamically* — like fluids.[19] The strength or rigidity of the solid will no longer significantly determine the velocity of the disturbance front. The mere density of the solid will become a dominant parameter. Transverse waves no longer propagate through the material, and the elasto-plastic double pulse does not form. At still

---

17 "At least two" because other wave phenomena can propagate along the surface of the objects (such as Rayleigh waves and Love waves). See *Impact dynamics*, p2, and also *Impact cratering*, pp29–33

18 *Impact cratering*, pp33–7. See also pp100–1, Rosenburg, Z: "The dynamic response of ceramics to shock wave loading", pp73–105, Brebbia, C A; Sanchez-Galvez, V; editors: *Shock and impact on structures* (Southhampton: Computational Mechanics Publications, 1994)

19 *Impact dynamics*, pp3–5

higher levels of energy, material within the contact zone explosively vaporises.[20]

The physical (and causal) upshot of these various phenomena can be illustrated by considering impacts when a small projectile deforms or penetrates a large solid. In the low-energy elastic regime, the local effects produced by the impact are strongly coupled to the overall deformation of the target structure — so the entire target should be causally relevant to the result of the impact. As the impact velocity rises, the response of the overall structure becomes *less* important to the result of the impact. It is secondary to the behaviour of a small local zone around the point of impact (typically around 2–3 projectile diameters in extent), and the significant factors in the local zone will include velocity of impact, geometry, material composition, strain rates and local plastic flow. At higher velocities, the hydrodynamic properties of the local zone dominate the plastic effects, and at still higher velocities things explode in the zone.[21] In these cases we want to say that dominant causal influence on the effects of the impact is no longer the entire target but rather the impact zone, and the nature of the relevant causal couplings can vary considerably.

So what causal morals should we take from this? First, in a comparatively simple case of physical dependence — impact — it is clear that the common sense objects involved in the interaction are not necessarily the relevant structures we should be telling our causal stories about. For example, in a high-velocity penetration of a target by a projectile the small impact zone of the target, not the entire target, is going to be the most significant causal determinant of the result of the impact.

In the case of the two objects colliding producing plastic deformations, the physical forces felt at the contact surfaces at any instant will be predominantly coupled to the material within the zone of plastic deformation, not to the material within the wider contact zone defined by the elastic deformation. The plastic zone should be the significant causal determinant of the result of the impact on the contact surfaces (or indeed things between the surfaces at the points of contact), not the entire object.

---

20 *Impact dynamics*, p156
21 *Impact dynamics*, pp156–7

Finally, and the crucial point with regard to the bonking of Frank, in a case where the impact or related relevant process has a finite duration, since physical involvement (and hence causal involvement) in an impact propagates through objects at a finite velocity, that impact or related process may be over and done with before all the interacting common sense objects have had a chance to become physically involved. Any material in the objects outside the contact zone at the point at which the impact ceases is physically irrelevant to the impact or related process — it virtually might as well have been made of titanium, or custard, or not have been there at all.[22] Thus, if causal dependence tracks physical dependence, the material outside the contact zone should not be causally relevant to the impact and related processes.

This may seem a long-winded way of making a rather simple point. Of *course* our common sense intuitions may lead us astray about putative instances of causation, and of *course* we should advise our assignments of causal relations with our scientific understanding of the world. Jonathan Bennett puts the case rather nicely:

> Objection: "But do you deny that when an explosion causes a fire, the explosion emits force, pushes things around, acts as the elbow in the ribs?" Yes, I do. When an explosion causes a fire, what happens is that molecules bump into other molecules, increasing their velocity to the point where they react rapidly with the ambient gases, etc. The idea that the pushing is done not by the molecules but by the explosion is just the afterglow of ignorance about what an explosion is.[23]

---

22 "Virtually" does indicate a genuine equivocation. The properties of the matter or lack of it near the *boundary* of the contact zone can influence the behaviour of the matter within the contact zone. Consider the meniscus which forms at the interface of liquids and air. The surface tension at the meniscus will affect the properties of the liquid near the surface. Specifically, it affects the behaviour of disturbances propagating through the liquid near the surface (which are precisely the disturbances which interest us). The precise change of behaviour near the interface will be a function of the material composition of the materials on either side of the interface. So the growth of the contact zone in a rock after an impact will be affected to some extent by a nearby material boundary *outside* the contact zone in the rock, and the degree to which the boundary influences the contact zone will depend on the physical characteristics of the boundary (rock/titanium, rock/custard or rock/void).
We can avoid these problems by adopting the same strategies often used by materials scientists and engineers — limiting our attention to impacts between *semi-infinite* bodies, where semi-infinite means that there is no nearby boundary which affects the physical behaviour of the contact zone (semi-infinite bodies under this definition can be quite small). The examples we shall consider will involve collisions between semi-infinite bodies.

23 pp22–3, Bennett, Jonathan: *Events and their names* (Indianapolis: Hackett Publishing Company, 1988)

Well, the point of our long-winded discussion of the dynamics of impact is that when we emerge from the afterglow of ignorance about collisions, we discover that the sophisticated Lewis-style analysis no longer assigns relations of causal dependence in cases of impacts involving single common-sense objects.

Let's return to Frank, standing on the surface of the Earth waiting to be crushed by the impact of the Moon. Let's now cash out the crushing of Frank using our informed understanding of the dynamics of impact. The process starts when the first layer of duct tape touches the top of Frank's head, continues as Frank is progressively flattened, and ceases when there is nothing left of Frank to be crushed. If Frank is 2 m tall, the process will have a short, finite duration — around 0.028–0.18 μs.[24]

Since the Moon is not a perfectly rigid body, the initial contact zone involved in Frank's crushing will be smaller than the entire Moon — it will be the outermost layer of duct tape directly over Frank. This contact zone will expand at a finite velocity until Frank ceases to be crushed. How big will the contact zone grow? Let's make some generous assumptions. We can assume the longest crushing duration: 0.18 μs. We can assume that the growth of the contact zone will be governed by an elastic longitudinal wave, not a shock wave, but we will assume the speed of sound in the material of the Moon is very high, thus making the contact zone grow as quickly as possible. So let us assume that the speed of sound in the Moon is the same as in solid carbon: 18 350 ms^{-1} (nearly four times faster than the speed of sound in steel). But even using these generous estimates, by the time that Frank has been crushed the contact zone has only grown to about 3.3 m in radius. The rest of the Moon (a roughly spherical object with a diameter of 3476 km) is physically irrelevant to the crushing[25].

But in *that* case, the counterfactual analysis must treat the crushing of Frank as a case of overdetermination — to be precise, a case of late preemption. If the initial contact zone had not been there to crush Frank, the rest of the Moon would have crushed him a fraction of a microsecond later. Even though the

24 If we treat the Moon's impact as a large asteroidal impact, the range of possible velocities of impact is 11.2–72.8 kms^{-1} (*Impact cratering*, p205). The upper and lower limits on crushing time are merely the range of times it would take for the Moon to close a 2 m gap at these velocities.

25 Of course we should also be telling a similar story about the involvement of the Earth.

contact zone of the Moon and the rest of the Moon are contained within the same common sense object, they are as causally independent with regard to the crushing process as individual rocks within a shower of rocks would be with regard to a window-breaking event. Franks' crushing *wasn't* caused after all! The upshot is that if we stop being naive about the objects involved in collisions and try to key objecthood to our best understanding of physical dependence, the sophisticated Lewis-style counterfactual analysis does an even worse job of fulfilling our basic desideratum than the naive analysis. It doesn't pay to emerge from the afterglow of ignorance about causation.

Here's a "Lewisian" response to the problem of single-object overdetermination. So the analysis breaks down in cases of deep impact — what matter? We can let lunar collisions and their like be problem cases, but what relevance does this have to everyday run-of-the-mill causation? The causal story told about such cases as the bonking of Frank may seem odd, but we should just swallow this the odd result given that the counterfactual analysis does such a splendid job of dealing with ordinary-size examples of causation. Spoils to the victor!

Unfortunately, the same principles of physical dependence that say the Moon crushing Frank must be a case of single object overdetermination also say that many ordinary-size phenomena must be cases of single-object overdetermination.

Consider a standard window. Rather than breaking it by striking it with a standard rock, we will strike it with a standard long rod, with a length several times larger than its diameter. Like the crushing of Frank by the Moon, we can treat the breaking of the window by the long rod as a process of finite duration: the window is whole when the rod first touches it, it undergoes some process of structural failure, and when the window is perforated the breakage is complete.[26]

---

[26] Long rods are frequently used as impacting objects in experimental and theoretical examination of impact. By virtue of being long, the propagation of stress disturbance through the rod can be treated as propagation of disturbance through a semi-infinite body. By virtue of being narrow by comparison with length, the stress disturbance in the bar can be adequately modelled as a one-dimensional process. (See *Impact dynamics*, pp2–5 and pp30–2 for discussions of the interplay between theoretical models and experimental analysis; see pp160–1 for pictures of long-rod penetrators in action.)

For an account of what perforation means by comparison with penetration, see *Impact dynamics*, pp155–7, see also *Impact*, pp240–2.

Suppose the glass is 1 cm thick, and is about as strong as ordinary window glass. The rod is made of steel, is 1 cm in diameter, and strikes the glass at 100 ms^{-1}. At this speed of impact a steel ball bearing 1 cm in diameter would have sufficient momentum to break the window. But more than a ball-bearing-sized piece of the rod may be brought to bear on the window while the breaking event is still in progress. How much of the rod could actually have some physical relevance to the window-breaking event under these conditions?

The duration of the event is 10^{-4} s. The speed of sound in steel is 4877 ms^{-1}, so the "contact zone" would have propagated approximately 49 cm down the rod before the breakage had been completed. So if the long rod was about 55 cm long, 6 cm of the rod would be physically irrelevant to the breakage and should not be considered part of the relevant object within the impact event. At higher impact velocities, the contact zone will be proportionately shorter: 24 cm at 200 ms^{-1}, and 12 cm at 400 ms^{-1}. If the rod was made of a different material with a different characteristic speed of sound, the contact zone sizes would be different again. For a copper rod, with a speed of sound of 3353 ms^{-1}, the contact zone sizes would be 34 cm at 100 ms^{-1}, 17 cm at 200 ms^{-1} and 8 cm at 400 ms^{-1}.

Suppose that all six cases involve 55 cm-long rods. The simple counterfactual analysis says in all six cases the rod striking the window causes the window to break — but for the rod striking the window, the window would not have broken. But this analysis wrongly ascribes physical involvement to the entire rod, which is simply not the case. Greater or lesser amounts of the rods are simply not physically relevant to the window being broken. But if we admit that only part of the rod could be involved in each breakage, the counterfactual analysis provides no causal story in any of the six cases — if the physically involved part of the rod had not broken the window, the remaining part of the rod would have. All six cases are actually instances of single-object overdetermination. Single-object overdetermination is everywhere, not just in deep impact.

Let's recap and sum up. In simple, paradigmatic cases of physical dependence such as impacts, Lewis-style counterfactual analyses of causation fail to assign relations of causal dependence due to overdetermination problems. The failure is ontological rather than epistemological. It isn't that we can't know that causation is going on in these cases. Under the basic Lewis-style analysis it simply isn't *true* that there is causation going on.

These failures are evidence that causal dependence fails to track physical dependence under a Lewis-style analysis. This conclusion is supported by the overdetermining dumbbell *gedankenexperiment*, in which a Lewis-style analysis assigns causal dependence in one set of physical circumstances and does not in another, even though there are no relevant physical differences between the two cases. The overdetermination problem is being generated by the common-sense individuation of objects within events.

We could try and bring causal dependence back into line with physical dependence by giving up common-sense individuation of objects within events in favour of keying object individuation to physical involvement. But this just results in many more overdetermination problems arising than under the naive Lewis-style analysis. Consider the bonking of Frank. So long as we can claim that the entire common sense object "the Moon" is physically involved in Frank's crushing, a Lewis-style analysis can say that Frank's crushing was caused. But when we acknowledge that this claim is physically implausible, modify it in accordance with our best relevant understanding of physical dependence (the dynamics of colliding solids), and claim that only a small part of the Moon is physically responsible for the crushing of Frank, we immediately have a case of single-object overdetermination. In fact, single-object overdetermination crops up all over the case; ironically, more so in cases where we are absolutely certain there is a causal story to be told than elsewhere.

So it does not seem to be the case that we can treat the Lewis-style counterfactual analysis "as if" it was intended to fulfil our basic desideratum. Causal dependence simply does not track physical dependence in such an analysis. Some other analysis of causation is required.

Transference

Prehistory and general considerations

Although Lewis-style counterfactual analyses may represent the current philosophical orthodoxy on causation, a recurrent notion has been that causation resolves to the transference of energy or momentum. The suggestion has been "in the air" more than "in the print", but several attempts have been made to formulate an analysis over the past 140 years.

Modern transference analyst David Fair notes that the "necessary and sufficient conditions" analysis of causation presented by Curt Ducasse seems to be on the right track towards a transference analysis.[27] Yet Ducasse explicitly criticises the transference analysis in the 1951 book *Nature, mind and death*[28] and somewhat more elliptically in an earlier work, *Causation and the types of necessity*.[29] Ducasse's target appears to be Charles Mercier, a respected psychiatrist and part-time philosophical gadfly. Mercier's 1916 book *On causation with a chapter on belief*[30] provides much of the philosophical groundwork for Ducasse's own analysis, with the notable difference that Mercier thinks the relation between cause and effect should be cashed out in terms of transferences of energy.[31] So it seems that Ducasse's analysis resembles transference analyses because it originally *was* a transference analysis.

Mercier appears to have stolen the notion of transference from the eighth edition of John Stuart Mill's *A system of logic ratiocinative and inductive*,[32] where Mill criticises a transference analysis proposed by Alexander Bain in Bain's 1870

---

27 pp224–31, Fair, David: "Causation and the flow of energy", pp219–50, *Erkenntnis* vol 14 (1979)

28 pp138–42, Ducasse, Curt J: *Nature, mind and death* (La Salle: The Open Court Publishing Company, 1951)

29 Ducasse, Curt J: *Causation and the types of necessity* 2nd edition (New York: Dover Publications, 1969). Originally published in 1924.

30 Mercier, Charles: *On causation with a chapter on belief* (London: University of London Press, 1916)

31 For example, *On causation with a chapter on belief*, pp80–2

32 pp348–53, Mill, John Stuart: *A system of logic ratiocinative and inductive*, Robson, J M; editor: *The collected works of John Stuart Mill* Volumes VII–VIII (Toronto: University of Toronto Press, 1974)

work *Logic*.[33] Ducasse's criticism of the transference analysis seems to be (at least initially) based on discussions by Ernst Mach of the relation between the conservation of energy and causation in his book *History and root of the principle of conservation of energy*[34] and his later article "On the principle of the conservation of energy".[35] Mach also identifies a precursor of the transference analysis presented by Wilhelm Wundt in 1866.[36]

A more detailed discussion of the prehistory of the transference analysis is beyond the scope of this thesis.[37] We shall be concerned with the modern version of the analysis reinvented by Jerrold Aronson in two 1971 papers, "On the grammar of cause" and "The legacy of Hume's analysis of causation",[38] and developed by Aronson, Fair and Adrian Heathcote during the 1970s and 1980s.[39]

The transference analysis clearly seems to be a good candidate for an analysis of causation which fulfils our basic desideratum — energy and momentum transfers are species of physical dependence, so causal dependence resolves to a species physical dependence. But not all physical dependence may be causal dependence under the modern analysis.

Fair notes that some instances of causal dependence, such as "stoppings", are not to be interpreted as energy transfers. For example, I press a switch on the monitor of my computer, and the screen goes blank. I appear to have caused the screen to go blank, and it seems the screen's going blank physically depending on my flicking the switch, yet energy and momentum was not transferred from the switch to the screen. Fair still thinks such phenomena are causal — it should best be interpreted as causation by virtue of raising a

---

33 See the introduction to Bain, Alexander: *Logic* 2nd edition (London: Longmans, Green & Co, 1895). Bain's analysis could fairly be said to also anticipate the conserved quantity analyses of Dowe and Salmon.

34 Mach, Ernst: *History and root of the principle of conservation of energy* (Chicago: The Open Court Publishing Company, 1911). Originally published in 1872.

35 Mach, Ernst: "On the principle of the conservation of energy", pp137–85, Mach, Ernst: *Popular scientific lectures* 3rd edition (Chicago: The Open Court Publishing Company, 1898)

36 *History and root of the principle of conservation of energy*, pp39–40

37 For a more detailed discussion see Smith, Tony: "Energy, psychiatrists and mysterious chemicals: the lost history of the transference analysis", MA thesis (Chapel Hill: University of North Carolina at Chapel Hill, forthcoming)

38 Aronson, Jerrold: "On the grammar of cause", pp414–30 *Synthese* vol 22 (1971); Aronson, Jerrold: "The legacy of Hume's analysis of causation", pp135–56, *Studies in history and philosophy of science* vol 2 (1971)

39 We will not discuss the similar analysis developed by Hector-Neri Castañeda.

"potential barrier" which stops energy transference from taking place along the wires leading from the electricity source to the monitor.[40]

There are at least two ways we could interpret Fair. One is to assume that all physical dependence is associated with causal dependence. In that case, the transference analysis needs to be supplemented by a "potential barrier" analysis. Since we have no satisfactory account of potential barriers above and beyond an enumeration of different phenomena we currently think count as potential barriers, the "supplemented" transference analysis is incomplete and programmatic in character. (This is Fair's conclusion.[41]) Alternatively, we might assume that not all physical dependence is associated with causal dependence — causal dependence involves energy transfers and some other dependence resolves to raising and lowering potential barriers. In effect, we could claim there are two or more different brands of cement of the universe, causation and something else. Therefore, although the transference analysis might provide a satisfactory account of causal dependence, the full analysis of physical dependence is still programmatic and incomplete.

Since our basic desideratum does not require all physical dependence to resolve to causal dependence or vice versa, we shall not discuss potential barriers. We shall limit our attention to the core notions of energy and momentum transference. Does causal dependence track physical dependence under this analysis (or portion of an analysis)?

Rather surprisingly, it *doesn't*. We shall try and develop a best possible version of the transference analysis, and see whether it surmounts various objections to the transference analysis. We shall conclude that while the analysis seems to give a satisfactory account of causal interactions, it fails to give a satisfactory account of the connections between different interactions.

---

40 "Causation and the flow of energy", pp244–5
41 "Causation and the flow of energy", pp248

Gasking on manipulability

Aronson presents the basic structure of the modern transference analysis in "On the grammar of cause" as an alternative to the manipulability analysis of causation presented by Douglas Gasking.[42]

Gasking argues that Humean regularity analyses of causation are incomplete. The Humean argues that A causes B means that events of type B regularly follow events of type A. Yet much of our causal talk involves instances of simultaneous causation — a feature of event causes another feature of the same event. For example, we want to say that the cause of an iron bar glowing is that the iron is currently at a temperature of 1000 C. Which feature is cause, and which feature is effect?

One approach might be to deny that the explanatory relationship between the glowing bar and the temperature of the bar is causal at all. We might reconstruct such an approach from David Lewis's "Events':

> Why did Xanthippe become a widow? Because she was married to Socrates at the time of his death. (Noncausal.) Because Socrates was made to drink hemlock. (Causal ...) Why did Fred talk or walk then? Because he talked (noncausal) and he did that because he had just heard a joke he couldn't keep to himself (causal).[43]

Similarly, we might say that the bar is glowing because it at a temperature of 1000 C but deny this is a causal explanation. The causal explanation might be that it had been in a furnace for the previous 15 minutes.

But Gasking's approach is to accept that the relationship between the temperature and the glowing is causal and then attack the inadequacy of the regularity analysis of causation in these contexts. Gasking presents an alternative analysis of causation, which links a statement about the causes of something with "a recipe for producing it or for preventing it":

> When we have a general manipulative technique which results in a certain sort of event A, we speak of producing A by this technique. (Heating things by putting them

42 Gasking, Douglas: "Causation and recipes", pp479–87, *Mind* vol 64 (1955)
43 p269, Lewis, David: "Events", pp241–69, Lewis, David: *Philosophical papers volume II* (Oxford: Oxford University Press, 1986)

> on a fire.) When in certain cases application of the general technique for producing A also results in B we speak of producing B by producing A. (Making iron glow by heating it.) And in such a case we speak of A causing B but not *vice-versa*. Thus the notion of causation is essentially connected with our manipulative techniques for producing results. Roughly speaking: "A rise in the temperature of iron causes it to glow" means "By applying to iron the general technique for making things hot you will also, in this case, make it glow".[44]

So causation statements are properly *manipulation-by-agents* statements. The use of the word "cause" in contexts where there are no agents to speak of, such as the sentence "gravity causes unsupported objects to fall" involves, in Gasking's words, a "sophisticated extension from its more primitive and fundamental meaning".[45]

The most plausible interpretation of Gasking's "sophisticated extension" is that statements such as "gravity causes unsupported objects to fall" are polite fictions. We use the term "cause", here in the same way as we use the term "design" in statements such as "the human body contains number of design flaws and compromises". We know what is meant by the latter statement — some structures in the human body do not function as well as some different structures would in the same body. But — various heretics excepted— we do not mean by such a statement, even by implication, that the human body was designed poorly by its designer. Similarly, the use of the term "compromise" in this context should not be taken as meaning that certain features of the body appear in modern humans as a result of protracted bouts of negotiation in smoke-filled rooms. So any causal statements about non-agent-dependent relations between events are at best enlightening metaphors and at worst literally false and misleading locutions.

The Russell-Quine thesis

So under a manipulability analysis, what role should causal statements such as "gravity causes unsupported objects to fall" play in natural sciences such as physics, chemistry or biology? One sensible answer to the question is *none*. The manipulability analysis is a road to what we might call the Russell-Quine thesis

44 "Causation and recipes", p483
45 "Causation and recipes", p487

— that the notion of cause plays no worthwhile role in science, or at least a mature science. We can find a strong statement of the thesis in Bertrand Russell's "On the notion of cause":

> … I wish, first, to maintain that the word "cause" is so inextricably bound up with misleading associations as to make its complete extrusion from the philosophical vocabulary desirable; secondly, to inquire what principle, if any, is employed in science in place of the supposed "law of causality" which philosophers imagine to be employed …[46]

A slightly more tolerant version appears in Willard Van Orman Quine's "Natural kinds":

> We have noticed that the notion of kind, or similarity, is crucially relevant to the notion of disposition, to the subjunctive conditional, and to singular causal statements. From a scientific point of view these are a pretty disreputable lot. The notion of kind, or similarity is equally disreputable …[47]
>
> In general we can take it as a very special mark of the maturity of a branch of science that it no longer needs an irreducible notion of similarity and kind. It is that final stage where the animal vestige is wholly absorbed into the theory. In this career of the similarity notion, starting in its innate phase, developing over the years in the light of accumulated experience, passing then from the intuitive phase into theoretical similarity, and finally disappearing altogether, we have a paradigm of the evolution of unreason into science.[48]

So for Russell, causal talk in science is just rot. For Quine, causal talk is rot that makes useful fertiliser for the eventual flowering of a scientific discipline. For Gasking, such talk of "causes" is part of the province of "popular science": interesting rot talked by educated laymen and "some scientists in their less strictly professional moments."[49]

Aronson agrees with Gasking that the regularity analysis fails to determine causal asymmetry in cases of simultaneous causation. Yet Aronson cannot

46 p1, Russell, Bertrand: "On the notion of cause", pp1–26, *Proceedings of the Aristotelian Society* vol 13 (1912)

47 p167, Quine, Willard Van Orman: "Natural kinds", pp159–70, Boyd, Richard; Gasper, Philip; Trout, J D; editors: *The philosophy of science* (Cambridge, MA: MIT Press, 1991)

48 "Natural kinds", p170

49 "Causation and recipes", p487

agree with Gasking's manipulability analysis of causation, because the analysis has a "pernicious outcome"[50] — the Russell-Quine thesis. So a further analysis of causation is required in order to save both the causal and scientific status of statements such as "gravity causes unsupported bodies to fall" — the transference analysis.

Aronson argues that anthropomorphic locutions are not the fundamental form of causal statements. Rather, borrowing a concept from J L Austin, the word "cause" is a *dimension word* — "the most general and comprehensive term in a whole group of terms of the same kind, terms that fulfil the same function"[51] — for simple transitive verbs such as "knock", "push" or "pull". In other words, statements such as:

> John knocked the book on the floor

can be transformed into:

> John caused the book to be on the floor

with only a slight loss of specificity of meaning. If cause is a dimension word for these kind of verbs, then the fundamental sense of causal statements cannot be that they are manipulation exemplars, because:

> ... the notion of manipulation is connected with concepts that serve to *modify* intelligent behaviour; but transitive verbs can be used without being modified by adverbs of manner such as "skilfully", "delicately", "intentionally", etc.[52]

Aronson thus concludes that expressions involving the "mechanical" transitive verbs are our exemplars of causal expressions.

This line of argument, though quite clear and evocative, is a tactical mistake on Aronson's part. His debate with Gasking appears to be *semantic*, a disagreement over the actual *sense* of causal locutions in everyday talk. Gasking thinks the correct analysis of the meaning of "cause" is manipulation by agents; Aronson, agent-independent mechanical processes.

---

50 "On the grammar of cause", p414
51 "On the grammar of cause", p417
52 "On the grammar of cause", p419

In fact, Aronson is concerned with an *ontological* failure of the regularity analysis. The reason why "the heat of the bar makes it glow" is causal and "the glowing of the bar makes it hot" is not causal is not that the latter statement is linguistically aberrant. Rather, there is a *real* causal asymmetry between being at a certain temperature and glowing; an ontological feature of causation which is not captured by the regularity analysis and which appears to be independent of manipulation by agents. (Aronson is forced to continually reiterate in later works that questions of ontology, not grammar are the thrust of "On the grammar of cause".[53])

So what is this asymmetric ontological feature of causation underlying the appearance of mechanical transitive verbs in our causal talk? Simply put, it is the transference of a physical quantity, such as energy, momentum, heat or velocity from and to objects. In the case of the iron bar, high levels of energy in the bar are dissipated by the emission of photons. So a quantity of energy (heat in the bar) is transferred away from the iron bar as radiation (the glow from the bar), and thus we can say the heat of the bar caused the bar to glow. But no energy is transferred from the radiation to the bar, so we cannot say the glowing of the bar caused the bar to be hot. (In contrast, Gasking's short explanation for the asymmetry is that human beings have a manipulative technique for making iron glow by heating it, but not for making it hot by glowing it.[54])

The transference analysis of causation does not provide a direct route to the Russell-Quine thesis — causal statements such as "gravity causes unsupported objects to fall" refer to the sorts of physical dependence treated of in the sciences. In the case of a body falling in vacuum towards the Earth, energy is being transferred from the Earth's gravitational field to the body, which manifests in the body as an increased level of kinetic energy. Under the transference analysis, scientific causal statements are not *metaphorically* causal — when true, they are as causal as our everyday statements about things being made to happen, such as "heating an iron bar to 1000 C causes it to glow".

---

53 pp294–5, Aronson, Jerrold: "Untangling ontology from epistemology in causation", pp293–305 *Erkenntnis*, vol 18 (1982) and pp249–50, Aronson, Jerrold: "Conditions versus transference: a reply to Ehring", pp249–57 *Synthese* vol 63 (1985)

54 "On the grammar of cause", p427

We say that the transference analysis *does not provide a direct route* to the Russell-Quine thesis rather than *does not lead* to the thesis, because of the problem of Quine. Quine quickly rejected his dismissal of causation in "Natural kinds" — that causal talk in science was a simple instance of kind talk and would be rejected along with all the other disreputable kind talk as soon as a science had reached an adequate level of sophistication — in favour of what appears to be a transference analysis of causation in *The roots of reference*:

> Of these two wayward idioms, the causal and the dispositional, the causal is the simpler and the more fundamental. It may have had its prehistoric beginnings in man's sense of effort, as in pushing. The imparting of energy still seems to be the central idea. The transfer of momentum from one billiard ball to another is persistently cited as a paradigm case of causality. Thus we might seek a simpleminded or root notion of causality in terms of the flow of energy. Causes and effects are events such that all the energy in the effect flowed from the cause.[55]

This certainly looks like a rehabilitation of the notion of cause in scientific talk by Quine, especially when he goes on to say:

> Let us sort out the good and the bad features of this notion of cause. A possible objection is that it is too special, applying only to physics. My answer is materialistic. Causality is a relation of events, and all events, mental and social ones included, are a matter ultimately of the action of physical forces upon particles. My concern here is different from Hume's; his was with the epistemological basis for a causal relation, while mine is with the ontological nature of the causal relation as an object of scientific theory. All will agree, materialists and others, that causal efficacy within the material world, at any rate, is compounded of microphysical forces, despite our incapacity to single out all those components in every particular case.[56]

Hence we find Aronson citing Quine as a proponent and defender of the transference analysis.[57]

Yet Quine can hardly be called a great fan of causation:

> A third ... objection, and one that I can share, is that my appeal to energetic world lines is not sophisticated enough. On what basis can an earlier and later bit of work be

---

55 p5, Quine, Willard Van Orman: *The roots of reference* (La Salle: Open Court, 1973)
56 *The roots of reference*, p6
57 "Untangling ontology from epistemology in causation", pp293

associated as two manifestations of one and the same continuing bit of energy? The very distinction between matter and energy wavers in modern physics, and even the notion of the identity of an elementary particle from moment to moment has fallen on evil days, what with quantum jumps. Now I take this consideration to suggest simply that a notion of cause is out of place in modern physics. Nor can this come as a surprise.[58]

Really, Quine still holds the Russell-Quine thesis. If we must talk of cause, best to talk of it in terms of energy transfers — but better would be not to talk of it at all! Rather than being an analysis of causation, Quine's transference account is a step along the way to a dissolution of causal notions in the sciences. Thus, we find critics of the transference analysis such as Douglas Ehring claiming that Quine objects to the transference analysis.[59] Even Fair states on the basis of Quine's statements that rather than being a proponent and defender of the transference analysis Quine "has suggested (and rejected) explaining our ideas of causality in terms of energy flow."[60] We can contrast this with Aronson's position:

> Since the notion of "cause" is so related to transitive verbs, I think those who maintain that "cause" should be removed from scientific and philosophical discourse are, *ipso facto*, committed to the removal of transitive verbs — which occur often enough in the sciences — as well.
>
> ... However, even if such a task were successful in "eliminating" transitive verbs and "cause" from our vocabulary, it seems that scientists (and others) would continue to speak of the various sources of quantitative change. Physicists would still insist on talking about energy and momentum transfer; and, as I have pointed out above, this talked is wedded to conservation principles. For these reasons, alone, removal of "cause" (and its more specific counterparts) from science and philosophy is a price, I believe, that scientist and nonscientist alike would not be willing to pay.[61]

58 *The roots of reference*, p6

59 p255, Ehring, Douglas: "The transference theory of causation", pp249–58, *Synthese* vol 67 (1986)

60 "Causation and the flow of energy", p237

61 "On the grammar of cause", p428–9

Aronson's analysis

Lets look at the ontological details of the Aronson's transference analysis. Aronson's first move is to distinguish between *natural* changes in objects and *unnatural* changes. A natural change is a change an object undergoes when it is left to its own devices, independently of other influences. Such changes are not causal. An unnatural change is a change resulting from an external influence. Such changes may be causal.[62]

Aronson provides an example of such a natural change: the inertial persistence of motion of objects in the absence of external influences. Such objects change their position, but such changes are not causal. Similarly, Aronson notes:

> ... Galileo regarded circular motion about the center of the Earth to be "entirely natural and self explanatory". Newton would reject a request for a causal explanation of constant linear motion and Einstein would not search for the cause of motion along a geodesic.[63]

This provokes a question: what do scientists' attitudes about the self-explanatory nature of some phenomena have to do with whether those phenomena are causal or not? The answer lies in Aronson's desiderata for developing the analysis — providing a causal basis for a theory of scientific explanation.

In an earlier paper, "Explanations without laws",[64] Aronson reacts against Hempel's deductive-nomological (D-N) theory of explanation. Simply put, a D-N explanation is a modus ponens of an initial condition on an instantiation of a general law. Let Ea be some statement describing an event in need of explanation. Let L be a law, a universal generalisation of the form $\forall x(Cx \rightarrow Ex)$. Let Ca be some initial condition. So, where $Ca \rightarrow Ea$ is an elimination instance of $\forall x(Cx \rightarrow Ex)$, the explanation takes the form of the inference:

62 "On the grammar of cause", pp420–2
63 "On the grammar of cause", p421
64 Aronson, Jerrold: "Explanations without laws", pp541–57, *Journal of philosophy* vol 66 (1969)

(1) Ca

(2) $Ca \rightarrow Ea$

(3) $\therefore$ Ea

Aronson argues that not all explanations take this form. Simple statements containing transitive verbs — such as "John knocked the book on the floor" are also explanatory in this context, without being lawlike. "On the grammar of cause" can therefore be seen as providing both a justification why such statements are explanatory — they are causal — and an argument why causal locutions are not simply instances of lawlike regularities of succession.

So given that the purpose of Aronson's transference analysis is to provide some basis to a theory of explanation, it follows that those phenomena which do not require explanation need not be assumed to be causal. Natural changes are "self-explanatory". So for the purposes of Aronson's theory of explanation, his analysis of causation need not treat them as causal.

(Of course, this is hardly a knock-down argument that natural changes are *not* causal. Later on we shall discuss arguments by Phil Dowe[65] that an adequate analysis of causation does require such natural changes to be causal.)

Having drawn the distinction between natural and unnatural changes, Aronson then gives three conditions which must be fulfilled for causation to take place:

(1) In "A causes B", "B" designates a change in an object, a change which is an *unnatural* one.

(2) "In "A causes B", at the time B occurs, the object that causes B is in contact with the object that undergoes the change.

(3) Prior to the time of the occurrence of B, the body that makes contact with the effect object possesses a quantity (eg velocity, momentum, kinetic energy, heat, etc) which is transferred to the effect object (when contact is made) and manifested as B.[66]

---

65 Dowe, Phil: "What's right and what's wrong with transference theories", pp363–74, *Erkenntnis* vol 42 (1995)

66 "On the grammar of cause", pp421–2

These conditions could readily be generalised so that A or B refer to several objects. A favourable characteristic of this analysis is that the overdetermination problems which lead us to reject the counterfactual analysis simply do not arise.

Consider the overdetermining dumbbell. In the connected case energy or momentum is transferred from the dumbbell to the window, and the window undergoes an unnatural change, so the window breakage is causal. In the disconnected case, energy and momentum is transferred from the rocks to the window, and the window undergoes an unnatural change, so the window breakage is *still* causal. In the case of the bonking of Frank we can conclude that Frank's being crushed (an unnatural change in Frank) is caused by the contact zone of the Moon (because energy and momentum from the contact zone is transferred to Frank). The rest of the Moon outside the contact zone does not cause Frank's being crushed, because no energy is transferred from that part of the Moon to Frank. The fact (or rather the counterfactual circumstance) that the rest of the Moon would crush Frank if the contact zone was not there has no bearing on the causal status of Frank's being crushed. Everything appears to be as it should be — so far, so good for the transference analysis.

The transference analysis (as discussed so far) claims that there is far less causation going on in the universe than a Lewis-style counterfactual analysis does. To illustrate the point, let's borrow Quine's "paradigm case of causality" and consider a white billiard ball striking a red billiard ball at rest. But to avoid complicating factors such as the causal involvement of a billiard table, let's assume that the red is hanging motionless *in*, and the white is moving at a constant velocity *through*, interstellar space. This is illustrated in figure 4.

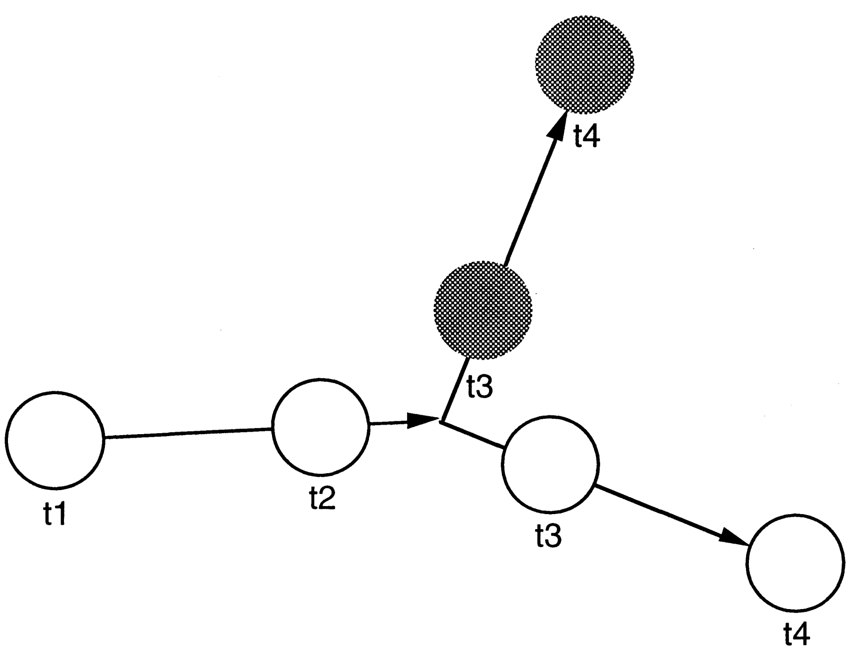


Figure 4 — billiard balls and causation

Both the moving white and the stationary red trace out spatiotemporal regions in the universe which possess certain properties (such as whiteness in the case of the white, redness in the case of red, and in the case of both — we hope — being made of some kind of ivory substitute). So any section of the spatiotemporal path of either ball counts as an event under a Lewis-style counterfactual analysis. So any two contiguous but non-overlapping spatiotemporal regions along the path of, say, the white (before the white hits the red), will be causally related. With reference to figure 4, we would conclude, that the white's behaviour at t_1 was a cause of the white's behaviour at t_2 — *ceteris paribus*, but for the white moving at a certain velocity at t_1, it would not be moving at a certain velocity at t_2. Similarly, the red's behaviour at t_1 would be a cause of the red's behaviour at t_2, even though the motionless hanging of the red in interstellar space amounts to what Lewis describes as a thoroughly uneventful course of events.[67] Of course, the collision with the white with the red will be a cause of the red's behaviour at t_3 and the white's behaviour at t_3. And as per our considerations of the circumstances at t_1 and t_2, the link between the white at t_3 and the white at t_4 will be causal, and the link between the red at t_3 and the red at t_4 will be causal.

---

67 "Events", pp260–1

Someone adopting the interpretation of events in which events are *changes*[68] but who also adopted a Lewis-style counterfactual analysis of causation might take a different view of this scenario. She might hold the link between the red at t_1 and the red at t_2 is not causal since the red undergoes no changes over this period — thus there are no distinct events for there to be a causal relation between.

Under the transference analysis presented thus far, the only causation going on in this scenario is at the point of collision. The only changes that the red or the white are undergoing between t_1 and t_2 are inertial persistences of motion (or perhaps in the case of the red inertial persistences of lack of motion), which are paradigmatic examples of natural changes. But causation only involves objects undergoing unnatural changes.

A digression on relata

Is it a good thing or a bad thing that Aronson's transference analysis holds the there is less causal dependence in the world than the counterfactual analysis would have us believe? Does this point make one analysis comparatively better than the other? The answer we want to give here is that this discrepancy will only matter in so far as it affects the way that the analyses fulfil our basic desideratum. It would not matter that the transference analysis says there is far less causal dependence about, so long as that causal dependence shows up *in the right places*. At this point, we have a *prima facie* case that under the transference analysis causal dependence does show up in the right places and thus we should prefer the transference analysis.

At this point, we should examine an argument that these considerations don't matter a damn when it comes to comparing the relative worth of the two analyses. The argument hinges on the fact that causal dependence under a Lewis-style counterfactual analysis is a relation between *events*, and causal dependence under the transference analysis is a relation between *objects*. So if it is the case that causation is properly a relation between objects, the counterfactual analysis is trivially false; if between events, the transference

---

68 As per, for example, Ducasse, Curt J: "On the nature and the observability of the causal relation", pp125–36, Sosa, Ernest; Tooley, Michael; editors: *Causation* (Oxford: Oxford University Press, 1993).

analysis is trivially false. The question of which is a better analysis resolves to the question of the true relata of the causal relation.

This causal relata "problem of comparison" could be sidestepped fairly easily if we can show that there is no relevant difference between events and objects in causal contexts. But the problem of comparison is supported by both our intuitions that events and objects are ontologically rather different beasties, and the way we talk about events and objects.

Events have locations and durations, objects just have locations. Similarly, events have spatial parts and temporal parts. Consider a battle for control of three hills that lasts for three days. The battle can be broken up into three spatial parts (we can refer to the battles for each of the hills), and can be broken up into three temporal parts (we can refer to the first day of the battle, the second, and the third). On the other hand, objects seem to have spatial parts, but no temporal parts: the *whole* of an object exists at any time the object is present, not just a temporal part of the object; and when we refer to an object we refer to the whole object, not just a temporal part of that object. We can borrow an example of this from Bennett:

> Someone who says "the [Trans-Canada Highway] is one the other side of town" means that a certain (spatial) part of it is there, but if someone says that Vermeer's "Girl in a red hat" is in the National Gallery, we don't take this as a *façon de parler* for the statement that a temporal part of the picture is there, some of its earlier parts being in Delft.[69]

But might this ontological distinction between objects and events be being overplayed? For example, Andrew Newman argues that events are just Fregean objects, in that they occur only once and are not predicated of anything.[70] Thus, any causal relation between events is trivially a causal relation between objects. But this won't do, because the counterfactual analysis we are considering asserts that events are properties of spatiotemporal regions. So events are predicated of regions, and thus these events just *aren't* Fregean objects, they are Fregean concepts. The "concept correlate" of an event — the object we end up referring to when we try to refer to an event — would therefore be a Fregean

---

69 *Events and their names*, p114
70 p530, Newman, Andrew: "The causal relation and its terms", pp529–50, *Mind* vol 98 (1988)

object, but that is no real help. For Newman's thesis to work *events* rather than *event-correlates* need to be objects.

A better response is that our common sense view of objects is just out of date — a relic of our view that time is peculiarly different from space. This view leads us to think that it might be useful to draw a special distinction between things which are extended only through the three dimensions of space (objects) and things which also have an explicit time index (events). But surely we can take the various facts of general and special relativity to show that there is no important distinction between any of the four dimensions of spacetime. Hence if there is any profit to be made from talking of objects at all, we should really consider them as having four-dimensional extension — like events, they are related to four-dimensional spatiotemporal zones; like events they are the sort of things that should be said to have temporal parts.

Ontologically, there are several ways we might cash out this four-dimensional view of objects: one suggestive version is supplied by Bennett:

> ... we can be helped to understand the notion of a thing *in* space if we if we analyze it in terms of qualitative variation *of* space. The basic idea is that for there to be an atom in a given region of space is for that region to be *thus* rather than *so*. This project of understanding the contents of space in terms of the attributes of spatial regions is neutral with respect to time: it could be deployed in terms of regions at instants or regions throughout periods. But ... if for *any* reason we are thinking of the atom as being temporally extended and thus as having temporal parts, then the natural and perhaps inevitable procedure is to analyze the notion of an atom in terms of attributes of *spatiotemporal zones*.[71]

On such a view, objects are pretty much like Lewisian events — perhaps, just *are* events. The only distinction between the events of the Lewis-style counterfactual analysis and objects is that events are properties of spatiotemporal regions, but need not be every property of their particular region, whereas the natural interpretation of an object taken as a qualitative variation of a spatiotemporal region is that the object is *every* qualitative variation in that region. Hence objects are (or are related to) a particular subset of events: those that are full contents of regions, or, to borrow some more

71 *Events and their names*, p117

terminology from Bennett, concrete events.[72] So the problem of comparison seems to be defused.

If we adopt a different analysis of events, we can still defuse the problem. Suppose we held that events are changes. Then we find ourselves arguing that events can not be objects for the following reason due to Donald Davidson:

> Occupying the same portion of space-time, event and object differ. One is an object which remains the same object through changes, the other a change in an object or objects. Spatiotemporal areas do not distinguish them, but our predicates, our basic grammar, our ways of sorting do.[73]

But if events are changes in objects, then there are few distinctions between Aronson's object relata and an event relata. Consider Aronson's first two conditions:

> (1) In "A causes B", "B" designates a change in an object, a change which is an *unnatural* one.
>
> (2) "In "A causes B", at the time B occurs, the object that causes B is in contact with the object that undergoes the change.[74]

On the literal reading of condition (1), effects just are events — they are changes in objects.

The question then is what *A* designates. Is it an object? In this case the transference analysis has mixed relata (objects and events), and the problem of comparison with an analysis which only incorporates events may still stand. But consider the situation where A itself has been caused. As good an example as any is "heating an iron bar to 1000 C causes it to glow". The hot iron being hot is thus a change in the iron bar and (A) thus designates an event. On this reading the apparent object relata of Aronson's transference analysis reduces to an event relata and there is no causal relata problem of comparison with the counterfactual analysis. Obviously, there may be problems for this reduction in instances of first causes — instances where A designates something uncaused.

---

72 *Events and their names*, p103

73 p176, Davidson, Donald: "Reply to Quine on events", pp172–6, LePore, Ernest; McLaughlin, B; editors: *Actions and events: perspectives on the philosophy of Donald Davidson* (Oxford: Basil Blackwell, 1985)

74 "On the grammar of cause", pp421–2

But any analysis of causation with event relata that takes events to be changes in objects will also have peculiar relata problems in statements of the form "A causes B" where A is a first cause. So there doesn't seem to be any special problem of comparison related to cases of first causes.

What happens if we take a less natural reading of Aronson's first condition? Perhaps what Aronson meant to say was "... 'B' designates *an object undergoing a change* ...". This would bring the language of condition 1 into line with that in condition 2, which is full of talk of objects and objects undergoing changes. In this case, following Davidson's distinction, we might take the transference analysis as genuinely having a non-event relata (objects rather than changes), and the problem of comparison arises again.

But this doesn't seem to amount to much of a problem after all. Our earlier result still holds: except in the case of first causes, A and B in statements of the form "A causes B" will refer to objects undergoing changes. But it is impossible for there to be an object undergoing a change without there being a change in an object. So in all standard causal situations events supervene on objects undergoing changes: every causal sequence described in terms of a set of events could be given a parallel description in terms of a set of objects undergoing changes.

This line of argument is followed by Fair in his presentation of the transference analysis:

> The causal relation of ordinary language apparently takes as arguments events, physical objects, (instantiated) properties, actions, mental states, and facts ... We must reduce this ontology so described to an ontology of physical objects described in terms of physical magnitudes in order to be able to apply the physics ...
>
> The ontological aspect of this reduction is often fairly straightforward; we concern ourselves with the physical objects *comprising* what are described as events, actions, mental states, and facts and with the objects *exemplifying* what are described as properties.[75]

[75]"Causation and the flow of energy", p233

So the distinction between object relata and event relata is innocent. It is just a matter of *terminological* choice whether we present our analysis in terms of events, objects, or both.[76]

The problem of comparison is easily circumvented: in the worst-case scenario we just translate all the object talk into event talk (or vice versa) and see how the analyses square off in particular cases. And as we have seen, the transference analysis so far appears to be improvement over the counterfactual analysis when it comes to fulfilling our basic desideratum.

Developing the transference analysis

Some work needs to be done to tighten up the basic transference analysis supplied in "On the grammar of cause". First, we should be more precise about which quantities are being transferred. Aronson notes that the transference analysis is associated with conservation principles:

> When change is thought of in quantitative terms, it becomes quite natural to ask for the origin of that change (or new quantity). So in seeking the cause of a change that takes place, we are essentially seeking its source or contributor. But this is another way of saying that quantities (eg momentum and energy) within a system are conserved.[77]

If this is the case, then velocity cannot be one of the transferred quantities. Consider a billiard ball striking another object and stopping. If velocity were conserved, that object would move away at the velocity of the billiard ball, regardless of the mass of the object. This is not what we observe in collisions. What is conserved in collisions is momentum: the product of mass and velocity of objects involved in the collision.

---

76 This argument resembles a view of Reichenbach's reported by Anthony Quinton — that every ordinary singular statement in which reference is made to an individual can be interpreted in terms of objects or events. For example, "George VI was crowned in Westminster Abbey in 1937" (where the property of being crowned at a certain time and place is ascribed to George VI) could equally well be reported in terms of "event-properties" ascribed to events, such as "An event that happened in Westminster Abbey in 1937 has the event property of being a crowning of George VI". See pp198–9, Quinton, Anthony: "Objects and events", pp197–214, *Mind* vol 88 (1979)

77 "On the grammar of cause", pp426–7

Another criticism is due to Fair. Under the transference analysis, the asymmetry of the causal relation is due to the transference of a quantity *from* the cause *to* the effect. If we are indiscriminate about what sort of quantities are being transferred, we will find the asymmetry pointing in the wrong direction. Define the quantity "negenergy" as the energy lost in an interaction. The direction of causation governed by transference of this quantity is from effect to cause.[78] We follow Fair in supposing that the relevant quantities for transference are energy and momentum (both linear and angular), which *prima facie* are transferred from cause to effect, and are also governed by conservation principles. Thus we should modify Aronson's conditions for causation:

(1) In "A causes B", "B" designates an unnatural change in an object

(2) In "A causes B", at the time B occurs, the object that causes B is in contact with the object that undergoes the change

(3) Prior to the time of the occurrence of B, the body that makes contact with the effect object possesses a quantity (energy or momentum) which is transferred to the effect object (when contact is made) and manifested as B.

We shall call a transference analysis which fits these modified conditions an Aronson-Fair analysis.

A further amendment also seems advisable. The issue in question is the physical scale at which causation occurs. The conditions obeyed by an Aronson-Fair analysis refer to unnatural changes in *objects* and energy or momentum transfers from one *object* to another *object*. Yet objects vary dramatically in physical scale. An electron is an object. A billiard ball is an object. Frank is an object. The Moon is an object. A galaxy is an object. The universe is an object. All these objects undergo unnatural changes.[79] So all these objects enter into causal relations — presumably, the same sort of causal relations. So Aronson-Fair, as developed so far, is what we will call a *scale-invariant* analysis of

78 "Causation and the flow of energy", p239

79 Arguably, even the universe. Some cosmological theories posit that our universe is structurally stabilised by being joined by wormholes to other universes. Energy levels in the various universes are balanced out by energy transfers through the wormholes. If these theories are correct, then our entire universe undergoes unnatural changes — it is interacting with the other universes.

causation. Scale-invariant means simply that the sorts of causal relations hold between the relata of the analysis at every physical scale the relata can occur at.

But scale-invariance surely poses a problem for an analysis of causation in which causal dependence tracks to physical dependence. Few if any physical phenomena manifest themselves at such a maximally wide variety of physical scales as, for example, the scattering of an electron off an electron and the scattering of a galaxy off another galaxy. Is the same sort of energy or momentum transference *really* taking place in the case of the electrons and the case of the galaxies?

We are faced with a dilemma: either one of these cases is causal and the other isn't (which we want to deny), or the notion of energy and momentum transference Aronson-Fair appeals to is massively disjunctive (which is also unattractive).

Well, perhaps the notion of energy and momentum transfer is not all that disjunctive. Maybe we could get by with *two* separate notions, drawn from two sources familiar to the educated layperson: quantum theory and classical relativistic theory. At the scale of the electron scattering off the electron, energy and momentum transfers are governed by the rules of quantum physics: in this case, primarily by quantum electrodynamics (the fundamental theory of the electromagnetic interaction). At the scale of the galaxies colliding, the energy and momentum transfers are governed by relativistic mechanics.

(At the scale of billiard balls and Frank we typically apply the rules of Newtonian mechanics, but acknowledge that what is really governing the transferences is relativistic mechanics. The use of Newtonian mechanics is driven by purely epistemic motives — it is a usefully precise approximation in most cases involving billiard balls and people, and it is considerably more convenient to use than relativistic mechanics. We do not use it because of an ontological motivation: for example, that at the scale of billiard balls Newtonian rules apply, and relativistic rules do not.)

What seems more likely is that we have to make do with *one* notion of energy and momentum transfer: the notion drawn from quantum theory. This conjecture is argued for by Adrian Heathcote in his 1989 paper "A theory of causality: causality = interaction (as defined by a suitable quantum field theory)".[80]

Heathcote supports the general principles of the Aronson-Fair transference analysis, but believes there are problems about which notion of energy or momentum transference is being appealed to in the analysis, and some ambiguity over the scale that causation really takes place at.[81] Heathcote presents the case that an adequate analysis of causation needs to fulfil certain desiderata: that there should be no causal paradoxes (or that the fabric of spacetime is causally "well-behaved"), that causal influence does not propagate faster than the speed of light, and that causal influence always flows from cause to effect, from past to the future. In a formidable feat of erudition, Heathcote then sets out to demonstrate that an analysis of causation in which causal interaction reduces to the physical interactions defined by a suitable quantum field theory fulfils these desiderata (or at least does not obviously contradict them).[82]

We will not argue for Heathcote's position. We don't feel *qualified* to argue the point. We shall just accept Heathcote's conjecture that it is possible to produce an Aronson-Fair-style transference analysis in which the notion of energy transference appealed to is drawn from quantum theory.

But what of the relativistic notion of energy transference we would like to adopt for macroscopic causal phenomena such as billiard balls, Frank and galaxies? Heathcote, following Douglas Currie, argues that such a notion of energy transference can not provide a satisfactory analysis of causation:

> Take a theory of relativistic classical particles (not based on a [quantum field theory]); one might think that it would be possible to have the particles interact by collisions. Since the theory is relativistic we would expect that all of the intuitive requirements that we placed on causality would hold, ie, that causal propagation would be into the *interior* of the *forward* light cone. One would then not need to tie causality to field

80 Heathcote, Adrian: "A theory of causality: causality = interaction (as defined by a suitable quantum field theory)", pp77–108, *Erkenntnis* vol 31 (1989)

81 "A theory of causality: causality = interaction (as defined by a suitable quantum field theory)", p84

82 "A theory of causality: causality = interaction (as defined by a suitable quantum field theory)", pp82-99. In fact, our discussion will show in passing that the third desideratum is actually violated by the quantum energy/momentum transference model.

> theories in particular. However, though the relativity gives the correct schematic restrictions on causality to avoid paradoxes, the particle nature of the theory turns out to prevent a proper definition of causality. The relevant theorems were proven by Douglas Currie in a paper in 1963[83] ... — a paper which deserves to be better known by both philosophers and mathematicians. Essentially Currie showed that in the relativistic Hamiltonian formulation of classical particle mechanics particles cannot interact — their world lines are all straight, they cannot be accelerated and hence their velocities are constant. In short relativistic Hamiltonian theory can only apply to free particles. Hence one cannot think of relativistic particles affecting one another, therefore there is no causality. This is a very surprising conclusion. The only way out is to go over to a quantum theory of fields.[84]

Once again, we shall not argue for this position. We shall just accept that it is the case. Rather unfairly — since the failure to substantiate the point is ours — we shall refer to the thesis that the only suitable notion of energy and mass transference is that provided by quantum theory as the *Heathcote conjecture*.

We believe that an adequate Aronson-Fair transference analysis must respect the Heathcote conjecture.[85] There is no disjunction between the notions of energy or momentum transference at different scales, there is only quantum energy/momentum transference.[86] Since transference only occurs at the

---

83 The reference given by Heathcote is to Currie, Douglas: "Interaction *contra* classical relativistic Hamiltonian particles mechanics," pp1470–88, *Journal of mathematical physics* vol 4 (1963)

84 "A theory of causality: causality = interaction (as defined by a suitable quantum field theory)", p103

85 And, probably, so does Aronson. See "Conditions versus transference: a reply to Ehring", p250, where Aronson claims that "... the transference model is committed to reducing complex, "macroscopic" cases of making things happen to combinations of fundamental processes of transference." On the following page, the examples given of these fundamental processes are the exchange of messenger particles as governed by the forces described by quantum field theories.

86 It makes most sense to talk of energy/momentum transference rather than energy transference or momentum transference once we adopt the quantum notion of transference. Under quantum theory, the momentum of a particle is generally related to the energy of a particle by the equation:

$$E^2 = m_{rest}^2c^4 + p^2c^2$$

where "E" is energy, "m_{rest}" is the rest mass of the particle, "p" is momentum and "c" is the velocity of light. But momentum is possessed by all entities in the quantum realm, including wavelike objects with zero rest mass such as photons. The momentum of such an entity is determined by the equation:

quantum scale, an adequate Aronson-Fair analysis is not scale-invariant. Causation properly takes place only at the quantum scale.

Does this fling us back onto the other horn of our dilemma, that the collision of galaxies (and indeed the bonking of Frank and the gentle clack of billiard balls on green baize) is not causal under an adequate Aronson-Fair analysis? We say no. The solution must be to accept that when we refer to an everyday unnatural change, such as a white striking a red, we are not directly referring to a single, simple causal relation as our commonsense talk of changes and objects might lead us to believe. The causal dependence will be the sum of a very complex set of quantum-scale transferences.

First consider the case of the scattering of an electron from another electron. Two electrons approach each other, each with some range of momentum and energy.[87] The electrons then interact via the electromagnetic interaction. As a first approximation, this involves the exchange of a virtual photon (the quantum of the electromagnetic field) between the two electrons. Two electrons then move apart from one another, with different energy/momenta than the two approaching electrons.

The situation is more complex in the case of the billiard balls because billiard balls are not solely made up of electrons. The atoms comprising the molecules

---

$$p = \frac{h}{\lambda}$$

where "h" is Planck's constant and " λ " is the wavelength of the photon (or other such wave). In such objects, the relation between energy and momentum is even simpler:

$$E=pc$$

Thus any transference of energy at this scale will also involve a transference of momentum, and it seems pointless to continue drawing a distinction between the transference of energy and the transference of momentum.

87 We cannot be more precise than this. The Heisenberg uncertainty principle:

$$\Delta x \Delta p > h/4\pi$$

(where Δx is the uncertainty in position of a particle, Δp is the uncertainty in momentum and h is Planck's constant) dictates that if we precisely specify the position of a particle such as an electron then we cannot precisely specify the momentum of that particle. This is not due to technological limitations on our capabilities to observe the electron's position. There simply is no fact of the matter about the electron's precise momentum within a small enough area of space. (See pp54–7, Davies, Paul: *The forces of nature* (Cambridge, UK: Cambridge University Press, 1979) comprising the billiard balls contain many other electrically-charged particles, such as protons, which also interact with the electrons via the exchange of virtual photons. We will only mention the electrons in the description of the white striking the red — this will be sufficient to demonstrate the complex flavour of the dependence.

At the initial "point" of interaction at the initial "instant" of contact of the white and the red, the electrons in the outer surface of the red and the outer surface of the white exchange virtual photons. The resultant energy/momenta of electrons in the outer surface of the red and the white brings these electrons into closer proximity with electrons deeper inside the balls and further exchanges of virtual photons take place. The resultant energy/momenta of electrons in the deeper levels brings them closer to electrons at even deeper levels, and so forth. A cascade of energy/momentum transfers occurs through the electrons in the red.

Behind this initial wave of energy/momentum transfers, more and more energy/momentum is being pumped into the red through the point of contact with the white by the continued close proximity of the electrons in the white to the electrons in the red. A similar state of affairs is pertaining in the white. Eventually, enough energy/momentum is transferred by the electromagnetic interactions between the electrons in the balls in order that the statistical sum of energy/momenta of the electrons within the balls carries the balls out of contact.

Note that an Aronson-Fair analysis which respects the Heathcote conjecture must relax conditions 2 and 3, insofar as those conditions require the relevant objects to be in contact while causation is taking place. What is instead required is that the objects are exchanging quanta, such as virtual photons. This is not a bad thing for the analysis. Consider the north poles of two magnets being pushed together. The magnets repel each other, but do not need to come into contact with each other for repulsion to take place. If we were to strictly apply the contact clauses of the conditions, we would not be able to say that this was a causal phenomenon. If we specify that causation is the linking exchanges of the quanta of the electromagnetic field — in this case, virtual photons — the analysis still allows this phenomenon to be causal, which seems correct.

So it *is* possible to give an account of the macroscopic causal phenomena we observe purely by reference to the quantum notion of energy/momentum transference: we merely have to refer to many, *many* separate instances of transference. Even if the Heathcote conjecture is true, the causal status of macroscopic physical interactions is preserved.

In fact, by limiting our notion of energy and momentum transference to quantum energy/momentum transference we gain a much clearer causal analysis of macroscopic interactions than we would otherwise. Consider the bonking of Frank. If we try to interpret this example with an Aronson-Fair analysis which appeals to a Newtonian or relativistic notion of energy or momentum transference, it would make sense to treat the relevant objects (Frank, the Moon and the Earth) as if they were single mass points. Hence we would conclude that the *entire* Moon and the *entire* Earth play a causal role in Frank's crushing. But for causal dependence between the Moon and Frank to track physical involvement between the Moon and Frank, causal involvement should be limited to the contact zones of the Moon and Frank (while there still is something that we could dignify with the name Frank). It isn't clear why this should be the case under an Aronson-Fair analysis which does not accept the Heathcote conjecture. There is no such problem if we limit the notion of energy and momentum transference to that of quantum theory. The contact zones propagating through the relevant objects *just are* the quantum-scale cascades of energy/momentum transfers between charged particles inside the Moon and Frank.

So, in summary, an adequate transference analysis of causation is most likely to be an Aronson-Fair analysis which respects the Heathcote conjecture.

A (mostly) surmountable problem

Unfortunately, even an Aronson-Fair analysis which respects the Heathcote conjecture faces some significant problems. Some can be overcome; some we can be less sanguine about.

The first version of a problem which can *mostly* be overcome is originally raised by Fred Dretske and is discussed by Aronson in "On the grammar of cause". In ordinary language we often say things cause other things to stop. This might seem to cause problems for a transference analysis. If a motor scooter runs into a wall and is stopped, what did the wall transfer to the scooter?

Even at this early stage of development of the analysis — when many more physical quantities than energy or momentum might instantiate a causal transfer — Aronson finds lacking the answer that inertia or rest are transferred. Inertia or rest don't seem to be the sort of quantities that could be possessed or passed on. At first Aronson argues that the motor scooter hitting the wall is an instance of causation, but the actual direction of causation is opposite to that suggested by language: the scooter actually transfers kinetic energy to the wall in the form of heat and deformation in the wall. But then Aronson argues that the wall also transfers momentum to the scooter. In fact, in cases of collisions between two bodies, A and B, causal relations come in pairs: one runs from cause A to effect B, and another from cause B to effect A.[88]

It seems easier to understand this defence in an adequate Aronson-Fair analysis. Go back to the white striking the red. As a working hypothesis it seems reasonable to say that the virtual photons that instantiate the transference of momentum between the electrons at the point of contact could just as readily have originated from an electron in the red as they do from an electron in the white.[89] Similarly, throughout the virtual photon cascades in the red and the white, some virtual photons will be originating from the deeper electrons, and some from the shallower electrons. So it really does appear that we have quantum transferences proceeding in both directions, which we can interpret as summing to two macroscopic transferences, one from the white to the red and one from the red to the white. If the truth of our working hypothesis is granted, then Aronson's solution appears to be correct.

John Earman extends the basic Dretske problem.[90] Earman notes that the transference analysis attempts to give an objective basis to the asymmetry of causation which is independent of anthropomorphic manipulation. Therefore, the analysis should be observer independent. But special relativity shows us that the direction of transference in collisions *is* observer dependent.

---

88 "On the grammar of cause", p424

89 If this seems like an unnecessary equivocation, it is because we shall soon argue that it is *not* reasonable to accept this hypothesis.

90 Earman, John: "Causation: a matter of life or death", pp5–25, *Journal of philosophy* vol 73 (1976)

We can illustrate this with our billiard balls. Consider an observer moving in the same frame of reference as the white before impact. She will observe a stationary white, which a red moves towards and knocks away. For that observer, the red has transferred momentum to the white. Consider another observer moving in the frame of reference of the red before impact. He will see the white move up and knock the red away. For that observer, momentum will be transferred from the white to the red. Hence whether the white caused the red to move or the red caused the white to move will be observer dependent.

There are several responses that transference analysts have or could make to this example. One might be to question whether observer dependence really does conflict with there being an objective basis to causal asymmetry. Similarly, Fair's response is to bite the bullet and accept Earman's contention that the asymmetry of momentum and energy transfers is observer-dependent in relativistic contexts. But Fair argues that this phenomenon supports the transference analysis, since intuitively it seems that causal asymmetry is also observer-dependent in these contexts.[91]

Appeals to intuition are tempting, but intuition is ever a fickle guide to such matters. Heathcote and Aronson prefer to argue against Earman's contention. Heathcote does not deal with the objection directly, but appears to think that the problem will be solved by adopting the correct notion of energy/momentum transference.[92] Aronson's response is along the same lines as his solution to the Dretske problem: causal relations come in pairs, and a body at rest is quite capable of transferring energy or momentum to a body in motion.[93] We think that so long as our working hypothesis about the origins of virtual quanta in quantum interactions is true, Aronson's reply is correct for an adequate Aronson-Fair analysis.

The basic idea behind Dretske and Earman's objections appears to be that in the case of macroscopic bodies, the classical momentum of a body (p) is given by the equation:

$$p = mv$$

91 "Causation and the flow of energy", p240

92 "A theory of causality: causality = interaction (as defined by a suitable quantum field theory)", p84

93 "Untangling ontology from epistemology in causation", pp298–9, and pp57–9, Aronson, Jerrold: *A realist philosophy of science* (New York: St Martin's Press, 1984)

where "m" is the body's mass and "v" is the body's velocity. In the cases of the wall struck by the scooter, the white in the white's frame of reference and the red in the red's frame of reference, the objects mentioned have a velocity of zero, so by the equation they have a momentum of zero. If so, how could they transfer momentum to something else?

Well, the answer is supplied by the quantum picture of the electromagnetic interaction. Electrons in any motion state are surrounded by a swarm of virtual photons.[94] Just because one electron in an interaction is more or less at rest and the other is more or less in motion, it doesn't mean that the virtual photon transferring momentum between the electrons comes from the electron more or less in motion. The macroscopic sum of these exchanges add up to there being a causal dependence of the body in motion on the body at rest as much as there is a dependence of the body at rest on the body in motion.

So we appear to have successfully fended off the Dretske-Earman objections and defended Aronson's contention that there are paired causal dependences in instances of collisions. But there is a bug in our argument. Our working hypothesis suggests that virtual photons transfer energy/momentum *from* one electron *to* another. There is a temporal directionality of transfer that underlies the causal asymmetry of cause and effect. Yet there is no such directionality involved in the exchange of a *virtual* photon in an electromagnetic interaction (nor in the exchange of any *virtual* particle mediating one of the other quantum interactions).

Some quantum interactions do display the required sort of directionality for Aronson's causal asymmetry. Consider the glowing of the heated iron bar. Here excited electrons decay from higher to lower energy orbitals[95] around the iron atoms. This produces *real* photons which carry energy/momentum away from the iron bar, and then knock electrons in the retina of our eyes from lower energy orbitals into higher energy orbitals. Enough of this sort of thing produces the macroscopic phenomenon of the hot iron bar glowing. Real

---

94 *The forces of nature*, pp124–7

95 An orbital is a likely position that an electron can be found in around an atom. It is not a precise location: an electron in any orbital could be found anywhere in the vicinity of the atom, except, typically, at the centre of the nucleus (provided restrictions of the Pauli exclusion principle are respected). But different orbitals dictate different volumes of space that the electron is more or less likely to be in. Electrons more between orbitals by virtue of emission or absorption of photons.

photons have enough of an independent existence from the objects that emit or absorb them for us to profitably describe as moving from or to objects. Hence we can refer to them as being "emitted from" or "absorbed by" without doing too much damage to the language we are describing the phenomena with.

Virtual photons, on the other hand, do not lead such an independent existence. Their lifetime is governed by the Heisenberg uncertainty principle and can be calculated by the formula:

$$\Delta t \approx \frac{h}{\Delta E} \approx \nu^{-1}$$

where " ν " is the frequency of the photon.[96] In this time, the virtual photon is able to move a distance equivalent to one of its own wavelengths before it is either captured by another charged particle (producing the electromagnetic interaction we have been discussing within the billiard balls) or is reabsorbed by the emitting particle (producing the phenomenon of a charged particle interacting with its own electromagnetic field, known as "self interaction").

Unfortunately, over this timescale it is impossible to determine which charged particle is "emitting" and which charged particle is "absorbing" the virtual photon.[97] As per the constraints imposed by the Heisenberg uncertainty principle on knowing the precise location and the precise momentum of a quantum-scale particle, this failure to distinguish which particle emits the photon and which absorbs the particle is not due to technological constraints on our ability to conduct precise enough observations to distinguish the difference. There simply is no fact of the matter about which particle emits and which particle absorbs the virtual photon.[98]

So the grounds for Aronson's contention that in cases of macroscopic collisions we find equidirected pairs of asymmetric causal relations are undermined. An adequate Aronson-Fair analysis should be more inclined to

---

96 *The forces of nature*, pp124–6

97 *The forces of nature*, pp127–8

98 Another way of expressing this point is that over small distances photons have a propensity to travel faster or slower than c (c is interpreted as the *average* speed of light in a vacuum). See p96, Feynman, Richard: *QED: the strange theory of light and matter* (Princeton: Princeton University Press, 1985).

state that there is a symmetrical causal relation between the two colliding objects.[99]

It might be objected that some of the interactions between the charged particles in the billiard balls might involve real photons. These interactions are truly directional, and could sum to equidirected asymmetric causal dependences. (By this line of argument, there are *three* causal relations afoot in a collision: two equidirected asymmetric dependences and a symmetrical interdependence.) Yet asymmetry would still not be guaranteed in causal dependence. Consider causal phenomena such as radioactive decay,[100] which are based on the weak interaction rather than the electromagnetic interaction. This interaction is not mediated by exchanges of photons, but rather by three other bosons: the W^+, W^- and Z^0. In our universe, the interaction is almost exclusively mediated by virtual versions of these bosons. Real versions of these particles are extremely unlikely to appear of their own accord, and are extremely difficult to produce by artificial methods.[101] So virtually all radioactive decay involves causally symmetric dependence.

So we are forced to retrench the adequate Aronson-Fair analysis slightly: it does not provide quite the account of causal asymmetry in simultaneous causation that it was intended to provide. Is this a major problem?

Perhaps not. We still have an account of causal asymmetry in the paradigm case of the hot iron causing a glow, because this involves real photons. We have a solution to the Dretske-Earman examples, because the existence of a symmetrical causal relation between rest object and moving object gives as good a grounds for the rest object having causal influence on the moving object as the existence of an equidirected pair of asymmetric causal relations.

Aronson might be concerned that such a retrenched analysis would not provide an adequate basis for a theory of scientific explanation. We want to say that the hot iron explains the glowing, but we do not want to say that the

---

99 So to some extent Heathcote's desiderata for an analysis of causation are violated by the quantum notion of energy/momentum transference. In some causal interactions — those involving virtual quanta — there is no fact of the matter about whether causal influence propagates into the future or the past.

100 Cited by Aronson as a causal phenomenon: see "Conditions versus transference: a reply to Ehring", p251

101 For a good account of the difficulties involved, see Watkins, Peter: *Story of the W and Z* (Cambridge, UK: Cambridge University Press, 1986)

glowing explains the heat of the iron. The existence of a causal asymmetry between the heating and the glowing provides rather a nice ground for the explanatory asymmetry. If most causal dependence turns out to be symmetrical, might we not lose our account of explanatory asymmetry?

Well, consider the cases where the causal dependence is symmetrical. The red and the white cause each other to move. But in this case we are inclined to believe that the impact of the red explains the motion of the white as much as the impact of the white explains the motion of the red. Consider the weak interaction. Does the transmutation of a proton into a neutron explain the nearby transmutation of an electron into a neutrino, or vice versa? Either explanation seems fine. Given that explanatory symmetry is manifest in these cases, perhaps causal symmetry doesn't amount to much of an problem for Aronson on these grounds.

Bouncing Frank

The next problem is considerably less clearly overcome by an adequate Aronson-Fair analysis, and certainly not by an Aronson-Fair analysis which does not respect the Heathcote conjecture. The argument is that an Aronson-Fair analysis gives an inadequate account of "causal connectivity" for an analysis of causation in which causal dependence tracks physical dependence.

Let us consider a similar problem case to the bonking of Frank. This time we are interested in the *bouncing* of Frank. As in the case of the bonking of Frank, the Moon (securely wrapped in duct tape) strikes the Earth. This time, rather than standing under the impact point, Frank is on the other side of the Earth, directly opposite the point of impact, steadily leaping up and down.

Now clearly the Moon striking the Earth counts as an instance of causation on an Aronson-Fair analysis which does not respect the Heathcote conjecture. Both bodies undergo unnatural changes, both bodies are in contact, and both bodies transfer energy and momentum to one another. Frank's bouncing counts as an instance of causation. The surface of the Earth is holding Frank up, preventing him from sinking to the centre of the Earth under the influence of the Earth's gravitational pull. The soles of Frank's boots are doing a similar job, preventing the Earth from sinking to the centre of Frank. At the beginning of Frank's leap, the Earth and Frank are in contact. When Frank leaves the ground, a small quantity of energy and momentum has been transferred to Frank by the Earth, and a near equal small quantity of energy and momentum has been transferred to the Earth from Frank.

Now consider two cases. In the first, an observer notes that Frank's last leap occurs after the Moon has struck the Earth, but just *before* the growing contact zone resulting from the impact has encompassed the part of the Earth that Frank is leaping from. In the second, the observer notes that Frank' last leap occurs *just fractionally after* the contact zone encompasses the point from which Frank is leaping, ("Just fractionally after" is intended to capture the circumstance that there is still a reasonably contiguous surface of the Earth for Frank to be leaping from, and the additional impulsive loading supplied by the Earth to Frank is not so great that it interferes with the leaping process by, for example, breaking Frank's legs.)

The physical circumstances of Frank's last leap will be considerably different in each case. In the first case, so long as Frank is putting pretty much the same effort into the leap as in the previous leaps, the distance leapt vertically will be pretty much the same as in each of the previous leaps. In the second case, the Earth that Frank leaps from contributes much more momentum and kinetic energy to Frank than it did in the previous leaps. Effectively, the Earth acts like a giant springboard. Frank will travel considerably further vertically upwards on his final bounce. (He will not travel further downwards. Unlike a normal springboard, the Earth — or, as is more likely, the remains of the Earth — will be rapidly rising up behind Frank and will swat his shattered body from the sky fairly soon after he begins to fall back.)

So we have a dramatic difference between these two instances of the bouncing of Frank. It is fairly clear why there is a difference in height in Frank's last bounce in each case. In the big last leap case the impact of the Moon on the Earth is physically contributing to the height of Frank's leap by adding to the impulsive loading supplied by the Earth on Frank, and in the normal leap case, the Moon's impact *does not* physically contribute to the impulsive loading supplied to the Earth. To follow the basic desideratum that causal dependence tracks physical dependence, we should insist on saying that the impact of the Moon is causally *connected* to Frank's last bounce in the second case, and not causally connected in the first.

Unfortunately, the basic conditions of an Aronson-Fair analysis don't allow us to say this. No explicit guidance is provided on the circumstances under which it is correct to say "given A is causally connected to B and B is causally connected to C, then A is causally connected to C". (For brevity's sake, from now on we will symbolise "is causally connected to" as "$\rightarrow$".)[102]

There may be implicit guidance on this issue. One natural interpretation of the Aronson-Fair conditions is that A is *never* causally connected to C, because the B that is being appealed to in "A $\rightarrow$ B" is different from the B that is being appealed to in "B $\rightarrow$ C". Consider the bouncing of Frank. The analogue of "A $\rightarrow$ B" is the Moon striking the Earth, transferring momentum and energy and producing rather violent changes in the Earth — the various impact phenomena we have discussed in earlier sections. The analogue of "B $\rightarrow$ C" is the Earth transferring momentum to Frank in the course of his leap. So in the first instance B refers to *a change in an object* (the impact phenomena experienced by

---

102 Lest we be accused of multiplying philosophical terminology unnecessarily, note that a failure of causal *connectivity* is not the same thing as a failure of causal *transitivity*.

Under a transitive analysis of causation (such as a Lewis-style counterfactual analysis), if we have the causal relation A $\rightarrow$ B and the causal relation B $\rightarrow$ C, then we have the causal relation A $\rightarrow$ C. It has been argued that causal relations can fail to be transitive in this manner. To paraphrase an example recently used by Douglas Ehring, let A be Davidson putting potassium salts in a fireplace, let B be a purple fire in the fireplace a little later, and let C be the death of Elvis as a result of the fire still later on. Davidson's actions cause the purple fire, so we have A $\rightarrow$ B. The purple fire causes Elvis's death, so we have B $\rightarrow$ C. So by transitivity we have A $\rightarrow$ C — Davidson's actions caused Elvis's death. But Ehring argues that this is absurd. All Davidson did was make the flames purple. Elvis would have burned to death even if the fire wasn't purple. So Davidson isn't a cause of Elvis's death (*Causation and persistence: a theory of causation*, pp76–7). Michael McDermott provides a similar case: My dog bites off my right forefinger. The next day I set off a bomb by pressing the detonator button with my left forefinger — but I would have used my right forefinger if I had a right forefinger to use. The dog bite caused me to use my left forefinger. Pressing the button with my left forefinger caused the bomb to explode. But the dog bite isn't a cause of the explosion. (See McDermott, Michael: "Redundant causation", pp523–44, *British journal for the philosophy of science* vol 46 (1995).)

Neither of these putative cases of failure of transitivity involve failures of connectivity. The potassium salts *did* make the flames purple. The purple flames *did* burn Elvis to death. My right forefinger *was* missing when I pressed the detonator button. What we have is rather a failure of *salience*: the purpleness was present in the flames but was not at all efficacious in bringing about Elvis's death; the finger was missing, but this was not efficacious in bringing about the explosion. Indeed, failures of transitivity presuppose that connectivity is satisfied in a given situation. If the purple fire had burnt out and a subsequently lit non-purple fire had killed Elvis, we would not be worried about a failure of transitivity (similarly, if the dog had bitten off my finger, but someone else had pressed the button the following day). Under these non-connective circumstances, the question of transitivity does not even arise.

We will not insist on adequate transitivity. *Je n'ai pas temp. Je n'ai pas temps.* But we will insist that an analysis of causation must be adequately connective.

the Earth) and in the second instance B refers to *the object itself* (the Earth). Given this ambiguity between references to B, we conclude that there is no causal connection between A and C under these circumstances, or indeed under any circumstances, since such an assertion would involve some ambiguity over the nature of the intervening B.

This interpretation of connectivity renders Aronson-Fair style transference analyses *inadequately connective*. We would conclude that in both cases of the bouncing of Frank the impact of the Moon would play no causal role in the height of Frank's last leap. But clearly there is a difference in each case. This difference is also clearly linked in one case to a physical contribution of the impact of the Moon to the leap and in the other case to a lack of such a contribution. So causal dependence comes adrift from physical dependence. So the basic desideratum that transference analyses are supposed to preserve is violated.

This interpretation of causal connectivity in Aronson-Fair is closely linked to the interpretation of causal relata in Aronson-Fair analyses which states that causal relations have mixed relata: the antecedent of the relation is an object, the consequent of the relation is a change in the object. But as we have argued earlier, we need not adopt this interpretation of causal relata. *Both* A and B in "$A \rightarrow B$" could be interpreted as changes in objects, or, equivalently, as objects undergoing changes. The same is true of B and C in "$B \rightarrow C$".

So under this interpretation of relata, we appear to have rescued causal connectivity in the case of the bouncing of Frank in which the contact zone of the Moon-Earth collision has spread to Frank's jumping-off point. The Moon (A) is causally connected to the Earth (B), which we can treat of as an object undergoing change (the various impact phenomena). The Earth (B), an object undergoing change in the form of the impact phenomena is causally connected to Frank (C), another object undergoing change (the leaping). There is no ontological ambiguity in the nature of B in the two circumstances "$A \rightarrow B$" and "$B \rightarrow C$", so there is no bar on "$A \rightarrow C$".

Unfortunately, such an interpretation makes Aronson-Fair analyses *excessively* connective. The fact that the B term in $A \rightarrow B$ and $B \rightarrow C$ refers to one and the same change in an object (or one and the same object undergoing one and the same change) *cannot* be sufficient to establish $A \rightarrow C$ if we wish to preserve the basic desideratum that causal dependence tracks physical dependence. For the purposes of demonstration, suppose that such a fact *is* sufficient. Now consider the case of the bouncing of Frank where the contact zone of the Earth-Moon impact has not reached Frank's jumping-off point. Clearly we have $A \rightarrow B$ (the Moon-Earth causal connection) and $B \rightarrow C$ (the Earth-Frank causal connection). B in each case is one and the same object undergoing one and the same change (the Earth undergoing various impact phenomena). So, by hypothesis, the impact of the Moon *is* causally connected to Frank's last leap. But in this case there is no physical contribution made by the impact of the Moon. Once again, causal dependence has come adrift from physical dependence. The appeal to the presence of one and the same intervening object undergoing one and the same change is too blunt an instrument to adequately establish connectivity between different causal dependences in a transference analysis.

The moral is that the basic Aronson-Fair analysis lacks the resources to give the correct answers to questions of causal connectivity. The case of the bouncing of Frank shows us that in some cases a later physical dependence is connected to an earlier physical dependence, and in some cases it is not. On one reading of Aronson-Fair, there are *no* corresponding causal connections between these dependences. On the other plausible reading of Aronson-Fair, there are *too many* causal connections. So as it stands, Aronson-Fair fails the connectivity requirement for an analysis of causation motivated by the basic desideratum that causal dependence should track physical dependence.

To rehabilitate Aronson-Fair we need to provide an explicit account of causal connectivity that captures the difference between the two cases of the bouncing of Frank. To preserve our basic desideratum that causal dependence tracks physical dependence, this account should be sensitive to the physical differences between the two cases. The physical difference in this case is very clear. In the case of Frank's big last leap, the spreading contact zone of the impact of the Moon with the Earth has reached Frank's jumping-off point before Frank's last leap. In the case of Frank's normal last leap, the contact zone has not reached the jumping-off point before Frank's last leap.

Can we rescue the situation by moving to an Aronson-Fair analysis which respects the Heathcote conjecture? As we saw earlier, such an analysis seems to tie the progressive causal involvement of macroscopic bodies in collision to the progressive propagation of impact phenomena through those bodies, which appears to be precisely what we need to solve the connectivity problem. But our concern is that this might trade the macroscopic connectivity problem for a quantum-scale connectivity problem.

It seems plausible to us that when electrons are in a particulate mood, they behave like they are single mass points. Intuitively, we would claim causal influence must always be able to propagate through a point, connecting an interaction on one side of the point with an interaction of the other side of the point. Yet the factors which underlie interaction in these circumstances, such as the Heisenberg uncertainty principle, suggest that there might not be any such fact of the matter about the quantum objects connecting the interactions. Worse still, if the electron is in a *wavy* mood, and not behaving like a single mass point, would a disturbance in one area of the waveform be immediately connected with a disturbance somewhere *else* in the waveform? If not, the electron — and other similar quantum particles — might be subject to the same connectivity problems as any macroscopic object.

We are acutely aware that a satisfactory answer to these problems might involve a satisfactory solution to the various measurement problems of quantum theory. Given that we are in no position to offer such a solution (and as a matter of general principle look somewhat askance at philosophers who *do* offer such solutions), we are reluctant to claim that an Aronson-Fair analysis which respects the Heathcote conjecture can provide a satisfactory solution to the connectivity problems.

The genidentity of energy/momentum

Perhaps there is a relatively simple solution to the connectivity problem. In another 1971 paper by Aronson, "The legacy of Hume's analysis of causation", Aronson places another constraint on causal dependence.

> But ... is there an element to the causal relation in addition to mere sequence of events? The answer is "Yes!" — the causal relation is more than sequence in that the cause and effect are not only objects or events which are constantly conjoined, but the cause object [*A*] possesses a quantity (momentum, energy, force, etc) which is transferred to the effect object [*B*] ... *A*'s momentum was transferred to *B* at the instant

> it made contact with *B*. Since such a process involves numerical identity of that particular quantity throughout the sequence of events, it can be seen that, ontologically, this type of causation is more than mere sequence of events.[103]

"Numerical identity" is intended to mean something more than that an equal amount of the quantity appears in one form or another throughout the relevant process.[104] Aronson means that the additional momentum acquired by B is *one and the same* momentum lost by A. This is the same notion of identity, or "genidentity" appealed to in considerations of personal identity. Aronson's arguments throughout "The legacy of Hume's analysis of causation" re-emphasise that this is the notion of identity being used. The genidentity requirement is preserved in more advanced versions of the analysis — Fair thinks this genidentity of energy and momentum is essential to causal connection, and is even prepared to specify the conditions for this genidentity:

> Explicit definitions allow their specifications for particular objects at a time. A closed system over some time interval of such a class of objects is one which can be surrounded by a surface across which no gross energy or momentum flows over that interval. That is, the partial derivatives of energy and momentum are zero over the boundary for all objects. As long as we can be sure that the system remains closed, the conservation laws tell us that both quantities will be conserved relative to the system for the duration that the objects possess them and in all interactions within the system. Assuming we can identify these objects themselves through time, various specifications of closed systems allow the identification of energy and momentum through time in both coarse and fine-grained ways.[105]

This may be a very important move on Fair's part. Genidentity is normally associated with things rather than properties of things or quantities possessed by things. If Aronson's notion that a *quantity* such as energy/momentum can be genidentical cannot be sustained, perhaps we can retreat to Fair's notion that the relevant identity considerations resolve to the genidentity of the objects possessing this quantity. The notion of genidentity of energy/momentum is

---

103 "The legacy of Hume's analysis of causation", p145

104 Fair appears to think that Aronson means something different here: that *all* the energy transferred away from the cause object must end up in the effect object. ("Causation and the flow of energy", p239) Really, Aronson is talking about the identity of the energy, not its exact quantity.

105 "Causation and the flow of energy", p234

also retained in later work by Aronson (but Aronson still holds in these later works that the *quantity* is genidentical).[106]

So how does this help with connectivity problems? Well, we now seem to have an ontological criterion for distinguishing between the causal status of the two cases of the bouncing of Frank. In the case of the normal last leap, the *Moon's* energy/momentum has not arrived at the leaping-off point in time to be transferred to Frank. The only quantity being transferred to Frank is the *Earth's* energy/momentum (to be scrupulously precise, the energy/momentum of a *part* of the Earth). Hence the lunar impact is not causally connected to Frank's last leap. In the case of the big last leap, some of the Moon's energy/momentum *has* reached Frank's jumping-off point, manifest in the form of the edge of the spreading contact zone. So some of the Moon's energy/momentum is transferred to Frank, and the lunar impact is causally connected to Frank's last leap. So if the coherence of the notion of the genidentity of energy/momentum can be established, an adequate Aronson-Fair analysis might be shown to be adequately connective.

Might be shown. Arguably, the genidentity criterion still leaves an adequate Aronson-Fair analysis inadequately connective. Assume the genidentity criterion does hold in our universe. Now consider our electron-only model of macroscopic interactions: electrons exchange virtual photons which alter the energy/momenta of the electrons, bringing the electrons into closer contact with other electrons so that further exchanges take place, producing a cascade through the macroscopic body containing the electrons. Yet the electrons at the contact point need not exchange all and only all their energy/momenta in the initial interaction. So the energy/momenta exchanged between electrons in the outer surface of one object and the deeper electrons need not entirely be energy/momenta from electrons in the other object. The amount of energy/momenta from the other object will, more likely than not, be diffused as the energy/momenta exchange cascade propagates through the object. In the case of Frank's big last leap, so many such exchanges will take in the cascade that it is extremely likely that *none* of the original energy/momenta will be manifest in the fringe of the contact zone that reaches Frank's jumping-off

106 For example: "The genidentity of these quantities throughout a physical interaction dictates that they can only be transferred or exchanged from one part of a system to another when interaction takes place." ("Conditions versus transference: a reply to Ehring", p249)

point. So even if the genidentity of energy/momentum holds, an adequate Aronson-Fair analysis still may not be adequately connective.[107]

Perhaps this bug could be fixed. But before bothering to do so, we would want to have some assurance that the genidentity of energy/momentum really is a feature of our universe. What grounds do we have for supposing that it holds?

Fair's identity conditions suggest that the genidentity of energy/momentum is associated with conservation principles. Aronson is more explicit — genidentity of energy/momentum is *required* by the conservation of energy and momentum.

Aronson considers an explanation (by Feynman again) why apparent instances of non-simultaneous action at a distance such as the electromagnetic or gravitational interactions do not involve violations of the conservation of energy or momentum.[108] The Sun attracts the Earth which moves towards the sun. The Earth thereby acquires momentum. For momentum to be conserved, the sun must also acquire an equal and opposite momentum. If there was some delay in the Earth acquiring momentum, then the momentum acquired by the Sun would not be cancelled out, and for some time interval momentum would not be conserved. Yet there *is* some delay before the Earth acquires that momentum. The Feynman explanation is that during the interregnum between the Sun acquiring the momentum and the Earth acquiring an equal and opposite momentum, the momentum of the sun is cancelled out by an equal and opposite momentum possessed and propagated through the gravitational *field* of the Sun. Aronson draws the following conclusion:

---

107 Such examples are raised by Ehring in the context of the transitivity of causation: the initial energy supplied in an initial interaction in a chain of interactions may be entirely dissipated by the time the last interaction takes place. So under the transference analysis the initial interaction is not be a cause of the final interaction. But by transitivity, the initial interaction should be a cause of the last interaction ("The transference theory of causation", pp256–7).
Aronson's response ("Conditions versus transference: a reply to Ehring", p254) is to accept that transference is not transitive (an issue we are neutral on), but deny that this is a problem, since causation is not transitive (another issue we are neutral on). However, Aronson's acceptance of Ehring's point that energy might be dissipated is tantamount to accepting that the transference analysis is not adequately connective in cases such as the bouncing of Frank (which we are prepared to be disturbed about).

108 "The legacy of Hume's analysis of causation", p146

> Note how … one and the same quantity — *viz*, momentum — is transmitted at 186,000 miles a second from its source to the object affected, and that at each stage of this transmission the momentum occurs in various forms. In other words, instead of saying the momentum of one object is annihilated and the same amount is recreated in another at a later time — thus, violating conservation principles — we speak of the momentum of a particle being converted into field momentum and then, again, into particle momentum. In this way, momentum is conserved. But if we accept Hume's doctrine of identity being a fiction while admitting of the existence of momentum, a violation of the conservation of momentum would occur each time a causal relation occurred in a mechanical situation.[109]

Even more succinctly (in "On the grammar of cause", commenting on the same Feynman passage):

> This suggests a definite connection between our notion of "cause" and conservation principles. The ancient maxim, *Ex nihilo, nihil fit* tells us that if a quantity comes into being, there must be some source of that quantity.[110]

In other words, energy or momentum that shows up in one part of a causal situation is energy or momentum from some other part of the system. It is a natural move then to suppose that the relevant energy and momentum is *one and the same* energy and momentum from the other part of the system.

Ancient maxims are all very well and good. But does our current best understanding support the notion of the genidentity of energy/momentum? D Dieks argues that even in classical physical theories the energy and momentum conservation principles are merely global constraints on the total energy and momentum in a system. No genidentity of quantities is required.[111]

Another cause for concern is that quantum energy/momentum transferences appear to be associated with (sufficiently brief) *violations* of the conservation of energy/momentum. For example, consider this discussion of the electromagnetic interaction by Paul Davies:

---

109 "The legacy of Hume's analysis of causation", p147

110 "On the grammar of cause", p426

111 See p88, Dieks, D: "Physics and the direction of causation", pp85–110, *Erkenntnis* vol 25 (1986)

> … in order to create an electron-positron pair, a gamma ray photon needs an energy at least equal to their combined rest mass (about 1 MeV). However, if the photon has less energy than this, *virtual pair creation* can still occur. What this means is that the photon suddenly turns into an electron-positron pair, which then annihilates a brief moment later and creates another photon. Because it all happens in a microscopic region of space the particles are also allowed to suspend the law of conservation of momentum (recall [the Heisenberg uncertainty principle] $\Delta x \Delta p \approx h$), so the annihilation process only recreates a single photon rather than the two or more that must appear after the annihilation of a *real* electron-positron pair.[112]

Or consider this passage by Peter Watkins:

> When considering time intervals familiar to us in daily life, the conservation of energy is exactly obeyed. However, when we consider extremely short time intervals the uncertainty principle requires that there be a large uncertainty in the energy so energy conservation can be violated for short periods. The "books" have to be balanced over longer periods but large amounts of energy can be "borrowed" if the time interval is short.[113]

In the case of the electromagnetic interaction, this "borrowing" takes the form of virtual photons. But where does this borrowing take place from? Not necessarily from an "originating" electron or a "destination" electron as *Ex nihilo, nihil fit* would lead us to believe. Energy/momentum conservation is being briefly violated: there is more energy/momentum involved than just that of the two electrons. This suggests that the conservation of energy/momentum does not necessitate the genidentity of transferred quantities of energy/momentum.

This only establishes that it is *consistent* with our understanding of the conservation principles and quantum physics in general that these quantities are not genidentical. Dieks makes the stronger claim that the genidentity of these quantities is *not compatible* with our understanding of quantum theory.[114] The example used is Planck's law for the distribution of energy in black body radiation:

---

112 *The forces of nature*, p124
113 *Story of the W and Z*, p11
114 "Physics and the direction of causation", pp88–91

> ... it is essential that the elements of energy possess no identities of their own. That is, the interchange of two energy elements must be considered as not changing the situation in any way; the physical states before and after the interchange are literally identical. This is not just one possible way of viewing things. If it is supposed that the states *are* different, although it may not be ascertainable experimentally, the statistics used in the derivation become different and this leads to a wrong form of the radiation law (*viz*, Wien's law instead of Planck's law). The mere ascription of an identity to the energy elements therefore entails consequences which are observably wrong.[115]

Unfortunately, Dieks' exposition doesn't make it clear that these "energy elements" are the same things as the supposedly genidentical quantities of energy/momentum of an adequate Aronson-Fair analysis. No matter: Dowe expands on this point, arguing that genidentity is not the sort of relation that could apply to any quantity such as energy/momentum.

The idea is simple; we have already alluded to it. Genidentity is a notion that applies to stand-alone things, such as people. If quantities of energy/momentum are such things, then those quantities might qualify for genidentity. But these quantities are not stand-alone things. They are possessed by things (electrons, photons, and so forth), but they do not stand alone as things in their own right. These quantities are best interpreted as *properties* of things. It makes no sense to talk of the genidentity of properties over time.[116]

The only thing that could rescue the notion of genidentity of energy/momentum is Fair's reduction of the genidentity of energy/momentum to genidentity of the objects possessing energy/momentum. But there is a big bad bug in this tactic. Under an adequate Aronson-Fair analysis, energy/momentum transference takes place at the quantum scale. The objects that properly possess the quantities of energy/momentum are quantum-scale entities such as electrons, photons and quarks. The genidentity of objects *fails* at this physical scale.

Dieks *states* this is the case, noting that:

---

115 "Physics and the direction of causation", p89
116 "What's right and what's wrong with transference theories", pp368–71

> For instance, if two electrons are scattered it is impossible to tell which one of the incoming electrons is the same as a chosen one of the pair of outgoing electrons. Even stronger, not only is it impossible to answer the question experimentally, but the mere supposition that the answer is possible (that is, the supposition that the electrons possess an identity of their own) leads to contradictions. This peculiar situation, which is in strong conflict with classical intuitions, obtains quite generally. Also, if there is no interaction between the electrons, quantum theory tells us that there is no *genidentity* of the various particles. This means that if we make repeated measurements of the positions of *n* electrons we find at each measurement *n* places; but according to quantum theory it is not permitted to regard any two places in subsequent measurements as belonging to the path of one and the same particle.[117]

Wesley Salmon also makes this point as part of a criticism of Dowe's closely-related conserved quantity analysis of causation.[118]

Given that this is the crucial problem for the adequate Aronson-Fair analysis, we shall give a brief account of why we know genidentity fails at the quantum physical scale in this universe.

First consider a "classical" case involving two billiard balls in the absence of any other objects. We observe the white to have a certain momentum at position w and the red to have a certain momentum at position x. Then a little later we observe the white to have a certain momentum at position y and the red to have a certain momentum at position z. How did the billiard balls move from w and x to y and z in order that the white and the red have the momentum observed at y and z?

Well, there are a number of possible paths the balls could take. The white could have proceeded directly from w to y and the red could have proceeded directly from x to z. Alternatively, the white and the red could have moved towards each other, collided, and moved apart, eventually arriving at positions y and z. A moment's reflection shows that there are an infinite number of such possible paths involving collisions.

---

117 "Physics and the direction of causation", p89

118 See p468, Salmon, Wesley: "Causality and explanation: a reply to two critiques", pp461–77, *Philosophy of science* vol 64 (1998)

The balls will only take one of these possible paths. We could even determine which unique path was taken by the red and the white by examining the momentum measurements taken at the four positions w, x, y and z. Finally, we can be certain that whatever paths were taken, the white does not end up at z and the red does not end up at y, because we observe the ball at y to be white and observe the ball at z to be red.

Now let's consider the energy/momentum of two electrons, observed at w and x and then at y and z. First, we will insist that one of the electrons observed at either w or x is genidentical to one of the electrons observed at either y or x, and the other electron observed at either w or x is genidentical to the other electron observed at either y or z. Our equivocation signals an immediate epistemological problem. Electrons, unlike billiard balls, cannot be distinguished on the basis of their colour, or other properties. So it is not impossible that the electron at w could be genidentical to the electron at z (contrary to the case of the billiard balls).

Once again, there are an infinite number of possible paths the electrons could take to reach y and z. The electron at w could proceed directly to y, and the electron at x could proceed directly to z. Or the electrons at w and x could proceed in some other directions but exchange a virtual photon and still end up at y and z respectively. Or along the way the electrons could have exchanged three virtual photons, and so forth. To complicate things further, the exchanged virtual photons could have briefly formed virtual electron-positron pairs between the "times" of "absorption" and "emission" by the real electrons.

Of course, due to the epistemological problem of distinguishing the electrons, it is possible that the electron at w could have proceeded directly to z and the electron at x proceeded directly to y, and so forth.

So which possible path did the electrons take? Trying to determine this by analogous means as in the case of the billiard balls ("reading off" the actual path from the original momentum measurements at w and x) is thwarted by an ontological problem. For a start, the uncertainty principle tells us that there is no precise fact of the matter about the energy/momenta of the electrons at w and x or y and z. In fact, determining a unique path that the electrons take is also ruled out by the uncertainty principle — it would effectively involve precisely specifying the energy/momenta of the electrons at every point of their "trajectory" on the way to y and z.

Given these considerations, it should not be surprising that it is only a probabilistic matter that the two electrons at w and x end up at y and z *at all*. To be precise, given the presence of two electrons at w and x we can calculate an *amplitude* that two electrons will be found at y and z. (An amplitude is the square root of the probability of a quantum event.)[119]

The differences between the classical motion of the billiard balls and the quantum motion of the electrons do not end here. The amplitude of the electrons turning up at y and z is not equal to the amplitude of *one* of the possible paths the electrons can take. The actual amplitude is given by the *sum* of the amplitudes of all the possible paths the electrons could take: proceeding directly to y and z, exchanging a virtual photon along the way, exchanging two virtual photons, etc.[120]

Now we come to the failure of the genidentity assumption. If there are two distinct electrons which maintain their identity, then there are two distinct events which could underlie the energy/momenta measurements which we make at y and z. In one event, the electron at w has ended up at y and the electron at x has ended up at z. In the other, the electron at w has ended up at z and the electron at x has ended up at y. The probability of electrons ending up at y and z will thus be determined by adding the probability of one event to the probability of the other — the square of the sum of the amplitudes of the possible paths of electrons travelling from w to y and x to z is added to the square of the sum of the amplitudes of the possible paths of electrons travelling from w to z and x to y.

But this is not consistent with the predictions of quantum electrodynamics (the theory of the electromagnetic interaction). We find that the probability of

---

119 See *QED: the strange theory of light and matter*

120 The possible path that contributes the most to the overall amplitude of the electrons moving to y and z is the path in which the electrons proceed directly to y and z. The other paths contribute less to the sum of amplitudes because they involve junctions or couplings between particles — such as an electron emitting a virtual photon. Every such coupling reduces the amplitude of that particular path by a factor of approximately 10^2. The amplitude of the possible path involving the exchange of a single virtual photon involves two such couplings, so the amplitude of that path is approximately 10^4 times smaller than the direct path. (*QED: the strange theory of light and matter*, pp90–7)

electrons ending up at y and z with particular energy/momenta is given by the square of the sum of the amplitudes of *all* possible paths of electrons from w and x to y and z, including both those in which the electron at w ends up at y and those in which the electron at w ends up at z.[121] Now we have a simple modus tollens. If the electron at w *is* genidentical with one of the electrons at y or z and the electron at x *is* genidentical with the other electron at y or z (regardless of whether or not we could *tell* which electron is which) we could make certain predictions about the probabilities of electrons ending up at y and z with particular energy/momenta. These predictions turn out not to be the case. So it is not the case that genidentity holds for electrons, since the probability of electrons turning up at y and z turns out to be the square of the sum of the paths, not the sum of the squares. This result can be generalised to cover all other quantum scale objects.

So Fair's notion of genidentity of transferred quantities fails. The genidentity of transferred quantities of energy/momenta cannot be reduced to the genidentity of the objects that those quantities are exchanged between, because an adequate Aronson-Fair analysis respects the Heathcote conjecture. Thus the objects that quantities of energy are being exchanged by are objects at the quantum physical scale, such as electrons. But the genidentity relation fails at this physical scale — if it did hold, the world would be observably different from the way it actually is.

Following Dieks and Dowe, we also see that the genidentity relation fails to coherently apply to quantitative properties of objects. So it is not the case in this universe that causation involves the transference of genidentical quantities of energy/momenta. If this is the case, an adequate Aronson-Fair analysis fails to supply an adequate solution to the problem of connectivity. If an analysis of causation fails to provide an adequate account of connectivity, causal dependence under that analysis fails to track physical dependence. So, rather

121 To be accurate, this would only be the case if electrons were not polarised. In the actual situation, where electrons do have a polarity, we *subtract* the amplitude of the possible path along which the electron at w proceeds directly to z (and x to y) from the amplitude of the path in which the electron at w proceeds directly to y (and x to z). This amplitude is then added to the *difference* between the amplitude of the electron at w proceeding to y exchanging a single virtual photon along the way with the electron at x proceeding to z, and the amplitude of the electron at w proceeding to z and exchanging a single virtual photon along the way with the electron at x proceeding to y, and so forth. So when we talk of the actual amplitude being the square of the sums, this is shorthand for the square of the sums and the differences.

surprisingly, the best possible version of the transference analysis of causation fails to fulfil our basic desideratum.

Where to from here? There is an analysis of causation which prima facie provides an excellent account of causal connectivity — the *process* analysis of causation, most famously developed by Wesley Salmon. Let us investigate whether this analysis of causation is supportable and can fulfil our basic desideratum.

Processes — Salmon₁

Interaction and connection, production and propagation

An intriguing feature of our discussion of the failure of the transference analysis has been the ontological relocation of the mystery of causal connections. In manipulability and counterfactual analyses (and other analyses which interpret causation in terms of the necessary and sufficient conditions for an event to be brought about, as given by Ducasse and J L Mackie) causal relations connect more or less persistent things in sequences. The mystery that these analyses seek to solve is the nature of the "linking relation". The transference analysis provides quite a plausible account of what constitutes the linking relation — exchange of virtual photons or other such virtual particles carrying energy/momentum from thing to thing. The mystery of causal connections turns out to be the nature of the *linking-thing* connecting the causal interactions. We have seen that denying that any such thing connects causal interactions produces an analysis which is inadequately connective. Allowing any object undergoing a change to provide the link produces an analysis which is excessively connective. And we cannot maintain in an Aronson-Fair analysis that this link is the genidentity of quantities of energy/momentum.

Given that the failure of the transference analysis to adequately fulfil our basic desideratum derives from its failure to provide a satisfactory account of linking-things, it seems likely that an adequate analysis of causation would need to give an explicit account of causal linking-things as well as causal linking-relations. The obvious candidates for such an analysis are the various process analyses of causation developed by Phil Dowe and Wesley Salmon.

The defining attitude of a process analysis is the conviction that we need to account for two separate features of causation. Firstly, there is the power or efficacy that causation has to influence states of affairs, which Salmon refers to as the productive aspect of causation:

> When we say that the blow of a hammer drives a nail, we mean that the impact produces penetration of the nail into the wood. When we say that a horse pulls a cart, we mean that the force exerted by the horse produces the motion of the cart.[122]

Secondly, there is a propagative, or persistent quality to causation which connects these productive episodes:

> Experiences which we had earlier in our lives affect our current behaviour. By means of memory, the influence of these past events is transmitted to the present. A sonic boom makes us aware of the passage of a jet airplane overhead; a disturbance in the air is propagated from the upper atmosphere to our location on the ground ... As all these examples show, what happens at one place and time can have significant influence upon what happens at other places and times. This is because causal influence can be propagated through time and space.[123]

Clearly, the propagation of contact zones through the Moon (in the case of the bonking of Frank) and the Earth (in the case of the bouncing of Frank) are also instances of this sort of causal propagation.

A process analysis thus generally posits two basic causal mechanisms in its ontology: causal interactions, which constitute the productive episodes; and causal processes, which constitute the linkages between productive episodes.[124] (As we shall discuss presently, in Salmon's first process analysis the productive episodes are not solely constituted by causal interactions — hence the slight caveat "generally".)

Clearly, some of the versions of the transference analysis that we have discussed in the previous chapter could readily be classified as process analyses.[125] These transference analyses also provide convenient examples of the problems an adequate process analysis of causation needs to overcome.[126]

---

122 p285, Salmon, Wesley: "Causality: production and propagation", pp285–301, Salmon, Wesley: *Causality and explanation* (New York: Oxford University Press, 1998)

123 "Causality: production and propagation", pp286

124 p248, Salmon, Wesley: "Causality without counterfactuals", pp248–60, Salmon, Wesley: *Causality and explanation* (New York: Oxford University Press, 1998)

125 Dowe frequently claims this: see, for example, p179, Dowe, Phil: "Process causality and asymmetry", pp179–96, *Erkenntnis* vol 37 (1992) and p195, Dowe, Phil: "Wesley Salmon's process theory of causality and the conserved quantity theory", pp195–216, *Philosophy of science* vol 59 (1992)

126 As before, an "adequate" analysis is an analysis in which causal dependence tracks physical dependence.

The first challenge is that the account of processes needs to be adequately connective. One version of the Aronson-Fair analysis which does not assert the genidentity of quantities of energy or momentum would formally qualify as a process analysis. Transference of quantities of energy or momentum is the account of causal interactions; being one and the same interposing object undergoing one and the same physical change is the account of causal processes. Yet the case of the bouncing of Frank shows us that this account of processes is excessively connective, so it is not an adequate process analysis.

The second challenge is that the account of processes needs to be based on some feature of *this* world. An Aronson-Fair analysis which asserts the genidentity of transferred quantities of energy or momentum formally qualifies as a process analysis. Transference of quantities of energy or momentum provides the account of causal interactions; transmission of genidentical quantities of energy or momentum from interaction to interaction provides the account of processes. Arguably, genidentity of transferred quantities would provide an adequately connective account of processes. Unfortunately, genidentity does not apply to quantitative properties, and genidentity of properties can not be reduced to genidentity of the relevant objects the quantities are being transferred between, because genidentity does not apply to the relevant objects in this universe. So this process analysis also fails.

The third challenge — which is also the challenge most frequently discussed by process analysts — is that not every "thing" which might legitimately be said to connect two causal interactions should be considered to be a causal connection between the interactions. Intuitively, the gist of this problem is quite easy to grasp. For example, consider a standard rock being hurled through the air and colliding with a standard window. Intuitively, the rock flying through the air seems to qualify as a connection between causal interactions (in this case, someone hurling the rock, and the window breaking). The flying standard rock should be a causal process under a process analysis of causation. But suppose the sun is shining while the rock is being hurled through the window. The rock casts a shadow on the rock-thrower's hand. As the rock travels through the air, the shadow travels down the thrower's body and along the ground towards the window. Just before the rock hits the window, the shadow travels up the wall the window is located in, and on to the window. When the rock hits the window, the shadow is being cast on the point of contact of the rock with the window. The shadow, like the rock, is a persistent thing. The shadow connects the same two causal interactions as the rock. But intuitively we *wouldn't* want the shadow to count as a causal connection between the hurling of the rock and the breaking of the window. The shadow shouldn't be a causal process. Some filtering conditions are required to distinguish the causal connections from connections like the shadow.

There are several ways we could block the possible causal status of a shadow under a process analysis, although perhaps there is not much of a difference between each move. First, we could deny that the shadow *really* is a thing, it is the absence of a thing — indeed, only a relative absence of a thing. It is a reduction in illumination in a localised area. If a shadow isn't a thing, *a fortiori* it couldn't be a linking thing.

The idea has some merit, but is rarely pursued. The reasons appear to be semantic: whether or not shadows fall under the ontological category of things, we quantify over shadows in ordinary language. The shadow of the gnomon on a sundial has a localised position, a direction, and we can draw inferences from our observations of it; we talk of it in much the same way as the hands on a clock, which usually count as things. Eliminating the causal connectivity of shadows by denying their metaphysical thinghood just doesn't seem worth the ensuing disruption of ordinary language. But Dowe does comes very close to using this line of argument:

> Another problem Fair faces is Salmon's objection that it is not possible to distinguish energy transmission from a regular appearance of energy giving the example of a spotlight "spot" moving along a wall. However, one must question whether energy can legitimately be ascribed to a spot. The energy referred to is manifested by the segment of *wall*, and the segment of wall does not move.[127]

Similarly, we might argue that the features of the shadow are actually features of the objects along the shadow's path.

Another strategy is to accept that shadows are things, but deny that they are processes. There is a certain tactical danger in employing this move. The term "process" in the sense used in Salmon and Dowe's process ontologies is mostly a term of art: any intuitive understanding of the term derives from its frequent

---

127 "Wesley Salmon's process theory of causality and the conserved quantity theory", p214

characterisations in terms of our everyday understanding of things moving through space and persisting through time. If we were to draw a sharp distinction between processes and everyday things such as shadows, we might find ourselves wondering if we understood "process" at all.

The standard move is therefore to accept that the shadow is a process, but deny that it is a *causal* process. The standard parlance is that the shadow is a *pseudo* process. This causal process/pseudo process distinction is probably equivalent to either a thing/not-a-thing or an any-old-thing/process distinction, but stubs fewer intuitive toes.[128]

So, an adequate process analysis of causation is going to have to provide a detailed account of the distinction between pseudo processes and causal processes — indeed, independently of our desideratum Dowe argues that this is the chief task of the process analyst of causation.[129] At the very least, the various different accounts of the causal/pseudo distinction will provide us with a taxonomic device for distinguishing between the various process analyses.

Salmon$_1$ — basic principles and prehistory

In the previous chapter, we said we were surprised that the best formulation of the transference analysis failed to fulfil our basic desideratum. This surprise was largely due to the dearth of credible alternative bases for an adequate analysis of causation. If an analysis of causation which reduces causal dependence to energy/momentum transfers fails to track physical dependence, what on Earth would?

But perhaps there is a credible alternative. Consider this comment by Davies on the apparent violations by virtual particles of the temporal asymmetry of cause and effect:

---

128 On the other hand, see p324, Dowe, Phil: "Causality and conserved quantities: A reply to Salmon", pp321–33, *Philosophy of science* vol 62 (1995), where Dowe argues that the causal/pseudo distinction can be cashed out in terms of varieties of objects for the "conserved quantity" process analysis.

129 See Dowe, Phil: "Causation, causal processes", *Stanford encyclopedia of philosophy* (http://plato.stanford.edu/entries/causation-process/, last modified 9 December 1996)

> If the time order of the emission and absorption of virtual quanta cannot be measured does it not imply something similar to instantaneous transfer of information between the particles? Could we not use virtual photons to send real messages faster than light? The answer is no. Nature has arranged for quantum theory and relativity to co-exist with wonderful consistency. To convey information using electromagnetic signals we need at least one whole wave to encode it. However, we have seen that virtual photons cannot travel beyond a distance of one wavelength, so they are useless for messages. Only real photons can carry information. Apparent causality violation by virtual photons is a Pyrrhic victory against relativity, because in the absence of information transfer, the difference between cause and effect is transcended.[130]

Of course, this observation is no help to Aronson, Fair or Heathcote with regard to establishing an asymmetric cause-effect dependence in quantum interactions. Under the transference analysis, causal dependence does not reduce to information transfer but rather to energy/momentum transfer. Quantum theory therefore tells us that under such an analysis there is no temporal fact of the matter about what is cause and what is effect in quantum scale interaction. This is hardly a Pyrrhic victory.

But the observation does suggest another plausible mechanism by which causal dependence could track physical dependence. Suppose that causation is correctly analysed in terms of the flow of information — a popular notion in the physical sciences.[131] If information is physically instantiated, then causal dependence between things might track physical interactions between things, because physical interactions would be the means by which information is transferred between things. Similarly, a satisfactory account of what kinds of physical things can genuinely instantiate information might solve the problem of connectivity and the other challenges faced by a process analysis of causation.[132]

---

130 *The forces of nature*, pp127–8

131 The view is by no means restricted to Davies. See, for example, the discussion of the causal consequences of Bell's inequality in pp162–76, Pagels, Heinz: *The cosmic code: quantum physics as the language of nature* (Harmondsworth: Penguin, 1984).

132 Even at this stage of formulating such a programme, a potential problem in fulfilling our basic desideratum should be apparent. Physical dependence is associated with virtual quanta. Yet according to Davies virtual quanta cannot carry information. So if a particular instance of physical dependence involves virtual quanta, that dependence would not be causal, since causal dependence involves the transfer or transmission of information, neither or which can be associated with virtual quanta.

The first process analysis of causation we shall discuss — historically, the first of two distinct process analyses developed by Wesley Salmon — is just such an information transfer analysis. The demarcation between causal and pseudo processes is cashed out in terms of the ability to transmit a *mark*, where a mark is the sort of thing which could be used as a signal. Similarly, causal interactions are taken to be instances of *marking*.

This first process analysis (which we shall refer to as "$Salmon_1$") is developed and outlined in a raft of articles and books beginning with "Causal and theoretical explanation"[133] in 1975 and concluding with "Causal propensities: statistical causality versus aleatory causality" in 1990.[134] As might be expected of such a work in progress, the role and nature of the critical concepts in the analysis evolve considerably.[135] The "mature" version of $Salmon_1$ is consequently very rich and complex.

Since some light is cast on the dark corners of $Salmon_1$ by examination of the development of Salmon's thought on causation (and the development of Salmon's thought is an interesting topic in its own right), we will give a detailed semi-historical presentation of the evolution of $Salmon_1$. There are also technical benefits to such an approach. Many of the complexities arise from attempts to fill lacunae in the analysis, or responses to criticisms of earlier versions of such attempts. A historical analysis of $Salmon_1$ therefore serves as a useful description of what a credible process analysis of causation *could not involve*. Any such description will be useful for our interests — determining whether some variety of process analysis our basic desideratum.

The prospects for our interests seem bright: Salmon eventually claims to have located causal dependence in physical dependence:

---

133 Originally titled "Theoretical explanation", but reprinted as Salmon, Wesley: "Causal and theoretical explanation", pp108–24, Salmon, Wesley: *Causality and explanation* (New York: Oxford University Press, 1998). All page references are to the latter source.

134 Originally appearing as Salmon, Wesley: "Causal propensities: statistical causality versus aleatory causality", pp95–100, *Topoi* vol 9 (1990) and reprinted in pp200–7, Salmon, Wesley: *Causality and explanation* (New York: Oxford University Press, 1998). All page references are to the latter source.

135 Arguably, $Salmon_1$ comprises three separate process analyses — each analysis distinguished by successive attempts to give a satisfactory account of causal interactions. Since all three analyses share the mark criterion, and therefore are all information transfer analyses, we prefer to refer to them as successive stages of the development of $Salmon_1$. This better serves to distinguish Salmon's early thoughts on process analyses from his later development of a process analysis, $Salmon_2$, which is not an information transfer analysis.

> … I have sought a physical relation that constitutes a causal connection; I have suggested that causal processes fill the bill … causal connections exist in the physical world, and not just in our minds. Moreover, causality is neither logical nor metaphysical; causality is physical — it is an objective part of the structure of our world.[136]

If this could be shown to be true, then our desideratum will be fulfilled — or at least a process analysis will not be subject to the problems of a Lewis-style counterfactual analysis, where one set of physical dependences proves to be causal and another not causal, even though there are no relevant physical differences between the two sets of circumstances.

Yet although Salmon thinks that it is a good thing that causal dependence tracks physical dependence (indeed, *reduces* to physical dependence), this is not the fundamental desideratum driving the development of $Salmon_1$ or even $Salmon_2$. Like Aronson, Salmon is primarily interested in developing a theory of scientific explanation — roughly speaking, that all scientific explanation supervenes on the causal structure of the world. Since one of the goals of science is to provide an explanation for all physical dependences, all physical dependences should be causal dependences.

Of course, a theory of explanation need not be causal. Hempel's deductive-nomological theory is an example. Salmon's original conception of scientific explanation is also acausal — indeed, it might be well be described as an attempt to banish loose talk of causes from scientific explanations.

Like Aronson, Salmon also reacts against Hempel's deductive-nomological theory of explanation, where an explanation of an event is a modus ponens of an initial condition on an instance of a lawlike generalisation linking those conditions with that event. Salmon also reacts against Hempel's inductive-statistical theory, in which, roughly speaking an explanation is an inductive inference *from* an initial condition and a strong statistical correlation between the event to be explained and the initial condition *to* a high probability that the event occurs. Salmon's alternative "statistical relevance" (S-R) theory of explanation is that something explains something else when the explaining

---

136 p24, Salmon, Wesley: "A new look at causality", pp13–24, Salmon, Wesley: *Causality and explanation* (New York: Oxford University Press, 1998)

thing stands in a particular kind of statistical relevance relation to the thing needing to be explained:

> Adopting the S-R approach, we begin with an explanatory question in a form somewhat different from that given by Hempel. Instead of asking, for instance, "Why did x get well within a fortnight?" we ask, "Why did this person with a cold get well within a fortnight?" Instead of asking "Why is x a B?" we ask, "Why is x, which is an A, also a B?" The answer — at least for preliminary purposes — is that x is also a C, where C is *relevant* to B within A. Thus we have a prior probability $P(B|A)$ — in this case, the probability that a person with a cold (A) gets well within a fortnight (B). Then we let C stand for the taking of vitamin C. We are interested in the posterior probability $P(B|A.C)$ that a person with a cold who takes vitamin C recovers within a fortnight. If the prior and posterior probabilities are equal to one another, the taking of vitamin C can play no role in explaining why this person recovered from the cold within the specified period of time. If the posterior probability is not equal to the prior probability, then C may, under certain circumstances, furnish part or all of the desired explanation.[137]

A major problem for the S-R approach is that not every statistical correlation of this sort is explanatory. For example:

> A rapidly falling barometric reading is a sign of an imminent storm, and it is *highly correlated* with the onset of storms, but it certainly does not *explain* the occurrence of a storm.[138]

Now it is startlingly easy to identify what is going on in this case if we are allowed to use causal talk in our explanations. The falling barometric reading does not *explain* the onset of the storm because the reading does not *cause* the storm. The correlation between the storm and the barometric readings arises because each phenomenon has a common cause — a drop in atmospheric pressure. So the drop in atmospheric pressure explains both the storm and the barometric reading. Unfortunately, there is no room for such causal talk in the S-R theory as roughly outlined.[139]

---

137 pp34–5, Salmon, Wesley: *Scientific explanation and the causal structure of the world* (Princeton: Princeton University Press, 1984)

138 *Scientific explanation and the causal structure of the world*, p43

139 And also in its full formal development — Salmon notes that the S-R theory in its first presentation is "purely statistical and acausal". See *Scientific explanation and the causal structure of the world*, p191

Yet the S-R theory has some resources to deal with this problem. Let A be a series of days in some locale. Let B be the occurrence of a storm. Let C be a recent sharp drop in the barometric reading. We find that C is relevant to B within A — ie:

$$P(B|A.C) \neq P(B|A)$$

Yet we do not want to count C as an explanation for B. We can escape this problem by noting that where D is a drop in atmospheric pressure, we find that:

$$P(B|A.C.D) = P(B|A.D)$$

In words, it is statistically irrelevant to the occurrence of a storm in the presence of a sharp drop in atmospheric pressure that the drop in pressure was *measured by a barometer*. So the barometer reading is not explanatory after all. In Salmon's terminology, the drop in pressure "screens off" the barometric reading.

On the other hand, the drop in atmospheric pressure is not screened off by the barometric reading. We find:

$$P(B|A.C.D) \neq P(B|A.C)$$

Barometers sometimes break and register a fall independently. So the drop in pressure is statistically relevant to the onset of the storm, and explains the onset of the storm.[140]

Given that the screening-off relations are at hand for the S-R approach, it might be argued that we do not need extra causal talk to supplement our explanations. It might be argued further that causal relations *reduce*, upon analysis, to statistical relations between events. Indeed, this is Salmon's early approach to causation. So in a sense there is a $Salmon_0$ analysis of causation — that talk of causal relations can be analysed in terms of statistical relations, or possibly *replaced* by talk about statistical relations — although Salmon himself never explicitly works out such an analysis in any detail.[141]

---

140 This presentation follows closely that given in *Scientific explanation and the causal structure of the world*, pp43–4

141 Salmon expounds the hope that such an analysis can be carried out (perhaps along the lines of the analysis presented by Reichenbach in Reichenbach, Hans: *The direction of time* (Berkeley: University of California Press, 1956)) in p55 and p81 of Salmon, Wesley:

The relationship between causal dependence and statistical dependence undergoes a sea change in the 1975 paper, "Causal and theoretical explanation". Here Salmon attempts to augment the S-R approach to explanation — he feels something needs to be said about causal relevance as well. The precise nature of Salmon's project is curiously unclear. We find in the introduction to the article:

> I shall agree from the outset that *causal relevance* (or causal influence) plays an indispensable role in scientific explanation, and I shall attempt to show how this relation can be explicated in terms of the concept of statistical relevance ... The theme of this essay will be the centrality of certain kinds of statistical relevance relations in the notions of causal and theoretical explanation.[142]

The implication appears to be that causal relevance reduces to statistical relevance. Yet the body of the article suggests a diametrically opposed programme, as does the conclusion:

> According to the present account, statistical dependencies are explained by, so to speak, filling in the causal connections in terms of spatiotemporally continuous causal processes ... Causal or theoretical explanation of a statistical correlation between distinct types of events is an exhibition of the way in which those regularities fit into the causal structure of the world — an exhibition of the causal connections between them that give rise to the statistical-relevance relations.[143]

This ambivalence as to the "direction of reduction" is not just a mark of "Causal and theoretical explanation being located on the cusp of Salmon_0 and Salmon_1. The ambiguity persists throughout Salmon_1. Generally Salmon appears to be claiming that causal relevance *gives rise* to statistical relevance, but sometimes causal interactions are characterised *in terms of* statistical relationships, and sometimes all causal phenomena seem to be fundamentally statistical. These later ambiguities in Salmon_1 may partially amount to residues of Salmon_0.

Statistical explanation and statistical relevance (Pittsburgh: University of Pittsburgh Press, 1971) and reiterates that this was his first hopes for an analysis of causation in *Scientific explanation and the causal structure of the world*, pp44–5 and "Causal propensities: statistical causality versus aleatory causality", p201.

142 "Causal and theoretical explanation", pp109

143 "Causal and theoretical explanation", p122

A second potential source of confusion is the close relationship between statistical correlations in observational data and the probabilistic behaviour of things. The sea change in $Salmon_1$ is that statistical correlations between observations no longer constitute causation; they constitute evidence that causation is going on. But the underlying causal mechanisms can themselves be irreducibly probabilistic.

Salmon's favourite example is Compton scattering of light from electrons. A high-energy photon is absorbed by an electron. The "excited" electron then emits another high-energy photon. Although the total energy of the emitted photon and post-emission electron will equal the total energy of the incident photon and preabsorption electron, there is no way, even in principle, of predicting with certainty the individual energies of the emitted photon and post-emission electron.[144]

Salmon's interpretation of such cases is that the underlying causal mechanisms themselves are probabilistic (not just our observational data). To be precise the mechanisms possess *propensities* — real tendencies, or dispositions, for particular situations to come about. So in a case of Compton scattering we might say that the excited state of the electron did not determine the energy of the emitted photon — rather, the excited electron had a range of propensities to emit a photons, and in this case the propensity to emit a photon of this particular energy was actualised.

The statistical frequency data we use to determine relations of statistical relevance between probabilities of events are generated by the strength of the propensities of the underlying causal mechanisms.[145] Returning to the barometer example, we find that:

$$P(B|A.D) \neq P(B|A)$$

where the probabilities are long-run frequencies derived from meteorological observations (our statistical data). The long-run frequencies, however, are produced by the strong propensity for storms to occur on days where there has

---

144 See, for example, pp133–4, Salmon, Wesley: "Why ask 'why'?", pp125–41, Salmon, Wesley: *Causality and explanation* (New York: Oxford University Press, 1998).

145 See p14, Salmon, Wesley: "Dynamic rationality: propensity, probability, and credence", pp3–40, Fetzer, James; editor: *Probability and causality: essays in honor of Wesley C Salmon* (Dordrecht: D Reidel Publishing Company, 1988)

been a sudden drop in atmospheric pressure, and the weaker propensity for storms to occur on any old day.

So in the mature $Salmon_1$ statistical correlations really do reduce to the underlying causal set-up — to precise, to the strengths of the underlying propensities. But often as not the only information on the strength of the underlying propensities that we have is drawn from our observation of the generated long-run frequencies in statistical data. So although statistical correlations are *generated* by causal mechanisms, frequently the nature of the underlying causal mechanisms can only be *analysed* in terms of the statistical correlations (which gives the impression that the causal mechanisms resolve to the statistical correlations).[146]

A further source of potential confusion is Salmon's desideratum for developing process analyses. Salmon's primary goal is *neither* the reduction of causation to statistical correlation *nor* the reduction of statistical correlation to causal mechanisms. The prize is a theory of scientific explanation. The basis for this theory is the S-R approach. $Salmon_1$ is to be interpreted as an *augmentation* to S-R, not a *replacement* for S-R. Both statistical relevance and causal relevance are required for satisfactory explanations: both have fundamental significance for Salmon. Salmon makes this point explicitly in a late paper, "Causality and explanation: a reply to two critiques"[147]:

> My main motivations in working out a theory of causal explanation as a successor to my theory of statistical explanation were the convictions, first, that causality is an essential ingredient in scientific explanation, and second, that causal relations cannot be explicated wholly in terms of statistical relations. These points still seem sound. In [*Scientific explanation and the causal structure of the world*] I characterised scientific explanation as a two-tiered structure, consisting of statistical relevance relations on one level and causal processes and interactions on the other ... I would now say (1) that statistical relevance relations, in the absence of connecting causal processes, lack

146 This is not to say that our understanding of propensities *only* comes from our frequency observations (see "Dynamic rationality: propensity, probability, and credence", p25), nor that propensity weighting must slavishly follow observed frequencies. Our theoretical understanding of a particular phenomenon may allow us to make propensity predictions. These predictions may then be checked against the observed frequencies (if such observations are available). As with any such experimental set-up, if the observed frequencies do not match those predicted by our theoretical weighting of propensities, there may well still be much to do before the theory is discarded.

147 Strictly speaking, this 1998 paper refers to $Salmon_2$ rather than $Salmon_1$.

> explanatory import and (2) that connecting causal processes in the absence of statistical relevance relations, also lack explanatory import. In various discussions I have focused on (1) to the virtual neglect of (2) … [T]his was a mistake. Both are indispensable.[148]

Salmon's comments here about causal explanation being a successor to statistical explanation might seem to vitiate the point we are trying to make. But Salmon clearly states in *Scientific explanation and the causal structure of the world* (the mature version of $Salmon_1$) that his aim is to *supplement* S-R, and that at a "fundamental level" S-R provides important insights into the nature of scientific explanation.[149] It seems that Salmon was at least partially of his $Salmon_2$ frame of mind during the exposition of $Salmon_1$. Hence some confusion may arise over whether causal relevance is more fundamental than statistical relevance (or *vice versa*) because both are being occasionally treated (at least in spirit) as fundamental — but fundamental to *explanation*, not one another.

Common causes

Let's return to "Causal and theoretical explanation". Salmon's concern is instances of "common causes", such as the sudden fall in barometric reading and the subsequent storm. Salmon notes that in such cases (letting A be the storm, B be the prevailing weather conditions, and C the barometric reading) both the reading and the prevailing weather conditions are statistically relevant to the storm:

$$P(A|B) > P(A) \text{ and } P(A|C) > P(A)$$

Yet we find that the prevailing weather conditions "screen off" (or render statistically irrelevant) the barometric reading from the storm, ie:

$$P(A|B.C) = P(A|C)$$ [150]

On the other hand, the barometric reading does not screen off the weather conditions from the storm. Note that we can tell an exactly parallel story about the statistical relevance of the storm to the *barometer*. We find:

---

148 "Causality and explanation: a reply to two critiques", pp475–6
149 *Scientific explanation and the causal structure of the world*, p45
150 "Causal and theoretical explanation", pp120–1

$$P(B|A) > P(B) \text{ and } P(B|C) > P(B)$$

Yet the prevailing weather conditions screen off the storm from the barometer:

$$P(B|A.C) = P(B|C)$$

We can conclude that the barometric reading *doesn't* explain the storm, and the storm *doesn't* explain the barometric reading, but the common cause — the prevailing weather conditions — *does* explain the reading and *does* explain the storm.

The point of departure from $Salmon_0$ is that rather than taking the above statistical relevance relations as explicating our loose talk of "common causes", Salmon asks *why* these statistical relevance relations hold in such cases. The answer is that the statistical relevance relations track causal relevance, which is to be understood as the presence or absence of causal processes connecting the events. A causal process connects the prevailing weather conditions (the common cause) to the barometric reading, and a causal process connects the weather conditions to the storm, but no causal processes connect the storm and the barometric reading.

At this stage of development of $Salmon_1$, causal processes are described relatively informally. The basic requirement for a *process* is that it is a spatiotemporally continuous entity. The characterisation of a *causal* process is given mostly by example:

> There are many causal processes in this physical world; among the most important are the transmission of electromagnetic waves, the propagation of sound waves and other deformations in various material media, and the motion of physical objects. Such processes transpire at finite speeds no greater than that of light; they involve the transportation of energy from one place to another, and they can carry messages.[151]

The "transportation of energy" observation both resembles a transference analysis of causation and foreshadows the $Salmon_2$ conserved quantity analysis.[152] Yet Salmon thinks the ability to carry messages is more fundamental to whether a process is causal or not. Drawing on an idea by Reichenbach, the

151 "Causal and theoretical explanation", p114
152 As Salmon retrospectively notes in *Causality and explanation*, p123

ability to carry messages is cashed out in terms of the ability to transmit a *mark*.[153] Once again, the characterisation of marks is given by example:

> ... [We] can mark a beam of light by placing a red filter in its path. A beam of white light, encountering such a filter, will lose all of its frequencies except those in the red range, and the red color of the beam will thus be a mark transmitted onward from the point at which the filter is placed in its path.[154]

Pseudo-processes can be distinguished from causal processes by their inability to transmit information:

> In the context of relativity theory, it is essential to distinguish causal processes, such as the propagation of a light ray, from various pseudo-processes, such as the motion of a spot of light cast on the wall by a rotating beacon. The light ray itself can be marked by the use of a filter, or it can be modulated to transmit a message. The same is not true of the spot of light. If it is made red at one place because the light beam creating it passes through a red filter at the wall, that red mark is not passed on to the successive positions of the spot. The motion of the spot is a well-defined process of some sort, but it is not a causal process. The causal processes involved are the passages of light rays from the beacon to the wall, and these can be marked to transmit a message. But the direction of message transmission is from the beacon to the wall, not across the wall.[155]

So many of the signal features of $Salmon_1$ have been introduced in "Causal and theoretical explanation": we have an account of common causes, explanatory statistical relevance typically tracks causal relevance, causal mechanisms are to be understood in terms of causal processes, and causal processes are demarcated from pseudo processes by their ability to transmit marks. As yet, however, there is nothing more than an informal account of what it means for a mark to be transmitted, and there is no account of causal interaction. Without an account of interaction, at this stage of development $Salmon_1$ is not truly a process analysis of causation.

Nor is it much help in solving our connectivity problems. Consider the bouncing of Frank. On the information we have been given so far, both the

153 "Causal and theoretical explanation", p114
154 "Causal and theoretical explanation", p114
155 "Causal and theoretical explanation", p114

contact zone of the lunar impact spreading through the Earth and the Earth itself would count as processes. So even in the case of Frank's normal last leap, where the contact zone has not reached Frank's jumping-off point, there seems to be a connecting causal process between the impact and the leap. We need some account of why the spreading contact zone is the relevant causal process to invoke to explain why there is no connectivity in the case of the normal last leap.

Although this point needed to be addressed, $Salmon_1$ *still need not* have developed into a true process analysis from this point. Salmon's main goal is merely to provide an account of the causal underpinnings of the statistical relevance relations relevant to scientific explanation. It could be argued that by providing an account of "direct cause" statistical relevance (connection of two circumstances by a causal process) and "common cause" statistical relevance (connection by independent causal processes to a common circumstance), he has already provided enough of an account of causation to support his theory of explanation. Of course, $Salmon_1$ *does* develop into a true process analysis with a detailed account of causal interaction. But "Causal and theoretical explanation" does not appear incomplete without such an account.

"At-at"

Salmon provides a more formal characterisation of causal processes in the 1977 paper "An 'at-at' theory of causal influence".[156]

Once again we find that the line of demarcation between causal processes and pseudo processes is that causal processes transmit marks and pseudo processes do not. The paradigm example of marking a causal process is again placing a red filter in the path of a beam of light.[157]

For the first time in $Salmon_1$, instances of marking — such as interposing a red filter in the beam of light — are identified with causal interactions.[158] No

156 Originally appearing as Salmon, Wesley: "An 'at-at' theory of causal influence", pp215–24, *Philosophy of science* vol 44 (1977), and reprinted in pp193–9, Salmon, Wesley: *Causality and explanation* (New York: Oxford University Press, 1998). All page references are to the latter source.

157 "An 'at-at' theory of causal influence", p197

158 "An 'at-at' theory of causal influence", p195

detailed account of causal interactions is provided, but the as-yet-primitive notion plays a crucial role in the explication of causal processes.

This explication is extremely simple. A mark that has been introduced into a process by a single interaction at point A is transmitted to point B if and only if it occurs at B and at all stages of the process between A and B without additional interactions.[159] Consider the beam of light. It is marked by a causal interaction with the red filter — it becomes red. The beam of light remains red from then onwards without further interactions. So the beam of light is a causal process. Now consider a moving red spot of light cast by the red beam of light on a wall. The red spot, as noted in "Causal and theoretical explanation"[160], is a well-defined process of some sort — it certain seems to be something more than a piece of spatiotemporal junk — and displays redness like the beam of light. Yet the redness of the spot is produced by an interaction between the beam of light and the wall at every point of the trajectory of the spot across the wall. So according to the "at-at" theory the spot does not transmit a mark, and is not a causal process.

The term "at-at" is due to the parallel between this account of transmission of marks causal propagation and Russell's "at-at" theory of motion. To move from A to B is simply to occupy the intervening points at the intervening moments — being *at* the intervening points *at* the corresponding instants.[161] Similarly, transmission of a mark by a process from A to B involves the mark being *at* the intervening positions *at* the corresponding times in the absence of any further interaction.[162]

Now $Salmon_1$ *prima facie* provides an account of why Frank's big last leap is causally connected to the lunar impact and the normal last leap is not. The edge of the spreading contact zone of the lunar impact is a mark — a stress disturbance in the structure of the Earth. In the case of the big last leap, we find this stress disturbance at every intervening position at the corresponding times between the surface of the lunar duct-tape wrapping and Frank. In the case of the normal last leap, there is a spatiotemporal gap. So in the case of the big last

159 "An 'at-at' theory of causal influence", p197
160 "Causal and theoretical explanation", p115
161 "An 'at-at' theory of causal influence", p198
162 Salmon gives a rather nice account of the parallel case of transmission of marks in "Why ask 'why?'?", p130

leap, the Moon transmitted a mark to Frank, and in the normal last leap it did not.

Causal interactions = conjunctive forks

Although we might not have needed an account of causal interactions after "Causal and theoretical explanation", certainly something needs to be said *now*. Otherwise the account of processes provided by Salmon is simply incomplete. The first important contribution on this issue is Philip von Bretzel's suggestion in "Concerning a probabilistic theory of causation adequate for the causal theory of time"[163] that we understand causal interactions in terms of the analysis of common causes.

The von Bretzel paper is concerned primarily with an amendment to Reichenbach's probabilistic analysis of causation, and the use of such an analysis to give a causal account of the asymmetry of temporal direction. Yet many of the features of Reichenbach's analysis of causation carry across to $Salmon_1$ — cashing out the pseudo/causal process distinction in terms of mark transmission, for example — so many of von Bretzel's points can be transferred to $Salmon_1$.

Outlining the details of Reichenbach's analysis of causation is, fortunately, beyond the scope of this thesis. In brief, von Bretzel's concern is that a crucial theorem used by Reichenbach to establish a temporal direction from his causal analysis — the principle of the local comparability of time order — is explicated in terms of temporal concepts, and hence is unacceptably circular.[164]

The problem before von Bretzel, then:

163 von Bretzel, Philip: "Concerning a probabilistic theory of causation adequate for the causal theory of time", pp173–90, *Synthese* vol 35 (1977), reprinted pp385–402, Salmon, Wesley; editor: *Hans Reichenbach: logical empiricist* (Dordrecht: D Reidel Publishing Company, 1979). All page references are to the latter source.

164 "Concerning a probabilistic theory of causation adequate for the causal theory of time", p389

> … is how to define a non-temporal way of ordering the events of the causal net so that causes may be distinguished from effects. To do this it is necessary to introduce the concept of a conjunctive fork.[165]

A conjunctive fork (the concept and terminology once again originates with Reichenbach[166]) is a state of affairs composed of three distinct events A, B and C where C is a common cause of A and B. It is precisely the same type of common cause we have been discussing in the case of the barometer reading and the storm.

A conjunctive fork is characterised by the following probability relations:

(1) $P(A.B|C) = P(A|C) \times P(B|C)$

(2) $P(A.B|\overline{C}) = P(A|\overline{C}) \times P(B|\overline{C})$

(3) $P(A|C) > P(A|\overline{C})$

(4) $P(B|C) > P(B|\overline{C})$ [167]

We can derive from these relations the following inequality:

$$P(A.B) > P(A) \times P(B)$$

The import of this inequality is that the existence of a common cause as defined by the conjunctive fork makes the joint occurrence of two distinct events much more likely than would otherwise be anticipated.[168] If, for example, sudden drops in barometric readings (A) and storms (B) were independent phenomena, the probability of a storm occurring just after an abrupt drop in barometric reading (P(A.B)) would be equal to the product of P(A) and P(B). In fact P(A.B) is much larger than that, because A and B are *not* completely independent phenomena. They have a common cause — the drop in atmospheric pressure.

---

165 "Concerning a probabilistic theory of causation adequate for the causal theory of time", p390

166 *The direction of time*, p159

167 I have used the notation given by Salmon in "Causality: production and propagation", pp158–9. The von Bretzel notation ("Concerning a probabilistic theory of causation adequate for the causal theory of time", p390) follows Reichenbach's presentation in *The direction of time*, p159 and the notation used by Salmon throughout *Statistical explanation and statistical relevance*.

168 "Concerning a probabilistic theory of causation adequate for the causal theory of time", p391

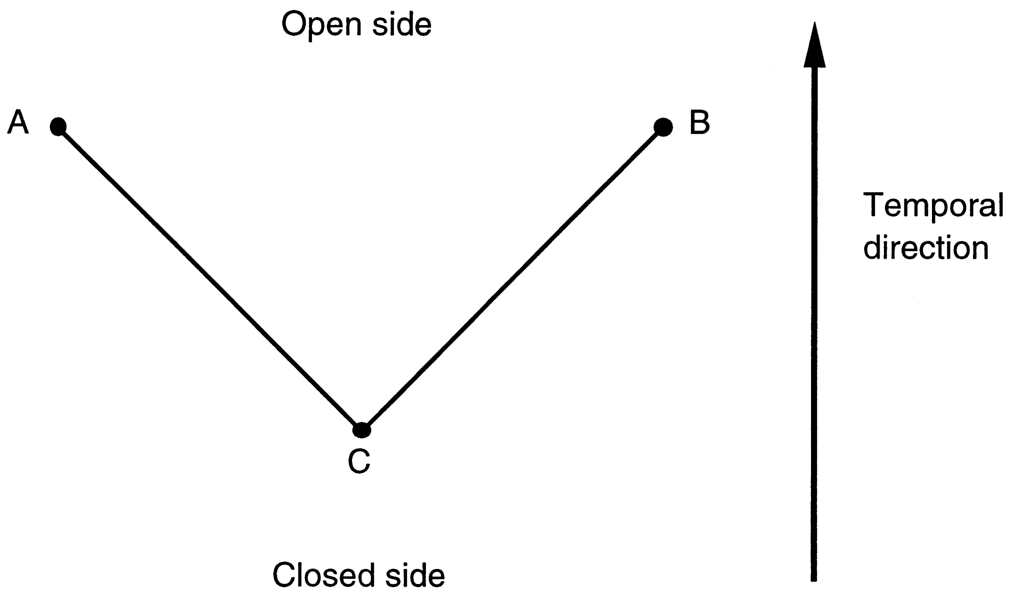
We can also derive from the four relations the previously-discussed screening-off relations, ie:

$$P(A|B.C) = P(A|C)$$

and:

$$P(B|A.C) = P(B|C)$$

The specification of direction of time follows rather simply. A conjunctive fork comprising events A, B and a common cause C are *open on one side* and *closed* on the other side. In any such fork C precedes A and B — the direction of time points to the "open" side of the fork. This is much easier to show than to say, and is demonstrated in figure 5.



Conjunctive forks determine temporal direction by being "open" to the future

Figure 5 — determination of the asymmetry of time by a conjunctive fork

Another way of expressing this point is that there is no "principle of the common effect" parallel to the "principle of the common cause" as defined by the conjunctive fork. Triads of events "closed to the future" but "open to the past", such as two independent events A and B with a common effect E, simply do not form conjunctive forks.[169] (The argument, presumably, is that there are no such configuration of events with probability values that fulfil the conditions for a conjunctive fork.[170]) This is shown in figure 6.

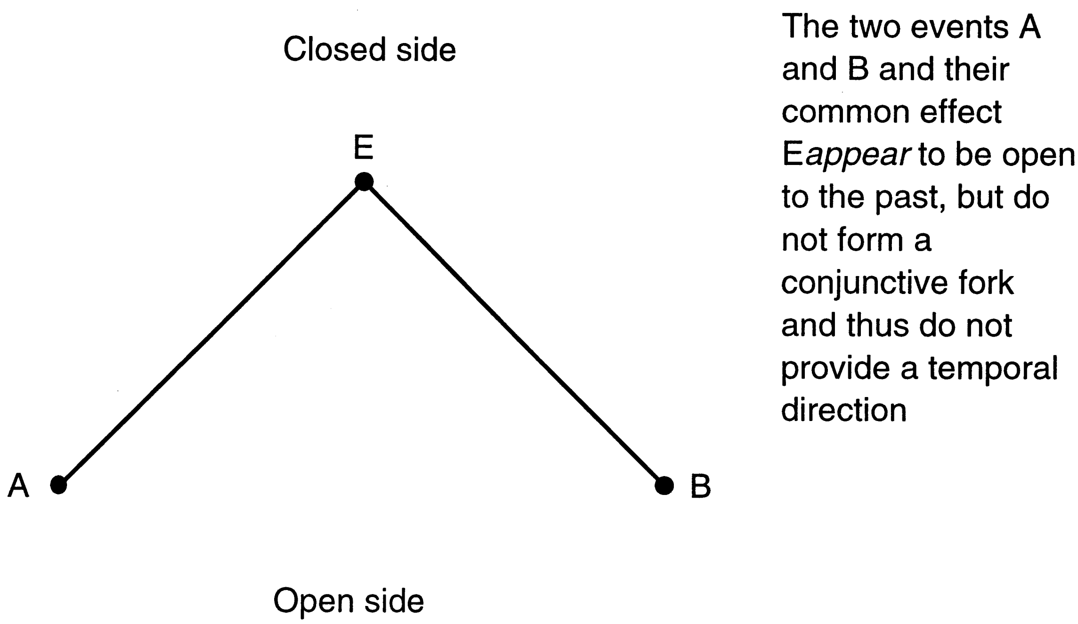


Figure 6 — common effects do not determine temporal directions

On the other hand, Reichenbach and von Bretzel allow that in the case of two events A and B with both a common effect E *and* a common cause C, the events A, B and E *do* form a conjunctive fork. But such conjunctive forks are not open to the past — they are closed by the common cause C. So by von Bretzel's criterion, neither the conjunctive fork A, B and E (which is closed on both sides) nor the conjunctive fork A, B and C (which is closed on both sides) would provide a temporal direction. Figure 7 illustrates this situation.

---

169 Originally a conjecture by Reichenbach, reaffirmed by Salmon in "Causality: propagation and production", p292.

170 This appears to be the line taken by Salmon in *Scientific explanation and the causal structure of the world*, p164. It just does not seem to be the case that in our world common effects conspire to produce statistical correlations in prior independent causes. The asymmetry thesis, therefore, should be understood as an empirically derived conjecture about the world.

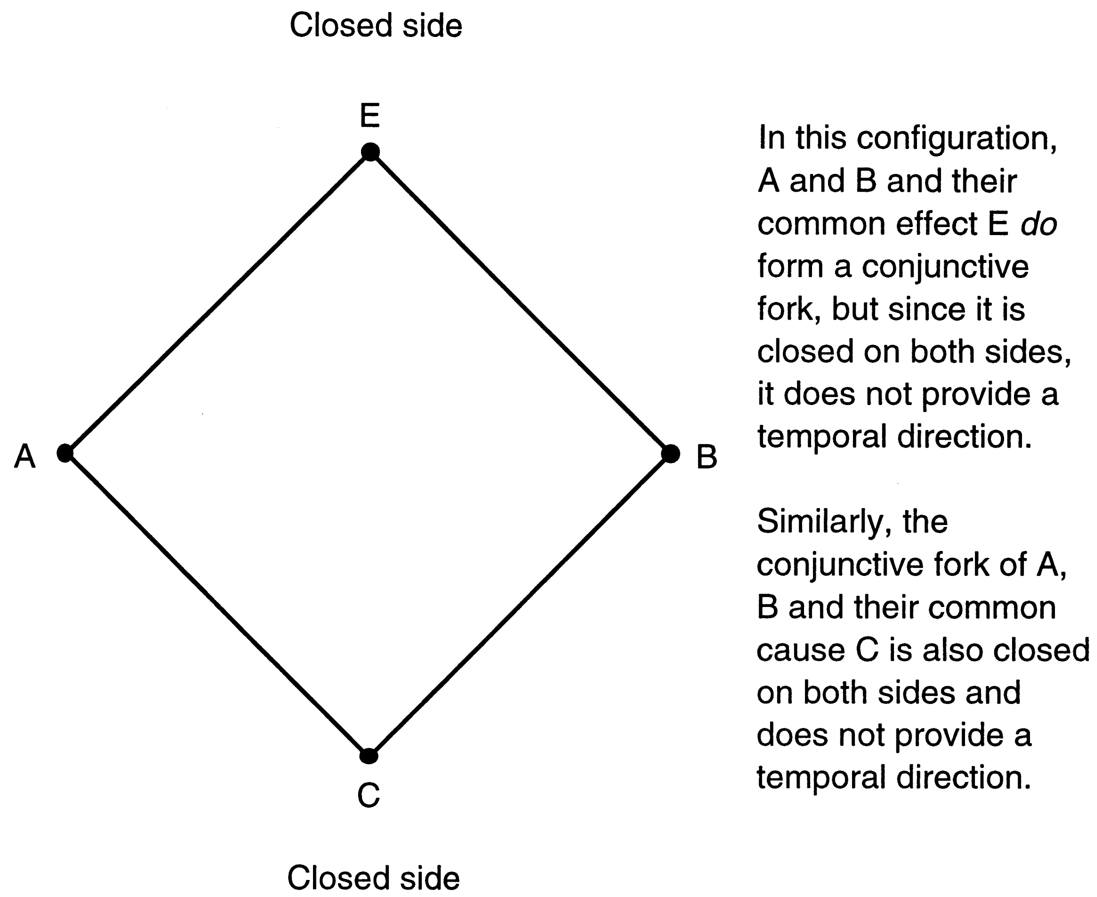


Figure 7 — "closed" common effects do not provide a temporal direction

In later papers, Salmon makes a stronger conjecture: even in such tetrads of events as illustrated in figure 7, two events and a common effect do not form a conjunctive fork. A corollary of this conjecture is that any conjunctive fork, whether it is closed by a common event or not, points in the same temporal direction.[171] No proof of the conjecture is provided in the paper, however, and to the best of our knowledge no such proof has *ever* been published.

A hitherto unstated point about conjunctive forks is that the causal connections between the triad of events — the "arms" of the forks — are

171 See, for example, "Causality: propagation and production", p292–3. Indeed, this may not be the sort of conjecture which is amenable to *logical* proof. Like Reichenbach and von Bretzel's asymmetry theses, Salmon's conjecture is empirical. As he notes in *Scientific explanation and the causal structure of the world*, pp166–7, probability values can be assigned to a tetrad of events A, B, C and E such that both A, B and E and A, B and C do form conjunctive forks. Yet there don't seem to be any examples of such a double fork in *this* world. Salmon suspects that such an arrangement of events may violate some "basic physical principle", but does not know what this principle might turn out to be.

processes. Indeed the processes need to be causal processes rather than pseudo processes. The rationale is never stated in any detail by von Bretzel, but would appear to be based on the fact that pseudo processes (such as Salmon's example of the moving spot of light) can propagate faster than the speed of light. If such processes transmitted causal influence, then it would be the case that causal influence from an event could be transmitted to events which happened earlier than the event. In that case we could find correlations between events associated with their common effects — there would be conjunctive forks which are open to the past. Hence the asymmetry of temporal direction could not be based on the asymmetry of conjunctive forks — because causal forks would not be asymmetrically directed.

Hence von Bretzel needs to find a criterion to distinguish between causal processes and pseudo processes in *Reichenbach's* analysis of causation (von Bretzel does not explicitly deal with Salmon's at-at criterion[172]). He notes that the typical move in such circumstances is to invoke a genidentity criterion.[173] To use Salmon's example of the rotating beacon and the spot cast on the wall, we might say that the beacon is a causal process because it involves the same object at all stages of the process. The moving spot does not involve the same object at every stage: it comprises many different bits of wall.[174] However, von Bretzel finds the notion of genidentity problematic, and prefers to try to give a criterion in terms of the conjunctive fork and Reichenbach's principle of causal betweenness.

Discussion of Reichenbach's analysis is beyond the scope of this thesis so we will simply state some of von Bretzel's observations and conclusions rather than his rationales. Much like Salmon in "An 'at-at' theory of causal influence", von Bretzel notes that pseudo processes such as the moving spots are continuously generated by mechanisms other than themselves: in von Bretzel's words, they are derivative from causal processes — in the case of the spot,

---

172 This is rather odd. Von Bretzel appears to have been a student of Salmon's at the University of Arizona, cites Salmon several times in the paper, and the paper is published the year after "An 'at-at' theory of causal influence" has been accepted for publication — von Bretzel *must* have been aware of at-at.

173 Unlike Aronson or Fair, von Bretzel cashes out genidentity as holding between events containing the same object (see "Concerning a probabilistic theory of causation adequate for the causal theory of time", p396).

174 As we shall discuss in a later section, Dowe uses such a genidentity criterion to distinguish causal processes from pseudo processes in his conserved quantity process analysis.

derivative from the beams of light produced by the rotating beacon. This is demonstrated in figure 8.

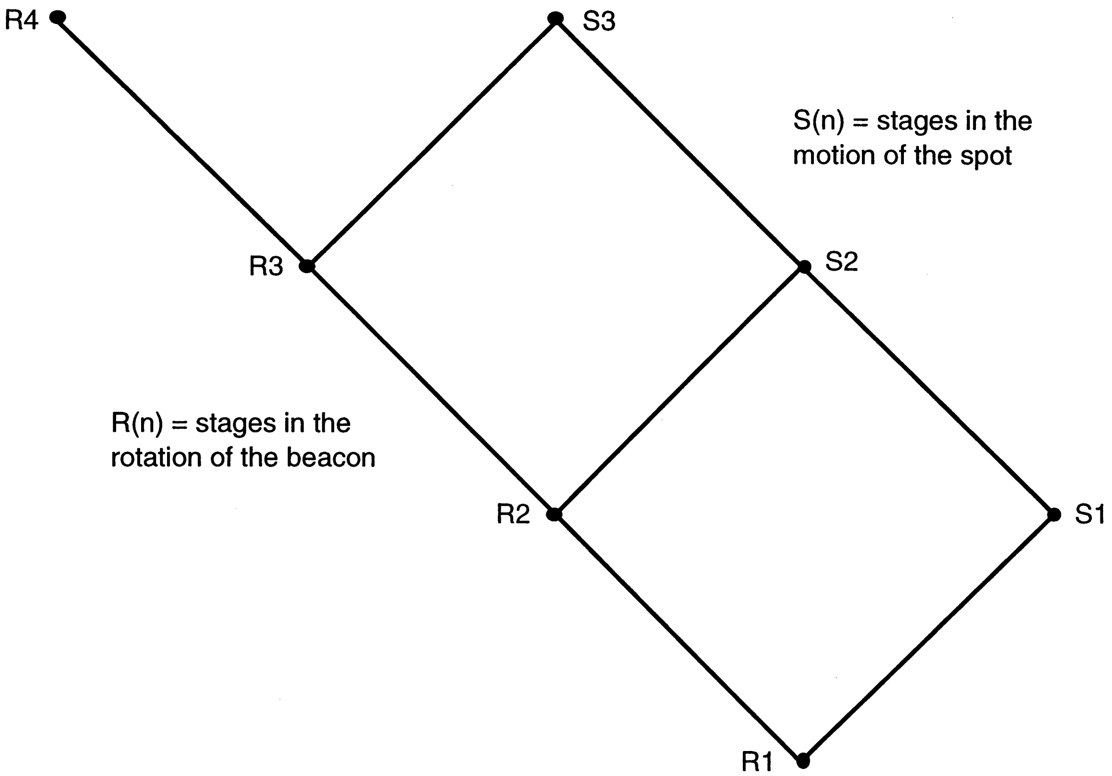


Figure 8 — pseudo process of the moving spot is derived from the causal process of the rotating beacon

The R-series of events is a causal process (stages in the rotation of the beacon). The S-series is a pseudo process (stages of the motion of the spot). The lines (R_1,S_1), (R_2,S_2), (R_3,S_3) are also causal processes (beams of light connecting the spot and the beacon).

So? The key assertion is that the forks (R_2,S_1,R_1), (R_3,S_2,R_2), (R_4,S_3,R_2) are *conjunctive* forks. R_1, R_2 and R_3 are all common causes. They are also precisely what we would want to identify as causal interactions — a beacon releasing light. Indeed this is precisely the same sort of causal interaction as occurs in our original example drawn from Gasking and Aronson of the glowing hot iron bar. It is no surprise that von Bretzel concludes:

> ... causal interactions form conjunctive forks, and consequently, common causes. If there were no interactions, there would be no common causes, and therefore, no explanations for coincident statistical relevancies and no direction to either causal

processes or time. The events which comprise the nodes of interaction between different causal processes are therefore quite basic.[175]

Causal interactions ≠ conjunctive forks

This conjunctive fork analysis of causal interactions obviously has great potential for transfer to Salmon_1. Since the conjunctive fork is simply a statement of the common cause phenomenon discussed in "Causal and theoretical explanation", it looks like Salmon_1 can be provided with an account of causal interaction to support the at-at account of processes (*and* a causal theory of time) without having to posit any additional theoretical bells and whistles. Salmon_1 probably did take this form for some time — indeed, Salmon frequently notes that von Bretzel alerted him to the direct relationship between causal interactions and "causal forks".[176] But no paper on Salmon_1 is ever published to this end. The beautiful theory of at-at processes plus conjunctive fork causal interactions is spoiled by a couple of nasty little facts.

One of the facts is identified by Bas van Fraassen in *The pragmatics of explanation*:[177]:

> Salmon's point of departure is Reichenbach's principle of the common cause ... This means that a correlation of simultaneous values must be explained by a prior common cause C of events A and B. Salmon gives two statistical conditions that must be met by a common cause C of events A and B:
>
> (a) $P(A.B|C) = P(A|C) \times P(B|C)$
>
> (b) $P(A|B.C) = P(A|C)$ "C screens off B from A."
>
> If $P(B|C) \neq 0$ these are equivalent and symmetric in A and B.
>
> ... To assume Reichenbach's principle to be satisfiable ... is to rule out all genuinely indeterministic theories. As example, let a theory say that C is invariably

175 "Concerning a probabilistic theory of causation adequate for the causal theory of time", p400

176 For example, see *Scientific explanation and the causal structure of the world*, pxi, and "Why ask, 'why?'?", p141.

177 van Fraassen, Bas: "The pragmatics of explanation", pp143–50, *American philosophical quarterly* vol 14 (1977)

followed by one of the incompatible events A, B, or D, each with probability $\frac{1}{3}$. Let us suppose the theory complete, and its probabilities irreducible, with C the complete specification of state. Then we will find a correlation for which only C could be the common cause, but it is not. Assuming that A, B, D are always preceded by C and that they have low but equal prior probabilities, there is a statistical correlation between $\phi = (A \vee D)$ and $\psi = (B \vee D)$ for $P(\phi|\psi) = P(\psi|\phi) = \frac{1}{2} \neq P(\phi)$. But C, the only available candidate, does not screen off ϕ from ψ: $P(\phi|C.\psi) = P(\phi|\psi) = \frac{1}{2} \neq P(\phi|C)$ which is $\frac{2}{3}$. Although this may sound complicated, the construction is so general that almost any irreducibly probabilistic situation will give a similar example.[178]

Salmon gives a rather more accessible description of the problem in "Causality: production and propagation":

> Consider a simple example. Two pool balls, the cue ball and the 8-ball. lie on a pool table. A relative novice attempts a shot that is intended to put the 8-ball into one of the far corner pockets; but given the positions of the balls, if the 8-ball falls into one corner pocket, the cue ball is almost certain to go into the other far corner pocket, resulting in a "scratch". Let A stand for the 8-ball dropping into the one corner pocket, let B stand for the cue ball dropping into the other corner pocket, and let C stand for the collision between the cue ball and the 8-ball that occurs when the player executes the shot. Assume that the probability of the 8-ball going into the pocket is $\frac{1}{2}$ if the player tries the shot, and that the probability of the cue ball going into the pocket is also about $\frac{1}{2}$. It is immediately evident that A, B and C do not constitute a conjunctive fork, for C does not screen A and B from each other. Given that the shot is attempted, the probability that the cue ball will fall into the pocket (approx $\frac{1}{2}$) is *not* equal to the probability that the cue ball will go into the pocket given that the shot has been attempted and that the 8-ball has dropped into the other far corner pocket (approx 1).[179]

Intuitively, the collision of the cue ball with the 8-ball is pretty clearly both a causal interaction *and* a common cause of A and B. So not all common causes or causal interactions could be conjunctive forks.

---

178 "The pragmatics of explanation", p146 (I have harmonised van Fraassen's notation with the version we have been using to date.)

179 "Causality: production and propagation", pp293–4. See also *Scientific explanation and the causal structure of the world*, pp168–9, and a different explication using the "irreducibly probabilistic" example of Compton scattering in "Why ask 'why?'?", pp133–4.

Deviating from the strict chronological sequence of development of $Salmon_1$, we can note a further problem with the thesis that causal interactions are conjunctive forks: not every triad of events fulfilling the statistical conditions of the conjunctive fork turns out to be an instance of a common cause, let alone a causal interaction.

Salmon cites a case raised by Ellis Crasnow. A woman usually arrives at her office at 9am and makes herself a cup of coffee. But sometimes she arrives promptly at 8am and her secretary has already made her coffee for her. On precisely these latter mornings she is met at the office by an associate who works at a different office. That the coffee is made for her at the office (A) and the associate shows up those mornings (B) is a coincidence that needs to be explained. It could be noted that on ordinary mornings she catches the 8am bus (C) and on the meeting mornings she takes the 7am bus ($\overline{C}$). Plausibly, the events A, B and C fulfil the four statistical conditions for a conjunctive fork, yet C is clearly not a cause of either A or B.[180]

It might appear that $Salmon_1$ has more resources to deal with Crasnow's example than van Fraassen's problem. Both Salmon and von Bretzel argue that true conjunctive forks involve causal connections between the triad of events in the form of causal processes. Hence the bus ride can be ruled out as being a true common cause in Crasnow's example — it simply doesn't possess the right connections. Yet such a strategy rules out conjunctive forks as causal interactions. Interactions would be partially defined in terms of causal processes, but under the at-at theory processes are defined in terms of causal interactions. The circularity looks vicious.

180 This version of Crasnow's example is inspired by the 1981 paper "Causality: production and propagation", pp293, but Crasnow examples first make their appearance in p217, Salmon, Wesley: "Probabilistic causality", pp50–74, Pacific philosophical quarterly vol 61 (1980), reprinted in pp208–32, Salmon, Wesley: *Causality and explanation* (New York: Oxford University Press, 1998). All page references are to the latter source.

Causal interactions = interactive forks

The van Fraassen example is intended to show that common causes cannot provide a basis for scientific explanation — a step along the way to the conclusion that science contains no explanations at all.[181] Salmon draws a different moral — understandably, given that $Salmon_1$ is intended to support a theory of scientific explanation. There must be more than one kind of "causal fork".[182] This notion is first pursued in the 1977 paper, "Why ask 'why?'?".

Again, Salmon's primary aim in "Why ask 'why?'?" is to develop his theory of scientific explanation. The focus is on the philosophical importance of the principle of the common causes as a basis for scientific explanation. (Note the dramatic change of emphasis from $Salmon_0$. In *Statistical relevance and statistical explanation*, common causes are more-or-less problem cases — now they are virtually the rock the theory is built on.)

Salmon claims that to understand common causes we need to distinguish between causal interactions and causal processes. The initial characterisation is by example: the transmission of light from one place to another or the motion of a material particle to one place or another is a causal process; collisions of billiard balls, or absorption or emission of photons, are causal interactions. Salmon notes that interactions are the sort of things that we tend to describe as events. (Given the earlier account of interactions being the means by which marks are introduced, Salmon is probably thinking in terms of a "change" interpretation of events.) In a further elaboration to the analysis of interactions, Salmon notes that interactions are also associated with the beginning and end of the propagation of a process — not just instances of modification of processes. Causal processes are distinguished from pseudo processes by the ability of causal processes to transmit causal influence, which Salmon interprets as the transmission of marks. The transmission of a mark is to be understood in terms of the at-at theory.[183]

At this point it might be expected that Salmon would give an account of causal interactions, perhaps beginning with an explanation of why conjunctive

181 "The pragmatics of explanation", p149

182 *Scientific explanation and the causal structure of the world*, pxi

183 "Why ask 'why?'?", p130–1

forks can't do the job. But since Salmon's aim is supporting his theory of explanation, the next stanza in the paper is framed as a discussion of "certain configurations of processes with special explanatory import"[184] — common causes. The overall strategy is to give an account of direct explanatory relevance (in terms of direct causal linkages by processes) and then give an account of explanatory correlations in terms of spatiotemporal configurations of processes where two states of affairs are not directly linked by a single process.

Using a semi-historical presentation, Salmon first treats of Reichenbach's characterisation of common causes in terms of conjunctive forks, as described by the four statistical conditions given in our discussion of von Bretzel. An example of a conjunctive fork type of common cause would be increased rates of leukemia amongst soldiers exposed to nearby atomic bomb blasts. This is to be explained in terms of a large number of *different* radiation or fallout processes being produced by the explosion, propagating outwards and then interacting with different soldiers. Screening off occurs in this circumstance. The incidence of leukemia in some soldiers exposed to the blast is statistically relevant to the incidence of leukemia in other soldiers exposed to the blast, yet the incidence in other soldiers is screened off by the blast.

Salmon notes van Fraassen's problem that not all common causes fulfil the conditions of the conjunctive fork, and provides a concrete example in Compton scattering. In Compton scattering, the sum of the energies of the emitted photon (E_1) and the post-emission electron (E_2) will equal the total energy of the excited photon (E). Suppose $P(E_1)$ and $P(E_2)$ are 0.1. If Compton scattering formed a conjunctive fork, $P(E_1.E_2)$ would equal 0.01. But $P(E_1.E_2)$ equals 0.1 — because the electron will have energy E_2 if and only if the energy of the photon is E_1.[185] Clearly in such cases the common cause does not screen the effects off from one another, because:

$$P(E_2|E) \neq P(E_2|E_1.E)$$

Salmon's conclusion is that there are two kinds of causal forks: conjunctive forks which are exemplified by the leukemia case; and *interactive* forks which

184 "Why ask 'why?'?", p131

185 "Why ask 'why?'?", p132–3. Our expression of this example differs slightly from Salmon's for purposes of explanatory brevity — Salmon treats E as the energy of the incident photon, which is only approximately equal to the sum of the energies of the emitted photon and the post-emission electron.

are exemplified by Compton scattering, or our earlier-discussed pool shot. The hitherto primitive notion of causal interaction is to be understood as the incidence of an interactive fork.

The initial characterisation given for the interactive fork is statistical. The interactive fork is a triad of events A, B and C connected by causal processes and fulfilling a similar set of statistical conditions to the conjunctive fork, except that the condition:

$$(1) \quad P(A.B|C) = P(A|C) \times P(B|C)$$

is replaced by the condition:

$$(1) \quad P(A.B|C) > P(A|C) \times P(B|C)$$ [186]

A further characterisation is also hinted at in terms of mutual modifications of two processes at a spatiotemporal intersection:

> What we want to say, very roughly, is that when two processes intersect, and both are modified in such ways that the changes in one are correlated with the changes in the other — in the manner of an interactive fork … — we have a causal interaction.[187]

In later papers, Salmon discards the statistical characterisation of interactive forks (and thus causal interactions). Dowe notes that Salmon "has never stated any reasons" for this move.[188] We shall soon discuss one apparent reason given in "Causality: production and propagation". Salmon claims elsewhere to have given an account of some of his concerns in "Causal propensities: statistical versus aleatory causality".[189] But we think the main problem with the statistical characterisation of the interactive fork is the same circularity problem suffered by the thesis that conjunctive forks are causal interactions. Suppose interactive forks were statistical correlations. The statistical frequency data from which we derive these correlations are generated by the underlying configurations of causal processes. Yet under the at-at theory causal processes are defined in

---

186 "Why ask 'why?'?", p133–4 and pp140–1

187 "Why ask 'why?'?", p135

188 "Wesley Salmon's process theory of causality and the conserved quantity theory", p206. Dowe proceeds to give four reasons of his own why the statistical characterisation is untenable.

189 "Causality without counterfactuals", p252

terms of causal interactions. So defining, or even *characterising*, causal interactions in terms of statistical correlations looks dangerously circular.

At this point of the development all the basic structures of Salmon$_1$ are in place: an at-at account of causal processes, an interactive fork account of causal interactions, and conjunctive forks. Further papers in Salmon$_1$ tend to concentrate on improving the characterisations and definitions of these concepts.

Digression — against conjunctive forks

Before outlining the main developments in latter papers, we will examine the mature notion of the conjunctive fork in Salmon$_1$. Why is it there? It plays no role in the description of processes, it plays no role in the description of interactions. Nor is it clear that any process analysis of causation needs the conjunctive fork as a primitive notion.

It is tempting to argue that the presence of the conjunctive fork in Salmon$_1$ is purely a matter of historical contingency — The early Salmon$_1$ is concerned with common causes as characterised by the conjunctive fork, so the conjunctive fork maintains a vestigial presence in the latter Salmon$_1$. But this would be more than a little harsh. The conjunctive fork maintains rather a lot of causal functionality.

The most significant function is to provide a direction to causation. As Reichenbach, von Bretzel and Salmon all note, conjunctive forks are only open to the future if they are open at all. Interactive forks display no such temporal asymmetry. This is, as we have noted before, an empirical observation. To use an example provided by Salmon, if we encounter a pattern of symmetrical ripples on a pond, we find that it has been produced by some happenstance at the centre of the pond — like a dropped pebble — the influence from which propagates out to the edges of the pond. We do not find a multiplicity of happenings at the edges of the pond propagating influence towards the centre of the pond. We find a common cause may produce a myriad of correlated effects, but we do not find a collection of such correlated causes conspiring to produce a common effect. But consider Compton scattering. We may find a photon of energy E_1 impacting an electron of energy E_2 which then emits a photon of E_3, leaving the electron with energy E_4, where $E_1+E_2=E_3+E_4$. We could readily find instances of Compton scattering where the incident photon and electron have energies of E_3 and E_4 respectively, and the emitted photon and electron have energies of E_1 and E_2 respectively.[190] There is no such cause and effect asymmetry in the interactive fork parallel with that of the conjunctive fork.[191] Thus conjunctive forks provide a convenient means by which $Salmon_1$ can be used to provide a causal theory of time.[192]

The conjunctive fork also supplies Salmon with an answer to whether causation is a general relation holding between classes of events, or a singular relation holding between particular events. Under a Humean regularity analysis, causation is a general matter. For there to be *regular* conjunction of like circumstances, there must be more than one instance of conjunction of these circumstances. Any singular causal statement such as "A caused B" is therefore a disguised general statement: "things of the A kind cause things of the B kind", or somesuch.

---

190 This is clearly a different kind of temporal symmetry from that which we discovered in an adequate Aronson-Fair analysis. In that case there was no fact of the matter in an interaction between two electrons about whether a virtual photon was absorbed or emitted by a particular electron. Salmon seems to be merely asserting something parallel to the view that in the case of our two electrons proceeding from w and x to y and z that the measurements of energy/momenta at w and x could just as readily be found at y and z and *vice versa*.

191 "Why ask 'why?'?", p141

192 There is some reason to suppose that Salmon also takes the asymmetry of the conjunctive fork as a mechanism for determining the *temporal* asymmetry of explanation. We find in p345, Salmon, Wesley; Salmon, Merrilee: "Alternative models of scientific explanation", pp333–46, Salmon, Wesley: *Causality and explanation* (New York: Oxford University Press, 1998) (originally published in 1979) a hint that the temporal asymmetry of explanation is to be understood in terms of a causal theory of explanation and the temporal asymmetry of causation.

We find a much more explicit connection in pp174–5 of the 1991 paper Salmon, Wesley: "Explanatory asymmetry: a letter to Professor Adolf Grünbaum from his friend and colleague", pp164–77, Salmon, Wesley: *Causality and explanation* (New York: Oxford University Press, 1998) . Here, citing Reichenbach's principle of the common cause, he notes that the view seems to be correct that the "anisotropy of time, anisotropy of causality, and anisotropy of explanation all go together." Salmon must here be thinking in terms of the conjunctive fork. It is never *quite* clear whether Salmon thinks that the asymmetry of conjunctive forks might provide a solution to the other traditional problems of explanatory asymmetry — why does the height of the tower explain the length of the shadow but the length of the shadow not explain the height of the tower? We have a suspicion that such a strategy might underlie the criticism of van Fraassen in Salmon, Wesley; Kitcher, Philip: "Van Fraassen on explanation", pp178–90, Salmon, Wesley: *Causality and explanation* (New York: Oxford University Press, 1998). But there is no direct evidence in the text.

Salmon notes that a singular causal process can transmit causal influence between (roughly speaking) a single cause and a single effect. Such singular statements as "A caused B" are therefore not disguised generalisations under $Salmon_1$. However, the conjunctive fork is defined in terms of statistical frequencies and thus refers to sets of causes and effects. The conclusion is that the debate over singular versus general causation is therefore explained by causation being a partly singular and partly general phenomena: the former aspect provided by causal processes, the latter by conjunctive forks.[193]

Despite these useful features of the conjunctive fork, it does not seem to be either clearly conceived or thoroughly at home in the ontology of $Salmon_1$. We find in "Why ask 'why?'?" and later work such as the 1981 paper "Causality: production and propagation" that conjunctive forks are to be associated with the *productive* aspect of causality. Conjunctive forks play a role in the production of independent causal processes from a special set of background conditions.[194] Here Salmon is clearly thinking about cases like the atomic bomb blast producing radiation processes. Yet if anything could be said to produce these processes, it must be a set of causal interactions: nuclear fission, nuclear fusion, decay of radioactive byproducts, radiation from unstable isotopes of elements produced by irradiation from the preceding sources, and so forth. The conjunctive fork does not appear to be a certain kind of basic causal mechanism which underlies a certain set of statistical relevance relations. At best it is a description of a certain *configuration* of basic causal mechanisms which has useful explanatory import.

Similarly, not all instances of conjunctive forks as described by Salmon involve production of processes. To adapt one of Salmon's examples, consider Bill and Stephen who go for a walk through the woods and for lunch sample the local fungi. Later that day, both Bill and Stephen come down with similar horrendous stomach cramps. It seems highly improbable that both chaps would

---

193 See *Scientific explanation and the causal structure of the world*, p182 and "Causality: production and propagation", p300. This general character of statistical correlations may also indicate another reason why Salmon moves away from a statistical characterisation of the interactive fork. A causal interaction involving a spatiotemporal intersection of two processes undergoing persistent modification at the point of intersection seems a thoroughly *singular* set of affairs. Yet the statistical characterisation of the interactive fork seems precisely as *general* as that of the conjunctive fork. See also Dowe's "Wesley Salmon's process theory of causality and the conserved quantity theory", p207.

194 "Causality: production and propagation", p298, see also *Scientific explanation and the causal structure of the world*, p179

independently suffer stomach cramps on the same afternoon. We naturally suspect that Bill's cramps (A) and Stephen's cramps (B) have a prior common cause which makes it much less improbable that B and C should come about on the same afternoon — the lunch of pretty mushrooms (C).[195] Salmon claims that this set of affairs forms a conjunctive fork. We are inclined to agree with him. Yet the causal processes here must certainly be Bill and Stephen. Bill and Stephen do not pop into existence as a result of the lunch of mushrooms.

Intriguingly, this "walk in the woods" variety of conjunctive fork displays a much less convincing degree of temporal asymmetry than the "atomic blast" variety. It may be inconceivable that perfectly circular ripples propagate from the edges of the pond to a point in the centre. It does not seem so enormously odd that two friends with tummy troubles might go for a walk in the woods, both eat a mushroom with antinausea effects, and stroll away feeling much more healthy. Or, to keep the symmetry more exact, the two friends might both stop to vomit and then feel much better once their stomachs had been emptied.

Let's pursue this latter kind of "common cause" a little further. Consider a closely-related example. A doctor is treating a patient for acute gastrointestinal discomfort. Suddenly, the doctor realises she has seen these symptoms before. Checking the patient records, she finds that she has treated five different patients with these symptoms over the past 10 years. Each patient has presented in autumn. Each patient has been an outdoorsy type. Pursuing a sudden hunch, she checks a guide to the local fauna, and discovers that the *Amanita reichenbachia* mushroom sprouts in the local woods during autumn, and if eaten produces precisely the symptoms she has been treating her patients for.

Would this set of circumstances constitute a conjunctive fork? It is tempting to say no: the five patients do not "sup together out of a common pot",[196] nor do the patients presenting for treatment amount to simultaneous events for which we search for a common (meaning one-and-the-same) cause.[197] Such a construal would describe the second mushroom poisoning case as five separate patient presentations (A, B, C, D and E) connected by five different causal process to

[195] "Causality: production and propagation", pp289–90
[196] "Causality: production and propagation", p299
[197] *The direction of time*, pp158–9

five different *A reichenbachia* ingestions (F_1, F_2, F_3, F_4 and F_5) with no fork structure.

But there is something profoundly unsatisfying about this move. If we were to treat F as a *type* of event — eating *A reichenbachia* — then, so long as we don't add in Salmon and von Bretzel's extra requirement that the relevant events are connected by causal processes, there seems no question that any triad like (A,B,F) would be any less a conjunctive fork than Bill and Stephen falling ill after their walk in the woods. In addition, it seems that an adequate account of general causation (which Salmon suggests conjunctive forks provide) would count *A reichenbachia* as a genuine general causal correlation.

The *A reichenbachia* poisoning case seems to be an inverse version of Crasnow's example. In Crasnow's example, existence of probabilities fulfilling the conditions of the conjunctive fork produce spurious causal correlations. Causal processes are introduced to the analysis of the fork to eliminate the spurious correlations. In *A reichenbachia*, probabilities exist which fulfil the conjunctive fork, and we have what seems to be a clear causal correlation between independent events, yet when we introduce the causal processes the correlation is made spurious.

We can find *A reichenbachia* analogues for many other classic examples of the conjunctive fork. Salmon gives the example of two students handing in identical papers. We can exclude the possibility that they copied from each other. We infer there is a common cause: a paper in a sorority or fraternity file, for example. The file paper, and the two students' papers form a conjunctive fork.[198] Let's extend this example by noting that it is becoming more and more common for papers to be posted on the internet. It is entirely common for internet sites to be mirrored by other internet sites. Two students hand in identical papers. We think we can confirm that they did not copy from each other: one lives in Argentina, one lives in Turkey. Then we discover that they have internet access. We determine from traffic records that they could not have contacted each other, but discover that they have both found the paper on the internet. But they copied the paper from different mirror sites. Treat the two mirrored papers as the same paper type, and we have a conjunctive fork. Add the causal processes in and the correlation is spurious.

---

198 "Causality: production and propagation", p289

The problem with these two cases is that types of causes look like they should produce genuine causal correlations as much as single causes. Let's call this the principle of the *like* cause. Suppose we wish to rule out such cases as conjunctive forks. We do so by asserting that the effects in a conjunctive fork. must be linked by causal processes to a single token cause — a *common* cause. One danger of such a strategy is that it would seem to rule out classical examples of conjunctive forks as merely being *like* causes. Bill and Stephen may sup from the same pot, but is that a truly common cause? Do they have to eat from the same mushroom? Do they have to swap mouthfuls of the same mushroom?

It could also be argued that the strategy of ruling out like causes as conjunctive forks is incoherent. Remember that the conditional probabilities cited in the four conditions for the fork are derived from long-run frequencies. Can data for such probabilities be generated without appealing to cause types?

In the case of the atomic blast, it looks like it might be possible. Consider condition 1:

$$P(A.B|C) = P(A|C) \times P(B|C)$$

So long as a bomb blast (C) involved exposure of enough servicemen, it seems that we can derive the conditional probabilities of being a serviceman contracting leukemia (A) given exposure, being another serviceman contracting leukemia (B) given exposure, and two different servicemen contracting leukemia given exposure (A.B), by reference to a single bomb blast.

Clearly, however, the probabilities are not derived from reference to a *single* serviceman A and a *single* serviceman B. There seems something *ad hoc* about restricting probability derivations to token causes when we are forced to make these derivations from type effects.

Now consider the barometer example. Here it is surely the case that the conditional probabilities of a drop in barometric reading (A) given a drop in atmospheric pressure (C), a storm (B) given a drop in atmospheric pressure, and a drop in barometric reading *and* a storm given a drop in atmospheric pressure have been derived by reference to many different drops in atmospheric pressure as well as many different barometric readings and storms. In fact we might argue that this kind of derivation of probabilities from many different "trials" involves better data acquisition methods than the atomic blast example, and we should have better confidence in the values of the probabilities that we derive in this case. There is definitely something dissonant about restricting conjunctive forks — via interpolation of causal processes — to cases of token causes, when the statistical relations that the fork must obey have been derived by reference to type causes.

Let's summarise what we think has gone wrong: Reichenbach's concept of the conjunctive fork does not provide an analysis of the principle of the *common* cause, it provides a characterisation of the principle of the *like* cause. The conjunctive fork describes a statistical correlation between type causes and type events, not all of which are common causes.[199] Unfortunately, it fails as a criterion for determining whether a general causal relation holds between types of events because it holds in cases of spurious causal connections, such as Crasnow's example. When the definition of the conjunctive fork is expanded in Salmon$_1$ to include reference to connecting causal processes, and thus rule out Crasnow examples, genuine causal correlations such as *A reichenbachia* are rendered spurious. So even when amended, the conjunctive fork fails as a characterisation of general causation.

We have noted that the conjunctive fork does not appear to be a basic causal mechanism in Salmon$_1$. We have also noted that even in a world like ours only some types of conjunctive fork appear to be associated with irreversible processes (this association being taken as evidence for conjunctive forks only being open to the future). This casts serious doubt on the conjunctive fork forming the basis of a causal account of either temporal asymmetry or temporal explanatory asymmetry. At best, we might say that conjunctive forks are *accidentally* associated with asymmetries.

Our conclusion is that the conjunctive fork is an optional extra to the Salmon$_1$ process analysis of causation. Arguably, Salmon also comes round to this opinion. In the late Salmon$_1$ paper "Causal propensities: statistical causality versus aleatory causality" (1990), he argues that satisfactory accounts of causation place primary emphasis on causal mechanisms (processes and

---

199 It is a shame that Reichenbach died before presenting *The direction of time* for publication. The problem we have identified might well have been identified in the dialogue between author, editor and referees, and no doubt addressed by Reichenbach in the final text.

interactions) "but does not disdain statistical regularities". Conjunctive forks are clearly identified as a form of statistical regularity.[200] Placing secondary emphasis on conjunctive forks surely shows *some* disdain for their ontological status. Perhaps the key to the continued presence of the conjunctive fork in $Salmon_1$ is, as ever, Salmon's ultimate goal of a theory of explanation. Statistical regularities — including the conjunctive fork — may not have enormous ontological significance, but they do have great explanatory significance. Indeed, as suggested in "Causality and explanation: a reply to two critiques", they may have explanatory status which is not reducible to the underlying ontology.

Ontological interactive forks

Let's return to the development of $Salmon_1$. In the 1981 paper "Causality: propagation and production" we find a reiteration of the earlier themes. Causal processes are associated with the propagation of causal influence, which is explained in terms of the at-at theory of transmission of marks. Causal interactions are interactive forks.

The probability relations associated with the interactive fork are given (as per "Why ask 'why?'?"), but the actual characterisation of the interactive fork is no longer statistical. Salmon claims only that *many* cases of causal interactions are associated with the probability relations[201] and in a supporting footnote, Salmon also remarks that the probability relations do not form part of the definition of either causal interactions or interactive forks.[202]

Salmon's diffidence is partially explained by the phenomenon of the *perfect* fork. A perfect fork is a causal fork where the conditional probabilities of A given C or B given C are either 0 or 1. As Dowe puts it, such forks are deterministic limits of causal forks.[203] These forks are significant because, for example, in the case where:

$$P(A.B|C) = P(A|C) \times P(B|C) = 1$$

200 "Causal propensities: statistical causality versus aleatory causality", p207
201 "Causality: production and propagation", p295
202 "Causality: production and propagation", p301
203 "Wesley Salmon's process theory of causality and the conserved quantity theory", p198

the fork may represent a limiting case of either a conjunctive fork *or* an interactive fork (even though such a fork violates condition 1 of the "Why ask 'why?'?" statistical characterisation of the interactive fork).[204] If the interactive fork can not be distinguished from the conjunctive fork on purely statistical grounds it can not be given a purely statistical characterisation.[205]

The new characterisation is ontological. The ontological characterisation of the interactive fork locates the fork at spatiotemporal intersections of two processes. Yet there is more to an interactive fork (and hence causal interaction) than mere intersection. Firstly, the intersection must be between two casual processes, not two pseudo processes nor a causal process and a pseudo process. Secondly, the two intersecting processes acquire a persistent modification at the point of intersection which persists beyond the point of intersection.[206] A space-time diagram of such a causal interaction would take the form of an X, and the interactive fork can thus be referred to as an "X-type" interaction.[207]

Strictly speaking, where Salmon refers to "spatiotemporal intersections of processes", we should substitute "spatiotemporal intersections or interfaces of processes". The exclusive talk of intersections appears to stem from Salmon's favourite examples of causal processes. When the beam of white light passes through the red filter and is marked by becoming red (an example of a causal interaction), the obvious intuitive claim is that there is a spatiotemporal region jointly occupied by both the beam of light and the filter. So there is at least a *prima facie* case for this interaction involving an intersection of processes. But consider another "macroscopic" example of a causal interaction: the white billiard ball striking the red. Clearly we should class this as a causal interaction under $Salmon_1$, yet clearly — apart, perhaps, from some surface mingling of the contact zones — we never find the white inside the red or the red inside the

---

204 "Causality: production and propagation", pp297–8

205 The perfect fork is significant only as a problem case in $Salmon_1$. Since conjunctive forks determine temporal direction and interactive forks do not, one must determine (by other means than statistical considerations) whether a perfect fork is a deterministic limit of a conjunctive fork or an interactive fork before using it to determine a temporal direction.

206 We must be careful of loose talk here. "Beyond" implies there is a temporal direction to the interactive fork, but as Salmon notes, interactive forks display no temporal asymmetry. Strictly speaking, we should say that two processes intersect and each process has different characteristics on each side of the intersection.

207 The term X-type is actually introduced by Salmon in the later (1984) work *Scientific explanation and the causal structure of the world,* p181, but it is clear from the characterisation of the interactive fork in "Causality: production and propagation" that they are X-type interactions.

white. It is much more accurate to describe this interaction as involving a spatiotemporal interface rather than an intersection.

The interface interpretation is even more warranted at the quantum physical scale. Consider Compton scattering. When the electron "absorbs" the incident photon, it is *not* the case that there is a photon sloshing around inside the electron like water in a sponge. The "post-absorption" electron is simply a more energetic electron that the "pre-absorption" electron. Similar circumstances obtain in beta decay, where a proton "emits" a virtual W boson which then decays into a positron and a neutrino (and the proton becomes a neutron). W bosons do not lurk hideously in nucleons waiting to be emitted, and electrons and positrons do not lurk hideously within the bosons waiting to be emitted. They only exist on the other side of the "interface" with the "originating" particle.[208]

Let's apply this version of $Salmon_1$ to the bouncing of Frank. Consider Frank's big last leap. This can be expressed in terms of three processes and two interactions. The Moon (process one) strikes the Earth (process two). A contact zone, a persistent modification in kinetic energy and momentum distributions propagates through the Moon, and another through the earth (interaction one). The modification in the Earth is transmitted to Frank (process three). Frank is modified by gaining a great deal of momentum and kinetic energy and the surface of the Earth is modified by losing some energy and momentum to Frank (interaction two).

It may appear, given the example we have chosen, that $Salmon_1$ converges with the transference analysis at this point. Indeed, in "Causal and theoretical explanation" we have already seen that most, if not all causal processes transmit energy, and in "Causality: production and propagation" Salmon notes that most causal interactions — perhaps all — involve momentum and energy transfers from process to process, and cites Fair's transference analysis as providing an illuminating account of the role of energy and momentum transfer in causation.[209] Salmon also suggests that all fundamental physical interactions can be interpreted in terms of the interactive fork.[210] Such a

---

208 To the extent that the uncertainty principle allows us to talk meaningfully of interfaces in these cases.
209 "Causality: production and propagation", p294 and supporting footnote p301
210 "Causality: production and propagation", pp298–9

programme seems entirely consonant with the Heathcote conjecture, that causal interaction is properly associated with the interactions described by an appropriate quantum field theory.

Yet Salmon still holds that the defining feature of causal interactions and processes is the introduction and propagation of marks. As we saw in an earlier quote from Dowe, Salmon (with Fair in mind) argues that processes should not be associated with transmission of energy, because there is no way of distinguishing between the transmission of quantities of energy (in a beam of light, for example) and the regular appearance of energy (in a moving spot of light being cast on a wall by the same beam of light). In other words, a process analysis in which processes are defined as transmissions of energy does not adequate distinguish between causal and pseudo processes.[211] Conversely, the edge of the expanding contact zones need not qualify as a mark simply on the basis of its energy density. It is also a deformation in the crystalline structure of the Earth. We could also cash out all the interactions in the case of the bouncing of Frank in terms of modifications of crystalline structure, not just modifications in energy/momentum density.

Note that this new account of causal interactions locates interactions in the same place as a basic Aronson-Fair analysis. A spatiotemporal intersection of two processes corresponds fairly closely to the point of contact of two

---

211 The objection is neither relevant nor correct. Under a process analysis, if two circumstances are connected by a causal process, then one circumstance could causally depend on the other regardless of whether or not we could *tell* that the two circumstances were causally connected. If we consider an adequate Aronson-Fair analysis to be a variety of process analysis, it would fail to have an adequate causal/pseudo process distinction if it did not have an ontological criterion for distinguishing between regular appearance and transmission of energy. In fact, Aronson-Fair *does* have such a criterion: transmission involves the propagation of genidentical quantities of energy/momenta and regular appearance does not. The energy/momentum of the moving spot is constituted by many different bits of energy/momenta, so it is a pseudo process not a causal process. The problem for Aronson-Fair is that its particular criterion does not hold in this universe. To use the distinction we drew at the beginning of this chapter, Aronson-Fair fails the second challenge for an adequate process analysis rather than the third challenge.

It even seems, *prima facie*, that we could ontologically distinguish between the transmission of energy/momentum and mere regular appearance of energy/momentum by adopting an at-at theory of energy/momentum transmission. Photons propagate along light beams in the absence of local causal interactions, so the light beam is a causal process. The photons in the spot appear as a consequence of continual local interactions between the beam and the wall, so the spot is a pseudo process. Indeed, in the mature form of $Salmon_2$ beams of light qualify as causal processes for precisely this reason.

interacting middle-sized objects. But it locates interactions in a different place to an Aronson-Fair analysis which respects the Heathcote conjecture.

Consider our electron-only model of physical interaction. In the case of two electrons repelling each other, an adequate Aronson-Fair analysis locates the causal interaction in the exchange of the virtual photon. Effectively, the interaction is located in an intervening *process*. In the case of a collision between macroscopic objects such as the white and the red, the interaction is located in the cascades of virtual photon exchanges spreading through the contact zones of the objects. Under Salmon$_1$, there should be two interactions in the single electron case: one at each point of *interface* between the virtual photon processes and the electron processes (these interfaces correspond to the couplings we discussed in our disproof of genidentity at the quantum scale). Similarly, in the macroscopic case we can locate interactions at each coupling of a virtual and/or real quantum-scale object in the cascade. But we could also legitimately locate the interaction at a *single* point — the point of impact of the white and the red. Thus Salmon$_1$ is another scale-invariant analysis of causation: like a basic Aronson-Fair analysis, but unlike an *adequate* Aronson-Fair analysis.

Both locations of causal interactions in Salmon$_1$ seem inappropriate in an analysis of causation which tracks physical interaction. To begin with, according to the Heathcote conjecture physical interactions can only take place at the quantum scale. In that case, how can *simple* causal interactions (interactions which need not be reduced into the sum of myriad quantum interactions) take place at macroscopic scales? Similarly, the account of causal processes in Salmon$_1$ is also scale invariant: we can treat the spreading contact zone as a single processes, or a complex network of quantum-scale processes and interactions. But under the at-at theory of causal propagation, a single process only propagates in the absence of local interactions. Surely the cascade of quantum-scale interactions amount to local interactions. In that case, we seem obliged to claim that the macroscopic entity that the quantum-scale circumstances sum to is not a genuine process,

In addition, locating the causal interactions at couplings between quanta turns the phrase "causal interaction" in Salmon$_1$ into a term of art. Insofar as we understand interaction, it is by reference to physical interaction. By locating causal interactions at couplings we go beyond our best fundamental understanding of physical interactions. At best, when we designate these couplings as "interactions" we are applying a metaphoric extension of our common sense notion of interactions.

A brief story should serve to demonstrate this point. Suppose we are asked to participate in a demonstration of interaction. We are shown a white and a red billiard ball rolling towards one another, striking with a gentle "clack" and rebounding away. We are asked to draw a quick sketch of what we have seen: so we draw something like figure 9.

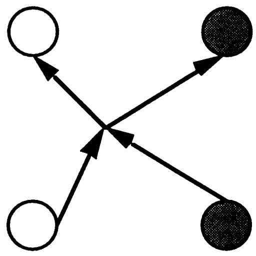


Figure 9 — artist's impression of macroscopic interaction

We are asked to indicate where the interaction took place. We hem and haw a little over whether the points of contact between the balls and the billiard table count as interactions, but then indicate that the interaction between the balls takes place at the intersection of the arrows. This corresponds with the $Salmon_1$ view that interactions take place at intersections or interfaces of processes.

Then we are asked why the interaction took place in the way it did: why the balls interacted in the way they did, rather than passing straight through each other. We might say: because the balls are hard solid objects. Then we are asked to give a more fundamental account of the macroscopic interaction. Knowing a reasonable amount about chemistry and other quantum theory, we say that the material of the red (and of the white) is quite tightly bonded together, and there are few half-filled orbitals on the surface of the balls which would allow the balls would stick together. Hence the predominant factor in the physical interaction is the repulsive electromagnetic interaction between the electrons in the objects. Virtual photons are exchanged, and so on.

And this is the point at which the analysis in terms of fundamental interactions stops, with the statement that virtual quanta are exchanged. The "emission" or "absorption" of virtual quanta at couplings is not a more fundamental physical interaction underlying the exchange. The existence of couplings is a significant factor in fundamental interactions, as we saw in our discussion of genidentity at the quantum scale. But it is only a crucial *part* of physical interaction, not an interaction itself.

The reason why we are tempted to think a coupling is a physical interaction is that it occupies an analogous position in a quantum-scale interaction to the position where we commonsensically locate interactions in cases of macroscopic interaction: spatiotemporal intersections or interfaces of objects (or processes). But this is a metaphoric extension of common sense rather than an application of our best understanding of physical interaction. Thus, by locating causal interactions at spatiotemporal intersections, the phrase "causal interaction" has come adrift from best understanding and fundamental application of the term "interaction".

So how might a Salmon_1-style analyst reply to these charges? We think the best response must be along the lines that these criticisms are only telling against a energy/momenta transference process analysis, where causal dependence genuinely has to reduce to sums of quantum-scale physical dependence and interactions are properly associated with processes rather than intersections of processes. But Salmon_1 is an *information* transfer analysis, to which the criticisms do not apply.

Information is manifest at every physical scale: marks can range in scale from increased energy levels of quantum particles to chalk marks on billiard balls, to billboards, to the configuration of lights marking airport runways, to the configurations of town lighting observed from Earth orbit. Although the macroscopic processes and interactions may be composed of many quantum-scale processes and interactions, frequently the relevant form of mark being transmitted or transferred is manifest at the macroscopic level, so we can limit our attention to a single macroscopic process and interaction rather than the complex network of quantum-scale processes.

Salmon is sympathetic to the view that most processes might strictly speaking not be processes but networks of processes. Yet he still argues for some scale invariance in the manner we have suggested:

> When it comes to practical investigation of actual processes pragmatic considerations determine the level of analysis. For some purposes the motion of a molecule of a gas between collisions with other objects (other molecules, Brownian particles, or walls of containers) may be considered a single causal process; for other purposes the motion of a baseball from a bat to a window (in spite of innumerable collisions with molecules in the atmosphere) may be regarded as a single causal process. I think we gain greater philosophical insight into causality by operating at a rather rarefied theoretical level, recognising, of course, that we must often descend from such abstract heights when it comes to practical investigations.[212]

Similarly, it need not matter that interaction is a term of art when applied to quantum-scale couplings between quanta. So long as there is a plausible definition of the term which does not explicitly cash it out as a species of physical dependence, the usage is permissible. Arguably, Salmon$_1$ supplies such a definition — causal interaction is information exchange between processes. *Prima facie*, this sort of phenomena could take place at a spatiotemporal intersection of any process: no deviant account of physical interaction is being invoked.

We think that even these best rejoinders fail. Positing the scale invariance of causation on the basis of the pragmatics of investigation of the world produces an external problem for Salmon$_1$. The problem is that Salmon$_1$ is intended to provide an objective basis for a theory of explanation. An explanation is objectively justified by appeal to an underlying configuration of causal processes and interactions, which are supposed to be objective, physical features of the world. If these processes and interactions turn out not to be objective features of the world but rather are artifacts of our explanatory requirements, then the justification of explanations are other explanations. The objective basis of justification promised by Salmon$_1$ just isn't delivered.

So the decoupling of relevance of the quantum-scale network of processes and interactions from the macroscopic processes and interactions required by a scale-invariant information transfer process analysis can't be cashed out by a failure of *explanatory* relevance. What seems to be needed is a failure of *physical* relevance between the complex quantum-scale process-interaction cascades and

---

212 "Causality and explanation: a reply to two critiques", p464. Note that these observations are in the context of Salmon$_2$, rather than Salmon$_1$, but the sentiments could readily be transferred to the earlier analysis.

the associated simple macroscopic process-interaction configurations. But this just isn't the case in our universe.

Consider a whole series of bouncing-of-Frank experiments. The strength of the lunar impact is kept constant over the series of experiments, as are the properties of each Frank. We alter each trial bouncing by changing the chemical structure of the Earth. This will affect the rate at which the mark introduced into the Earth by the lunar impact propagates to Frank's jumping-off point. So in some of the trials Frank will have a normal last leap, and in some of the trials Frank will have a big last leap. The propagation of the macroscopic mark (the spreading contact zone) through the Earth is found to depend sensitively on the changes in chemical structure. But these changes largely amount to changes in the number and packing of electrons within the Earth — in other words, the quantum-scale process-interaction cascade *is* physically relevant to the macroscopic processes and interactions. Hence the interactions within the

Consider the $Salmon_1$ interaction between a virtual photon and an electron. At the coupling, information is either being passed from the photon to the electron, or from the electron to the photon, or is being exchanged between the photon an the electron. Yet as we saw in the Davies excerpt in the first section of this chapter, information can only be encoded by real quanta, not virtual quanta. So no such information transfer or exchange could take place. It appears that $Salmon_1$ could only assign causal status to a coupling between real quanta. So if we are appealing to a standard notion of physical information in $Salmon_1$, this renders many physical interactions acausal — for example, virtually all radioactive decay in the known universe, which is mediated by the exchange of virtual W^-, W^+ and Z^0 bosons.

In fact, it is very difficult to avoid the involvement of virtual quanta couplings even in quantum-scale interactions which apparently involve only real quanta. Consider the annihilation of a real electron and a real positron producing real photons. One perfectly good interpretation of this physical interaction is that the real photons are virtual photons from the cloud surrounding the electron and the positron which are "liberated" by the annihilation of the electron and positron.[214] Since virtual photons were involved at some stage of this set of circumstances, a breakdown in information transfer occurs and no causation takes place. Consider Salmon's favourite example of Compton scattering. We can regard the emission of the real photon as being the result of energy from the excited electron "promoting" into real status one of the virtual photons from the cloud around the electron.[215] Once again no standard information transfer could be taking place since there is a virtual photon involved. Conceivably, virtually all couplings could be acausal.

Let's recap and summarise. We are looking for a process analysis of causation in which causal dependence tracks physical dependence. In the light of the Heathcote conjecture, this version of $Salmon_1$ doesn't seem to fit the bill. Firstly, is inappropriately scale-invariant. Secondly, physical interactions are correctly located *in* processes rather than *at* couplings.

The defence is to claim that $Salmon_1$ is an information transference analysis. Physical information is scale invariant, and information exchanges are located

214 *The forces of nature,* p128
215 *The forces of nature,* p128

at couplings, so causal interactions should be located at couplings. But this defence fails. Our best understanding of *physical* information is that it is not scale invariant. "Macroscopic" processes and interactions are not actually processes and interactions in their own right, they are complex networks of quantum-scale processes and interactions. So the level at which information is genuinely being transmitted and transferred is at the quantum scale. But if this is the case, then causal interactions simply can't be located at couplings, because physical information would then be being exchanged between real quanta and virtual quanta, and virtual quanta can not encode physical information.

There are two possible upshots of all this. Firstly, this version of $Salmon_1$ is simply wrong, because it is in conflict with our understanding of physical information. Secondly, the notion of information appealed to in order to defend against the charge that causal interaction is a term of art *is itself* a term of art, since the relevant information certainly could not be *physical* information. So unless either the notion of information transference is abandoned in $Salmon_1$ or some satisfactory definition of the term of art is supplied, $Salmon_1$ will not fulfil our basic desideratum.

Counterfactual $Salmon_1$

The final major stage of development of $Salmon_1$ occurs in the 1984 book *Scientific explanation and the causal structure of the world*. The standard verities are affirmed. The propagative aspect of causation is understood in terms of the at-at theory of transmission of marks by causal processes. The productive aspect of causation is explained in terms of conjunctive forks and interactive forks. Interactive forks are given an ontological characterisation in terms of X-type spatiotemporal intersections of processes involving the introduction of a mark into each of the processes. Causal interactions are interactive forks. Conjunctive forks provide the temporal asymmetry of causation.

Arguably Salmon also goes some way toward either abandoning the information transference basis of $Salmon_1$ or providing a substantive definition of the peculiar sense of information, although the definition is never applied explicitly to the notion of a mark. Marks are frequently referred to as being characteristics or modifications of structure of processes. In response to criticisms by Dowe of vagueness in these notions, in a later paper Salmon notes that these characteristics should be explicated in terms of another notion developed in *Scientific explanation and the causal structure of the world*:

> The key concept is that of an *objectively codefined class* [ODC] … which is explicated in terms of physically possible detectors attached to appropriate kinds of computers that receive carefully specified types of information. It is possible to ascertain, on the basis of local observations — detections — whether an entity possesses a given property at a particular time. Since, in scientific contexts, we often detect one property by observing another, it must be possible in principle to construct a computer to make the determination. For example, when we measure temperature by using a thermocouple, we actually read a potentiometer to detect an electromotive force (emf). The computer to which the explication refers must be able to translate the potentiometer reading into a temperature determination on the basis of laws concerning the electrical outputs of thermocouples, but without receiving information from other physical detectors. Notice that this explication is physical, not epistemic. This kind of definition would easily suffice to rule out properties such as … *being a shadow that is closer to the Harbour Bridge than to the Sydney Opera House* [Dowe], as well as properties such as grue.[216]

It is not clear that the definition of an ODC mandates scale invariance. (We shall not debate the point.) At first glance, the definition promises an account of how virtual quanta might transmit or exchange information. Virtual quanta could not carry physical information in the standard sense, because information cannot be encoded within a single wavelength of an entity, and virtual quanta propagate at most a single wavelength. Yet the definition of ODCs resolves "information" to possession of a particular property. It seems entirely plausible that virtual quanta possess properties.

Unfortunately, the definition of ODCs also requires some property of the virtual quanta to be observable. But virtual quanta are strictly unobservable. So virtual quanta don't possess ODC. Hence, information cannot be being exchanged across couplings involving virtual quanta. $Salmon_1$ still can't fulfil our fundamental desideratum, because the account of causal interactions is still appealing to an inadequate account of information transfer.

---

216 "Causality without counterfactuals", p251

Even if we could provide an adequate account of information transfer, there are a number of extremely difficult problems for the analysis. The first involves the requirement that causal interactions are located at X-type intersections of processes which persist through the interaction. Salmon reiterates his suggestion that all fundamental physical interactions might be understood in terms of the interactive fork. But most physical interactions at the quantum level do not involve X-type intersections of processes. For example, a hydrogen atom absorbing a photon exemplifies a λ -type intersection of processes. There are two "incoming" processes, one of which is annihilated at the coupling, and one "outgoing" process, which is a modified form of one of the "incoming" processes.[217] A hydrogen atom emitting a photon exemplifies a y-type intersection of processes. There is one "incoming" process and two "outgoing" processes, one of which is created at the coupling, and one of which is a modified form of the incoming process.[218]

Compton scattering, Salmon's favourite example of an interactive fork, actually involves such a λ -type intersection (absorption of the incident photon by the electron) linked by a process (the excited electron) to a y-type intersection (emission of a photon by the excited electron).[219] Salmon acknowledges the reductive programme is therefore contingent on an explication of the interactive fork in terms of these "simpler" λ -type and y-type intersections. Salmon is forced to leave this explication as an exercise for the reader.[220]

In fact, the requirements for the programme are even more complex. Consider the mutual annihilation of an electron and a positron. The energy released by the collision escapes in the form of photons. The energy is not carried away by a single photon. If this was the case, such an interaction arguably could be described in terms of the putative λ -type interactive fork. But this would seem to involve the claim that the photon is a modified electron process, which would be unwarranted special pleading.

---

217 As we noted earlier, since causal interactions are time-symmetric, such talk of "incoming and "outgoing" processes is merely a convenient figure of speech. We shall use scare quotes to indicate this.

218 *Scientific explanation and the causal structure of the world*, p181

219 As we noted earlier, the situation may be even more complex than this, involving intermediate stages such as the promotion of a virtual photon.

220 *Scientific explanation and the causal structure of the world*, pp181–2

One possibility is that two photons are produced. This corresponds in some ways to an X-type interactive fork, but it would again be unwarranted special pleading to argue that the two "outgoing" photons are modified versions of the two "incoming" electron processes.[221] The other possibility is that *three* photons are produced, producing a spatiotemporal intersection of *five* distinct processes: an "incoming" electron, an "incoming" positron, and three "outgoing" photons.[222] Let's call such interactions involving five (or more) processes *-type interactions.

Due to the existence of the *-type and the anomalous λ-, y- and X-type intersections, it seems that the "final theory" of the interactive fork needed for Salmon's programme has to be capable of dealing with interactions involving arbitrary numbers of "incoming" and "outgoing" intersecting processes and involving arbitrary numbers of creations, annihilations and modifications of processes. Indeed, a scale-invariant version of $Salmon_1$ also needs such a theory for causal interactions at the macroscopic scale. Consider a rock being hurled through the window. Here we have an intersection of two "incoming" processes (the rock and the window) and many, *many* outgoing processes (the rock and dozens of window shards), most of which were created by the interaction. This is close to a paradigm case of the *-type interaction.

The major innovation — and one of the most serious failings — introduced in *Scientific explanation and the causal structure of the world* is that the $Salmon_1$ accounts of mark transmission and causal interaction are given explicitly counterfactual formulations.

The modification to the at-at theory is motivated by a counterexample due to Nancy Cartwright. Consider the case of the rotating beacon. Suppose the moving spot of light on the wall is marked by a local interaction: being made red by a red filter being held up next to the wall. Suppose that a few nanoseconds later a red filter is placed over the lens of the beacon. So the spot is marked red by a single local interaction (the wall filter), and remains red

---

221 In fact, at the quantum scale *every* interaction involving an X-type intersection of processes will be anomalous. The claim that the "outgoing" processes are modified versions of "incoming" processes amounts to the claim that the outgoing processes are genidentical to the incoming processes. Yet as we have seen in our earlier discussion of the transference analysis and $Salmon_1$ there is no such genidentity at the quantum scale in our universe.

222 *The forces of nature*, pp108–10

without any further local interactions (because of the filter at the beacon). Cartwright concludes that under the previously developed account of transmission of marks (being at the intervening positions at the intervening times in the absence of additional interactions), the spot is transmitting a mark, and thus is not a pseudo process at all.

Salmon attempts to rule out such cases by introducing an explicitly counterfactual condition into the at-at theory of mark transmission, formally stated as principle MT:

> MT: Let P be a process that, in the absence of interactions with other processes, would remain uniform with respect to a characteristic Q, which it would manifest consistently over an interval that includes both of the space-time points A and B (A≠B). Then, a *mark* (consisting of a modification of Q into Q′), which has been introduced into process P by means of a single local interaction at point A, is *transmitted* to point B if [and only if] P manifests the modification Q′ at B and at all stages of the process between A and B without additional interventions.[223]

This modification to at-at is an extremely dangerous move on Salmon's part. Consider how MT operates in specific cases. In the case of the red filtered beam of light, marks are being transmitted by the beam because but for the interposition of the red filter, the beam would have remained white. We have a straightforward case of counterfactual dependence. In the case of the Cartwright example, marks are not being transmitted by the spot of light because but for the interposition of the red filter at the wall, the spot would *still* have turned red due to the filter at the beacon. We have a paradigmatic case of failure of counterfactual dependence due to late preemption. A standard Lewis-style counterfactual analysis treats late preemption as a problem case: a failure of the counterfactual analysis to establish causal dependence in cases where causal dependence ought to be occurring. MT actually *appeals* to late preemption to demarcate between cases of causal dependence and lack of causal dependence.

223 *Scientific explanation and the causal structure of the world*, p148. The "[and only if]" clause is a modification suggested by Elliott Sober. See "Wesley Salmon's process theory of causality and the conserved quantity theory", pp196–7.

In this mature version of $Salmon_1$ causal dependence is supposed to track physical dependence. Causal mechanisms are objective features of the universe: they are particular physical structures in the universe. The distinction between a causal process and a pseudo process should be that the causal processes are different kinds of physical structures from the pseudo processes. So if overdetermination problems such as late preemption serve to distinguish between causal processes and pseudo processes, then it must be the case that there is some relevant physical distinction between the physical circumstances in which counterfactual dependence occurs and the physical circumstances in which overdetermination problems arise. Yet the strange case of the overdetermining dumbbell demonstrates that overdetermination problems arise in sets of physical circumstances where there are no relevant physical differences from sets of physical circumstances in which straightforward counterfactual dependence occurs. So it can not be the case that $Salmon_1$ with MT tracks physical dependence.

If Cartwright's examples necessitate a counterfactual amendment to at-at, then at-at should be abandoned. But Cartwright's example actually does not necessitate the move to MT. The weak link in Cartwright's example is the assertion that the spot remains red without further causal interactions. Of course there are further interactions. The spot is produced by photons from the beacon scattering off the material of the wall. Each of these scatterings is a causal interaction under $Salmon_1$. In a case where these interactions cease — for example, if a particular piece of the wall is perfectly transparent to the light frequencies from the filtered beacon — not only would the spot not remain red, there would be no spot. If a portion of the wall is made of a material which perfectly absorbs visible frequencies of light and preferentially emits infra-red radiation when stimulated by absorption of visible light, then the spot persists in that area of the wall — as a "hot spot" — but it is not red.

We can generalise from this that all pseudo processes which, like the spot, consist of a continuous series of interactions are immune to Cartwright's criticism (because the required absence of local causal interactions never occurs). If a Cartwright example could be formulated using a pseudo process which is *not* a continuous string of interactions, then an amendment to at-at is required. In the absence of such an example Salmon may retain the original formulation of at-at.

Salmon also provides a counterfactual explication of causal interactions in *Scientific explanation and the causal structure of the world*, via the principle CI:

> CI: Let P_1 and P_2 be two processes that intersect with one another at the space-time point S, which belongs to the histories of both. Let Q be a characteristic that process P_1 would exhibit throughout an interval (which includes subintervals on both sides of S in the history of P_1) if the intersection with P_2 did not occur; let R be a characteristic that process P_2 would exhibit throughout an interval (which includes subintervals on both sides of S in the history of P_2) if the intersection with P_1 did not occur. Then, the intersection of P_1 and P_2 at S constitutes a causal interaction if:
>
> (1) P_1 exhibits the characteristic Q before S, but it exhibits a modified characteristic Q′ throughout an interval immediately following S; and
>
> (2) P_2 exhibits the characteristic R before S, but it exhibits a modified characteristic R′ throughout an interval immediately following S.[224]

The reference simply to processes rather than causal processes eliminates a worrying circularity from the earlier statistical and ontological characterisations of interactions. Causal processes were defined partly in terms of causal interactions, because the introduction of a mark is by means of a causal interaction. Yet causal interactions were defined in terms of spatiotemporal intersections of *causal processes*.

The reference in CI to *any* processes rather than just causal processes does not allow intersections of pseudo processes or intersections of causal processes and pseudo processes to qualify as interactions. Consider two moving spots of light projected on a white screen — paradigmatic examples of pseudo processes. One spot is green; one is red. The green moves from the top left corner to the bottom right, the red from the top right to the bottom left. They intersect in the middle, momentarily creating a yellow spot. Since the modification at the intersection does not persist beyond the intersection, this is not a causal interaction. Now suppose that the projectors are set up so that when the red and the green intersect the red changes direction and moves to the top left, and the green changes direction and moves to the bottom left. This *looks* like a causal interaction, but does not count due to the counterfactual condition. If the projector set-up was identical but for there being no red

224 *Scientific explanation and the causal structure of the world*, p171

projector, the green would still change direction. *Mutatis mutandis*, the red would still change direction. (Note that unlike the case of MT this is merely a straightforward case of absence of counterfactual dependence. CI does not appeal to overdetermination problems.) Salmon describes some more complex counterexamples of "collisions" of pseudo processes due to Patrick Maher and Richard Otte, but also argues that the counterfactual condition of CI rules that the cases are not causal interactions.[225]

Despite the apparent improvement in terms of elimination of the circularities of the earlier statistical and ontological characterisations of causal interactions, CI still only applies to the standard X-type interaction, and cannot account for interactions at anomalous X-type intersections, λ -type or y-type interactions, or *-type interactions.

A little further tidying up occurs in Salmon_1. The propensity interpretation of transmission of causal influence is supplied in the 1988 paper "Dynamic rationality: propensity, probability, and credence" and discussed in the 1991 paper "Causal propensities: statistical causality versus aleatory causality". We can now summarise the salient points of the mature Salmon_1:

1. Causal phenomena are explained in terms of two basic phenomena: propagation and production of (probabilistic) causal influence.

2. The causal mechanism underlying propagation of causal influence is the causal process. A *process* is a thing with some consistency of structure (such as a material object or a moving spot of light). A *causal* process is a process capable of transmitting a mark (a modification of structure). Mark transmission is defined by a counterfactual at-at theory (but perhaps need not be so defined).

3. The causal mechanisms underlying production of causal influence are conjunctive forks and interactive forks.

   (a) Conjunctive forks are configurations of causal processes which produce correlations between events obeying a specific set of statistical conditions. Conjunctive forks determine temporal asymmetry.

225 *Scientific explanation and the causal structure of the world*, pp171–4

(b) Interactive forks are X-type spatiotemporal intersections of processes where a modification of structure (a mark) is introduced into each process at the point of intersection, the modifications persist beyond the intersection, and the modification would not have occurred had the intersection not occurred. Interactive forks are causal interactions. Interactive forks are temporally symmetric.

4. Since characteristics and modifications of characteristics can occur at any scale, like the early Aronson transference analysis, $Salmon_1$ is a scale-independent analysis of causation.

Kitcher and Dowe against $Salmon_1$

We have noted that the mature $Salmon_1$ converges with the basic Aronson-Fair analysis (but not with an adequate Aronson-Fair analysis). Arguably, it converges even more dramatically with the counterfactual analysis.

Suppose we could demonstrate that all pseudo processes are continual strings of interactions. Then all pseudo processes would be immune to Cartwright's example, and Salmon need not provide an explicitly counterfactual version of the at-at theory. But even if we could do this, the account of causal interactions (CI) is still explicitly counterfactual. Since marks are produced by causal interactions, and markability demarcates causal processes from pseudo processes, the account of processes in $Salmon_1$ is still thoroughly counterfactual.

In fact, $Salmon_1$ is open to the charge that no fundamental work is being done in the analysis by the causal mechanisms postulated by the process ontology: causal dependence in $Salmon_1$ does not track physical dependence in the form of a network of interconnecting causal processes and interactions, but rather tracks relations of counterfactual dependence. This charge is laid by Phillip Kitcher in the 1989 article "Explanatory unification and the causal structure of the world":

> What is critical to the causal claims seems to be the truth of the counterfactuals, not the existence of the processes and the interactions. If this is correct then it is not just that Salmon's account of the causal structure of the world needs supplementing through the introduction of more counterfactuals. The counterfactuals are at the heart

of the theory, while the claims about the existence of processes and interactions are, in principle, dispensable. Perhaps these notions may prove useful in protecting a basically counterfactual theory of causation against certain familiar forms of difficulty (problems of preemption, overdetermination, epiphenomena, and so forth*). But, instead of viewing Salmon's account as based on his explications of process and interaction, it might be more revealing to see him as developing a particular kind of counterfactual theory of causation, one that has some extra machinery for avoiding the usual difficulties that beset such proposals. [*Kitcher's note: See Lewis (1973) ["Causation"], both for an elegant statement of a counterfactual theory of causation and for a survey of difficult cases.][226]

Indeed, the counterfactual formulation of the mature $Salmon_1$ is very similar to a Lewis-style counterfactual analysis. Consider the case of the moving spots of light on the screen. We determine by CI that no causal interaction has taken place by comparison of the actual case with either a single possible experimental set-up identical but for the presence of the red projector or a single possible experimental set-up identical but for the presence of the green. This is almost identical to Lewis's formulation in terms of comparison with the closest single possible world, or at least one of the joint closest possible worlds.

But if causal dependence is tracking counterfactual dependence in Salmon1, could causal dependence genuinely be said to be an objective physical feature of the world which could provide an objective physical basis to a theory of scientific explanation? Salmon is uncomfortably aware that the answer to this question is no — he has already expressed his reservations about the use of counterfactuals in *Scientific explanation and the causal structure of the world* due to the irreducibly pragmatic nature of counterfactuals:

> Consider the famous example about Verdi and Bizet. One person might say, "If Verdi had been a compatriot of Bizet, then Verdi would have been French," whereas another might maintain, "If Bizet had been a compatriot of Verdi, then Bizet would have been Italian." These two statements seem to be incompatible with one another. Their antecedents are logically equivalent; if however, we accept both conditionals, we wind up with the conclusion that Verdi would be French, that Bizet would be Italian, and they would still not be compatriots. Yet both statements can be true. The first person

[226] This summary by Kitcher of Kitcher's position is excerpted by Salmon in "Causality without counterfactuals', p253

> could be making an unstated presupposition that the nationality of Bizet is fixed in this context, while the second presupposes that the nationality of Verdi is fixed. What remains fixed and what is subject to change — which are established by pragmatic features of the context in which the counterfactual is uttered — determine whether a counterfactual is true or false. It is concluded that counterfactual conditional statements do not express objective facts of nature ...[227]

Salmon suggests a way out of this problem: scientists, or rather experimenters using scientific methods, determine by means of well-designed controlled experiments "which conditions are to be fixed for purposes of the experiment and which allowed to vary".[228] Such experimental measures are supposed to give an objective basis to the relevant counterfactual conditions. (We could view this way out as providing a distinction between Salmon$_1$ and a Lewis-style counterfactual analysis. The choice of relevant possible world is not being made on the basis of comparative similarity, but on the basis of the decisions of experimenters. On the other hand, Salmon may just be providing a criterion for determining which is the closest possible world.[229]) Yet, as Dowe points out, Salmon desperately needs to elaborate on why an otherwise pragmatic decision about fixing conditions becomes an objective decision when made by a scientist or experimenter.[230]

Salmon finds Kitcher's claim that Salmon$_1$ resolves to a Lewis-style counterfactual analysis "disconcerting", although he argues that the concentration on physical connections rather than the truth of counterfactual conditionals marks a fundamental distinction between Salmon$_1$ and a counterfactual analysis.[231] But this criticism and Kitcher and Dowe's reiteration of the problems of providing an account of "objective counterfactuals" (or objective grounds for determining which is the closest possible world) in combination with Salmon's own misgivings over counterfactuals is the deciding factor in favour of Salmon's rejection of Salmon$_1$ in favour of the Salmon$_2$ conserved quantity analysis of causation.[232]

---

227 *Scientific explanation and the causal structure of the world*, p149
228 *Scientific explanation and the causal structure of the world*, pp149–50
229 Kitcher argues for the latter interpretation in "Explanatory unification and the causal structure of the world", pp474–5
230 "Wesley Salmon's process theory of causality and the conserved quantity theory", p208
231 "A new look at causality", p18
232 "A new look at causality", pp18–20. See also "Causality without counterfactuals", p253, p259 and p260.

Smith against $Salmon_1$

We think the most telling criticism against the mature version of $Salmon_1$ is still the failure to give an adequate account of the sense of information being used. The definition of an ODC does not clearly mandate the scale-invariance of $Salmon_1$. Consequently, only quantum scale entities are genuine causal processes and only quantum-scale couplings are genuine causal interactions. At this scale, very few couplings involve X-type intersections of processes. Due to the breakdown of genidentity at this scale, even the X-type intersections do not qualify as causal interactions. An interaction involves the persistent modification of processes beyond the point of intersection. The notion of modification of a process requires the process before the modification to be one and the same process as the process after the modification. But there is no such fact of the matter about the genidentity of entities at this scale. Even if these problems could be solved, the location of causal interactions at couplings requires that information is being exchanged with virtual quanta. But the ODC interpretation of information resolves information to observable properties of processes. Since the properties of virtual quanta are unobservable, virtual quanta cannot carry such information. In effect, the only physical scale at which a causal interaction could genuinely be said to occur is a physical scale at which these causal interactions *could not* occur.

But suppose scale invariance was warranted by the ODC interpretation. In that case, the counterfactual qualities of $Salmon_1$ should not just cause concern over the objective nature of causation, as argued by Kitcher and Dowe. The counterfactual involvement mandates the rejection of the process/interaction ontology.

The debate over whether $Salmon_1$ should properly be referred to as a physicalist analysis with strong counterfactual elements (as per Salmon) or as a fundamentally counterfactual analysis with a few extra physical bells and whistles (as per Kitcher) probably resolves to a matter of taste. The crucial issue is that the process ontology does not help protect $Salmon_1$ from the "usual difficulties" of overdetermination, preemption and so forth. The process ontology produces many more such familiar difficulties. Of course, Salmon does not wish to abandon the analysis of causal dependence in terms of physical dependence. But suppose he did. Would it not be possible to accept

Kitcher's secondary observation — that $Salmon_1$ is a counterfactual analysis with special resources to deal with the problems of overdetermination, preemption and so forth?

Consider how the CI treats the case of the bonking of Frank. The relevant processes would be parts of Frank and parts of the Moon. In fact, they would be the expanding contact zones in Frank and the Moon — no other portions of the Moon or Frank are experiencing anything like a persistent modification over the course of the bonking on either side of the initial impact. We would be compelled to conclude that Frank's crushing did not depend on the entire Moon, just the contact zone of the Moon.

But if the 3.3 metre radius contact zone of the Moon had *not* been present when the Moon struck Frank, it is not the case that certain characteristics of the contact zone of Frank — such as rigidity, fleshiness, having rigid skeletal structure and so forth — *would not* have been altered. Frank would still have been crushed by part of the rest of the Moon. But the counterfactual criterion in CI requires that these characteristics of Frank would have been altered had that intersection or interface not taken place. Given that the counterfactual criterion is not satisfied it simply is not the case that the crushing of Frank *does not involve a causal interaction.*

We have an exact parallel between $Salmon_1$ and a Lewis-style counterfactual analysis in which assignment of object status is keyed to physical involvement. Cases of single-object overdetermination arise. Both analyses are more prone to Kitcher's "familiar forms of difficulty" than a standard Lewis-style analysis. By paying close attention to the physical connections involved in physical interactions, $Salmon_1$ is even less capable of assigning causal dependence in instances of obvious physical dependence than a standard Lewis-style analysis. It seems that the only way to turn $Salmon_1$ into a satisfactory counterfactual analysis is to reject the use of processes and interactions. But such a move amounts to the utter rejection of $Salmon_1$ in favour of a counterfactual analysis.

Processes — CQ and Salmon$_2$

Dreams of a final process analysis

Given the failure of the mature Salmon$_1$ to provide an analysis of causation in which causal dependence, what line of enquiry seems most likely to produce an analysis of causation which does fulfil our basic desideratum?

The fundamental problem with Salmon$_1$ is the failure to supply an adequate account of the information that is being transmitted by processes and transferred or exchanged at interactions. It is tempting to suppose that the correct avenue of approach is to develop a workable and plausible account of information and slot this in to Salmon$_1$.

But this would not be satisfactory. Even if we were to rehabilitate the notion of information the mature Salmon$_1$ is compromised by the counterfactual formulation of CI and MT. Contra Kitcher, it is not the case that Salmon$_1$ should be treated as a Lewis-style counterfactual analysis with extra tools (the process/interaction ontology) which allows it to deal more satisfactorily with overdetermination problems. The process/interaction ontology leaves Salmon$_1$ subject to more such problems than a standard Lewis-style counterfactual analysis — the counterfactual version of Salmon$_1$ is an even-less-than-usually-adequate counterfactual analysis of causation. Retaining the counterfactual formulation mandates the abandonment of the process/interaction ontology.

We have argued that the counterfactual amendment to the at-at theory of mark-transmission is not necessitated by the Cartwright example. So we could readily extricate the account of processes from counterfactual entanglement by retreating to the pre-MT version of the at-at theory. But we cannot similarly retreat to an earlier account of interactions in Salmon$_1$.

The earlier ontological characterisation of interactive forks is unacceptably circular. Causal processes are distinguished from pseudo processes by the mark criterion. Yet marks are introduced by interactions, and, under the ontological characterisation, interactive forks (which constitute causal interactions) are partially defined in terms of X-type intersections of causal processes.

Nor can we retreat to a statistical characterisation of causal interactions. Salmon's statistical characterisations of both the conjunctive and interactive forks also involve explicit mention of connecting *causal* processes, which is unacceptably circular. Another problem is that interactions are properly to be associated with the interactive fork, rather than the conjunctive fork. But the perfect fork can be a limiting case of either the interactive fork or the conjunctive fork. So causal interactions cannot be given a purely statistical characterisation.

So we cannot rehabilitate $Salmon_1$ merely by developing a new notion of information. Although a plausible $Salmon_1$ account of processes may be extricated from counterfactual entanglement, there is no satisfactory $Salmon_1$ non-counterfactual account of interactions. Some new ideas about causal interaction are required in order to produce a satisfactory non-counterfactual process analysis of causation.

Perhaps an old idea might suffice. We noted at the beginning of the preceding chapter that — despite other failings — the transference analysis seemed to provide an account of causal interactions that tracked physical interactions. We have seen that the mature $Salmon_1$ resembles a transference analysis. Salmon has already accepted that most causal interactions he wishes to associate with interactive forks — perhaps *all* such causal interactions — are associated with energy and momentum transfers.[233] Salmon speaks highly of Fair's transference analysis in this context. Perhaps the key to the conundrum is to identify the causal interactions with something like energy/momentum transfers or exchanges, rather than the marking of processes.[234]

Why "*rather* than the marking of processes"? Why not simply cash out "markings" in terms of energy/momentum transfers? Simply because "marking" implies some form of information transfer, and by cashing out "marking" thus we would be relying on yet another non-standard interpretation of information. Worse still, although energy/momenta transfers or exchanges presumably have something to do with transfer or transmission of

233 "Causality: production and propagation", pp294–5

234 See Salmon's discussion of this point in "Causality without counterfactuals", pp252–3

information, it is difficult to believe that one resolves to the other. Such an intuition underlies the quote from Davies at the beginning of the previous chapter asserting that causality is not violated by faster-than-light exchange of virtual quanta. Although energy/momenta exchanges may briefly exceed the speed of light, no information travels faster than the speed of light.

To put the point another way, the rationale for investigating the $Salmon_1$ information transfer analysis of causation was that it appeared to be a plausible *alternative* account of how causal dependence might track physical dependence, which might not be subject to some of the failings of the energy/momenta transference analysis. If information transfer is *not* significantly different from energy/momenta transfer, and our intention is to look for a way of overcoming the failures of the transference analysis, then we should attempt to deal directly with those failures rather than proceeding under the cover of darkness provided by a term-of-art interpretation of "information".

If we are rejecting the notion of information transfer or exchange at causal interactions, we should reject the notion that causal processes transmit information. As Salmon points out, the notion of marking is logically prior to the notion of a mark. If we deny that causal interactions fundamentally involve markings, we should deny that processes fundamentally involve transmission of marks.

Yet retaining *some* explicit notion of a causal process *prima facie* seems advisable. We concluded from our examination of the case of the bouncing of Frank that a basic Aronson-Fair analysis which did not assert the genidentity of energy/momenta was either inadequately casually connective or excessively causally connective. To produce an adequately causally connective analysis, we needed to key causal connectivity to the propagation of the lunar impact contact zone through the Earth. As we noted during our discussion of $Salmon_1$, an account of connection by causal process is tailor made for capturing adequate causal connectivity. Indeed, assuming a solution to the problem of dissipation could have been found the genidentity of quantities of energy/momenta would have supplied such an account of causal processes. Unfortunately, genidentity fails at the appropriate physical scale in this universe. Similarly, the challenge for the process analyses that supersede

$Salmon_1$ will be to identify a criterion for causal processes which is an actual feature of this world.

So the project that shows most promise of fulfilling our basic desideratum is a new process analysis of causation in which interactions are defined in terms of energy/momentum transfers or exchanges of the transference analysis (or some related mechanism). Two such analyses have been proposed: Dowe's conserved quantity analysis (which we will refer to as CQ) and Salmon's version of the conserved quantity analysis (which we have been referring to as $Salmon_2$).

Unfortunately, neither analysis can be regarded as satisfactory. Both analyses repeat the $Salmon_1$ error of locating causal interactions at couplings rather than in types of processes. The proposed role for conserved quantities in causal interactions does not seem to match the role that conserved quantities play in physical interaction. Indeed, physical interactions seem to be more closely associated with brief violations of certain conservation principles.

But most damningly, under the criteria proposed for processes there just don't seem to be any processes in this universe. Neither analysis can justify scale invariance, so properly speaking there are no processes except at the quantum physical scale. Dowe relies on a genidentity criterion for processes, but genidentity fails at this physical scale. $Salmon_2$ provides an at-at theory of processes. But at this physical scale few things appear to have enough spatiotemporal extension between interactions in order to warrant being referred to as processes. In the final analysis, either causation under $Salmon_2$ tracks explanatory dependence rather than physical dependence or $Salmon_2$ is merely a brief restatement of our best accounts of physical dependence.

Candidates for a final process analysis

In his 1992 paper "Wesley Salmon's process theory of causation and the conserved quantity theory", Phil Dowe tackles head-on the deficiencies of $Salmon_1$. Contra Kitcher, Dowe does not advocate that $Salmon_1$ should be construed as a counterfactual analysis of causation (which as we have argued mandates the abandonment of the process/interaction ontology). Rather, Dowe argues that $Salmon_1$ should be supplanted by a new, non-counterfactual process analysis.

Dowe begins the paper with an exposition of $Salmon_1$, and then proceeds to elucidate various deficiencies of the analysis. Although he does not present a historical analysis of the development of $Salmon_1$, as per our presentation in the previous chapter, Dowe does not restrict his criticism to the counterfactual form of the analysis. (For example, the statistical characterisations of causal interactions are examined and found wanting.)

We shall not give an detailed exposition of Dowe's arguments against $Salmon_1$. But the salient points of Dowe's attack include that the notions of process and interaction are interdefined in a dangerously circular manner (paralleling our concerns about the ontological and statistical characterisations of interactions). The counterfactual formulation of the principles MT and CI is out of place in an analysis which seeks to tie causation to objective matters of fact. The origin of processes cannot be analysed in terms of statistical relations by means of the conjunctive fork, and causal interactions cannot be analysed in terms of the statistical characterisation of the interactive fork (once again, paralleling our analysis). $Salmon_1$'s account of interactions is also limited to X-type intersections involving modification of two and only two causal processes, yet many interactions seem to be of the y-type and the λ -type.[235]

Dowe is also concerned by the vagueness of the critical terms "characteristic" and "structure", although his concern is not parallel to our concern over the justifiability of scale invariance. Indeed, CQ, Dowe's own process analysis is also intended to be scale invariant. Dowe's concern is that without some more precision in the definition of structure, an alteration in relational properties (a Cambridge change) might qualify as a change in characteristics of a process, and hence a causal interaction, when we would intuitively wish to deny their causal status.[236]

Dowe seeks to avoid these problems by presenting a new process analysis of causation, which we shall refer to as CQ (for "conserved quantities"). CQ is, as Salmon notes, beautiful for its simplicity:

> Definition 1. A causal interaction is an intersection of world lines which involves exchange of a conserved quantity.

---

235 "Wesley Salmon's process theory of causality and the conserved quantity theory", pp200–210

236 "Wesley Salmon's process theory of causality and the conserved quantity theory", p201

Definition 2. A causal process is a world line of an object which [possesses] a conserved quantity.[237]

A *world line* is the collection of points on a spacetime (Minkowski) diagram which represents the history of an object. A *conserved quantity* is any quantity universally conserved according to current scientific according to current scientific theories. Some conserved quantities are mass-energy, linear momentum, angular momentum, and charge. An *exchange* means that at least one incoming and at least one outgoing process manifest a change in the value of the conserved quantity. "Outgoing" and "incoming" are delineated on the spacetime diagram by the forward and backward light cones, but are essentially interchangeable. The exchange is governed by the conservation law. The intersection can therefore be of the form X, Y, λ or of a more complicated form. An *object* can be anything found in the ontology of science (such as particles, waves or fields), or commonsense.[238]

Several examples show how the definitions apply. The first is the bombardment of nitrogen with an alpha particle (a helium nucleus), producing an oxygen isotope and a proton (a hydrogen nucleus), which we can symbolise by the equation:

$${}^{4}_{2}He + {}^{14}_{7}N \rightarrow {}^{17}_{8}O + {}^{1}_{1}H$$

This is an example of an anomalous X-type interaction that is difficult to account for in $Salmon_1$. To describe this phenomenon in terms of mutual modifications of two processes beyond a point of intersection, we would be required to claim, for example, that the "outgoing" proton is a modified version of a nitrogen atom (or a helium nucleus). This seems unwarranted special pleading. But under CQ, by definition 2 each atom or nucleus is a causal process because they possess the conserved quantity *charge*, signified in the equation by the subscripts, which represent the numbers of protons in each nucleus. The total incoming charge is equal to the total outgoing charge, and at least one of the outgoing processes has a different charge from an incoming process, so the transmutation involves a causal interaction by definition 1.

237 "Wesley Salmon's process theory of causality and the conserved quantity theory", p210. The bracketed term "[possesses]", used in the later paper "Causality and conserved quantities: a reply to Salmon" replaces the original term "manifests". Dowe does not wish to give the impression that *observing* whether or not a process has a quantity underlies causation.

238 "Wesley Salmon's process theory of causality and the conserved quantity theory", p210

Y-type intersections are also causal under CQ. Consider the transmutation of radium into radon, symbolised by the equation:

$$^{226}_{88}Ra \rightarrow ^{222}_{86}Rn + ^{4}_{2}He$$

Both the "incoming" radium processes and the outgoing radon and alpha particle processes qualify as causal processes by virtue of possessing charge. The intersection involves an exchange of charge, so a causal interaction has taken place.[239]

Dowe claims that CQ is an improvement over an Aronson-Fair analysis by virtue of providing a workable account of processes and being more generally applicable:

> Using "conserved quantity" rather than "energy" is more general and more practical. Many common instances, such as collisions between cars or billiard balls, are more amenable to an analysis in terms of momentum rather than energy. Similarly, electrical causes can be more practically identified according to charge than by energy. Earlier we noted the difficulty of basing cause on energy on the grounds that it is not possible to identify a quantity of energy as being the same as an earlier quantity. This is avoided here because there is no notion of transference or transmission in the definitions.[240]

Note that CQ differs from $Salmon_1$ not only by giving a new account of causal interactions, but also abandoning the at-at account of processes. The new account is a genidentity criterion. We find in Dowe's "Causality and conserved quantities: a reply to Salmon":

> The theory could have been formulated in terms of *objects*: there are causal objects and pseudo objects — causal objects are those which possess conserved quantities, pseudo objects do not. Then a causal process is the world line of a causal object.[241]

And even more explicitly:

> ... the CQ theory identifies genuine causal objects according to the possession of certain properties at a time, and identifies genuine processes over time via the

239 "Wesley Salmon's process theory of causality and the conserved quantity theory", pp211–2
240 "Wesley Salmon's process theory of causality and the conserved quantity theory", p214
241 "Causality and conserved quantities: a reply to Salmon", p324

additional presumption of a relation of identity over time. Thus: an electron existing at a point in time is causal because it possesses charge, and its world line represents a causal process, assuming that the world line represents one and the same object over that time interval.[242]

Salmon considers Dowe's elimination of counterfactuals from the process analysis and the ability to designate a greater variety of spatiotemporal intersections as causal interactions to be a great boon for the process analysis project. In the 1994 paper "Causality without counterfactuals" he consequently renounces $Salmon_1$ in favour of his own version of CQ — $Salmon_2$.

There are several differences between the early version of $Salmon_2$ and CQ. Initially, Salmon believes that the process analysis should be cashed out in terms of *invariant* quantities rather than *conserved* quantities. (The value of a conserved quantity within a given physical system does not change over time; the value of an invariant quantity remains constant with respect to change of frame of reference.) In response to criticisms by Dowe and Christopher Hitchcock,[243] Salmon abandons this formulation in favour of conserved quantities in his most recent formulation of $Salmon_2$, presented in the 1998 paper "Causality and explanation: a reply to two critiques".[244]

But an issue that Salmon stands firm on is the rejection of Dowe's genidentity criterion for casual processes. Hence, in the latest version of $Salmon_2$ Dowe's definition 2 is replaced by two other definitions detailing a variant of the at-at theory of causal propagation. The mature version of $Salmon_2$ is:

(1) A causal interaction is an intersection of world-lines that involves exchange of a conserved quantity.

(2) A causal process is the world-line of an object that transmits a non-zero amount of a conserved quantity at each moment of its history (each spacetime point of its trajectory).

---

242 "Causality and conserved quantities: a reply to Salmon", p330
243 pp314–5, Hitchcock, Christopher: "Discussion: Salmon on explanatory relevance", pp304–20, *Philosophy of science* vol 62 (1995)
244 "Causality and explanation: a reply to two critiques", p472

(3) A process transmits a conserved quantity between A and B (A≠B) if and only if it possesses a fixed amount of this quantity at A and B and at every stage of the process between A and B without any interactions in the open interval (A,B) that involve an exchange of that particular conserved quantity.[245]

A notable distinction between CQ and $Salmon_2$ is thus that under CQ a single causal process may persist through a causal interaction, so long as an "incoming" process and an "outgoing" process are genuinely one and the same process, but under $Salmon_2$, a causal interaction signals the birth and death of every interacting process. By definition 3, the presence of an interaction dictates the absence of transmission at a particular spacetime point. But by definition 2, a process must transmit a conserved quantity at every spacetime point of its trajectory. So processes only extend from interaction to interaction, not through interactions.

Failings of the final process analyses

Do either CQ or $Salmon_2$ fulfil our basic desideratum? We think not.

Our first point, contra CQ but less so against $Salmon_2$, is that a conserved quantity analysis ought not be scale invariant as Dowe suggests. The argument is along similar lines to our critique of the scale invariance of $Salmon_1$.

How does an everyday, common-sense object come to possess *charge*? Charge is not (as far as we understand in our current scientific theories, which Dowe claims we are supposed to derive the notion from) a primitive or emergent property of common-sense objects. For example, we find that common-sense objects possess charge by virtue of being partly composed of charged particles such as ions. The ions possess a charge by virtue of being composed of unequal numbers of singly negatively charged particles (electrons) and singly positively charged particles (protons).[246] The electrons appear to possess charge as a primitive property. The protons possess charge by virtue of being composed of a multitude of fractionally charged quarks (the sum of these charges being equal to a charge of +1). The quarks appear to possess charge as a primitive property.

---

245 "Causality and explanation: a reply to two critiques", p462 and p468

246 Some other charged particles may contribute to the overall charge of the ion, but the same considerations applicable to protons and electrons will also apply to these particles.

So if possession or transmission of a conserved quantity by an object is the mark of a causal process, and common-sense objects do not possess conserved quantities in their own right but rather by virtue of being composed of quantum-scale entities possessing those quantities, it follows that common-sense objects should not be treated as processes in their own right, but rather as complex networks of quantum-scale processes.

As we indicated in the previous chapter, Salmon is much more sympathetic to the view that "common sense" processes are actually networks of distinct but connected processes,[247] although his concern seems to be based on the requirement in $Salmon_2$ that processes only extend from interaction to interaction, and common-sense processes are certainly involved in a great many internal and external interactions. As we noted, Salmon thinks there is a case for considering common sense objects "as if" they were processes in their our right: the pragmatic considerations governing explanation in some contexts make it considerably easier to conduct practical investigations of the world if we treat common sense processes as if they were single processes. But this defence won't do. As per $Salmon_1$, $Salmon_2$ is supposed to give an objective physical basis to a theory of scientific explanation. If the objective structures appealed to turn out to be artifacts of explanatory practice, then the theory of explanation is circular. At any rate, *ease* of explanatory practise ought not govern the physical scale at which processes exist. Ease of application of the process application is an epistemological issue separate from the ontological issue of the scale at which processes genuinely exist.

Secondly, we question whether the move to associate causation with conserved quantities constitutes any improvement over the association of causal dependence with quantities of energy/momentum of an Aronson-Fair analysis which respects the Heathcote conjecture. Indeed, we think the role of conserved quantities is misconceived in the conserved quantity analyses.

Dowe's argument that CQ is more generally applicable than a transference analysis because some examples of causation can be more *practically identified* by reference to conserved quantities such as charge rather than energy/momentum constitutes another confusion of the epistemology of causation with the ontology of causation. The point of moving to a process

[247] "Causality and explanation: a reply to two critiques", p464

analysis from an Aronson-Fair analysis is that the Aronson-Fair analysis fails to provide a satisfactory ontological account of processes. So the conserved quantity theory is supposed to be an ontological improvement. Similarly, what would amount to a satisfactory improvement of generality of application over the Aronson-Fair analysis is an increase in ontological applicability. To whit: certain processes which are not causal by virtue of possessing or transmitting energy/momentum but which *ought* to be causal processes would be causal by virtue of possessing or transmitting some other conserved quantity, and certain intersections of processes which are not causal by virtue of involving an exchange of energy/momentum but *ought* to be causal are causal by virtue of involving an exchange of some other conserved quantity.

In fact, we find that ontological generality speaks in favour of limiting our attention to energy/momentum. All quantum-scale entities which possess some other conserved quantity (for example, electric charge) also possess energy/momentum. But not all quantum-scale entities which possess energy/momentum possess charge. The proton, the W bosons, the Z^0 and the neutrino all possess energy/momentum. The proton and the W bosons are charged, and the Z^0 and the neutrino are not charged.

Similarly, all causal interactions under CQ and $Salmon_2$ which involve the exchange of the other conserved quantities also involve energy/momentum exchanges, but few (if any) of the other conserved quantities are always involved in causal interactions.

With respect to electric charge, consider the scattering of an electron from a neutrino (an instance of the weak interaction), which takes place by virtue of exchange of a virtual Z^0 between the electron and the neutrino ("exchange" being used in the sense of the transference analysis, or the sense used in quantum physics, that of the electron and the neutrino being connected by the virtual Z^0).[248] This involves *two* causal interactions under $Salmon_2$ or CQ, one located at the coupling of the electron with the Z^0 and one at the coupling of the neutrino with the Z^0. A change in value of charge takes place at the coupling of the Z^0 and the electron (since the electron has charge -1 and the Z^0 has no charge). Yet no such change in value of charge occurs at the coupling of the Z^0 and the neutrino (since neither the Z^0 nor the neutrino have any charge). But

248 *Story of the W and Z*, p49

changes in the value of energy/momentum possessed by each particle occur at each coupling.

So the move to associate causation with possession or transmission of conserved quantities in general, or with changes in value of conserved quantities in general constitutes no advance in ontological generality from an analysis which just focuses on energy/momentum. This suggests that conserved quantities should play some other role in an analysis of causal dependence (or perhaps in some other part of an account of physical dependence). We suspect that they are most at home in an account of *propensities*.

Some fundamental *physical* interactions occur in this universe, some do not. Protons interact with electrons via the electromagnetic interaction, but electrons do no interact with neutrinos via the electromagnetic interaction. But electrons *do* interact with neutrinos via the *weak* interaction. Is this just a de facto happenstance? Just an incidental fact about which couplings of particles have occurred up to now? No, it seems that the physical interactions that quantum-scale entities can be involved in are governed by the possession of particular conserved quantities, such as charge. The proton and the electron both possess electrical charge, so they can be connected by an exchange (in the physicists' sense) of a virtual photon. By virtue of possessing charge, the particles have a *non-zero propensity* to be subject to the electromagnetic interaction. The neutrino does not possess electrical charge, so it is never connected to a charged particle by exchange of a photon. By virtue of not possessing electrical charge, neutrinos have a propensity of zero to be subject to the electromagnetic interaction. But neutrinos possess other conserved quantities, such as weak hypercharge. Electrons also possess weak hypercharge. So electrons can be connected to neutrinos by the exchange of a virtual Z^0. By virtue of possessing weak hypercharge, the particles have a non-zero propensity to be subject to the weak interaction.

This alternative account of the role of conserved quantities other than energy/momenta in physical interactions and physical dependence can't be considered to be completely satisfactory as stated. Consider once again the example of "virtual pair production". A individual photon can briefly form a virtual electron-positron pair, which then re-coalesces into a photon. This is an instance of the electromagnetic interaction, so the photon must have a propensity to be subject to the electromagnetic interaction. Yet photons do not possess electrical charge.[249] Yet our account seems to be on a better track than CQ or Salmon$_2$. In the final analysis, if conserved quantities such as *charge* turn out to have anything to do with causal dependence, it seems much more likely that they should form part of the analysis of causal propensities, rather than causal production or propagation.

Thirdly, CQ and Salmon$_2$ follow Salmon$_1$ in locating causal interactions at couplings of processes. As can be noted from our preceding discussion of the role of conserved quantities in the process analysis, this creates a significant tension between our best accounts of physical interaction and the account of causal interaction. Physical interaction is normally associated with the presence of a connecting virtual quanta. But causal interaction is associated with the *couplings* between quanta. We saw that this identification with couplings in Salmon$_1$ hinged on an unsatisfactory interpretation of "information". The identification in CQ and Salmon$_2$ hinges on an unsatisfactory interpretation of "exchange".

Salmon and Dowe can't mean exchange in many ordinary senses of the word. Suppose David and Ed go out for a walk in the rain. They grumble about their respective umbrellas, and decide to swap. So they exchange umbrellas. Ed then walks along with David's umbrella, and David has Ed's. This type of exchange can not be being invoked in, for example, beta decay. Here a neutron emits a virtual W^- and transmutes into a proton (one exchange of the conserved quantity, electric charge). The virtual W^- then decays into an electron and an antineutrino (the second exchange of charge). Salmon and Dowe can't be

---

249 The other conserved quantities such as charge will also play a role in determining the number of "incoming" or "outgoing" processes at couplings. Because charge is conserved a real photon (with zero charge) has a non-zero propensity to convert into a virtual electron-positron pair (since the respective charges of the pair, -1 and +1, sum to zero) but a zero propensity to convert into a single electron or a single positron. Energy/momentum can also play some similar role in the analysis of causal propensities. In the case of the annihilation of a *real* electron-positron pair, the non-zero propensity that either two or three photons are formed and the zero propensity that one photon is formed is a consequence of the conservation of energy/momentum (but this set of propensities do not hold in the case of *virtual* electron/positron annihilation).

A propensity for a particular interaction to occur or not occur can also be generated by possession of a complex set of conserved quantities. A good example is the non-occurrence of weak neutral current decays of "strange" particles such as the K^0. (See *Story of the W and Z*, pp48–51 for a good discussion) literally saying that the electron now has the W^-'s charge. Otherwise they would be falling into the same trap as Aronson and supposing quantitative properties can be *genidentical*. Indeed, Dowe doesn't want to say this, and argues:

> The CQ theory is also noncommittal on the question of genidentity of physical quantities. The CQ theory simply requires that there is an exchange of some conserved quantity, where exchange is to be understood in terms of a change in the value of a quantity. The only sense of "identity" involved here is numerical equality.[250]

The sense of "exchange" appealed to is simply that a kind of cosmic bookkeeping standard is obeyed at the spatiotemporal intersections. The amount of the relevant conserved quantity possessed or transmitted by the "incoming" processes must be numerically *equal* to that of the "outgoing" processes, but a least one value of a conserved quantity possessed or transmitted by a process on each side must change.[251]

If *that* is the case, then we ought to find that the overall energy/momentum value — the quantity most generally ontologically associated with causal interactions — ought to be numerically equal on each side of a coupling. But this is not what our advanced understanding of physical interaction tells us. As we discussed earlier in the context of the transference analysis, production of virtual quanta is associated with *violations* of the conservation of energy/momentum. So the values aren't going to equal out across the couplings. Something is desperately wrong with the identification of causal interactions with "exchanges" in the CQ and $Salmon_2$ sense at couplings at the quantum physical scale.

It could be argued that these problems don't arise at quantum-scale couplings that only involve real processes. But as we saw in our earlier discussion of Compton scattering and matter-antimatter annihilation, it can be surprisingly difficult to find an example of such a coupling which can not be interpreted as involving virtual quanta in some way, which would

---

250 "What's right and what's wrong with transference theories", pp372–3

251 The "change" is important: it distinguishes causal interactions from mere spatiotemporal intersections of processes (for example, a neutrino meeting a photon). In these latter cases the relevant quantities are conserved on each side of the intersection, but no changes in values of quantities occur.

automatically involve a brief violation of the conservation of energy/momentum.

Since the signal, most general feature associated with physical interactions is the violation of the conservation of energy/momentum associated with the existence of virtual quanta, we think it is advisable that causal interactions should be associated with the *physicists'* notion of exchange at the quantum scale — being connected by an intervening virtual quanta. By identifying interactions with the entire virtual quanta, not just the couplings, we are claiming that causal interactions are not a basic causal mechanism *distinct* from causal processes, but rather are a *species* of causal process.

This might seem to vitiate both CQ and $Salmon_2$, but the analyses can be modified to deal with this reinterpretation of causal interactions very easily. The common ground of the analyses, definition 1, is converted into a definition of *couplings*:

> (1) A coupling is an intersection of world lines which involves a violation of the conservation of energy/momentum.

CQ's definition 2 could remain in its present form, although our earlier arguments suggest that "conserved quantity" should probably be replaced by "energy/momentum". The apparent contradiction of suggesting in definition 2 that energy/momentum is conserved but also suggesting in definition 1 that this conservation is violated can be avoided by following Dieks and Dowe in noting that energy/momentum is *globally* conserved. Before and after a physical interaction (before the creation and after the annihilation of virtual quanta), the value of this quantity remains the same.[252] During an interaction (while the virtual quanta are in existence) the conservation of energy/momentum is *locally* violated.

$Salmon_2$'s definition 2 could remain in its present form. Definition 3 could be modified as follows:

> (3) A process transmits a conserved quantity between A and B (A≠B) if and only if it possesses a fixed amount of this quantity at A and at B and at every stage of

---

252 "Physics and the direction of causation, p88; "What's right and what's wrong with transference theories", pp370–1

the process between A and B without any couplings in the open interval (A,B) that involve a violation of the conservation of energy/momentum.

The same caveat that applies to CQ's used of "conserved quantity" in definition 2 would apply to definitions 2 and 3 of $Salmon_2$.[253]

We don't think anything has been lost by our reinterpretation of causal interactions as a species of causal processes — in fact, something has been gained. If the quantum physical scale is the only physical scale at which processes could be claimed to really exist, then the process analysis now reflects the blurring of the boundaries between substances and interactions described by our best current theories of fundamental physics.

If causal interactions are a species of causal process, then the challenges for the conserved quantity analyses resolve to the challenge of providing an adequate account of causal processes. Unfortunately, our fourth objection to the conserved quantity analyses is that they do not provide such an account.

Are there any grounds for preferring the account of processes of CQ over $Salmon_2$ or *vice versa*? Dowe thinks CQ has the advantage, and provides a series of justifications in "Causality and conserved quantities: a reply to Salmon".[254] We think Dowe's most damaging objection to the $Salmon_2$ account relates to the requirement of the at-at theory that processes are spatiotemporally continuous. Dowe does not wish to rule out *a priori* the possibility of discontinuous processes.[255]

At the quantum physical scale, Dowe's objection seems to be well-founded. Consider the absorption and emission of light by an atom. An electron in the "ground state" (lowest energy) orbital around the atom absorbs a photon, and discontinuously jumps to a higher-energy orbital. The electron then emits a photon and discontinuously jumps to a lower-energy orbital. Given that there is

---

253 This alteration means that a corollary Salmon proposes to his definitions of causation no longer holds. The corollary — "When two or more processes possessing a given conserved quantity intersect (whether they interact or not), the amount of that quantity in the region of intersection must equal must equal the sum of the separate quantities possessed by the processes thus intersecting" (Causality and explanation: a reply to two critiques", p473) — is simply false if the relevant intersection is a coupling, since energy/momentum is not *locally* conserved (conserved on each side of a coupling).

254 "Causality and conserved quantities: a reply to Salmon", pp325–33

255 "Causality and conserved quantities: a reply to Salmon", p332

a spatiotemporal gap between the electron's being in the different orbitals in each case, $Salmon_2$ can not treat the motion of the electron as involving the transmission of causal influence. So the absorption of the first photon would not be a *cause* of the emission of the second photon, even though the electron could not have emitted without being moved to a higher level orbital (since an electron will only emit a photon if it jumps to a lower-energy orbital, and the ground state is the lowest-energy orbital). CQ *can* treat the absorption as a cause of the emission so long as the electron which emits the photon is genidentical with the electron which absorbs the photon.

But, of course, while this is a problem for the $Salmon_2$ account of processes, it is a catastrophe for the CQ account. As we discussed in our investigation of the transference analysis, it is nonsense to speak of quantitative properties of quantum-scale objects being genidentical. But it is also the case that *genidentity* fails for quantum-scale objects. If relations of genidentity did apply at the quantum scale, the universe would be observably different to the way we observe it to be. So the electron that emits the photon *can not be one and the same electron as the electron that absorbs the photon.*

This result *generalises.* As we have argued, the only scale at which processes *really* exist under a conserved quantity analysis is the quantum scale. The CQ criterion for being a causal process is based on a genidentity relation between objects. We can observe that genidentity fails at the quantum scale in this universe. So there are no real causal processes in this universe. Since causal interactions are a species of causal processes, there are no causal interactions in this universe. So CQ fails to fulfil our basic desideratum, because under CQ there is *no such thing* as causal dependence in this universe.

$Salmon_2$ is not quite so badly off. No genidentity relation is required by the at-at theory to account for transmission of causal influence by quantum-scale entities. For example, the claim that an electron transmits energy/momentum from one object to another does not presuppose the claim that one-and-the-same electron takes energy from one object and delivers energy to another object. Rather roughly speaking, all that is required is that at every point of a spatiotemporal trajectory between the two objects there is electron with such-and-such energy/momentum. This is in reasonable accordance with the practise in physics of defining a "current" of elementary particles: a rate of change of density of some particles (or property of particles) across a set of spatiotemporal points.

Yet there are still problems with the $Salmon_2$ conception of processes. Throughout the development of $Salmon_1$ and $Salmon_2$, we have been lead to understand processes in terms of macroscopic, significantly spatiotemporally extended objects, such as billiard balls or Frank. We understand interactions in terms of macroscopic phenomena such as collisions. When we examine the analyses closely, we conclude that these things aren't really "processes" or "interactions" — it is merely convenient for explanatory purposes to treat them as if they were processes and interactions. Actual processes and interactions would exist down at the quantum scale. We look down at the quantum scale and determine that interactions are best not treated as being specially different causal mechanisms from processes. So our justification for treating collisions *as if* they were actual interactions has disappeared. Collisions and other macroscopic "interactions" are *nothing like* actual interactions. So can we justifiably claim that macroscopic "processes" are anything like actual processes? That would seem to depend on whether there are any significantly spatiotemporally extended processes at the quantum scale.

Salmon claims there are such processes:

> Lest anyone believe that I have completely lost contact with physical reality in discussing extended causal processes without causal interactions, it should be noted that the mean free path of photons in intergalactic space is estimated in terms of very large numbers of light years. Such photons are causal processes, and the mean free path is the average distance between causal interactions with other processes.[256]

But this won't do. Consider first an electron moving aimlessly through intergalactic space. This does not form a single extended process. As we noted earlier, electrons constantly emit and reabsorb virtual photons ("self interaction"). Under $Salmon_2$, every coupling with a virtual photon will produce a new electron process. So we can only treat the electron *as if* it was a single significantly spatiotemporally extended process.

The same can be said of the *photon* moving through intergalactic space. Real photons continually convert into virtual electron-positron pairs which

256 "Causality and explanation: a reply to two critiques", p464

reconvert into single real photons. Every such conversion involves a coupling. so every such conversion involves the creation of a new process under Salmon_2. So we can only treat Salmon's intergalactic photons *as if* they are single significantly spatiotemporally extended processes.

So are *any* "actual" processes significantly spatiotemporally extended? Perhaps neutrinos might fit the bill. But this is by virtue of neutrinos being the least physically active of the known quantum entities, and the least likely to be involved in relations of physical dependence. There is something decidedly odd about treating significantly spatiotemporally extended macroscopic entities "as if" they are actual processes when the only "actual" processes they are *anything* like are our paradigm examples of entities which *do not stand in relations of physical dependence.*

So there is a significant tension in Salmon_2 between what the nature of processes is *supposed* to be and the nature of the *actual* causal mechanisms. Not just "common sense" processes are explanatory artifacts; *all* processes are explanatory artifacts. Given that Salmon_2 is based on the notion of a causal process, it appears that causal dependence in Salmon_2 resolves to explanatory dependence. So Salmon_2 does not fulfil our basic desideratum — it does not give an account of the cement of the universe, it gives an account of the cement of our explanatory structures. To put it another way, causal dependence does not track physical dependence, it tracks explanatory dependence.[257]

Could we escape this result by ditching the problematic notion of "processes"? Perhaps it is best to refer instead to something like a "current". Our picture of the causal structure of the world then becomes a complex network of coupled currents of quantum-scale real and virtual entities, where couplings occur because particular properties of some currents give those currents a non-zero propensity to couple with other currents. But this is no longer an analysis of *causal* dependence. It is a very brief, non-mathematical statement of our best theories of *physical* dependence.

We have fought our way to an analogue of Quine's criticism of the transference analysis: *if* we must talk of causation, best to talk of it in terms of

257 In which case, Salmon_2 does not fulfil Salmon's basic desideratum either. The process analysis is supposed to provide an objective, physical basis for scientific explanation. If processes are explanatory structures, Salmon's project is just circular.

energy transfers; but in the final analysis it is better to just talk of energy transfers. It seems that *if* we must talk of causal dependence, best to talk of it in terms of the processes of the conserved quantity analyses. But in the final analysis it is best to just talk of physical dependence.[258]

So we conclude that the conserved quantity analyses fail to fulfil our basic desideratum, that causal dependence should track physical dependence in this universe. Under our best understanding of CQ, we can observe that there are no actual causal processes in this universe, so there is no causal dependence in this universe. Under our best understanding of $Salmon_2$ *either* causal dependence does not track physical dependence but instead tracks explanatory dependence *or* $Salmon_2$ is not an analysis of causal dependence, but merely a brief, non-mathematical restatement of our best theoretical understanding of physical dependence.

258 And if $Salmon_2$ is just a brief summary of our best theories of physical dependence, $Salmon_2$ *still* doesn't fulfil Salmon's basic desideratum. The point of a theory of scientific explanation is to justify our best theories of physical dependence. If the basis of the theory of explanation is a brief statement of our best theories of physical dependence, then Salmon's project is still just circular.

Conclusion

We were looking for an analysis of causation in which causal dependence tracks physical dependence. None of the analyses we investigated fulfilled this basic desideratum.

We began by investigating a basic version of the current "orthodox" analysis of causation, the Lewis-style counterfactual analysis. The strange case of the overdetermining dumbbell shows that Lewis-style analyses do not track physical dependence. Rather, the analyses are sensitive to common-sense individuation of objects within events. The case of the bonking of Frank shows that if we try to solve this problem by keying object individuation to physical dependence, the "sophisticated" analysis produced is even less successful at fulfilling our basic desideratum than the "naive" analysis.

We turned our attention to the modern transference analysis of causation developed by Aronson, Fair and Heathcote, and constructed a best possible version of the analysis (an Aronson-Fair analysis which respects the Heathcote conjecture). The case of the bouncing of Frank demonstrates that even this version of the transference analysis fails to fulfil our basic desideratum, as the analysis does not give an adequate account of causal connectivity. Attempting to fix this flaw by asserting the genidentity of quantities of energy/momentum also fails because relations of genidentity do not apply to quantitative properties of objects, and genidentity fails when applied to quantum-scale objects (the scale at which energy/momentum exchanges actually take place in this universe).

In order to try and solve the problem of causal connectivity, we turned our attention to the process analyses of causation developed by Salmon and Dowe. We began with a semi-historical exposition of the development of $Salmon_1$. $Salmon_1$ fails due to its reliance on an unsatisfactory notion of information transfer and transmission, and the lack of a satisfactory account of causal interactions. Indeed, the mature form of $Salmon_1$ either resolves to a "sophisticated" Lewis-style counterfactual analysis or is at least subject to the same problems with fulfilling our basic desideratum as the "sophisticated" Lewis-style analysis.

We finally looked briefly at the two "conserved quantity" process analyses, CQ and $Salmon_2$. Applying a number of our earlier results to these analyses shows that these analyses also fail to fulfil our basic desideratum. Scale invariance fails for these analyses. So the only "actual" processes under these analyses exist at the quantum scale. So CQ fails because processes are defined by a genidentity relations across objects and genidentity fails at the quantum scale. The best reading of $Salmon_2$ fails because either causal dependence under $Salmon_2$ tracks explanatory dependence rather than physical dependence or $Salmon_2$ is not an analysis of causal dependence but merely a brief non-mathematical restatement of our best theoretical understanding of physical dependence.

So none of the analyses we have investigated fulfil our basic desideratum. Should we therefore embrace the Russell-Quine thesis, and call for the extirpation of causal talk from philosophical discourse?

Hardly. Our results only speak against analyses of causation in which causal dependence is *supposed* to track physical dependence. An analysis which explicitly does not seek to fulfil our desideratum, such as the manipulability analysis, is completely immune to our objections. Nor have we shown that *every possible* counterfactual, transference or process analysis fails to fulfil our desideratum. Lewis-style analyses are not the only counterfactual analyses of causation in town. Perhaps a satisfactory solution to the quantum measurement problem may be developed, in which case an Aronson-Fair analysis which respects the Heathcote conjecture might just be able to avoid the causal connectivity problem. A satisfactory account of processes might rehabilitate the process analysis. Or some completely different approach to an analysis of causation might do the trick.

But our results are still of some concern. The requirement that causal dependence tracks physical dependence seems so innocent, yet we have demonstrated that it is very hard for an analysis to fulfil this requirement. The casual contemporary practise of invoking causal dependence in philosophical theories *as if* causal dependence tracked physical dependence has to be called into question.

The following programme in the spirit of the Russell-Quine thesis suggests itself to us, and so we suggest it to philosophers in general. When a philosophical theory invokes causal dependence and assumes that causal dependence tracks physical dependence, ask yourselves two questions. Is an account of physical dependence all that is required for the theory to work? If so, replace all mention of causal dependence with the best appropriate accounts of physical dependence. Does the theory rely on features of a particular analysis of causation in order to work? If so, establish that causal dependence tracks physical dependence under that analysis. If neither project can be successfully carried out, commit the theory to the flames for it can contain nothing but sophistry and illusion.

Bibliography

Aronson, Jerrold: "Explanations without laws", pp541–57, *Journal of philosophy* vol 66 (1969)

Aronson, Jerrold: "On the grammar of cause", pp414–30 *Synthese* vol 22 (1971)

Aronson, Jerrold: "The legacy of Hume's analysis of causation", pp135–56, *Studies in history and philosophy of science* vol 2 (1971)

Aronson, Jerrold: "Untangling ontology from epistemology in causation", pp293–305 *Erkenntnis*, vol 18 (1982)

Aronson, Jerrold: *A realist philosophy of science* (New York: St Martin's Press, 1984)

Aronson, Jerrold: "Conditions versus transference: a reply to Ehring", pp249–57 *Synthese* vol 63 (1985)

Bain, Alexander: *Logic* 2nd edition (London: Longmans, Green & Co, 1895)

Bennett, Jonathan: *Events and their names* (Indianapolis: Hackett Publishing Company, 1988)

Davidson, Donald: "Reply to Quine on events", pp172–6, LePore, Ernest; McLaughlin, B; editors: *Actions and events: perspectives on the philosophy of Donald Davidson* (Oxford: Basil Blackwell, 1985)

Davies, Paul: *The forces of nature* (Cambridge, UK: Cambridge University Press, 1979)

Dieks, D: "Physics and the direction of causation", pp85–110, *Erkenntnis* vol 25 (1986)

Dowe, Phil: "Process causality and asymmetry", pp179–96, *Erkenntnis* vol 37 (1992)

Dowe, Phil: "Wesley Salmon's process theory of causality and the conserved quantity theory", pp195–216, *Philosophy of science* vol 59 (1992)

Dowe, Phil: "Causality and conserved quantities: A reply to Salmon", pp321–33, *Philosophy of science* vol 62 (1995)

Dowe, Phil: "What's right and what's wrong with transference theories", pp363–74, *Erkenntnis* vol 42 (1995)

Dowe, Phil: "Causation, causal processes", *Stanford encyclopedia of philosophy* (http://plato.stanford.edu/entries/causation-process/ last modified 9 December 1996)

Ducasse, Curt J: *Nature, mind and death* (La Salle: The Open Court Publishing Company, 1951)

Ducasse, Curt J: *Causation and the types of necessity* 2nd edition (New York: Dover Publications, 1969)

Ducasse, Curt J: "On the nature and the observability of the causal relation", pp125–36, Sosa, Ernest; Tooley, Michael; editors: *Causation* (Oxford: Oxford University Press, 1993)

Earman, John: "Causation: a matter of life or death", pp5–25, *Journal of philosophy* vol 73 (1976)

Effinger, George Alec: *The nick of time* (Sevenoaks: New English Library, 1987)

Ehring, Douglas: "The transference theory of causation", pp249–58, *Synthese* vol 67 (1986)

Ehring, Douglas: *Causation and persistence: a theory of causation* (New York: Oxford University Press, 1997)

Fair, David: "Causation and the flow of energy", pp219–50, *Erkenntnis* vol 14 (1979)

Feynman, Richard: *QED: the strange theory of light and matter* (Princeton: Princeton University Press, 1985)

Gasking, Douglas: "Causation and recipes", pp479–87, *Mind* vol 64 (1955)

Goldsmith, Werner: *Impact: the theory and behaviour of colliding solids* (London: Edward Arnold, 1960)

Mill, John Stuart: *A system of logic ratiocinative and inductive,* Robson, J M; editor: *The collected works of John Stuart Mill* volumes VII–VIII (Toronto: University of Toronto Press, 1974)

Pagels, Heinz: *The cosmic code: quantum physics as the language of nature* (Harmondsworth: Penguin, 1984)

Putnam, Hilary: "Is water necessarily H_2O?", pp54–79, Putnam, Hilary: *Realism with a human face* (Cambridge, MA: Harvard University Press, 1990)

Quine, Willard Van Orman: *The roots of reference* (La Salle: Open Court, 1973)

Quine, Willard Van Orman: "Natural kinds", pp159–70, Boyd, Richard; Gasper, Philip; Trout, J D; editors: *The philosophy of science* (Cambridge, MA: MIT Press, 1991)

Quinton, Anthony: "Objects and events", pp197–214, *Mind* vol 88 (1979)

Reichenbach, Hans: *The direction of time* (Berkeley: University of California Press, 1956)

Rosenburg, Z: "The dynamic response of ceramics to shock wave loading", pp73–105, Brebbia, C A; Sanchez-Galvez, V; editors: *Shock and impact on structures* (Southhampton: Computational Mechanics Publications, 1994)

Russell, Bertrand: "On the notion of cause", pp1–26, *Proceedings of the Aristotelian Society* vol 13 (1912)

Salmon, Wesley: *Statistical explanation and statistical relevance* (Pittsburgh: University of Pittsburgh Press, 1971)

Salmon, Wesley: *Scientific explanation and the causal structure of the world* (Princeton: Princeton University Press, 1984)

Salmon, Wesley: "Dynamic rationality: propensity, probability, and credence", pp3–40, Fetzer, James; editor: *Probability and causality: essays in honor of Wesley C Salmon* (Dordrecht: D Reidel Publishing Company, 1988)

Salmon, Wesley: "A new look at causality", pp13–24, Salmon, Wesley: *Causality and explanation* (New York: Oxford University Press, 1998)

Salmon, Wesley: "An 'at-at' theory of causal influence", pp193–9, Salmon, Wesley: *Causality and explanation* (New York: Oxford University Press, 1998)

Salmon, Wesley: "Causal and theoretical explanation", pp108–24, Salmon, Wesley: *Causality and explanation* (New York: Oxford University Press, 1998)

Salmon, Wesley: "Causality: production and propagation", pp285–301, Salmon, Wesley: *Causality and explanation* (New York: Oxford University Press, 1998)

Salmon, Wesley: "Causality without counterfactuals", pp248–60, Salmon, Wesley: *Causality and explanation* (New York: Oxford University Press, 1998)

Salmon, Wesley: "Causal propensities: statistical causality versus aleatory causality", pp200–7, Salmon, Wesley: *Causality and explanation* (New York: Oxford University Press, 1998)

Salmon, Wesley: "Explanatory asymmetry: a letter to Professor Adolf Grünbaum from his friend and colleague", pp164–77, Salmon, Wesley: *Causality and explanation* (New York: Oxford University Press, 1998)

Salmon, Wesley: "Probabilistic causality", pp208–32, Salmon, Wesley: *Causality and explanation* (New York: Oxford University Press, 1998)

Salmon, Wesley: "Why ask 'why'?", pp125–41, Salmon, Wesley: *Causality and explanation* (New York: Oxford University Press, 1998)

Salmon, Wesley: "Causality and explanation: a reply to two critiques", pp461–77, *Philosophy of science* vol 64 (1998)

Salmon, Wesley; Kitcher, Philip: "Van Fraassen on explanation", pp178–90, Salmon, Wesley: *Causality and explanation* (New York: Oxford University Press, 1998)

Salmon, Wesley; Salmon, Merrilee: "Alternative models of scientific explanation", pp333–46, Salmon, Wesley: *Causality and explanation* (New York: Oxford University Press, 1998)

Tait, P G: "The unwritten chapter on golf", pp80–1, Gratzer, Walter, editor: *A bedside* Nature (London: Macmillan Magazines Ltd, 1996)

van Fraassen, Bas: "The pragmatics of explanation", pp143–50, *American philosophical quarterly* vol 14 (1977)

von Bretzel, Philip: "Concerning a probabilistic theory of causation adequate for the causal theory of time", pp385–402, Salmon, Wesley; editor: *Hans Reichenbach: logical empiricist* (Dordrecht: D Reidel Publishing Company, 1979).

Watkins, Peter: *Story of the W and Z* (Cambridge, UK: Cambridge University Press, 1986)

Wesstein, Eric: "Roche limit" (http://www.astro.virginia.edu/~eww6n/physics/RocheLimit.html 1996–8)

Zukas, Jonas; Nicholas, Theodore; Swift, Hallock; Greszczuk, Longin; Curran, Donald: *Impact dynamics* (New York: John Wiley & Sons, 1982)